Crystallographic Identification of Metal-Binding Sites in *Escherichia coli* Inorganic Pyrophosphatase[†,‡]

Jussi Kankare,[§] Tiina Salminen,[§,||] Reijo Lahti,[||] Barry S. Cooperman,[⊥] Alexander A. Baykov,[#] and Adrian Goldman*[,§]

Centre for Biotechnology, SF-20521 Turku, Finland, Department of Biochemistry, University of Turku, SF-20500 Turku, Finland, Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6323, and A. N. Belozersky Institute of Physico-Chemical Biology, Moscow State University, Moscow 119899, Russia

Received November 6, 1995; Revised Manuscript Received February 7, 1996[⊗]

ABSTRACT: We report refined crystal structures of the hexameric soluble inorganic pyrophosphatase from *Escherichia coli* (E-PPase) to *R*-factors of 18.3% and 17.1% at 2.2 and 2.3 Å, respectively. Both structures contain two independent monomers in the asymmetric unit of an *R*32 cell. The difference between the structures is that the latter contains 1.5 Mg^{2+} ions per monomer. One metal ion binds to the "tight" metal-binding site identified by equilibrium dialysis studies, and is coordinated to Asp65, Asp70, and Asp102. The other metal ion, shared between two monomers at a hitherto unidentified metal-binding site in the dyad interface between trimers, is coordinated through water molecules to Asp26s and Asn24s from the two monomers. The hexamers with metal bound to them are more tightly associated than the ones without metal bound to them. Combined with our other mechanistic and structural data, the results suggest that, at high metal concentrations, E-PPase may bind at least 4.5 metals per monomer: two in the active site before binding substrate, two with substrate, and 0.5 in the dyad interface. Glu20 interacts via a water molecule with Asp70 and appears in the related yeast PPase structure (Heikinheimo, manuscript in preparation) to be involved in binding the second metal ion. Magnesium ion therefore stabilizes the hexamer form through both direct and indirect effects. The direct effect is by tighter association at the subunit interface; the indirect effect occurs because magnesium stabilizes the correct conformation of the loop between Glu20 and Ile32, a loop involved in trimer−trimer interactions. Our results thus provide a structural explanation for the solution studies that show that the E20D variant is partially hexameric and that the hexamer form can be stabilized by binding magnesium ion.

Inorganic pyrophosphatase (PPase,[1] EC 3.6.1.1) catalyzes the cleavage of pyrophosphate (PP_i) into orthophosphate (P_i). Pyrophosphate is a byproduct of DNA and protein synthesis; growing cells produce high levels of PP_i. By breaking down PP_i, PPases shift the equilibrium of these reactions toward synthesis. Pyrophosphatases are therefore essential constitutive enzymes present in all living cells (Chen *et al*., 1990; Lahti, 1983; Lundin *et al*., 1991) whose structures have been highly conserved over evolution (Cooperman *et al*., 1992; Kankare *et al*., 1994; Teplyakov *et al*., 1994). In addition, bacterial inorganic pyrophosphatases, with a subunit mass of 20 kDa, provide an attractively small protein for studies of energy transduction at the level of phosphoanhydride bond formation and breakdown (Herschlag & Jencks, 1990; Knowles, 1980).

PPases are strongly metal-dependent enzymes and require divalent cations, preferably Mg^{2+} (Cooperman, 1982), for catalytic activity. The role of the magnesium ions in their function is twofold: first, the preferred substrate is Mg_2PP_i and second, the enzyme binds metal ions as essential cofactors.

Detailed atomic information on bacterial PPases only became available in 1994 when three independent X-ray studies were published: a refined 2.0 Å structure of *Thermus thermophilus* pyrophosphatase (T-PPase) (Teplyakov *et al.*, 1994) and two independent *Escherichia coli* structures at resolutions of 2.5 Å (Oganessyan *et al.*, 1994) and 2.7 Å (Kankare *et al.*, 1994). The coordinates for all are now available from the Protein Data Bank (Bernstein *et al.*, 1977) with codes of 2PRD, 1IGP, and 1EIP, respectively.

In our earlier paper (Kankare *et al.*, 1994), we described the complex topology of E-PPase as a twisted β-barrel with excursions, the description of which is still valid (Kankare *et al.*, 1996). The structural core of the enzyme—the five-stranded highly twisted β-barrel—is formed by β1, β4, β5, β7, and β6; the barrel is capped on the top with α-helix B (Figure 1), and on the bottom with a loop between β-strands

† This work was supported by grants from the Finnish Academy of Sciences (Grants 1444 to A.G. and 3875 to R.L.), from the Suomen Tiedeakatemia (for J.K.), from the NIH (DK13212 and TW00407), from the International Science Foundation and the Russian Government (M2j000 and M2J300 to A.A.B.), and from the Russian Foundation for Basic Research (94-04-12658-a to A.A.B.).

‡ Coordinates have been deposited in the Brookhaven Protein Data Bank under ID codes 2eip (*apo* form) and 1IPW ($Mg_{1.5}$·E-PPase complex).

* Author to whom to address correspondence. FAX: 358-21-6338000. Telephone: 358-21-6338029. Email: Goldman@btk.utu.fi.

§ Centre for Biotechnology.

|| University of Turku.

⊥ University of Pennsylvania.

Moscow State University.

⊗ Abstract published in *Advance ACS Abstracts,* April 1, 1996.

[1] Abbreviations: E-PPase, *Escherichia coli* inorganic pyrophosphatase; Y-PPase, *Saccharomyces cerevisiae* pyrophosphatase, T-PPase; *Thermus thermophilus* PPase; $Mg_{1.5}$·E-PPase, the complex of E-PPase with 1.5 Mg^{2+} ions per monomer; rmsd, average root mean square deviation; P_i, inorganic phosphate; PP_i, inorganic pyrophosphate; MHDP, methane-1-hydroxy-1,1-diphosphonate, $O_3PHCOHPO_3$.

0006-2960/96/0435-4670$12.00/0

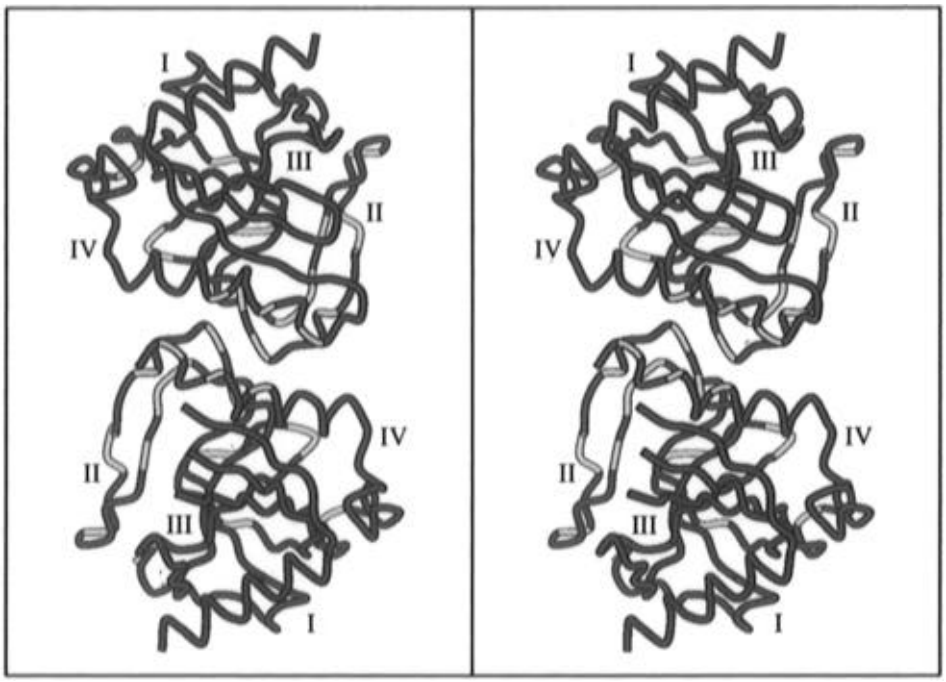

FIGURE 1: Fold of a dimer of *E. coli* PPase, shown in stereo and drawn using Molscript and Raster3D (Kraulis, 1991; Merritt & Murphy, 1994). The overall topology of the enzyme, as a highly twisted β-barrel with excursions, is emphasized. The barrel is in blue; the excursions (labeled with Roman numerals) are in red, and residues in the various oligomeric interfaces are in yellow.

5 and 6. Four longish excursions and a short loop extend from the barrel. It is the top surface of the β-barrel and the excursions that form the active site (Figure 1): 12 out of the 17 putative catalytically important active site residues probed by site-directed mutagenesis (Salminen *et al.*, 1995) reside on the excursions.

A detailed description of the structure at high resolution will be published elsewhere (Kankare *et al.*, 1996); in this paper, we focus primarily on the differences in the enzyme upon metal ion binding and how that affects the oligomeric structure of the enzyme.

EXPERIMENTAL PROCEDURES

Crystallization and Soaking Experiments. Overexpressed, purified E-PPase (Salminen *et al.*, 1995) crystals were grown as previously described (Heikinheimo *et al*., 1995). E-PPase crystallized in two different crystal forms in space group *R*32: a "short-axis form" with cell dimensions of $a = b =$ 111.4 Å and $c = 76.6$ Å (one monomer per asymmetric unit) and a "long-axis form" with cell dimensions of $a = b =$ 110.6 Å and $c = 154.7$ Å (two monomers per asymmetric unit). We obtained crystals from 16% PEG 400 and 140 mM $MgCl_2$ in 100 mM HEPES, pH 7.5 (Salminen, unpublished results), but they were too small for data collection. Long-axis form crystals (cell dimensions $a = b = 110.2$ Å, $c = 156.0$ Å) also grew from 0.8 M sodium citrate and 100 mM HEPES at pH ≈8.0. The conditions were sufficiently similar to encourage us to try soaking the citrate-grown crystals overnight in 16% PEG 400, 140 mM $MgCl_2$, and 0.5 mM inhibitor MHDP in 100 mM HEPES, pH 7.5. We hoped in this way to be able to bind Mg^{2+} to E-PPase; previous soaking and cocrystallization attempts had been unsuccessful, presumably because of the high concentration of salt. Although most of the crystals cracked, some did not. The ones that survived diffracted to 2.3 Å, and the crystal habit was the long-axis form with unit cell lengths of $a = b = 109.3$ Å and $c = 154.3$ Å.

Data Collection, Model Building, and Refinement. All the diffraction data were collected on an RAXIS IIc image plate detector (Sato *et al*., 1992) using Cu Kα radiation. The X-rays were generated using a Rigaku RU200 rotating anode X-ray generator (50 kV and 180 mA) and monochromated with a graphite monochromator. We used the programs DENZO and SCALEPACK (Otwinowski, 1993) for integrating the intensities, which were rather weak beyond 2.2−2.3 Å. We used the the program X-PLOR (Brünger, 1992b; Brünger *et al*., 1987) for structure refinement and map calculation. The models were fitted into the electron density using O (Jones *et al*., 1991). In all refinements the cutoffs used were lowest resolution 8 Å and data with $F_o > 2\sigma(F_o)$.

Identification of Magnesium Ion Bound to E-PPase. Metals could be identified as having bound to E-PPase on the basis of the following criteria. First, difference electron density maps calculated at 2.3 Å with coefficients of ($F_o^{soak} - F_o^{apo}$) and phases from the *apo*-E-PPase model showed two clear peaks: 10σ and 5σ (Figure 2). Such peaks had been missing from all previous electron density maps calculated from data collected on crystals grown in the presence of magnesium ion with or without inhibitor (Salminen, unpublished results). Second, the peaks were not seen when the concentration of Mg^{2+} in the soaking solution was lowered to 10 mM. Third, the coordination around the metal ion refined to roughly octahedral geometry (Figure 2). Fourth, the putative Mg^{2+} ions had quite low temperature factors. Finally, the magnesium−oxygen distances are within the accepted range for water molecules and carboxylates coordinated to magnesium ions: from 1.8 to 2.4 Å. As one of the two peaks (Figure 2) sits on the noncrystallographic 2-fold, its stoichiometry per monomer is 0.5, so the Mg-soaked crystals correspond to an $Mg_{1.5}$·E-PPase complex.

Structure Solution. In brief, the long-axis crystal form model, solved by molecular replacement from the short-axis model (Kankare *et al.*, 1996), was refined with X-PLOR. We kept the noncrystallographic monomers similar by using

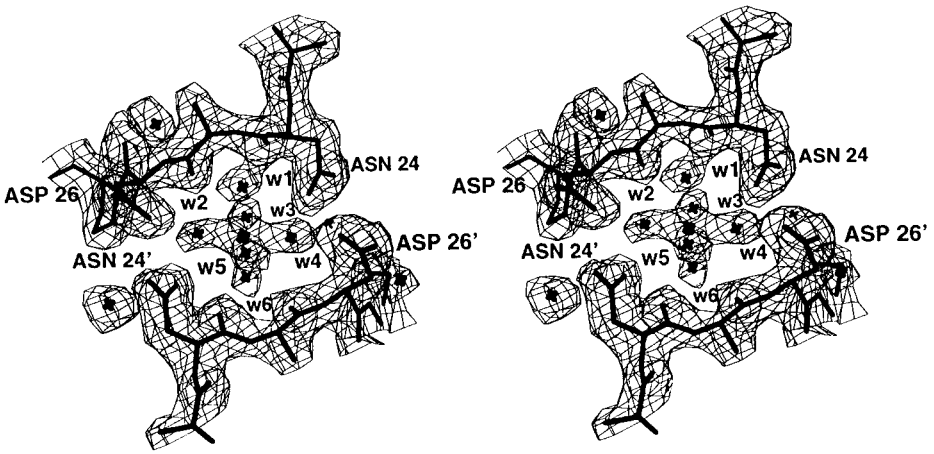

FIGURE 2: ($2F_o - F_c$) electron density map of the trimer−trimer interface region that includes Asn24 and Asp26, showing the density for one of the metal ions in the magnesium-soaked crystals. The electron density is shown in a fish-net representation, water molecules are shown as crosses, and the protein backbone is also shown. The cation is coordinated by six water molecules, w1−w6, which are in turn coordinated by Asn24, Asp26, and the backbone oxygens of Asn24 and Ala25. The noncrystallographically related residues are distinguished by primed and unprimed labels.

Table 1: Diffraction Data and Refinement Statistics

crystal form	resolution range (Å)	no. of reflections ($F_o > 2\sigma$)	no. of crystals	data completeness (%)	*R*-factor (on *I*) (%)
long *c*-axis	8.0−2.2	15787	2	82, 70 (2.3−2.2 Å)	8.2
Mg^{2+} soaked	8.0−2.3	13745	1	79.7, 67.9 (2.4−2.3 Å)	4.8

crystal form	*R*-factor (%)	*R*-free (%)	total no. of atoms	no. of H_2O molecules
long *c*-axis	18.3	23.4 (5)	2841	141
Mg^{2+} soaked	17.1	23.9 (8)	2805	103

crystal form	residues in favorable[a] Ramachandran regions (%)	rmsd[b] bonds (Å)	rmsd angles (deg)	rmsd dihedrals (deg)	av B^c (Å^2)
long *c*-axis	91.8	0.01	1.4	23.4	30.5
Mg^{2+} soaked	91.1	0.01	1.7	24.4	24.4

[a] The Ramachandran values were assigned with PROCHECK and are the percentage of residues in the most favorable regions; no non-glycine values were in the disallowed regions. [b] The deviations from ideality were calculated against the parameter set of Engh and Huber (1991). [c] This overall *B*-value calculation included only protein atoms.

the X-PLOR noncrystallographic symmetry restraint option with a weight of 300, but the monomers were rebuilt individually. We omitted 5% of the reflection data for use in the free *R*-factor calculation (Brünger, 1992a) and used the behavior of the free *R*-factor and simulated annealing omit maps (Hodel *et al*., 1992) to monitor refinement progress and correct errors. The restraints on the monomers were released at 2.2 Å resolution. The conventional *R*-factor converged to a value of 18.3% with a free *R*-factor of 23.4% at 2.2 Å resolution (Table 1). The final model contained 168 of the 175 residues in one of the monomers in the asymmetric unit (monomer II) and 173 of the 175 residues in the other (monomer I). Monomers II do not contain residues 98−99 and 146−148, which are in loop regions and apparently disordered, but monomers I do because of a packing interaction (Kankare *et al.*, 1996).

The refinement of the $Mg_{1.5}$•E-PPase structure was straightforward and produced no surprises. The refined long-axis *apo* model was used directly as the starting model for refinement. A typical refinement cycle was (1) visual inspection of ($F_o - F_c$), ($2F_o - F_c$), and ($3F_o - 2F_c$) electron density maps and manual model building with O (Jones *et al*., 1991), (2) conventional positional and in the later stages *B*-factor refinement with X-PLOR (Brünger, 1992b; Brünger *et al*., 1987), and (3) calculation of new maps and (1) repeated. Three simulated annealing refinement runs were done at 2.7, 2.5, and 2.3 Å, using protocols as in the X-PLOR manual (Brünger, 1992b). There was very little need for manual rebuilding in between these runs. The noncrystallographically related monomers were restrained with an X-PLOR noncrystallographic symmetry restraint weight of 300, and the restraints were released at 2.3 Å. Waters were added to the model once the resolution reached 2.5 Å by building into peaks in the ($F_o - F_c$) difference electron density maps that had heights greater than 3.5σ and that appeared to have reasonable hydrogen-bonding geometry. Waters were omitted if, in subsequent refinement runs, their temperature factor converged to a value greater than 60 Å^2. No atoms were assigned partial occupancy. The final *R*-factor converged to 17.1% and the free *R*-factor (8% omitted) to 23.9% (Table 1). The lower resolution of the data is, we believe, the reason for larger differences between free *R*-factor and conventional *R*-factor than occurred for the *apo* long-axis refinement.

The secondary structure elements were assigned with PROCHECK (Laskowski *et al*., 1993) and the hydrogen bonds with HBLUS (McDonald & Thornton, 1994). Structure analysis was done using O (Jones *et al*., 1991), WHATIF (Vriend, 1990), and X-PLOR (Brünger, 1992b). The images of the models were generated with Molscript (Kraulis, 1991), and the figure of electron density was with O (Jones *et al*., 1991). Superpositions were done with RgbSup (Paula Fitzgerald, unpublished program) because it allows one to *minimize* the number of residues in a superposition

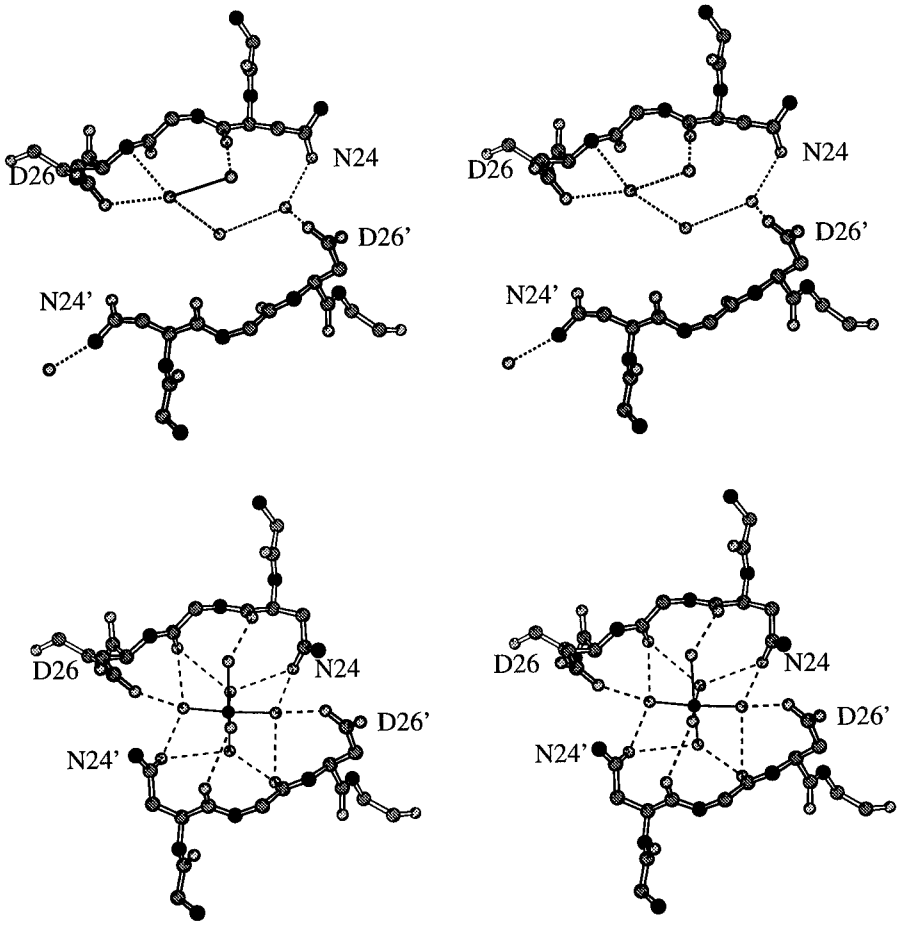

FIGURE 3: (a, top) Asn24−Asp26 interface region of *apo*-E-PPase. Water molecules are shown as balls and hydrogen bonds as dotted lines. The two monomers that contribute to the interface are distinguished by the primed and unprimed labels. (b, bottom) Asn24−Asp26 interface region of $Mg_{1.5}$·E-PPase showing the much more elaborate interface that forms when Mg^{2+} binds. Water molecules are shown as gray spheres, the magnesium ion is shown as a black sphere connected to its six ligands by solid lines, and hydrogen bonds as are shown dotted lines. The two monomers that contribute to the interface are distinguished by the primed and unprimed labels.

using a variable distance cutoff. This enables one to find small concerted changes against a relatively invariant core protein.

RESULTS

Oligomeric Contacts. The E20D, Y55F, H136Q, and H140Q variant proteins all dissociate into trimers much more readily than wild type (Baykov *et al.*, 1995; Salminen *et al.*, 1995; Velichko *et al.*, 1995; Volk *et al.*, 1996), but we have yet to find a single amino acid mutant that produces an E-PPase variant with a changed hexamer−dimer equilibrium. This suggests that E-PPase is best thought of as a dimer of trimers. In the long-axis crystal form discussed here, with two molecules in the asymmetric unit, the intertrimeric contacts occur around a noncrystallographic 2-fold axis that relates the two trimers of the physiologically relevant E-PPase hexamer (Kankare *et al.*, 1994); the intratrimeric contacts occur around the crystallographic 3-fold axis relating monomers in the trimer.

The principal intertrimeric contacts are between two helices A that are part of excursion IV (Figure 1) from 2-fold-related monomers (Baykov *et al.*, 1995; Kankare *et al.*, 1996). The interaction involves a three-center ionic and hydrogen-bonding interaction, His140−Asp143−His136′ (where His136′ is on the symmetry-related monomer). The interaction is buttressed with further hydrogen bonds (Kankare *et al.*, 1996).

Another important trimer−trimer interaction occurs between Asn24 and Asp26, part of excursion II. In this region, there are water-mediated hydrogen bonds between Asn24 and Asp26′ (but *not*, in our current model, between Asn24′ and Asp26) in the absence of metal ion (Figure 3a). However, the interface contains a site well-suited for binding divalent cation, which occurred when the crystals were soaked in high (140 mM) concentrations of magnesium ion (see Materials and Methods). The Mg^{2+} ion, which replaced a weakly bound water molecule (Wat51: temperature factor 43), showed clear density in the ($F_o^{soak} - F_c^{apo}$) difference electron density maps (Figure 3b) and was approximately octahedrally coordinated to six water molecules. Five of the distances were in the range 1.8−2.1 Å, while one was rather longer, at 2.7 Å. Under these conditions, the interface rearranges subtly (Figure 3). The side chains of the Asn24s reorient and become more ordered: their average temperature factor is 18 instead of 32.

As a consequence of these rearrangements, the trimers move relative to each other. When the *apo* monomer II model is superimposed on the Mg-containing monomer II model using RgbSup, the rmsd per α-carbon is 0.27 Å for 167 C_α superimposed, and when the *apo* monomer I model is superimposed on the Mg-containing monomer I model, the rmsd per α-carbon is 0.24 Å for 173 C_α superimposed. However, when monomers I are superimposed using the monomers II transformation matrix or *vice versa*, the rmsd per C_α is 0.49 Å.

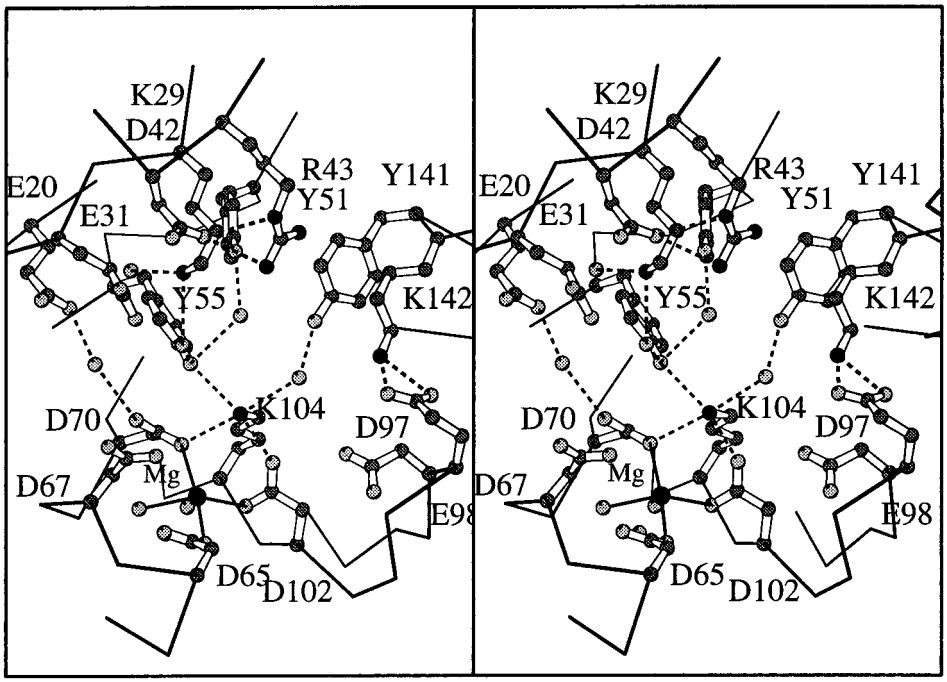

FIGURE 4: Active site of *E. coli* PPase showing the extensive network of hydrogen bonds (dotted lines). Carbon atoms are in dark gray, oxygen atoms in light gray, and nitrogen atoms in black. The Mg^{2+} ion is the larger black sphere connected to the ligands we currently see by solid lines. Parts of the C_α trace are also shown.

This difference is essentially due to a translation of monomer I toward monomer II by about 0.4 Å along *z* (translation vector 0.02 Å, 0.09 Å, −0.42 Å). The motion along −*z* corresponds to a tightening of the trimer−trimer interface and can also be seen in the change in accessible surface area as calculated in the program GRASP (Nicholls *et al*., 1991): in the *apo* structure, the surface area buried upon forming dimers from monomers is 1321 Å^2, whereas in the metal-soaked structure, it increases by 7% to 1414 Å^2.

Structure of the Active Site. The description below is based on the active site of the type I monomers for two, possibly interrelated, reasons: they have density for two loops, 98−99 and 146−148 (see Materials and Methods), and they bind Mg^{2+} more tightly, as evidenced by the temperature factor of the magnesium ion in the active site. The monomer I from the Mg^{2+}-soaked model could be superimposed on the equivalent *apo*-model monomer with an rmsd per C_α of 0.21 Å, indicating that the bound metals induced only slight changes in the overall structure. There are, however, specific changes in the active site geometry (see below).

The magnesium ion in the active site is directly coordinated to Asp65, Asp70, and Asp102 (Figure 4). It therefore corresponds to the tightly bound Mg^{2+} ion which has a binding constant of about 0.06 mM at pH 7.5 [calculated from Baykov *et al.* (1996)]. Consistent with the structure, the D65E and D102E variant proteins have lowered affinities for the tight metal-binding site (Hyytiä, manuscript in preparation). There is a large network of salt bridges and hydrogen bonds, some mediated by water molecules (Figure 4; Table 2). A three-center ionic interaction forms between Lys104, Asp70, and Asp102 that orients Asp70 and Asp102 so that they can bind the Mg^{2+} ion (Figure 4). Lys104 also forms a hydrogen bond with Tyr55, which is in turn hydrogen bonded to Tyr51.

Table 2: Closest Distances (in Å) between Active Site Side Chains of the Mg-Soaked Crystals

	K29	E31	R43	Y55	D65	D67	D70	D97	D102	K104	K142
E20	5.4	3.9	10.4	4.2	11.0	8.3	5.6	13.8	10.9	8.6	13.5
K29		3.9	3.7	4.2	11.2	8.7	7.5	10.3	9.1	7.1	7.5
E31			6.1	7.1	9.3	5.8	6.7	10.9	9.5	8.3	10.5
R43				8.5	12.3	9.4	10.4	9.1	9.7	9.1	4.5
Y55					8.4	7.9	3.3	8.8	5.8	3.0	9.5
D65						3.4	5.2	7.2	3.1	6.6	11.0
D67							5.4	8.5	6.9	7.7	10.6
D70								8.0	3.9	2.9	10.3
D97									4.1	6.4	4.3
D102										3.0	7.2
K104											8.1

Upon binding a Mg^{2+} ion, the active site undergoes a subtle rearrangement. When we used RgbSup (Paula Fitzgerald, unpublished program) with a cutoff of 0.4 Å to compare *apo* monomer I to metal-containing monomer I, 163 of 173 residues were below the cutoff with an rmsd of 0.21 Å per C_α. Ten residues were above the cutoff, with an rmsd of 0.49 Å per C_α. The Asp70 and Asp102 loops contain six of the nonmatching residues; two of the other four occur at the C-terminus, and there are two scattered unmatching residues (Pro27 and Lys34). The results for monomer II were similar. Consequently, the Asp65−Asp70 loop moves in toward the metal ion, as does the loop containing Asp102 (Figure 5). Although the average displacements are small, they are concerted.

The hydrogen-bonding network also rearranges. In the *apo* structure, Tyr55 is at the center of a network of hydrogen bonds that stabilize Asp70 (Salminen *et al.*, 1995) (data not shown). In the $Mg_{1.5}$·E-PPase structure, that role is now filled by Lys104 (Figure 4). This provides a plausible explanation of why the K104R variant shows reduced thermostability and increased Nile Red binding (Kankare *et al.*, 1996). Because of the extensive hydrogen-bonding network of the active site, mutating *any* residue can distort

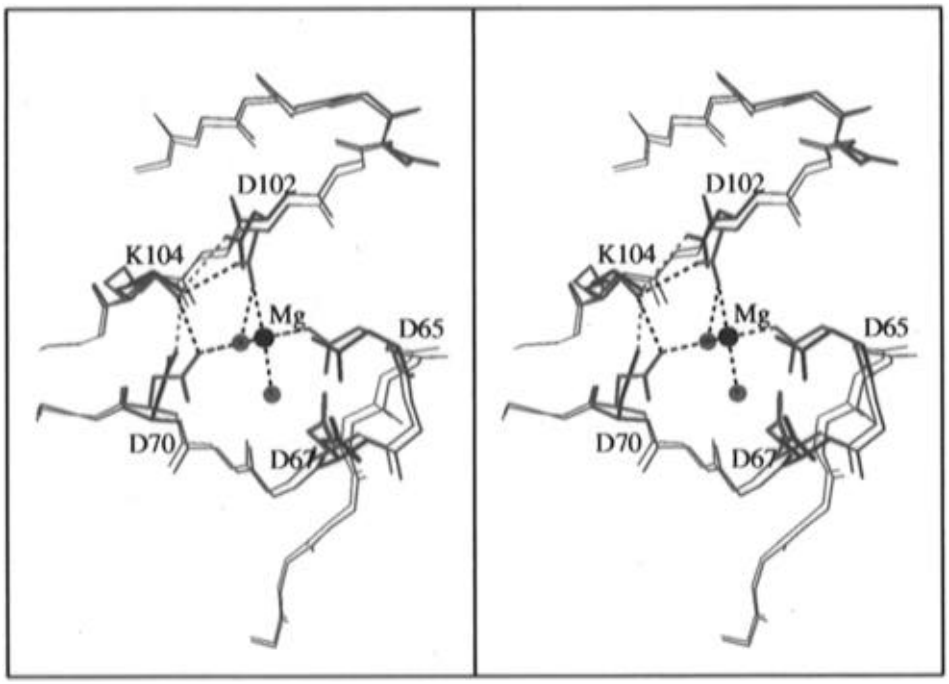

FIGURE 5: Superposition of two active site loops (60−71 and 94−105) of monomer I *apo*- and $Mg_{1.5}$·E-PPase. The superposition was derived using RgbSup (Paula Fitzgerald, unpublished program). The *apo* structure is in blue, and the Mg-soaked structure is in red. The magnesium ion (labeled) in the active site is in black, and the hydrogen bonds and Mg−ligand interactions are shown as dotted red or blue lines, as appropriate. The amino acids previously identified as important (see text) are labeled and their side chains shown as sticks. Not shown is a water molecule in the *apo* structure, which is in much the same position as the water molecule behind and to the right of the Mg^{2+} ion in the $Mg_{1.5}$·E-PPase model.

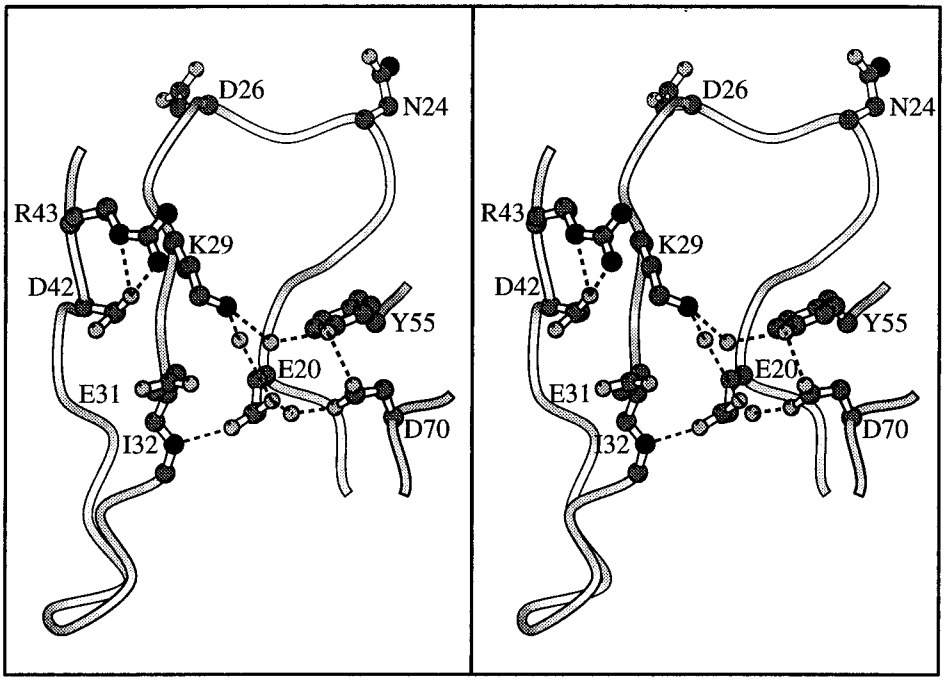

FIGURE 6: Loop between Glu20 and Ile32, drawn using the coordinates from monomer I of the $Mg_{1.5}$·E-PPase model. Also labeled and shown are other active site residues (Lys29, Glu31, Asp42, Asp70, Arg43, and Tyr55) and two residues from the trimer−trimer interface region (Asn24 and Asp26). Hydrogen bonds are shown as dotted lines, nitrogens are in black, carbons are in dark gray, and oxygens are in light gray.

the network, affect metal binding, and so affect the pK_a of the essential general base and acid (Baykov *et al.*, 1996).

In the $Mg_{1.5}$·E-PPase structure, Glu20 interacts with Asp70 through a water molecule (Figure 4). Furthermore, a hydrogen bond from the backbone amide of Ile32 to the carboxylate of Glu20 helps define the conformation of the loop in between (Figure 6) (Salminen *et al.*, 1995). As this region contains Asn24 and Asp26, part of the trimer−trimer interface (Figures 1 and 6), it is not surprising that the E20D variant has reduced hexameric stability. In our 2 Å structure of the Mn_2·Y-PPase complex (Heikinheimo, manuscript in preparation), a second metal ion is bound to Asp120, the equivalent of Asp70, but on the Glu20 side. Assuming this to be the weaker bound metal ion, the E20D variant would show somewhat weaker binding of the second metal ion (3-fold: Table 6; Volk *et al.*, 1996) because the spacing between

the carboxylate group and the metal ion would change. By mass action, magnesium binding could affect the conformation of the Glu20−Ile32 loop and thus the hexamer−trimer equilibrium.

Lys29, Arg43, and Lys142 appear to bind the substrate (Salminen *et al.*, 1995: Heikinheimo, manuscript in preparation: Hyytiä, manuscript in preparation). Teplyakov and co-workers (1994) located a sulfate ion in the active site which they argued bound in the same location in the active site as P_i would have. This sulfate formed ionic and hydrogen-bonding interactions with the T-PPase residues corresponding to E-PPase Lys29 and Arg43. Comparing the E-PPase active site to the picture of the T-PPase active site in Teplyakov *et al*. (1994), it is clear that the conformation of Arg43 in the two structures is different, further suggesting (see Materials and Methods) that neither sulfate, phosphate, nor the inhibitor MHDP is present in either the *apo*-E-PPase model or the $Mg_{1.5}$·E-PPase. Asp42, a residue that is not conserved between eukaryotic and prokaryotic PPases (Kankare *et al.*, 1994), may orient Arg43 correctly in the active site cavity and also appears to be different in the T-PPase and E-PPase structures.

The role of Glu31, Asp67, Asp97, and Glu98 is less clear from our models. The electron density for Glu31 was weak, whether Mg^{2+} was present or not. Glu98 was disordered in monomer II but not in monomer I, where the loop from 92 to 101 is stabilized by an intramolecular salt bridge between Glu98 and Lys142.

DISCUSSION

In a difference Fourier synthesis calculated with coefficients $F_o^{soak} - F_c^{apo}$ and using phases from our refined long-axis model, we were able to identify two strong peaks in places where Mg^{2+} could be expected to bind and with reasonable octahedral geometry. We were thus able to locate magnesium bound to the crystals—something that has hitherto not been achieved for a bacterial PPase (Salminen, unpublished observations; Teplyakov, personal communication).

In an accompanying paper, we report equilibrium dialysis experiments indicating the presence of three Mg^{2+} binding sites per monomer in the absence of substrate, with apparent dissociation constants of approximately 0.06−2.6, and 11 mM (Baykov *et al.*, 1996) at the pH (7.5) we employed in the metal-soaking experiments. As already mentioned, the Mg^{2+} that binds to Asp 65, Asp 70, and Asp 102 at the active site presumably corresponds to the site of highest affinity. This site is also seen in our Mn_2·Y-PPase structure (Heikinheimo *et al.*, in preparation) in which one of the Mn^{2+} ions is bound to the three Asp residues (115, 120, and 152) that align with E-PPase residues 65, 70, and 102 (Cooperman *et al*., 1992), in agreement with the results of Chirgadze *et al.* (1991).

However, unlike our Mn_2·Y-PPase structure, no second Mg^{2+} is seen at the active site. It is unclear why we do not see the second metal ion in the active site: a conformational change may occur upon binding that is incompatible with this particular crystal form. If this is so, it might explain why many of the crystals soaked in $MgCl_2$ cracked. Alternatively, a second metal ion may be bound at the active site, but at low occupancy, so that we cannot distinguish it from a water molecule. The reason for the absence of density for the inhibitor MHDP is easier to understand. The dimagnesium salt of MHDP has limited solubility, and so it precipitates in millimolar concentrations of magnesium (Baykov, unpublished observations). Consequently, no MHDP was available to bind to E-PPase.

There is no direct evidence linking the third Mg^{2+} binding site seen on dialysis to the active site, and it is tempting to speculate that it corresponds to the Mg^{2+} bound at the interface to Asn24 and Asp26, as reported here. The difficulty, of course, is the stoichiometry of 0.5, based on three Mg^{2+} per hexamer seen structurally, *versus* the apparent stoichiometry of a third full site seen on dialysis. Several possibilities exist. It may be that the third full site has nothing to do with the Asn24−Asp26 site and is not seen for the same reasons that the second site is not. Alternatively, the binding stoichiometry of 3 may actually reflect four sites having monomer stoichiometries of 1.0:1.0:0.5:0.5, and we are observing only two of these sites. Finally, it is possible that a stoichiometry of 2.5 sites/monomer plus some relatively weak, nonspecific binding could be interpreted as 3.0 sites/monomer.

This ambiguity might be resolved by dialysis experiments measuring Mg^{2+} binding stoichiometry to E-PPase variants at residue 24 or 26 that eliminate the Mg^{2+} binding site identified in this work, assuming no major structural changes on mutation. These variants could also be useful in precisely interpreting the effect of E20D mutation on hexamer stability (Volk *et al.*, 1996). Most important for such stability os the attachment of excursion II (residues 22−53) and excursion IV (residues 110−157) (Figure 1). The trimer−trimer interface around the 2-fold axis is composed of residues from excursion II and residues from excursion IV. Excursion II is bound to the barrel by hydrophobic contacts (Kankare *et al.*, 1996) and by a hydrogen bond between the carboxylate group of the barrel residue Glu20 and the backbone nitrogen of the residue Ile32 (Salminen *et al.*, 1995) (Figure 6). Raising the magnesium concentration could promote hexamer formation in the E20D variant in two ways. First, Y-PPase, with an active site organization that is identical (Heikinheimo, manuscript in preparation), binds a magnesium ion on the Glu20 side of Asp70. Increasing the magnesium concentration in the E20D variant could stabilize the native conformation of the Asp20−Ile32 loop and thus stabilize the excursion II interface. Second, the magnesium bound directly to Asn24 and Asp26 in the loop (Figure 3) creates a more tightly associated hexamer in our structure of wild-type PPase and would be expected to do the same in the E20D variant.

Our crystallographic results further suggest that the magnesium and pH dependence of the kinetics of wild-type and variant E-PPases (Baykov *et al.*, 1996: Hyytiä, manuscript in preparation; Volk *et al.*, 1996) might be at least in part accounted for by the extra 0.5 Mg^{2+} ion per monomer that we see in the interface region. Preliminary calculations with GRASP (Nicholls *et al*., 1991) (data not shown) suggest that the effect of binding a metal ion at the interface can be seen in the electrostatic potential of the active site.

Studies of additional PPase variants both in solution and by crystallography will clearly be important for unraveling detailed aspects of E-PPase structure and function. With refined structures of *apo*-E-PPase (Kankare *et al.*, 1996), $Mg_{1.5}$·E-PPase, and Mg_2·Y-PPase (Heikinheimo, manuscript in preparation) in hand, we are well placed to start studying the crystallography of PPase variants.

REFERENCES

Baykov, A. A., Dudarenkov, V. Y., Kasho, V. N., Käpylä, J., Salminen, T., Hyytiä, T., Cooperman, B., Goldman, A., & Lahti, R. (1995) *J. Biol. Chem. 270*, 30804−30812.

Baykov, A. A., Hyytiä, T., Volk, S. E., Kasho, V. N., Vener, A. V., Goldman, A., Lahti, R., & Cooperman, B. S. (1996) *Biochemistry 35*, 4655−4661.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F. J., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535−542.

Brünger, A. T. (1992a) *Nature 355*, 472−475.

Brünger, A. T. (1992b) *X-PLOR Version 3.1. A System for X-Ray Crystallography and NMR*, Yale University Press, New Haven, CT.

Brünger, A. T., Kuriyan, J., & Karplus, M. (1987) *Science 235*, 458−460.

Chen, J., Brevet, A., Fromant, M., Lévèque, F., Schmitter, J.-M., Blanquet, S., & Plateau, P. (1990) *J. Bacteriol. 172*, 5686−5689.

Chirgadze, N. Y., Kuranova, I. P., Nevskaya, N. A., Teplyakov, A. V., Wilson, K., Strokopytov, B. V., Arutyunyan, E. G., & Khene, V. (1991) *Sov. Phys. Crystallogr. 36*, 128−132.

Cooperman, B. S. (1982) *Methods Enzymol. 87*, 526−548.

Cooperman, B. S., Baykov, A. A., & Lahti, R. (1992) *Trends Biolchem. Sci. 17*, 262−266.

Engh, R. A., & Huber, R. (1991) *Acta Crystallogr. A 47*, 392−400.

Heikinheimo, P., Salminen, T., Cooperman, B., Lahti, R., & Goldman, A. (1995) *Acta Crystallogr. D 51*, 399−401.

Herschlag, D., & Jencks, W. P. (1990) *Biochemistry 29*, 5172−5179.

Hodel, A., Kim, S.-H., & Brünger, A. (1992) *Acta Crystallogr. A 48*, 851−858.

Jones, T. A., Zou, J. Y., Cowan, S. W., & Kjeldgaard, M. (1991) *Acta Crystallogr. A 47*, 110−119.

Kankare, J., Neal, G. S., Salminen, T., Glumoff, T., Cooperman, B., Lahti, R., & Goldman, A. (1994) *Protein Eng. 7*, 823−830.

Kankare, J., Salminen, T., Lahti, R., Cooperman, B. S., Baykov, A. A., & Goldman, A. (1996) *Acta Crystallogr. D* (in press).

Knowles, J. R. (1980) *Annu. Rev. Biochem. 49*, 877−919.

Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.

Lahti, R. (1983) *Microbiol. Rev. 47*, 169−179.

Laskowski, R. A., MacArthur, M. V., Moss, D. S., & Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.

Lundin, M., Baltscheffsky, H., & Ronne, H. (1991) *J. Biol. Chem. 266*, 12168−12172.

McDonald, I. K., & Thornton, J. M. (1994) *J. Mol. Biol. 238*, 777−793.

Merritt, E. A., & Murphy, M. E. P. (1994) *Acta Crystallogr. D 50*, 869−873.

Nicholls, A., Sharp, K. A., & Honig, B. (1991) *Proteins: Struct., Funct., Genet. 11*, 281−296.

Oganessyan, V. Y., Kurilova, S. A., Vorobyeva, N. N., Nazarova, T. I., Popov, A. N., Lebedev, A. A., Avaeva, S. M., & Harutyunyan, E. H. (1994) *FEBS Lett. 348*, 301−304.

Otwinowski, Z. (1993) in *Data collection and processing* (Sawyer, L., Isaacs, N., & Bailey, S. S., Eds.) pp 56−62, SERC, Daresbury Laboratory, Warrington, U.K.

Salminen, T., Käpylä, J., Heikinheimo, P., Goldman, A., Heinonen, J., Baykov, A. A., Cooperman, B. S., & Lahti, R. (1995) *Biochemistry 34*, 782−791.

Sato, M., Yamamoto, M., Imada, K., & Katsube, Y. (1992) *J. Appl. Crystallogr. 25*, 348−357.

Teplyakov, A., Obmolova, G., Wilson, K. S., Ishii, K., Kaji, H., Samejima, T., & Kuranova, I. (1994) *Protein Sci. 3*, 1098−1107.

Velichko, I., Volk, S., Dudarenkov, V., Magretova, N., Chernyak, V., Goldman, A., Cooperman, B., Lahti, R., & Baykov, A. (1995) *FEBS Lett. 359*, 20−22.

Volk, S. E., Dudarenkov, V. Y., Käpylä, J., Kasho, V. N., Voloshina, O. A., Salminen, T., Goldman, A., Lahti, R., Baykov, A. A., & Cooperman, B. S. (1996) *Biochemistry 35*, 4662−4669.

Vriend, G. (1990) *J. Mol. Graphics 8*, 52−56.

BI952637E