A Complete Structural Description of the Catalytic Cycle of Yeast Pyrophosphatase[†,‡]

Esko Oksanen,[§] Anna-Karoliina Ahonen,[§,||] Heidi Tuominen,[⊥] Vesa Tuominen,[§] Reijo Lahti,[⊥] Adrian Goldman,*[,§] and Pirkko Heikinheimo[§,||]

Structural Biology and Biophysics, Institute of Biotechnology, P.O. Box 65, University of Helsinki, FIN-00014 Helsinki, Finland, Department of Chemistry, University of Tromsø, N-9037 Tromsø, Norway, and Department of Biochemistry and Food Chemistry, University of Turku, FIN-20014 Turku, Finland

Received September 26, 2006; Revised Manuscript Received November 22, 2006

ABSTRACT: We have determined the structures of the wild type and seven active site variants of yeast inorganic pyrophosphatase (PPase) in the presence of Mg^{2+} and phosphate, providing the first complete structural description of its catalytic cycle. PPases catalyze the hydrolysis of pyrophosphate and require four divalent metal cations for catalysis; magnesium provides the highest activity. The crystal form chosen contains two monomers in the asymmetric unit, corresponding to distinct catalytic intermediates. In the "closed" wild-type active site, one of the two product phosphates has already dissociated, while the D115E variant "open" conformation is of the hitherto unobserved two-phosphate and two-"bridging" water active site. The mutations affect metal binding and the hydrogen bonding network in the active site, allowing us to explain the effects of mutations. For instance, in Y93F, F93 binds in a cryptic hydrophobic pocket in the absence of substrate, preserving hydrogen bonding in the active site and leading to relatively small changes in solution properties. This is not true in the presence of substrate, when F93 is forced back into the active site. Such subtle changes underline how low the energy barriers are between thermodynamically favorable conformations of the enzyme. The structures also allow us to associate metal binding constants to specific sites. Finally, the wild type and the D152E variant contain a phosphate ion adjacent to the active site, showing for the first time how product is released through a channel of flexible cationic side chains.

Soluble inorganic pyrophosphatases (PPases,[1] EC 3.6.1.1) catalyze one of the simplest phosphoryl transfer reactions, the hydrolysis of the pyrophosphate bond. Two structural families encode PPases (family I and family II PPases). The family I PPases are the most common; in this family, *Saccharomyces cerevisiae* (yeast) (Y-PPase) and the hexameric bacterial enzyme from *Escherichia coli* (E-PPase) are the best-studied. The catalysis of these enzymes is well-characterized (*1*−*6*), and structures exist along the reaction pathway (*1*, *7*, *8*). PPase catalysis requires three or four metal ions (*9*), depending on pH and PP_i concentration. In the yeast cell, the PP_i concentration and pH are such that the three-metal ion mechanism predominates (*9*), and the hydrolysis of the metal-complexed pyrophosphate is accelerated by a factor of 10^{10} in comparison to the uncatalyzed reaction (*10*). The catalysis is based on leaving group activation and effective nucleophile generation (*1*, *7*).

Six separate steps have been identified by kinetic analysis as shown in Scheme 1. The enzyme, which is activated by two metal ions (Scheme 1, A), binds one or two additional metal ions with the substrate (Scheme 1, B). The metal positions in the active site are termed M1−M4 in the proposed order of metal binding, where M1 and M2 are the activating metal ions and M3 and M4 bind with substrate (*7*) (Scheme 1). The number of metal ions involved depends on the pH, but the four-metal ion mechanism dominates at the crystallization pH (6.0). The enzyme undergoes an isomerization (Scheme 1, B → C), which is supported by both the enzymological (*4*, *11*) and structural data (*1*). The generation of the reactive nucleophile O_{Nu} between M1 and M2 is aided by a low-barrier hydrogen bond between D117 and the nucleophile (*1*). Since the protonation state of the nucleophile is not directly observed, we denote it O_{Nu}, as in ref *1*. After the hydrolysis step (Scheme 1, D), the two phosphate−metal complexes are released one by one (Scheme 1, E and F) and the enzyme returns to the two-metal ion form (Scheme 1, A).

[†] This work was supported by the Finnish Academy of Sciences (Grant 201611 to R.L.) and the ISB graduate school (to E.O.). A.G. and P.H. are members of Biocentrum Helsinki, and this work was supported by BCH. Data were collected at SNBL (BM01) at the European Synchrotron Radiation Facility (ESRF).

[‡] The coordinates and stucture factors have been deposited in the Protein Data Bank as entries 2IHP (wild type), 2IK0 (E48D), 2IK1 (Y93F), 2IK2 (D115E), 2IK4 (D117E), 2IK6 (D120E), 2IK7 (D120N), and 2IK9 (D152E).

* To whom correspondence should be addressed. Telephone: (358) 9 19158923. Fax: (358) 9 19159940. E-mail: adrian.goldman@helsinki.fi.

[§] University of Helsinki.

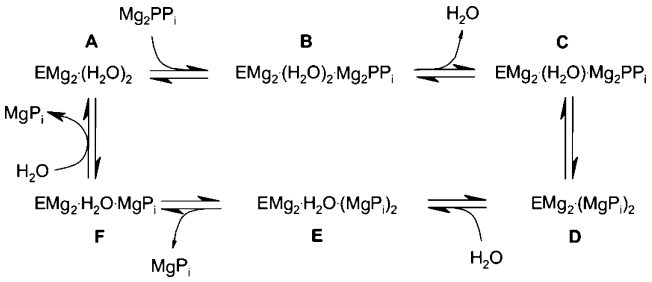
[||] University of Tromsø.

[⊥] University of Turku.

[1] Abbreviations: PPase, inorganic pyrophosphatase; Y-PPase, *Saccharomyces cerevisiae* inorganic pyrophosphatase; E-PPase, *Escherichia coli* inorganic pyrophosphatase; PP_i, inorganic pyrophosphate; P_i, inorganic phosphate; *wt*, wild type; RDS, rate-determining step; Tris, tris(hydroxymethyl)aminomethane; MPD, 2-methyl-2,4-pentanediol; MES, 2-morpholinoethanesulfonic acid; ESRF, European Synchrotron Radiation Facility; rmsd, root-mean-square deviation; LBHB, low-barrier hydrogen bond.

10.1021/bi0619977 CCC: $37.00
Published on Web 01/12/2007

Scheme 1 : Six-State Mechanism of Yeast Inorganic Pyrophosphatase



The rate-determining step (RDS) for the forward reaction with Mg^{2+} is hydrolysis of PP_i, whereas in the reverse, net synthesis direction, the RDS is not the condensation of the two P_i ions but the previous step (Scheme 1, C → B), which involves isomerization of the enzyme (*12*). Isomerization decreases the pK_a of the nucleophile from >7 to 5.85 (*11*), which is structurally explained by the two-water bridge collapsing into a single nucleophilic water between the M1 and M2 metal ions. Of the two phosphate moieties in PP_i, we define P2 as the electrophilic phosphate and P1 as the leaving group. The relative affinities of the P1 and P2 sites depend on the identity of the metal ion. In the presence of Mg^{2+}, P1 binds with higher affinity (*13*, *14*) and is released last, but the reverse is true (*5*) with Mn^{2+}.

In the earlier Y-PPase structures, the complexed metal ion is the less active (*1*, *7*) Mn^{2+} or Co^{2+} (I. P. Kuranova et al., unpublished PDB entry 1M38), each of which binds the phosphates with higher affinity (*14*) and affects the order of phosphate release (*13*, *15*). Nonetheless, we were able to describe structures corresponding to intermediates A and C−E in the proposed kinetic scheme (Scheme 1) (*1*). Intermediate A is represented by the Y-PPase·Mn_2 complex determined at 2.2 Å, with metal sites M1 and M2 filled [PDB entry 1WGI (*7*)]. The fluoride-inhibited enzyme at 1.9 Å resolution [PDB entry 1E6A (*1*)] contains four metal ions, unhydrolyzed PP_i, and a single fluoride ion replacing the nucleophile (intermediate C in Scheme 1). The product complex at 1.15 Å resolution [PDB entry 1E9G (*1*)] contains two conformations of the hydrolyzed product, which correspond to intermediates D and E (the latter being a hydrated rearrangement of D). B and F have not been seen.

We have mutated all the active site residues coordinating the substrate electrophilic group P2 or the metal ions that generate the nucleophile (*16*). The mutated enzymes have been previously functionally characterized (*4*, *16*, *17*), and the structure of the D117E product complex with Mn^{2+} has also been determined (*18*). We have now determined the structures of the wild type and all these mutants in the presence of Mg^{2+}. The differences between the Mg^{2+}- and Mn^{2+}-containing structures are very small, demonstrating the relevance of our earlier structures. However, the new structures in addition explore the ensemble of low-energy conformations accessible to the active site residues and demonstrate the existence of various enzymatic hydration states not yet directly observed by solution experiments. The current set of structures now allows us to describe all the catalytic states in Scheme 1 (A−F). We have used these structures to devise the first full structural model of PPase catalysis, including nucleophile generation and product release pathways, and alternate hydration intermediates.

EXPERIMENTAL PROCEDURES

The proteins were expressed and purified as described in ref *16*. The proteins were stored in 50% glycerol, 5 mM Mg^{2+}, and 75 mM Tris buffer (pH 7.2) at −20 °C, and prior to crystallization, the glycerol was washed away by once being diluted to 1:20 and once to 1:10 (1:200 together) with 60 mM MES (pH 6.0) and 10 mM Mg^{2+}. Proteins were crystallized at 4 °C in 8 μL sitting drops and in the presence of 5 mM Mg^{2+} and 1 mM PO_4^{3-} and a MPD concentration gradient from 16 to 19%. Thin plates formed within 2−4 weeks. The crystals were soaked in the cold room (4 °C) with 32% MPD, 30 mM MES, 10 mM Mg^{2+}, and 1 mM PO_4^{3-} a few minutes prior to being dipped into liquid nitrogen. Once aligned with the cryostream, the crystals were annealed once by blocking the cryostream to thaw the crystal and then recooling.

The data for the wild-type crystal were collected at Swiss Norwegian Beamline BM01 at the ESRF using a MAR image plate system (Mar Research). Data for E48D and D115E were collected on a Rigaku R3HU generator running at 50 kV and 100 mA and an R-AXIS IV^{++} image plate detector. The other data were collected on a Nonius generator using Cu Kα radiation at 40 kV and 170 mA on an R-AXIS IV detector. All data were integrated with Denzo (*19*) and scaled together with Scalepack (*19*) (DE48 and DE115) or Scala (*20*). Data collection and processing statistics are summarized in Table 1.

The structures were determined by molecular replacement with Molrep (*21*), using PDB 1HUK (Swaminathan et al., unpublished results) as a model. We also switched the nomenclature of the chains so that chain A belongs to the enzymatic state before catalysis and chain B to the product-containing form. The density was fitted with automated procedures using ARP/wARP (*22*) and model improvement in CCP4 (*20*). After this, the side chains with the zero-occupancy atoms were manually fit to the map in O (*23*), the active site area was inspected for the presence of the possible ligands, and all water residues with high *B*-factors were omitted, using a cutoff of 35 Å^2 for structures with a resolution 1.7 Å or better and 40 Å^2 for the lower-resolution structures. The refinement was continued in Refmac5 (*20*, *24*) through the ccp4i interface from CCP4 4.2.2, followed by a few cycles of solvent refinement in ARP/wARP and manual inspection of the map in O (*23*) or Coot (*25*). The quality of the models is described in Table 2.

The coordinates for the animation in the Supporting Information were prepared by linear interpolation between the structures listed in Figure 9 using LSQMAN (*26*). Pyrophosphate was modeled in the $D115E_A$ structure using Coot. The animation and other figures were made with PyMol (DeLano Scientific LLC).

RESULTS AND DISCUSSION

All of the eight crystal structures we report here (Table 1) are in space group $P2_1$ at resolutions between 1.5 and 1.9 Å. Because the concentration of phosphate and metal ion was insufficient to occupy both active sites fully during crystallization, the two monomers in the dimeric PPase adopt different conformations, and the active sites have different contents. One difference between the current crystal form and the $P2_12_12_1$ forms studied earlier (*1*, *7*) is that both open

Table 1: Data Collection and Processing Statistics

	YWT	E48D	Y93F	D115E	D117E	D120E	D120N	D152E
space group	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$
unit cell	$a = 51.92$ Å	$a = 51.71$ Å	$a = 51.77$ Å	$a = 51.58$ Å	$a = 51.76$ Å	$a = 51.65$ Å	$a = 51.65$ Å	$a = 51.77$ Å
$b = 93.55$ Å	$b = 93.20$ Å	$b = 93.54$ Å	$b = 93.37$ Å	$b = 93.23$ Å	$b = 93.19$ Å	$b = 93.00$ Å	$b = 93.22$ Å	
$c = 69.17$ Å	$c = 69.29$ Å	$c = 69.44$ Å	$c = 69.78$ Å	$c = 69.22$ Å	$c = 69.38$ Å	$c = 69.27$ Å	$c = 70.11$ Å	
$\beta = 99.66°$	$\beta = 99.47°$	$\beta = 99.88°$	$\beta = 99.86°$	$\beta = 99.58°$	$\beta = 99.62°$	$\beta = 99.33°$	$\beta = 100.01°$	
wavelength (Å)	0.8727	1.5418	1.5418	1.5418	1.5418	1.5418	1.5418	0.8000
resolution (Å)[a]	100−1.5 (1.53−1.50)	50−1.7 (1.76−1.70)	100−1.7 (1.74−1.70)	50−1.8 (1.86−1.80)	100−1.8 (1.85−1.80)	100−1.8 (1.85−1.80)	100−1.9 (1.94−1.90)	100−1.5 (1.55−1.50)
no. of unique reflections[a]	103964 (6946)	68800 (3877)	67570 (4572)	58354 (3832)	59558 (4078)	59922 (4148)	50505 (2671)	102428 (9238)
redundancy	7.4	3.5	4.1	3.2	3.5	4.1	3.5	8.3
completeness (%)[a]	99.8 (100.0)	96.8 (84.0)	94.4 (94.4)	96.8 (93.6)	99.3 (98.1)	99.9 (99.9)	99.3 (96.6)	97.7 (88.6)
R_{sym} (%)[a,b]	4.1 (32.1)	6.7 (25.3)	6.5 (34.6)	6.7 (34.4)	5.6 (32.3)	5.9 (38.8)	9.1 (54.2)	3.5 (15.4)
I/σ[a]	21.7(3.2)	21.8 (4.4)	7.9 (2.2)	18.5 (2.8)	9.8 (2.4)	9.5 (1.9)	7.7 (1.4)	33.2 (6.5)

[a] Values for the highest-resolution shell are given in parentheses. [b] $R_{sym} = \sum_{hkl}|I(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}I(hkl)$, where $\langle I(hkl)\rangle$ is the mean of the symmetry equivalent reflections of $I(hkl)$.

and closed forms are visible in a single structure. The monomer A active site is open; the bottom of the active site is occupied by two metal ions, M1 and M2, and the conformation of the protein allows the formation of a "two-water" bridge between them. This form of the enzyme has previously been described with Mn^{2+} in an orthorhombic crystal form of *wt*-Y-PPase (*7*) (A in Scheme 1). In the form described here, however, crystal contacts do not allow an open conformation for monomer B and its closed structure is very similar to that of both monomers in the second orthorhombic crystal form we (*1*, *18*) and others (*27*) described earlier. In the closed monomer, the loop of residues 103−115 is ordered and E109 forms an ion pair with K255 of a crystallographically related molecule; the loop is disordered in the open form. The active site of monomer B thus has to be in a closed conformation, but the conformation of monomer A appears to be determined by the concentrations of the metals and phosphate, which typically, but not always, led to open conformations. This, however, depended on the exact growth and soaking conditions. The active site conformations are (depending on the PPase variant) structurally equivalent to various intermediates along the reaction pathway.

Neither conformation is an artifact of crystal contacts. The monoclinic crystal form grows, however, at metal and phosphate concentrations such that both open and closed conformations are in equilibrium in the mother liquor. If either the open or closed form is dominant in solution, either the open or closed orthorhombic crystal forms will grow.

In what follows, monomer A describes open active site structures and monomer B closed ones. The monomers are denoted with a subscript following the variant (e.g., $E48D_A$) or residue number (e.g., $E48_B$ in monomer B).

The seven point mutations that were introduced are all of residues around the electrophilic phosphorus P2 and at the "bottom of the active site": the binding sites for M1, M2, and P2 and the nucleophilic water/hydroxide ion. The residues changed are Y93 in the P2 site, the metal-coordinating E48, D115, D120, and D152 residues, and the nucleophile-orienting/activating residue, D117. All the mutations are conservative in nature, so the mutated amino acid is chemically as close to the original as possible, except for D120, where structures of both E120 and N120 have been determined.

All crystals were cocrystallised with 5 mM Mg^{2+} and 1 mM P_i and soaked with 10 mM Mg^{2+} and 1 mM P_i prior to being frozen. The structures refined well (as described in Materials and Methods) with good stereochemistry and low *B*-factors (Table 2). The electron density maps were good, allowing Mg^{2+} ions to be identified by their geometry (Figure 1). Each active site contains one to four Mg^{2+} ions and zero to two phosphate ions (Table 3) depending on the mutation and possibly the exact soaking time.

Overall Structures. We analyzed the differences in the overall structures of the mutated proteins by superimposing the monomers on the corresponding chain of the wild-type enzyme. In most structures, the agreement with the wild type is good at the backbone level and the overall root-mean-square difference (rmsd) between the structures is from 0.1 to 0.28 Å on 218−263 superimposed C_α atoms. The few exceptions are predominantly in open monomer A. In D115E, D120E, D120N, and D152E, the monomer A backbone structure differs from that of the *wt* so that 87, 88, 178, and 141 C_α atoms, respectively, fall outside the 0.2 Å cutoff and the overall fit is 0.4−0.78 Å for all residues. The worst fit is for D120N, where the level of fit is similar to the fit between the *wt* open and closed monomers. In general, the monomer B conformation is more constrained and all the mutated structures adopt the same conformation as *wt* except for $D152E_B$, where the overall fit is 0.41 Å. All the other monomers fall within 0.18 Å of wt_B.

Conversely, the differences between the A (open) and B (closed) monomers are larger (Figure 2a). For instance, in *wt*, the rmsd of C_α atoms between the two monomers is 0.76 Å, and only 88 atoms fall inside the 0.2 Å cutoff. The areas with the biggest differences are the loops of residues 103−115 and 140−152, which coordinate M1 and M4. The two positions of M1 and the loop of residues 103−115 appear to be related to opening and closing the active site. Also, the α-helix of residues 180−195 and the loop of residues 196−208 include residues that assume a different conformation depending on the contents of the active site. The regions are involved in P1 binding. Below we describe the most salient features of each structure in turn, fitting them and earlier structures into an expanded catalytic scheme (Scheme 2), before discussing the results *in toto* and in the context of previous work.

Mg^{2+}-Containing wt Y-PPase Structure. In wt_A, the two Mg^{2+} ions at the base of the active site are 4.75 Å apart

Table 2: Refinement Statistics

	YWT	E48D	Y93F	D115E	D117E	D120E	D120N	D152E
D_{min} (Å)	1.5	1.7	1.7	1.8	1.8	1.8	1.9	1.5
total no. of unique reflections/ test set	103685/5172	68738/3472	67533/3417	58286/2933	59484/3002	59834/3024	46553/2370	102353/5150
(% size of the test set)	(5)	(5)	(5)	(5)	(5)	(5)	(5)	(5)
R_{work}/R_{free} [a,b]	0.162/0.180	0.159/0.199	0.167/0.190	0.175/0.217	0.162/0.190	0.167/0.203	0.154/0.210	0.162/0.182
	(0.175/0.208)	(0.247/0.318)	(0.310/0.356)	(0.293/0.304)	(0.233/0.279)	(0.229/0.265)	(0.252/0.294)	(0.181/0.210)
rmsd from ideality for bonds (Å)	0.011	0.015	0.013	0.016	0.013	0.015	0.015	0.009
rmsd from ideality for angles (deg)	1.445	1.531	1.536	1.565	1.410	1.470	1.503	1.283
no. of waters	670	782	505	657	437	520	736	762
Ramachandran plot, most favored	89.9% (435)	89.6% (429)	90.0% (433)	89.0% (431)	90.1% (435)	90.5% (430)	89.5% (434)	90.8% (443)
average *B*-factors (Å^2)								
main chain	15.3	16.9	24.1	20.6	25.0	24.8	27.3	15.0
side chain	16.6	18.5	26.0	21.7	26.7	26.3	28.3	17.5
all protein	15.9	17.7	25.0	21.2	25.8	25.5	27.8	16.2
Mg	21.1	18.5	22.5	30.6	24.5	42.2*	28.0	19.0
waters	26.6	28.5	32.4	31.9	31.1	32.7	38.5	27.4
all atoms	17.4	19.3	25.8	22.6	26.3	26.3	29.3	17.9

[a] Values for the highest-resolution shell are given in parentheses. [b] $R_{work} = \sum_{hkl}(|F_o(hkl)| - |F_c(hkl)|)/\sum_{hkl}|F_o(hkl)|$ for $F(hkl)$ not belonging to the test set. $R_{free} = \sum_{hkl}(|F_o(hkl)| - |F_c(hkl)|)/\sum_{hkl}|F_o(hkl)|$ for $F(hkl)$ in the test set.

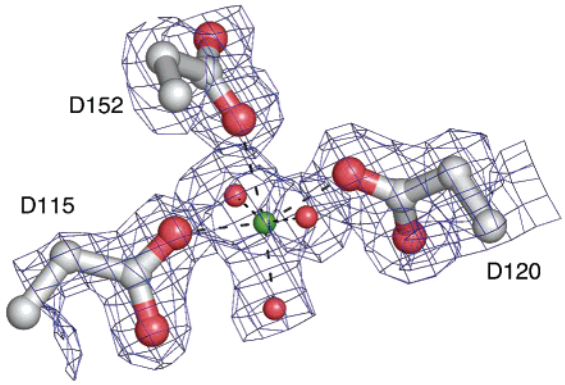


FIGURE 1: Electron density map showing the M1 site in E48D$_B$. The $2F_o - F_c$ map is contoured at 1σ. The Mg^{2+} ion is colored green.

(Figure 2b). The distance is slightly smaller than that in the Mn^{2+}-containing structure (4.92 Å) (*7*) presumably due to the difference in the size of the metal ions. The water structure and the conformations of the metal-coordinating residues are conserved, but the Mg^{2+} structure appears slightly more tightly packed, probably for the same reason. The metal and phosphate contents of wt_A (Table 3) indicate it represents the A intermediate (Scheme 2), analogous to the earlier Mn^{2+}-containing structure [PDB entry 1WGI (*7*)]. An additional phosphate ion, P_{exit}, is found in the putative product release channel, bound by $K73_A$, $R78_A$, and $K198_A$ (see below).

In wt_B, the two metals are now only 3.77 Å apart, and concomitant with this (*7*), the two-water bridge between them has collapsed to only a single water, the reaction nucleophile. In addition, the P1 leaving group phosphate and the associated metal ion, M3, are bound, matching the contents expected for the F intermediate (Scheme 2) in the Mg^{2+}-containing enzyme (*5*).

E48D Variant. In *wt*, E48 forms a hydrogen bond to one of the M2 coordination waters. In E48D$_A$, only M1 and M3 are present (Table 3). This suggests that the 150-fold decrease in K_{M2} due to this mutation (Table 4) (*4*) reflects not M2, but rather M3, binding. Waters at the M2 binding site rearrange slightly (with a maximum difference of 1.4 Å compared to that of *wt*), so only the waters hydrogen bonded to the backbone 94_A(O) and 118_A(O) atoms are as in *wt*. An additional water has appeared to replace $E48_A(O_{\epsilon 2})$. In E48D$_B$, the situation is in some ways very similar: $E48_B$-($O_{\epsilon 2}$) has been replaced with a water, and the M2 shell bonded by the backbone remains in place. Even though M2 is absent, E48D$_A$ is clearly in the open conformation and E48D$_B$ closed. These conformations can be distinguished because, as in wt_A, $D117_A$ and $M1_A$ are separated by two water molecules, while $D117_B$ is hydrogen bonded to a water directly coordinating $M1_B$ (Figure 3). The P1 phosphate is present in both active sites. This is not surprising, as P1 is not involved in closing the M1−M2 bridge. Therefore, E48D$_A$ is a model of the F′ intermediate and E48D$_B$ of the F intermediate (Scheme 2). M2 and M3 bind on the same side of the active site. In the variant, M3 is bound in both open and closed active sites, presumably to maintain charge balance.

Y93F Variant. In *wt*, Y93 makes a fairly short and tight hydrogen bond to one of the P2 oxygens (*7*) and a hydrogen

Table 3: Active Site Contents of Y-PPase and Possible Role in Scheme 2

	monomer A		monomer B		
	active site ligands	species in Scheme 2	active site ligands	species in Scheme 2	ref
wt	Mg1, Mg2, P_{exit}	A	Mg1, Mg2, Mg3, P1	F	this work
E48D	Mg1, Mg3, P1	F′	Mg1, Mg3, P1	F	this work
Y93F	Mg1, Mg2	A	Mg1		this work
D115E	Mg1, Mg2, Mg3, Mg4, P1, P2	E′	Mg1, Mg2, Mg3, Mg4, P1, P2	E	this work
D117E	Mg1, Mg2	A	Mg1, Mg2, Mg4, P2	C	this work
D120E	Mg1, Mg3, P1		Mg1, Mg3,[a] P1		this work
D120N	Mg1, Mg3, P1		Mg1, P1		this work
D152E	Mg2, P_{exit}		Mg2, Mg9		this work
wt (1WGI)[b]	Mn1, Mn2	A			*7*
wt (1E9G)[b]	Mn1, Mn2, Mn3, Mn4, P1, P2	E, D			*1*
wt (1E6A)[b]	Mn1, Mn2, Mn3, Mn4, F-PP_i	C			*1*
D117E (117E)[b]	Mn1, Mn2, Mn3, Mn4, P1, P2	E			*18*

[a] Likely two conformations. [b] The $P2_12_12_1$ crystal form has two equivalent active sites featuring the same intermediate of the reaction scheme.

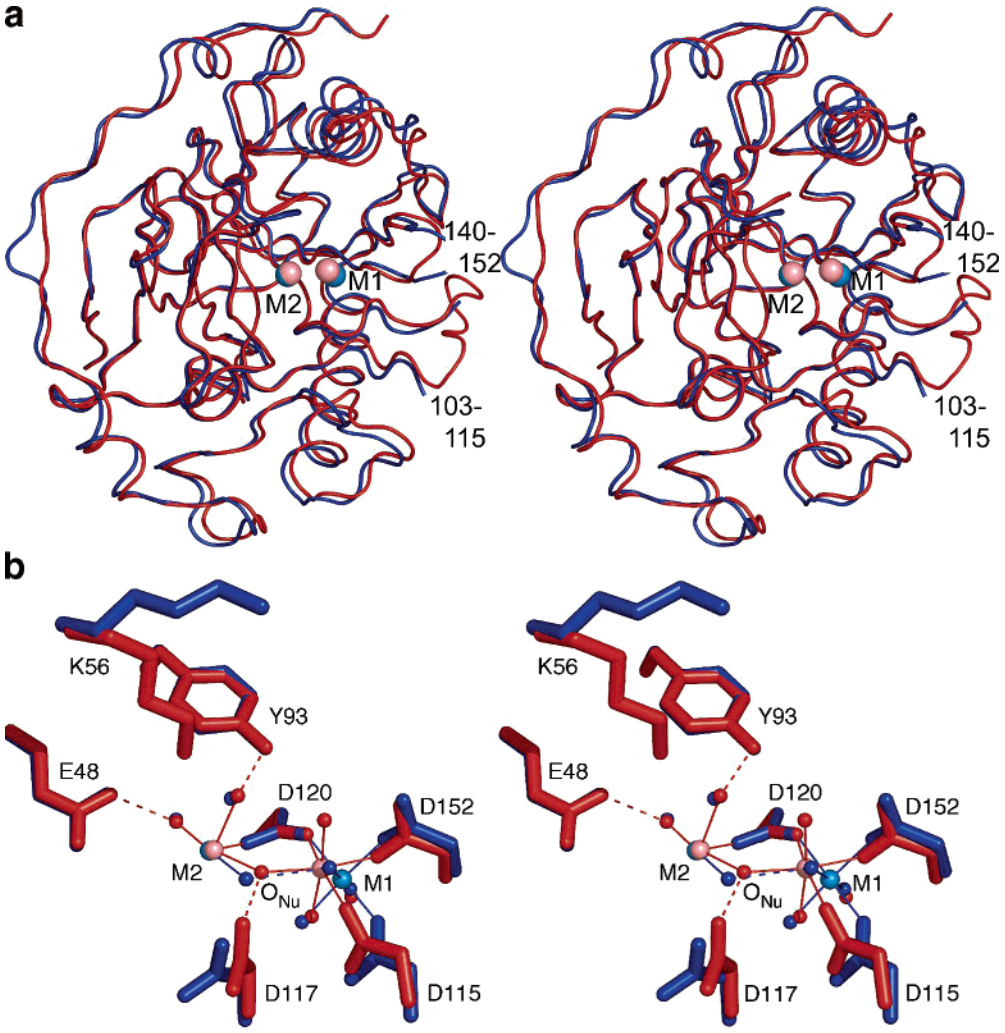
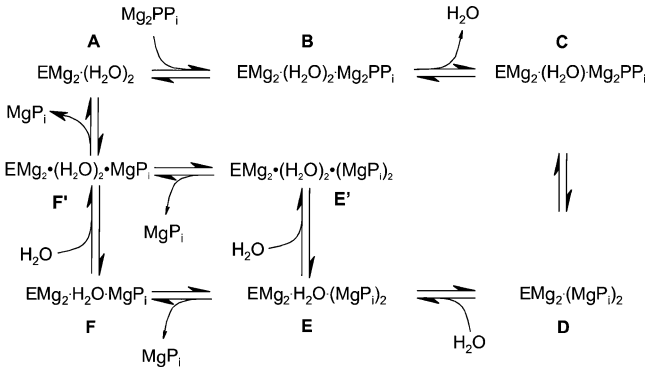


FIGURE 2: (a) C_α trace of wt_A (blue, metals light blue) superposed on wt_B (red, metals salmon), showing the main chain differences between the open (wt_A) and closed (wt_B) conformations. The major movements are in loops of residues 103−115 and 140−152. The active site metals are represented spheres. (b) Stereofigure of the active sites of wt_A (blue, metals light blue) and wt_B (red, metals salmon) superposed, showing the closing of the active site and the collapse of the two-water bridge that activates the nucleophile. Coordination bonds are represented with solid lines and hydrogen bonds with dashed lines.

bond to K154. Y93F has only 2.7% of the k_{cat}/K_m of *wt* (*17*). The simple explanation for the change would be that neither of these hydrogen bonds can form. The reality is significantly more complex. In the absence of substrate, $F93_A$ adopts a completely new conformation, binding in a cryptic hydrophobic cavity formed by $Y89_A$, $L122_A$, $W188_A$, and $F189_A$ (Figure 4). It replaces a buried water coordinated by the backbone of $I90_A$ and the $H91_A$ side chain; the $H91_A$ side chain rotates 180° around the $C_\beta-C_\gamma$ axis so that its $C_{\delta 2}$ atom faces into the hydrophobic pocket, while the $Y89_A$,

Scheme 2 : Revised Reaction Scheme, Based on the Data in This Paper



$H91_A$, and $F189_A$ side chains move back by <1 Å. In *wt*, a water molecule accepts a hydrogen bond from I90(N) and probably donates one to the W188 ring and to T185($O_{\gamma 1}$). Two waters fill the space occupied by the Y93 side chain, leaving the overall conformation and the active site hydrogen bonding pattern essentially unchanged (Figure 4). One would thus naively expect an active enzyme with little effect on k_{cat}/K_m or on the pK_a values on the enzyme. Indeed, the pK_a values measured before the substrate binds change by only a pH unit or so. However, k_{cat}/K_m changes to 3% of that of *wt* and pK_{ESH} from 7.7 to >10 on mutation. It appears that the mutation affects the behavior of the enzyme only when the substrate is bound.

This is indeed shown in the structure. M2 is not bound in $Y93F_B$ even though it is in $Y93F_A$ (Table 3 and Figure 4). This would be expected to reduce the catalytic efficiency through its effect on substrate binding and on nucleophile activation. Why does this happen even though all the M2-coordinating carboxylates are present? In $Y93F_B$, the closure of the active site squeezes $F93_B$ back into the *wt* Y93 position; motions of ~0.5 Å at $Y89_B$ and $F189_B$ close the cryptic hydrophobic pocket down. Consequently, the water network found in *wt* and in $Y93F_A$ can no longer form, disrupting the pattern of hydrogen bonds around K154. Wat28, which binds to $Y93_B$ and $K154_B$ in *wt*, is not present, nor can the Wat276−Wat433 network found in $Y93F_A$ form (Figure 4). Consequently, $K154_B$ now makes three hydrogen bonds: to $D120_B(O_{\delta 2})$, to a water, and to $D152_B(O)$. The interaction of $K154_B$ with $D152_B$, which includes an electrostatic interaction with the carboxylate group in addition to the hydrogen bond, affects the conformation of the M2 and M4 binding loop (residues 145−152). A water ligand for M2 and P2 is missing, reducing the affinity for M2 and P2. Consistent with the absence of M2, O_{Nu} shifts slightly toward M1. The changes are subtle but far-reaching.

D115E Variant. In the wild-type enzyme, D115 coordinates the first bound metal ion, M1, and the following residue, N116, is involved in the formation of a low-barrier hydrogen bond (LBHB) through its backbone (*1*). Preceding D115 in Y-PPase is a long and often disordered loop (Figure 2a). During the catalytic cycle, this loop moves and allows the D115-coordinated M1 and D117 to close the active site and generate the nucleophile (*7*). Upon mutation, K_m decreases somewhat, but k_{cat} is 1 order of magnitude lower (*17*) and uniquely reflects the chemical catalysis step (C → D), which is then rate-determining. As k_{cat} decreases, any decrease in K_m must be due to a decrease in the true K_d (A → B) or the equilibrium constant between intermediates B and C, i.e., the closing of the active site. In either case, the net binding affinity of substrate increases. Consistent with this, both open and closed conformations of D115E contain a full complement of ligands in the active site (Table 3).

Table 4: Metal Binding Affinities

	K_{M1} (μM)[a]	likely M	K_{M2} (μM)[a]	likely M
wt	0.7 ± 0.1		34 ± 6	
E48D	0.7 ± 0.1	M1	5000 ± 1000	M3
Y93F	0.7 ± 0.1	M1	60 ± 10	M2
D115E	0.8 ± 0.1	M1	400 ± 100	
D117E	1.4 ± 0.1	M1	170 ± 50	M2
D120E	24 ± 2	M1	2800 ± 700	M3
D120N	11 ± 1	M1	1400 ± 200	M3
D152E	2.8 ± 0.2	M2	1700 ± 500	M9

[a] Data taken from ref *4*.

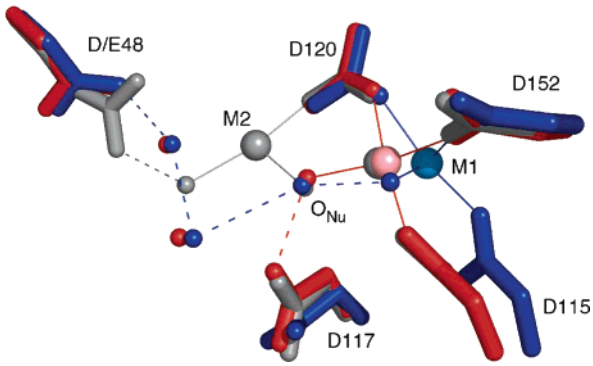


FIGURE 3: Active sites of $E48D_A$ (blue, metals light blue) and $E48D_B$ (red, metals salmon) superimposed on wt_B (gray), showing the loss of the M2 site and the difference in the hydrogen bonding network of the open $E48D_A$ and closed $E48D_B$. Even in the absence of M2, the position of O_{Nu} is conserved, allowing the classification of open and closed conformations. Coordination bonds are represented with solid lines and hydrogen bonds with dashed lines.

The open conformation of the $D115E_A$ active site contains the two-water bridge between $M1_A$ and $M2_A$ (metals separated by 4.4 Å) as seen in the resting enzyme complex (Scheme 2, A). The conformation of the $E115_A$ side chain is such that one carboxylate oxygen ($O_{\epsilon 1}$) coordinates $M1_A$, as in *wt*. The other carboxylate oxygen coordinates $M4_A$ via a water molecule. The longer side chain of $E115_A$ pushes back both $N116_A$ and the loop from residue 107 to 113, which is now disordered. Both active sites contain both phosphates P1 and P2, making $D115E_A$ a structural mimic of an overhydrated product release intermediate (E′) (Scheme 2). Consistent with this, other active site residues are in conformations typical of the "closed, bound" conformation. As in *wt* product complexes, $M4_A$ is coordinated only to D147 and D152, but at 5 Å, it is farther from $M1_A$ than in wt_A.

As the full complement of metals and phosphates is present, $D115E_B$ represents the E intermediate (Scheme 2 and Figure 5). The distance between $M1_B$ and $M2_B$ is 3.6 Å, shorter than those of wt_B (above) and the Mn−product complex structure (*1*) (3.7 Å). In addition, the $M1_B$−$M4_B$ distance is slightly shorter than in the *wt* product complex (PDB entry 1E9G). This may be because $E115_B$ coordinates both $M1_B$ and $M4_B$, in contrast to *wt*, where D115 coordinates M1 only. In both $E58_A$ and $E58_B$, the C_α−C_β bond rotates to an anti conformation. This change is not unique and may

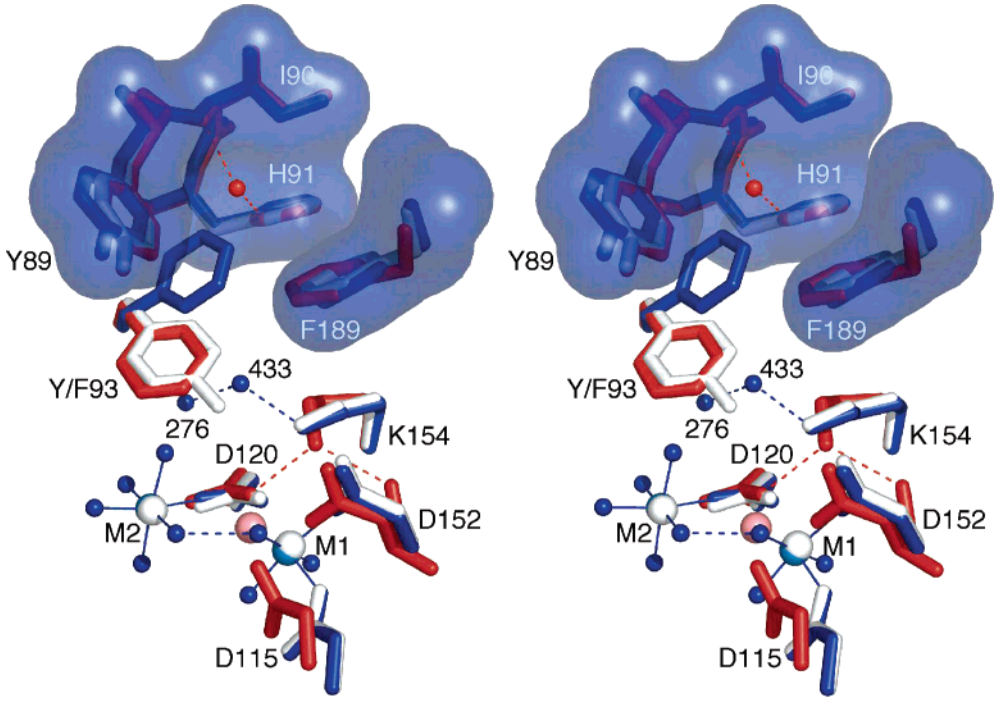
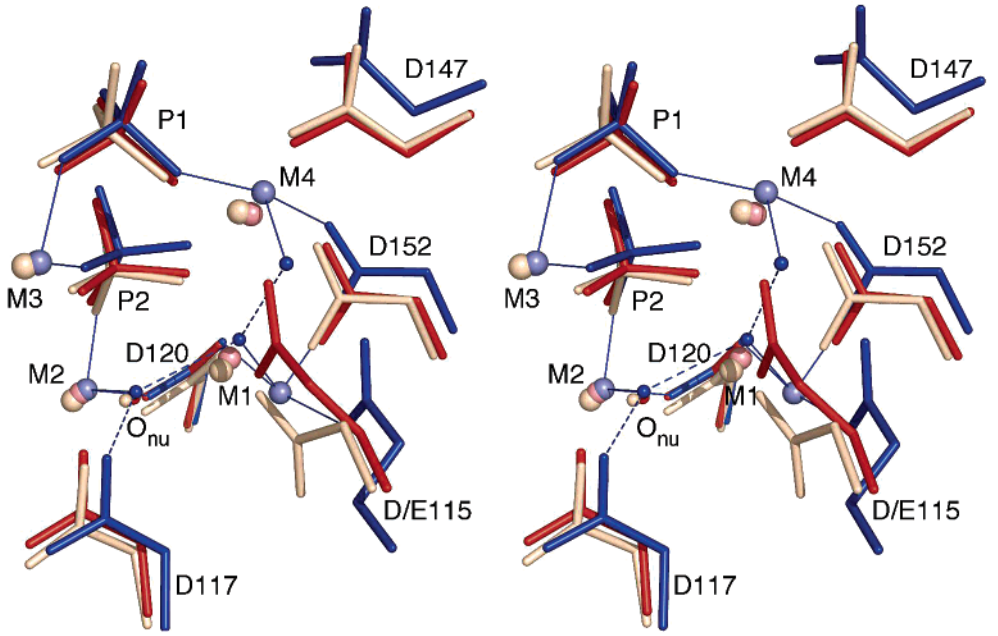


FIGURE 4: Structure of Y93F compared with *wt*. $Y93F_A$ (blue, metals light blue) is an open and $Y93F_B$ (red, metals salmon) a closed conformation. The open wt_A structure (white) shows that the hydrophobic pocket in *wt* is slightly smaller. The transparent surface shows the atomic surface of the hydrophobic pocket that forms in $Y93F_A$, allowing binding of $F93_A$ in an unusual conformation. In the $Y93F_A$ structure, two water molecules replace the F93 phenyl ring. Coordination bonds are represented with solid lines and hydrogen bonds with dashed lines.



35

FIGURE 5: Active sites of $D115E_A$ (blue, metals light blue) and $D115E_B$ (red, metals salmon) superposed on the high-resolution product complex 1E9G (wheat). Coordination bonds are represented with solid lines and hydrogen bonds with dashed lines.

be required for formation of the M3 site. It also occurs in the product (PDB entry 1E9G) and substrate (PDB entry 1E6A) complex structures but not in the *wt* structure reported here.

D117E Variant. In *wt*, D117 is hydrogen bonded to one of the M2 coordination waters and, in the closed conformation of the active site, forms a low-barrier hydrogen bond to the nucleophile O_{Nu} which increases the O_{Nu} nucleophilicity and orients a lone pair of electrons on O_{Nu} in an optimal position for the reaction (*1*). The effects of the mutation on K_m and k_{cat} are similar to those for D115; the substrate binds better, but the hydrolysis step is less efficient (*17*).

$E117_A$ has turned away from the two metal ions and forms a multicenter ion pair with $K61_A$ and $E63_A$ instead of hydrogen bonding to the nucleophilic O_{Nu}. Unlike the $D115E_A$ structure, the open active site has no bound

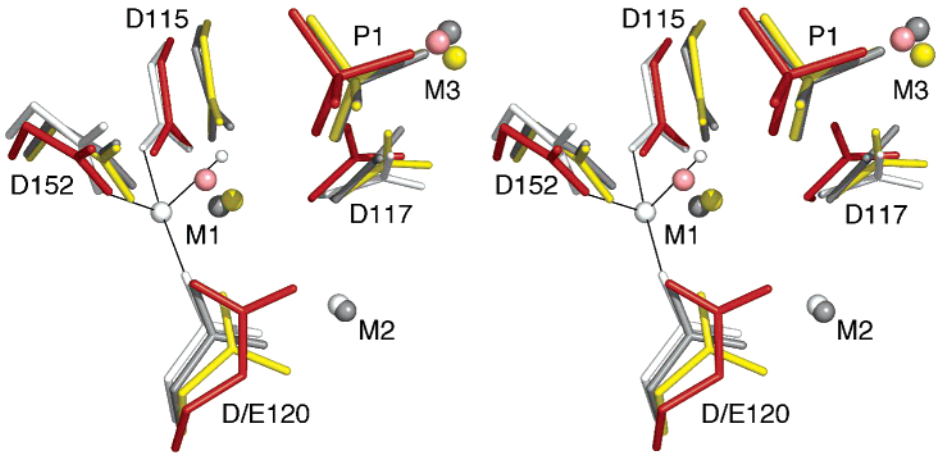


FIGURE 6: Structures of $D120E_A$ (red, metals salmon) and $D120N_A$ (yellow) compared with open (white) and closed (gray) *wt* conformations. In the D120E variant, the M1 binding residues are in a closed-type conformation yet M1 is between the M1 positions of the *wt* open and closed conformations. Metals are represented as spheres, coordination bonds with solid lines, and hydrogen bonds with dashed lines.

phosphate molecules. Hence, $D117E_A$ represents the A intermediate (Scheme 2). Consistent with this, R78 points toward the outside of the molecule but does not interact with D71 (*18*).

In the closed conformation, $E117_B$ is forced to adopt the space occupied by *wt* D117 and hence cannot interact with K61 and E62 as in monomer A. $E117_B(O_{\epsilon 1})$ and $O_{\epsilon 2}$ replace $D117(O_{\delta 2})$ and the nucleophilic water, leaving no space for the catalytic water to bind, as in our earlier Mn^{2+} structure (*18*). Both structures correspond to the C intermediate (Scheme 2). The major difference between them is that the P1 phosphate and M3 metals are missing from our Mg^{2+} structure, presumably due to the tighter binding of Mn^{2+} and P_i to the Mn^{2+}-containing enzyme (*5*).

D120E and D120N Variants. In all *wt* structures, D120 simultaneously coordinates both M1 and M2 at the bottom of the active site (*7*) (Figure 2b); not surprisingly, K_{M1} is more than 10-fold lower and K_{M2} 40−80-fold lower (Table 4) in the D120E and D120N variants. On the basis of the active site contents (Table 3), the measured K_{M2} values refer to M3, not M2, binding, as for the E48D variant. Intriguingly, the D120E variant has a stronger effect on metal binding (Table 4), while the D120N variant has a greater effect on the rate of hydrolysis; however, both are more than 1000 times less active than *wt* at pH 7.2 (*4*). The structures appear to provide an explanation for these results.

Both monomers in both variants bind M1, P1, and, all except $D120N_B$, M3 (Table 3). The M1 position is distorted in all monomers, and M2 is not seen at all. The huge decrease in the rate of hydrolysis further highlights the importance of M2 in generating the nucleophile.

In D120E, the backbone is distorted by ~1 Å; however, the position of $E120_A(O_{\epsilon 1})$ is almost the same as that of $D120_A(O_{\delta 2})$ in *wt*, and as in *wt*, it accepts a hydrogen bond from $K154_A$. $M1_A$ is bound, but in a position the same as neither the *wt* open nor *wt* closed form (Figure 6), but quite close to the position of an M1-coordinating water molecule in wt_A. The bond distances from protein ligands to M1 are consistent with considerably weakened M1 binding: 2.02−2.06 Å in wt_A and 1.99−2.20 Å in wt_B but 2.27−2.44 Å in D120E, with only square pyramidal rather than octahedral geometry. The $D120E_B$ active site is similar to the $D120E_A$ active site, but $M1_B$ is now very close to the wt_B position and the geometry improved. The lack of clear open and closed states of the active site correlates with large (0.3−0.7% of *wt*) changes in the net dissociation rate of PP_i in the D120N and D152E variants (*4*). This may explain why K_m changes by a factor of only 3 (*4*); substrate binding is not significantly affected, although catalysis is.

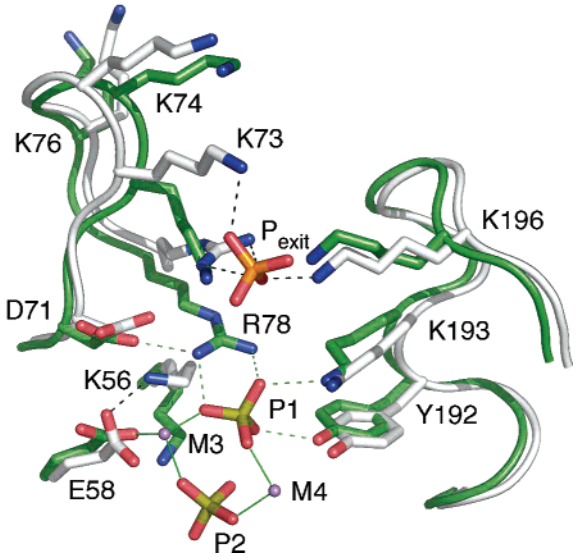


FIGURE 7: Superposition of wt_A (gray for carbon) and 1E9G alternate conformation A (green for carbon and lime for phosphorus) showing the waving motion of lysine side chains and the main chain motion of the respective loops in the product release channel and the coordination of the P_{exit} phosphate by K73, R78, and K196. Metals are represented as spheres, coordination bonds with solid lines, and hydrogen bonds with dashed lines.

Conversely, there is less backbone distortion in $D120N_A$; the backbone around N120 is essentially undistorted in comparison with that of wt_B. In comparison with $E120_A$, D120N binds M1 tighter (Table 4) even though the M1 ligand is now the amide N120 rather than the carboxylate E120. Consistent with tighter binding, D120N $M1_A$ has shorter protein−metal ion distances (2.04−2.36 Å) and better geometry and binds in the closed conformation position. The distance to the *wt* $M1_A$ closed position is ~0.3 Å, similar to that of the Mn^{2+}-containing product complex active site

(PDB entry 1E9G). It thus appears that the level of distortion at the M1 binding site, not the identity of the side chain at position 120, correlates with metal binding affinity.

D152E Variant. The large-scale distortions of the active site in the D152E variant are perhaps the most surprising results in the context of this study. They may happen because D152 is located at the end of the mobile loop of residues 146−152 (Figure 2a), so significant distortion is possible. Unlike the other D → E variants, none of the carboxylate oxygens of $E152_A$ and $E152_B$ overlap with the D152 carboxylates of *wt* (the closest approach is 1.17 Å to the carboxylate oxygen that does *not* bind M1). Consistent with this, M1 is absent in both monomers but M2 is present (Table 3); the K_{M1} binding constant (*17*) (Table 4) thus reflects binding to the M2 site. E152 in both monomers fills the empty M4 site and causes movement at the loops of residues 146−152 and 101−115; the distortion there, for instance, is larger than in the D120N variant. As M2 is present, the nucleophile−D117 system is present, demonstrating that M2 is necessary, but not sufficient, for nucleophile generation. Indeed, the increases in the enzyme pK_a are some of the largest (*17*), and the enzyme is less active than D120E, showing dramatic effects on k_{cat}, K_m, and k_{cat}/K_m (*17*). An additional phosphate ion, P_{exit}, is visible in the product release channel, bound to $K73_A$, $K198_A$, and $R78_A$.

In $D152E_B$, the nucleophile−D117 system is present, as in $D152E_A$, but in addition, we see both conformations of the $D117_B$ side chain and $N116_B$(O) as in the 1.1 Å structure of *wt* in the presence of product (*1*). This is not visible in the other structures described here; we believe that this reflects the release of constraints in the active site as a result of the absence of M1. There is also an additional metal ion (M9) outside the active site, coordinated through waters to E150; this may be due to the charge imbalance created through the absence of M1.

Roles of the Two Forms of the Active Site. The active site has two principal conformations: open and closed (*7*). In the open form, M1 and M2 are bridged by two water molecules, whereas in the closed form, only the activated nucleophile O_{Nu} remains between the metals. This is effected chiefly by ordering and closure of the loops involved in binding M1: residues 72−76, 103−115, and 147−152 (*7*) (Figure 2a). In addition, motions of K56 and R78 are probably involved in product release. This appears to be triggered by a change in the hydrogen bonding to D71 (*18*); in the closed conformation, D71 is hydrogen bonded to R78, thus allowing K56 and R78 less freedom of movement than in the open conformation.

In the presence of Mg^{2+}, the P2 phosphate is released first (*13*). We believe that this is driven by electrostatic repulsion. What then drives the release of the more tightly bound product phosphate P1? A channel of flexible positively charged residues could provide a low-energy pathway for its exit. A phosphate ion is indeed observed in this channel in both wt_A and $D152E_A$. In the open conformation, K56 and R78 are in a "waving" position, ready to pass the product phosphate along the chain of lysines formed by K73, K74, K76, K193, and K198 (Figure 7).

In the *wt* closed conformation, there is tight coordination of the substrate, a detailed match to the hydrogen bonding requirements of substrate and product, and efficient generation of the nucleophile. These requirements are not met well in the variants. Conversely, in the open conformation, the mutated residues have more space to adopt different conformations than in *wt* without causing large changes in the overall structure. The open (non-substrate-bound) form thus sometimes shows less change in kinetic parameters than when substrate is bound. An example is the Y93F variant. In $Y93F_A$, the phenyl ring of $F93_A$ binds to the hydrophobic core of the protein (Figure 4**),** but the active site geometry and solution properties (*4, 17*) change relatively little. In $Y93F_B$, it is forced back into the active site, and the binding properties of M2 and PP_i change significantly (Table 4) because the active site geometry is not properly conserved (Figure 4) (see below).

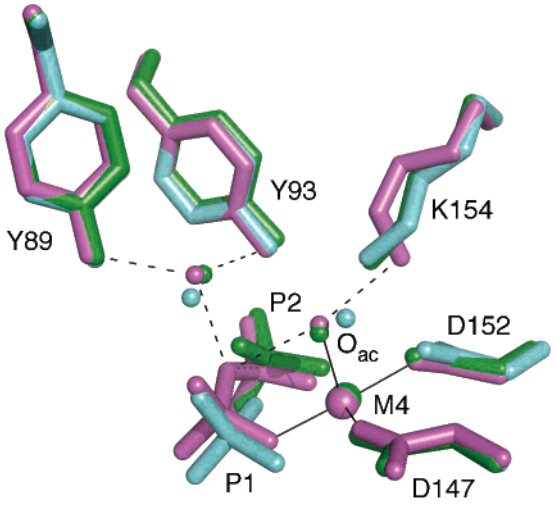


FIGURE 8: Potential general acid. The substrate complex [1E6A (*18*)] (magenta) is shown superposed with $D117E_B$ (green) and $E48D_B$ (cyan), showing the two conserved water molecules in the proximity of the bridging oxygen in PP_i. Coordination bonds are represented with solid lines and hydrogen bonds with dashed lines.

This remarkable chain of events stems from the loss of just two hydrogen bonds in comparison with the *wt* substrate complex: Y93−K154 and Y93−P2. Furthermore, the slight decrease in substrate binding affinity (*17*) might be due to the changes in the coordination of metal binding residues D120 and D152 and the binding of M2 and M4. It indicates that the key to catalysis in this enzyme is precise preorganization (*28*) and that, even though substrate binding can be unaffected as measured by K_m, the binding is no longer in a catalytically competent mode. Not surprisingly, a stronger effect is seen in the catalysis step itself (k_{cat} is 8% of that of *wt*), again emphasizing the key nature of the Y93−P2 hydrogen bond.

The previously determined structures containing substrate (PDB entry 1E6A) and product (PDB entry 1E9G) also include a rearrangement of the backbone and D117 residue, allowing formation of a short low-barrier hydrogen bond (*1*) between D117 and the nucleophile during catalysis (Scheme 2, between intermediates C and D). In the current structures, the rearrangement is not visible in the wild-type enzyme, which is already at the F intermediate of the catalysis, but is present in the structure of D152E monomer B, providing independent conformation of this rearrangement. The $D152E_B$ structure contains only the M2 metal in the active site, and even though the changes in the main chain conformation relative to that of *wt* are the largest for this variant, they are nonetheless rather small. The most marked are the movement

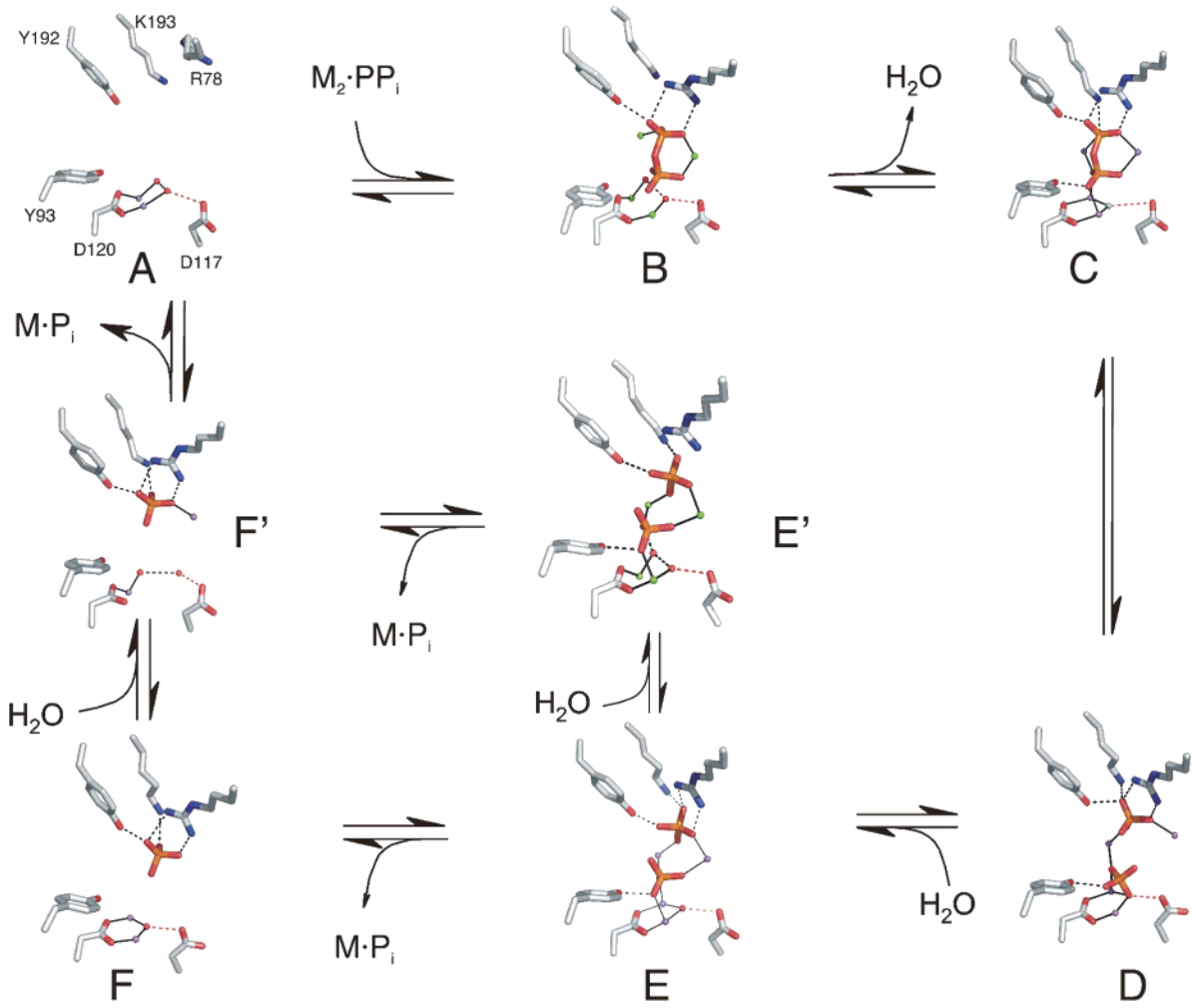


FIGURE 9: Structures of the active site in the various catalytic intermediates, labeled as in Scheme 2. Residues R78, Y93, D117, D120, Y192, and K193 (labeled in intermediate A), as well as metals, phosphates, and the two waters bridging M1 and M2, are shown. The structure for intermediate A is 1WGI (*7*), for B $D115E_A$ with PP_i modeled, for C 1E6A (*18*), for D 1E9G conformation B(1), for E 1E9G conformation A(1), for E′ $D115E_A$, for F wt_B, and for F′ $E48D_A$.

of N116(O) and the associated rotation of the peptide plane. The observation of this alternate conformation probably results from the increased flexibility due to the lack of M1.

Completion of the Catalytic Cycle. Of the intermediates in the revised catalytic scheme (Scheme 2), structural models were previously available for A and C−E (Table 3). The biggest change with respect to our previous structures, all with Mn^{2+}, is in the metal−ligand distances, allowing us to use relevant earlier structures in devising the structure-based reaction scheme below. We believe that the structures presented here allow us to produce a complete description of the expanded catalytic cycle (Table 3). Although the structures are of variant proteins, they correspond to expected hydration states and have the active site contents expected from solution data. With the Mg^{2+}-bound enzyme, P2 is released first, and in the structures presented here, the site is empty more often than not. This is consistent with earlier biochemical results (*5*) and provides support for the notion that the structures captured here represent typical states of the enzyme during catalysis. Consequently, we have used *wt* structures where possible to create a structural equivalent (Figure 9) of Scheme 2, which we have converted into a movie of catalysis (Supporting Information). In this figure, only intermediates B, E′, and F′ are from variant structures. $D115E_A$ formally corresponds to E′ (Scheme 2), but we have also used it for intermediate B by modeling pyrophosphate into this structure. The existence of such a species was suggested by F^- inhibition kinetics (*4*), and the $D115E_A$ structure confirms that the phosphate can bind at P2 even in the open conformation of the active site.

Metal Binding. We can now provide a structural assignment to our published equilibrium dialysis data on Mg^{2+} binding (*4*) (Table 4), because specific variants exhibit significantly altered active site contents. On the basis of this, the M2 site is destroyed in the E48D, D120E, and D120N variants. In each of these, monomer A contains only two Mg^{2+} ions in M1 and M3. As M1 is the only metal ion in $D120N_B$ (Table 3), it is logical to assume that the second metal binding constant (K_{M2}) of 1.4−5 mM refers to the binding constant for Mg^{2+} in the M3 site (Table 4). Similarly, in the $D152E_A$ structure, M1 is absent, so the binding constant of 2.8 μM for the first metal ion (K_{M1}) most likely refers to binding to the M2 site; the M2 binding constant decreases more than 10-fold when M1 is absent. (A 4-fold decrease would be expected by stochastical considerations.) It is harder to identify the step to which K_{M2} refers in the D115E variant, where all four metal sites are occupied in both monomers (Table 3), but it seems reasonable to suppose it is M2, as D115 is in the M1-binding site. If this is the case, then, on mutation, K_{M1} varies from 0.7 μM (E48D) to

24 μM (D120E), except for D152E, where M1 binding is presumably weaker than 5−10 mM, as the K_{M2} for D152E is presumably the "M9" site (Table 4) seen in the exit channel of $D152E_B$.

Similarly, K_{M2} varies from 34 μM (*wt*) to 400 μM (D115E) to more than 5−10 mM (D48E) in the presence of M1; in its absence, it is 2.8 μM (D152E). Finally, M3 binding is in the 1.4−5 mM range, ~10-fold weaker than in M2. As this spans a range of only 3-fold, we posit that, in *wt*, the M3 binding constant is also in the 1−5 mM range in the absence of substrate. It is striking that M4 binding has only been observed in the presence of substrate or product. This suggests that its binding constant in the absence of substrate is substantially weaker than that of M3, even though it has two direct protein ligands (D147 and D152) while M3 has just one (E58). We attribute this to the greater flexibility of the "M1 side" of the active site, in particular the loop of residues D147−D152 (Figure 2a). In the absence of substrate, the free energy cost of organizing this loop around M4 offsets the free energy gain of binding. Therefore, our earlier interpretation (*7*) that M3 was the modulating metal ion is incorrect; M4 plays that role.

Phosphate Binding and Catalysis. The phosphate in the P1 site can assume different orientations, even though it remains coordinated by the same residues (R78, Y192, and K193). When pyrophosphate is bound, the orientation of the P1 phosphoryl group is fixed by the covalent bond to P2, but when two phosphates are bound, P1 may orient like the leaving group of the substrate, where one oxygen points toward P2, or it may turn away to allow room for P2 to bind optimally. The first case is observed in the absence of M4 and P2 and in the immediate product complex (*1*), whereas the second case seems to be favored when both product phosphates are bound. This rotation of the phosphates is required for the transition between the immediate (D) and relaxed (E) product complexes (Figures 2b and 4).

It has been proposed that the hydrolysis of phosphodiester bonds is assisted by general acid catalysis at the bridging oxygen (*29*). In the substrate complex (PDB entry 1E6A), only two water molecules are found near the bridging oxygen. One is hydrogen bonded to Y93 and Y89, whereas the other coordinates M4 and forms a hydrogen bond to K154. These water molecules are conserved also in those variant structures that contain P1 ($E48D_B$) and P2 ($D117E_B$) in substratelike conformations. As the water between M4 and K154 is a metal ligand and close to a second positive charge, it would be a more likely candidate for the general acid O_{ac} (Figure 8). Interestingly enough, if the residue corresponding to K154 in E-PPase (K104) is mutated to isoleucine, the hydrolytic activity is destroyed but the reverse reaction, PP_i synthesis, still occurs (*30*). This is as expected, as biochemical data show that the rate-determining step in synthesis is not the bond-making step, but the isomerization from the closed form (C) to the open form (B) (*12*) (Scheme 2). A general acid or base at the bridging atom is thus not part of the RDS, as the isomerization takes place at the M1 site. Furthermore, the assistance of M4 in generating the general acid would be more important at low pH; i.e., we assign M4 as the "modulating metal ion" we have previously described (*9*). Finally, such an assignment is consistent with M4 binding more weakly than M3 (see above).

CONCLUSIONS

Although the effects of conservative active site mutations on PPase catalysis are not straightforward, they arise from changes in metal binding or modification of the preferred ground state structure. They also indicate, as shown above, the vast array of conformations available to the enzyme, such as the alternative conformation observed in $D152E_B$ (seen before in our high-resolution *wt* structure, PDB entry 1E9G) and the surprising changes seen in $Y93F_A$. Such information will allow the analysis of the correlated motions in the active site. We have completed the reaction scheme for Y-PPase (Figure 9) by determining structures for intermediates E′, F, and F′ and modeling intermediate B. These structures represent a minimal set of hydration states that must exist on the enzyme (Scheme 2). The alternative routes for product release and nucleophile regeneration cannot be distinguished on the basis of these kinetic data and may indeed even operate in parallel. The observation of these various states in the same crystal form suggests that they represent easily accessible low-energy conformations of the enzyme. Future work will focus on using these structures and kinetic data as input parameters for quantum-mechanical models of the free energy landscape of the reaction.

ACKNOWLEDGMENT

We acknowledge the help from Ed Hough, Ronny Helland, Ingar Leiros, and Ole-Andreas Andersen with data collection and the help from Michael Merckel and Erik Axel Vollan with computers. Provision of beamtime at SNBL (BM01) at the European Synchrotron Radiation Facilities (ESRF) is gratefully acknowledged. We also thank Alexander Baykov for helpful comments and Igor Fabrichniy for inspiring discussions.

SUPPORTING INFORMATION AVAILABLE

Animated movie based on a linear interpolation of our models of the catalytic intermediates shown in Figure 9. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Heikinheimo, P., Tuominen, V., Ahonen, A.-K., Teplyakov, A., Cooperman, B. S., Baykov, A., Lahti, R., and Goldman, A. (2001) Towards a Quantum-Mechanical Description of Metal Assisted Phosphoryl Transfer in Pyrophosphatase, *Proc. Natl. Acad. Sci. U.S.A. 98*, 3121−3126.
2. Cooperman, B. S., Baykov, A. A., and Lahti, R. (1992) Evolutionary conservation of the active site of soluble inorganic pyrophosphatase, *Trends Biochem. Sci. 17*, 262−266.
3. Halonen, P., Baykov, A. A., Goldman, A., Lahti, R., and Cooperman, B. S. (2002) Single-turnover kinetics of *Saccharomyces cerevisiae* inorganic pyrophosphatase, *Biochemistry 41*, 12025−12031.
4. Pohjanjoki, P., Fabrichniy, I. P., Kasho, V. N., Cooperman, B. S., Goldman, A., Baykov, A. A., and Lahti, R. (2001) Probing Essential Water in Yeast Pyrophosphatase by Directed Mutagenesis and Fluoride Inhibition Measurements, *J. Biol. Chem. 276*, 434−441.
5. Zyryanov, A. B., Pohjanjoki, P., Kasho, V. N., Shestakov, A. S., Goldman, A., Lahti, R., and Baykov, A. A. (2001) The electrophilic and leaving group phosphates in the catalytic mechanism of yeast pyrophosphatase, *J. Biol. Chem. 276*, 17629−17634.
6. Baykov, A. A., Cooperman, B. S., Goldman, A., and Lahti, R. (1999) Cytoplasmic inorganic pyrophosphatase, in *Progress in Molecular and Subcellular Biology* (Schröder, H. C., and Müller,

W. E. G., Eds.) Vol. 23, pp 127−150, Springer-Verlag, Heidelberg, Germany.
7. Heikinheimo, P., Lehtonen, J., Baykov, A., Lahti, R., Cooperman, B. S., and Goldman, A. (1996) The structural basis for pyrophosphatase catalysis, *Structure 4*, 1491−1508.
8. Samygina, V. R., Popov, A. N., Rodina, E. V., Vorobyeva, N. N., Lamzin, V. S., Polyakov, K. M., Kurilova, S. A., Nazarova, T. I., and Avaeva, S. M. (2001) The structures of *Escherichia coli* inorganic pyrophosphatase complexed with Ca^{2+} or $CaPP_i$ at atomic resolution and their mechanistic implications, *J. Mol. Biol. 314*, 633−645.
9. Baykov, A. A., and Shestakov, A. S. (1992) Two pathways of pyrophosphate hydrolysis and synthesis by yeast inorganic pyrophosphatase, *Eur. J. Biochem. 206*, 463−470.
10. Cooperman, B. S. (1982) The Mechanism of Action of Yeast Inorganic Pyrophosphatase, *Methods Enzymol. 87*, 526−548.
11. Belogurov, G. A., Fabrichniy, I. P., Pohjanjoki, P., Kasho, V. N., Lehtihuhta, E., Turkina, M. V., Cooperman, B. S., Goldman, A., Baykov, A. A., and Lahti, R. (2000) Catalytically Important Ionizations along the Reaction Pathway of Yeast Pyrophosphatase, *Biochemistry 39*, 13931−13938.
12. Baykov, A. A., Fabrichniy, I. P., Pohjanjoki, P., Zyryanov, A. B., and Lahti, R. (2000) Fluoride Effects Along the Reaction Pathway of Pyrophosphatase. Evidence for a second enzyme pyrophosphate intermediate, *Biochemistry 39*, 11939−11947.
13. Springs, B., Welsh, K. M., and Cooperman, B. S. (1981) Thermodynamics, Kinetics, and Mechanism in Yeast Inorganic Pyrophosphatase Catalysis in Inorganic Pyrophosphate: Inorganic Phosphate Equilibration, *Biochemistry 20*, 6384−6391.
14. Cooperman, B. S., Panackal, A., Springs, B., and Hamm, D. J. (1981) Divalent metal ion, inorganic phosphate, and inorganic phosphate analogue binding to yeast inorganic pyrophosphatase, *Biochemistry 20*, 6051−6060.
15. Zyryanov, A. B., Shestakov, A. S., Lahti, R., and Baykov, A. A. (2002) Mechanism by which metal cofactors control substrate specificity in pyrophosphatase, *Biochem. J. 367*, 901−906.
16. Heikinheimo, P., Pohjanjoki, P., Helminen, A., Tasanen, M., Cooperman, B. S., Goldman, A., Baykov, A., and Lahti, R. (1996) A site-directed mutagenesis study of *Saccharomyces cerevisiae* pyrophosphatase: Functional conservation of the active site of soluble inorganic pyrophosphatases, *Eur. J. Biochem. 239*, 138−143.
17. Pohjanjoki, P., Lahti, R., Goldman, A., and Cooperman, B. (1998) Evolutionary Conservation of Enzymatic Catalysis: Quantitative Comparison of the Effects of Mutation of Aligned Residues in *Saccharomyces cerevisiae* and *Escherichia coli* Inorganic Pyrophosphatase on Enzymatic Activity, *Biochemistry 37*, 1754−1761.
18. Tuominen, V., Heikinheimo, P., Kajander, T., Torkkel, T., Hyytiä, T. K. J., Lahti, R., Cooperman, B. S., and Goldman, A. (1998) The R78K and D117E Active-site Variants of *Saccharomyces cerevisiae* Soluble Inorganic Pyrophosphatase: Structural Studies and Mechanistic Implications, *J. Mol. Biol. 284*, 1565−1580.
19. Otwinowski, Z., and Minor, W. (1997) Processing of X-Ray Diffraction Data Collected in Oscillation Mode, *Methods Enzymol. 276*, 307−326.
20. Collaborative Computational Project Number 4 (1994) The CCP4 suite: Programs for Protein Crystallography, *Acta Crystallogr. D50*, 760−763.
21. Vagin, A., and Teplyakov, A. (1997) MOLREP: An automated program for molecular replacement, *J. Appl. Crystallogr. 30*, 1022−1025.
22. Perrakis, A., Morris, R., and Lamzin, V. S. (1999) Automated protein model building combined with iterative structure refinement, *Nat. Struct. Biol. 6*, 458−463.
23. Jones, T. A., Zou, J.-Y., Cowan, S. W., and Kjeldgaard, M. (1991) Improved Methods for Building Protein Models in Electron Density Maps and the Location of Errors in these Models, *Acta Crystallogr. A47*, 110−119.
24. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr. D53*, 240−255.
25. Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics, *Acta Crystallogr. D60*, 2126−2132.
26. Kleywegt, G. J., and Jones, T. A. (1996) Efficient Rebuilding of Protein Structures, *Acta Crystallogr. D52*, 829−832.
27. Harutyunyan, E. H., Kuranova, I. P., Vainshtein, B. K., Höhne, W. E., Lamzin, V. S., Dauter, Z., Teplyakov, A. V., and Wilson, K. S. (1996) X-ray structure of yeast inorganic pyrophosphatase complexed with manganese and phosphate, *Eur. J. Biochem. 239*, 220−228.
28. Cannon, W. R., Singleton, S. F., and Benkovic, S. J. (1996) A perspective on biological catalysis, *Nat. Struct. Biol. 3*, 821−833.
29. Admiraal, S. J., and Herschlag, D. (1995) Mapping the transition state for ATP hydrolysis: Implications for enzymatic catalysis, *Chem. Biol. 2*, 729−739.
30. Fabrichniy, I. P., Kasho, V. N., Hyytia, T., Salminen, T., Halonen, P., Dudarenkov, V. Y., Heikinheimo, P., Chernyak, V. Y., Goldman, A., Lahti, R., Cooperman, B. S., and Baykov, A. A. (1997) Structural and functional consequences of substitutions at the tyrosine 55 lysine 104 hydrogen bond in *Escherichia coli* inorganic pyrophosphatase, *Biochemistry 36*, 7746−7753.

BI0619977