Structure and Function Analysis of *Escherichia coli* Inorganic Pyrophosphatase: Is a Hydroxide Ion the Key to Catalysis?†

Tiina Salminen,‡,§,|| Jarmo Käpylä,‡,§,⊥ Pirkko Heikinheimo,§,|| Jussi Kankare,|| Adrian Goldman,|| Jukka Heinonen,§ Alexander A. Baykov,# Barry S. Cooperman,*,⊥ and Reijo Lahti*,§

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

Received July 25, 1994; Revised Manuscript Received October 17, 1994®

ABSTRACT: Using site-directed mutagenesis, we have completed replacing all 17 putative active site residues of *Escherichia coli* inorganic pyrophosphatase (PPase). We report here the production of 11 new variant proteins and their initial characterization, including thermostability, hydrophobicity, oligomeric structure, and specific activity at pH 8. Studies of the pH−rate profiles of 12 variants containing substitutions for potentially essential residues showed that the effect of the mutation was *always* to increase the pK_a of a basic group essential for both substrate binding and catalysis by 1−3 pH units. The D70E variant had the lowest activity at all pHs; the K29R, R43K, and K142R variants also had low k_{cat}/K_m values. The principal effect seen in the other variant proteins was higher and sharper pH optima; their pH-independent k_{cat} and k_{cat}/K_m values changed at most by a factor of 8. Our results suggest that the most likely candidate for the essential basic group affected by all mutations in the active site is a hydroxide ion stabilized by coordination to the essential Mg^{2+} ions. Analyzing our results using the structure recently obtained for *E. coli* PPase [Kankare *et al.* (1994) *Protein Eng. 7*, 823−830] led us to identify a group of residues, centered around Asp70 and including Tyr55, Asp65, Asp67, Asp102, and Lys104, that we believe binds the magnesium ions that are critical for the activity, possibly by stabilizing the essential hydroxide. Others, including Lys29, Arg43, and Lys142, are more spread out and more positively charged. They appear to be involved in binding substrate and product. Tyr55 is also a key part of the hydrophobic core of *E. coli* PPase; when it or residues that interact with it are conservatively mutated, there are changes in the overall structure of the enzyme as assayed by thermostability, hydrophobicity, or oligomeric structure.

The ubiquitous enzyme inorganic pyrophosphatase (EC 3.6.1.1) (PPase) plays an important role in energy metabolism, providing a thermodynamic pull for biosynthetic reactions, such as protein, RNA, and DNA synthesis (Kornberg, 1962). PPase may also have an important role in evolution in affecting the accuracy with which DNA molecules are copied during chromosome duplication (Lahti, 1983; Kukko-Kalske & Heinonen, 1985). PPase has been recently shown to be an essential enzyme both in bacteria (Chen *et al.*, 1990) and in yeast (Lundin *et al.*, 1991).

All known PPases require divalent metal ion, with Mg^{2+} conferring the highest activity. Mg^{2+} has a dual role; free magnesium ion activates the enzyme, whereas magnesium pyrophosphate is the substrate for soluble PPases (Cooperman, 1982; Lahti, 1983). PP_i can also complex with polyamines that can—at least in *Streptococcus faecalis* PPase—perform both of the roles usually played by magnesium (Lahti *et al.*, 1989; Labadi *et al.*, 1991).

The two best studied PPases are those from the yeast, *Saccharomyces cerevisiae*, and *Escherichia coli*. *E. coli* PPase is a homohexameric protein (Wong *et al.*, 1970) containing 175 amino acid residues per subunit (Lahti *et al.*, 1988). Its catalysis of PP_i•P_i equilibration has been analyzed in detail (Baykov *et al.*, 1990). The catalytic mechanisms of *E. coli* and yeast PPases are very similar (Cooperman *et al.*, 1992). Catalysis in the presence of Mg^{2+} proceeds via single-step direct phosphoryl transfer to water without a phosphorylated enzyme intermediate (Gonzalez *et al.*, 1984). The first P_i released from the enzyme contains oxygen from the solvent. Although yeast PPase binds up to four metal ions per subunit, it requires only three for activity (Cooperman, 1982); in contrast, *E. coli* PPase requires four (Baykov *et al.*, 1990).

A divalent metal ion binding cavity has been identified in yeast (Kuranova *et al.*, 1983; Terzyan *et al.*, 1984) and *E. coli* PPases (Kankare *et al.*, 1994). In yeast PPase, it contains 12 polar residues that could interact with bound metal ions and 5 basic residues that could interact with P_i, PP_i, or both (Kuranova *et al.*, 1983; Terzyan *et al.*, 1984). The 2.35 Å structure of the yeast PPase•Mn_3Pi_2 complex (Chirgadze *et al.*, 1991) showed that the geometry at the metal ions is distorted octahedral. The ligands to the three divalent cations are mainly Glu, Asp, and phosphate, and the phosphate ions have different occupancies, consistent with ^{31}P nuclear

† T.S. was supported by Grant 1009 from the Finnish Academy of Sciences to R.L. P.H. was supported by grants from Emil Aaltonen and from the Foundations of Turku University (Turun yliopistosäätiö) to Young Scientists. This work was also supported by a grant from the Finnish Academy of Sciences (11444) to A.G. and by NIH Grant DK13212 to B.S.C. and R.L. A.G. was a Lucille P. Markey fellow and the work was supported in part by the Lucille P. Markey Charitable Trust.

* Authors to whom to address correspondence (B.S.C.: FAX 215-898-2037; telephone 215-898-6330; email cooprman@pobox.upenn.edu. R.L.: FAX 358 21 633 6860; telephone 358 21 633 6845; email reila@sara.utu.fi).

‡ Equal first authors.

§ University of Turku.

|| Centre for Biotechnology.

⊥ University of Pennsylvania.

Moscow State University.

® Abstract published in *Advance ACS Abstracts,* January 1, 1995.

0006-2960/95/0434-782$09.00/0

magnetic resonance data (Hamm & Cooperman, 1978). Depending on the choice of alignment parameters, 14−17 of the putative active site residues described by Terzyan *et al.* (1984) are conserved in sequence alignments of *E. coli* and yeast PPases (Lahti *et al.*, 1990a; Cooperman *et al.*, 1992). Kankare *et al.* (1994) claim that no more than 14 are conserved in the structural alignment. Chemical modification studies of *E. coli* PPase have implicated Lys29 (Komissarov *et al.*, 1987), Glu98 (Raznikov *et al.*, 1992), and Trp149 (Kaneko *et al.*, 1993), as well as two unidentified tyrosines and histidines (Samejima *et al.*, 1988), as being important for enzyme activity. Site-specific mutagenesis studies and the recent 2.7 Å structure (Cooperman *et al.*, 1992; Kankare *et al.*, 1994) provide strong evidence that of this group only Lys29 is directly involved in catalysis, although at least some of the other residues are at or near the active site.

A minimal kinetic scheme, fully accounting for the overall catalysis by *E. coli* PPase of PP_i·P_i equilibration, has been put forward (Baykov *et al.*, 1990; Käpylä *et al.*, 1995). We seek to understand the basis for this catalysis, which accelerates the rate of PP_i hydrolysis by 10^{10} compared to that in solution (Cooperman, 1982). Having cloned and sequenced the *E. coli ppa* gene encoding PPase (Lahti *et al.*, 1988), we analyzed the conservation of functional residues between yeast and *E. coli* PPases (Lahti *et al.*, 1990a). The results made it possible to start probing the active site of *E. coli* PPase by site-directed mutagenesis. By making conservative replacements for 7 of the 17 putative active site residues, we showed that Tyr55, Glu98, Lys104, and Tyr141 are important for the structural integrity of the protein, whereas Asp97 and Asp102 are important for catalytic activity (Lahti *et al.*, 1990b, 1991).

In the present work, we used the *in vitro* mutagenesis technique to make conservative mutations of the 10 other putative active site residues, as well as of Lys34, which is conserved in all 8 known soluble PPases (Cooperman *et al.*, 1992). We purified the mutated PPases to homogeneity and characterized them by their thermostability, hydrophobicity, oligomeric structure, and specific activity at pH 8. pH−rate profiles were determined for the more interesting mutants. Finally, we analyzed both the data obtained on mutants we had made earlier (Lahti *et al.*, 1990b, 1991) and the mutants described here using the recently solved three-dimensional structure of *E. coli* PPase (Kankare *et al.*, 1994).

MATERIALS AND METHODS

Chemicals. Restriction endonucleases, T4 DNA ligase, T4 DNA polymerase, and alkaline phosphatase from calf intestine were from Boehringer Mannheim (Mannheim, Germany). (^{35}S)dATP (0.5 kCi/mmol) was from NEN Du Pont (Boston, MA). T4 polynucleotide kinase, DEAE Fast Flow ion-exchange resin, and Superdex 200 gel filtration resin were from Pharmacia (Uppsala, Sweden), sequenase was from USB (Cleveland, OH), and antibiotics, egg white lysozyme, and low-gelling temperature agarose (type VII) were from Sigma (St. Louis, MO). Nile red was from Eastman Kodak.

Bacterial Strains, Plasmids, and Culture Conditions. We used the following strains of *E. coli* as hosts: JM103 (Messing, 1983) for site-directed mutagenesis and CJ236 (Kunkel *et al.*, 1987) for preparing uracil-containing template DNA. We produced the PPase variants in MC1061/*YPPA1*(Δ*ppa*) (see below). We used pTP1, which contains the intact *E. coli ppa* gene (Lahti *et al.*, 1988), as the donor of the *ppa* gene we manipulated. The plasmids, pMAK705 (Hamilton *et al.*, 1989) and pYWT (Heikinheimo *et al.*, manuscript in preparation), were used to construct MC1061/*YPPA1*(Δ*ppa*). M13mp18 (Messing, 1983) was used as a vector in the mutagenesis, whereas pUC19 (Norrander *et al.*, 1983) was the vector for producing native and mutant PPases.

Strains were grown in 2× YT medium (Rodriguez & Tait, 1983) or on LA plates (Sambrook *et al.*, 1989) with 0.1 mg/mL ampicillin added as required. Phage transductions were plated on H-plates (Sambrook *et al.*, 1989).

DNA Manipulation and Mutagenesis. Restriction endonuclease digestions of the plasmids were performed according to the supplier's directions. Other DNA manipulations, including purification of plasmids, isolation of restriction fragments from low-gelling temperature agarose, dephosphorylation of the linearized vector with calf intestinal phosphatase, ligation, and transformation using $CaCl_2$, were carried out as described by Sambrook *et al.* (1989).

The *E. coli* PPase mutants were made by the site-specific mutagenesis method of Kunkel *et al.* (1987). To confirm the mutations and to ensure that no secondary mutations had arisen, the whole *ppa* gene was sequenced by the method of Sanger *et al.* (1977) using universal primer and four *ppa*-specific primers designed for this purpose (Lahti *et al.*, 1990b).

Construction of the MC1061/YPPA1(Δppa) Strain for Producing Background-Free Samples. The PPase variant samples we used previously were contaminated with a small amount (about 0.5% of total protein) of chromosomal-encoded wild-type PPase that complicated analysis (Lahti *et al.* 1991; Käpylä *et al.*, 1995). To produce background-free PPase variants, we replaced the chromosomal *ppa* gene in *E. coli* strain MC1061 (Sambrook *et al.*, 1989) with the *S. cerevisiae PPA1* gene (Kolakowski *et al.*, 1988) using a variation of the Hamilton *et al.* (1989) method so that a second gene is carried into the chromosome when the original is replaced. Into the *Sma*I site of pMAK705 (Hamilton *et al.*, 1989), we inserted a partially deleted, totally inactive fragment of the *E. coli ppa* gene (Lahti *et al.*, 1988), the *Pst*I−*Bgl*II fragment, which lacks the 5′ flanking region and the first 80 codons of the coding region. Next to this fragment we inserted the intact coding region of *S. cerevisiae* PPA1 (the *Eco*RI fragment of the pYWT plasmid; Heikinheimo *et al.*, manuscript in preparation) under the control of the *tac* promoter to ensure high levels of expression of *PPA1* in *E. coli*.

pMAK705 is temperature sensitive for DNA replication (Hamilton *et al.*, 1989); so, when MC1061-containing pMAK705 are grown at the nonpermissive temperature upon appropriate media, only bacteria in which the plasmid has already integrated into the chromosome survive (Hamilton *et al.*, 1989). At this point, with our modified pMAK705 plasmid, the yeast *PPA1* gene had been carried into the chromosome by the partially deleted *E. coli ppa* gene. There are two possible outcomes when the plasmid resolves out, depending on where the recombination takes place: either the plasmid acquires the partially deleted *E. coli ppa* gene and the yeast *PPA1* gene or else the chromosome does. The latter, MC1061/*YPPA1*(Δ*ppa*), is the strain we wished to obtain.

Table 1: Initial Characterization of *E. coli* PPase Variants

	aligned residues[a]							
	Y-PPase	E-PPase (sequence)	E-PPase (structure)	E-PPase variant	activity[b] (%)	decrease in thermostability[f]	increase in dye binding[f]	change in oligomeric structure[f]
putative active site residues[c]	E48	E20	E20	E20D	24	yes	yes	yes
	K56	K29[d]	K29	K29R	0.4	no	yes	no
	E58	E31	E31	E31D	13	yes	no	no
	R78[d]	R43	R43	R43K	15	no	no	no
	Y89[d]	Y51	Y51	Y51F[g]	64	no	no	no
	Y93	Y55	Y55	Y55F[b,g]	10	no	yes	yes
	D115	D65	D65	D65E	10	yes	no	no
	D117	D67	D67	D67E	2	no	no	no
	D120	D70	D70	D70E	0.01	no	no	no
	D147	D97	chain break	D97E[g]	20	no	no	no
				D97V[g]	0.5	no	no	no
	E148	E98[d]	chain break	E98V[g]	33	yes	yes	no
	E150[d]	G100	chain break	E101D[e]	110	no	no	no
	D152	D102	D102	D102E[g]	5	no	no	no
				D102V[g]	0.35	no	no	no
	K154	K104	K104	K104R[g]	6	yes	yes	yes
				K104I[g]	0.1	no	yes	no
	Y192	Y141	Y141	Y141F[g]	22	yes	no	no
	K193	K142	K142	K142R	19	no	no	no
	K198	K148	G147	K148R	100	no	no	no
other conserved polar residues	K61	K34	K34	K34R	25	no	no	no

[a] The sequence and structural alignments have been taken from Lahti *et al.* (1990a) and Kankare *et al.* (1994), respectively. E refers to *E. coli* and Y to yeast. [b] The activities are relative to that of the wild-type enzyme, which has 100% activity. The values presented correspond essentially to relative k_{cat} at pH 8.0. The values are background free, unlike those previously reported. [c] Placed at the active site on the basis of the crystal structure of yeast PPase (Terzyan *et al.*, 1984). [d] Placed at the active site by chemical modification studies [reviewed in Kankare *et al.* (1994)].[e] In a slightly modified sequence alignment, E101 in *E. coli* PPase aligns with E150 in yeast PPase (Lahti *et al.*, 1990a). [f] Yes means that the structural change induced by the mutation was drastic, and no means a slight or no change in structure. [g] Initial characterization of these variants, except for changes in oligomeric structure, has been presented previously (Lahti *et al.*, 1990a, 1991).

MC1061/*YPPA1(Δppa)* produced yeast, not *E. coli* PPase, based on the similarity of the isolated protein to the yeast, rather than *E. coli*, enzyme with respect to thermostability and pH−rate profile, elution from DEAE Fast Flow ion-exchange (see below) and Superdex 200 gel filtration columns, and subunit molecular weight, as estimated by SDS−PAGE analysis.

Purification of PPase Variants. We used the MC1061/*YPPA1(Δppa)* strain as a host for producing the PPase variants described here. The cells were grown as previously described (Lahti *et al.*, 1990b) and disrupted by French press in 25 mM Tris-HCl buffer, pH 8.4, containing 15 mM $MgCl_2$ (about 1 g of cells by wet weight/mL of buffer). After the cell debris had been removed by centrifugation, the supernatant was applied to a DEAE Fast Flow ion-exchange column to separate the plasmid-encoded *E. coli* PPase variants from the chromosome-encoded *S. cerevisiae* PPase. The latter did not bind to the column and appeared in the wash, while the former eluted from the column at about 0.2 M NaCl when a linear salt gradient of 0−0.5 M was used. After the ion-exchange chromatography, we increased the $MgCl_2$ concentration to 50 mM and heat-treated the samples for 20 min at 70 °C. Denatured impurities were pelleted, and the supernatant was dialyzed against 50 mM Tris-HCl buffer, pH 8.0, containing 10 mM $MgCl_2$. The purity of the enzyme preparations was checked by native and SDS−PAGE using PhastSystem (Pharmacia, Uppsala, Sweden) 10%−15% gradient gels, which were stained with Coomassie brilliant blue.

Enzyme Assay. PPase activities were assayed as previously described by Heinonen and Lahti (1981) or by continuous recordings of P_i liberation obtained with an automatic P_i analyzer (Baykov & Avaeva, 1981) as described by Käpylä *et al.* (1995). The method for calculating ionization constants and pH-independent parameters is also described in Käpylä *et al.* (1995). Protein concentration was determined using the Bradford (1976) assay.

Oligomeric State and Hydrophobicity. The fluorescence dye 9-(diethylamino)-5*H*-benzo[α]phenoxazin-5-one (Nile red) was used as a probe to monitor changes in surface hydrophobicity (Sackett & Wolff, 1987) of wild-type and mutant PPases (Lahti *et al.*, 1990b). Changes in quaternary structure were detected by the mobility of variant PPases on native PAGE PhastSystem 8%−25% gradient gels.

Thermostability. Protein was incubated at different temperatures for 15 min in the absence and presence of 5 mM magnesium at pH 8.0. The thermostability of the enzyme was assayed by measuring the percentage of activity left after incubation as a function of temperature, except for inactive variants, where the thermostability was assayed by native PAGE of the protein before and after heat treatment using PhastSystem 10%−15% gradient gels.

Structural Analysis. The structure of *E. coli* PPase was examined using the programs O (Jones *et al.*, 1991) and FRODO (Jones, 1978; Pflugrath *et al.*, 1985) running on Evans and Sutherland ESV-30 and PS390 graphics workstations.

RESULTS

Initial Characterization. Analysis of the conservation of functional residues between yeast and *E. coli* PPases suggested that there are 17 putative active site residues in *E. coli* PPase (Lahti *et al.*, 1990a). We started protein engineering on *E. coli* PPase by making conservative amino acid replacements in 7 of the 17 putative active site residues (Table 1) (Lahti *et al.*, 1990b, 1991). In the present work, the 10 remaining putative active site residues were replaced by oligonucleotide-directed mutagenesis, giving the variant proteins E20D, K29R, E31D, R43K, D65E, D67E, D70E, E101D, K142R, and K148R (Table 1). In addition to the

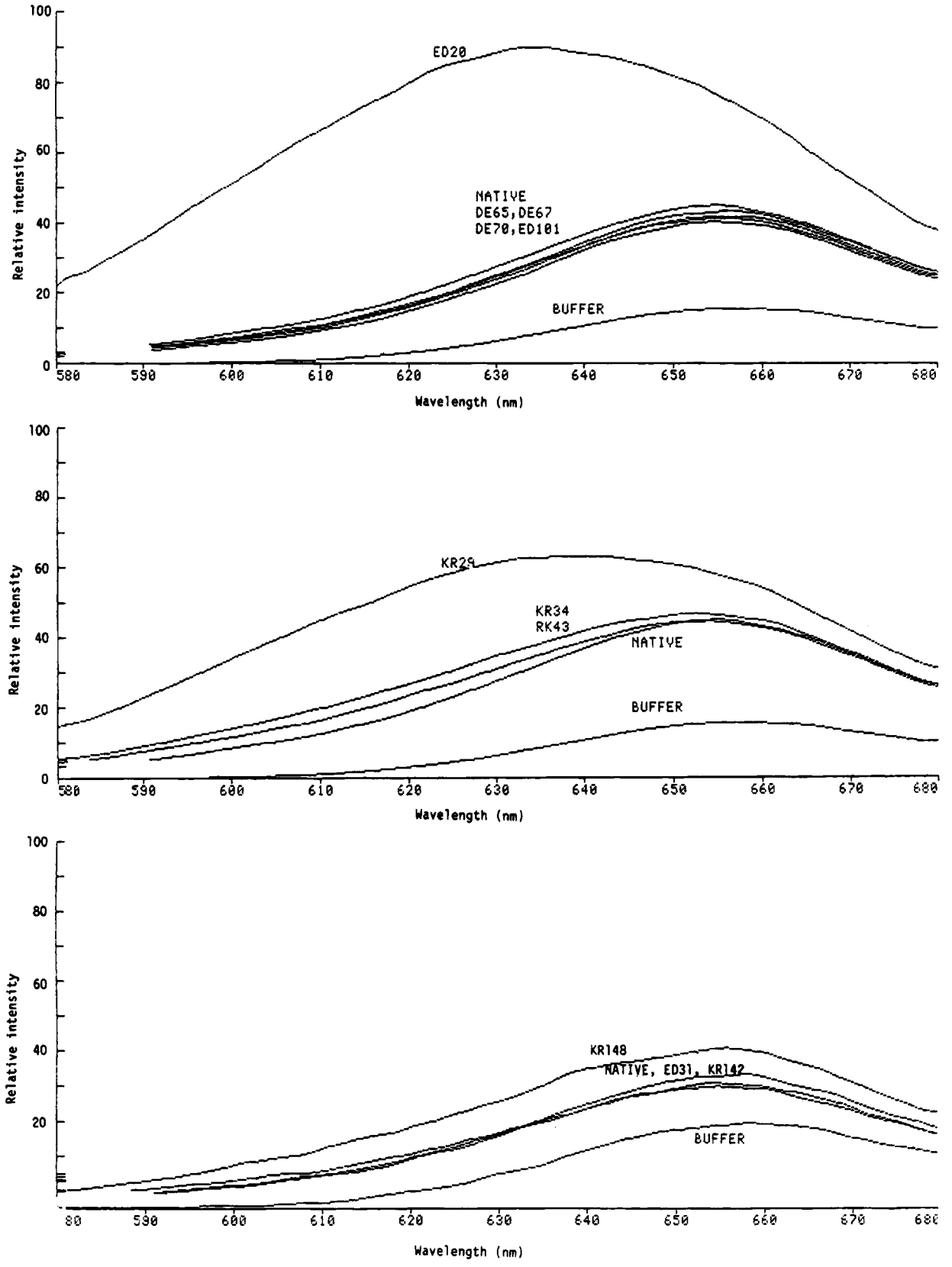

FIGURE 1: Fluorescence emission spectra of Nile red in the presence of the native and mutant PPases. The fluorescence emission spectra in the absence of proteins is also shown (buffer: 0.05 M Tris-HCl, pH 8.0). In each case the concentrations of proteins and Nile red were 0.5 mg/mL and 1 μM, respectively.

17 putative active site residues, Lys34 is conserved in all 8 known soluble PPases (Cooperman *et al.*, 1992) and so may also have an important role in catalysis. Therefore we replaced it with arginine (Table 1).

Using Nile red (Sackett & wolf, 1987) to monitor changes in surface hydrophobicity, we found large increases in the E20D, K29R, Y55F, E98V, and K104R/I variants (Table 1, Figure 1). In addition, the E20D, Y55F, and K104R

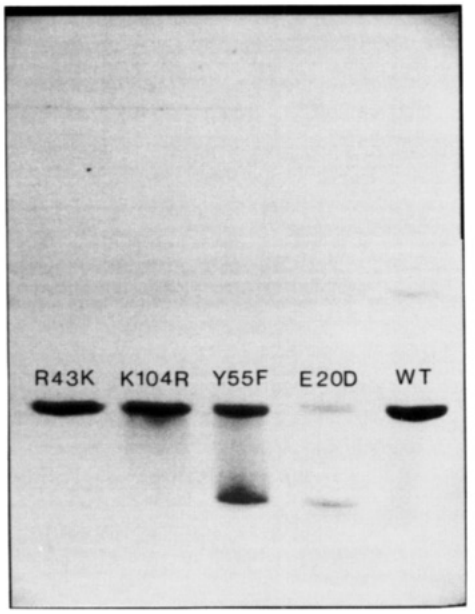

FIGURE 2: Mobility of *E. coli* PPase variants on native PAGE. PhastSystem gradient gels (from 8% to 25%) were used to monitor changes in quaternary structure.

substitutions led to partial dissociation of the hexamer characteristic of wild-type enzyme (Figure 2) while the E20D, E31D, D65E, E98V, K104R, and Y141F variants were all less thermostable than wild type (Table 1, Figure 3). The addition of Mg^{2+} increased the stability of wild-type and variant PPases (Figure 3).

Enzyme activity was measured at pH 8, somewhat to the acidic side of the optimum for *E. coli* PPase (Baykov *et al.*, 1990) and at a substrate (Mg_2PP_i) concentration much larger than K_m for wild-type PPase. The results (Table 1) present strong evidence that Glu98 is not essential for activity because its nonconservative replacement with Val results in a PPase with substantial activity. Furthermore, the less than 2-fold change in activity in the Y51F, E101D, and K148R variants suggests that Tyr51, Glu101, and Lys148 are also unlikely to be essential.

pH−Rate Profiles. Only those variants containing substitutions for potentially essential residues were analyzed further. This eliminated variants containing substitutions for Tyr51, Glu98, Glu101, and Lys148. In addition, the Y141F variant, which has a less than 5-fold reduction in activity (Table 1) and shows a change in structure, making results harder to interpret, was not studied further. pH−rate profile studies were carried out on the remaining 12 putative active site residues and the K34R variant. The results on the E20D variant will be reported elsewhere (Volk *et al.*, manuscript in preparation).

$k_{cat,app}$ and $k_{cat,app}/K_{m,app}$ for *E. coli* PPase and its variants display maxima as a function of pH [Figure 4; see also Käpylä *et al.* (1995)], permitting calculation of the apparent ionization constants for enzyme with substrate bound (pK_{ESH} and pK_{ESH_2}) and in the absence of substrate (pK_{EH} and pK_{EH_2}), as well as pH-independent values for k_{cat} and k_{cat}/K_m. Measurements were carried out at 20 mM Mg^{2+} concentration to minimize the effects of possible changes in Mg^{2+} affinity resulting from amino acid substitution (Käpylä *et al.*, 1995). Values of pK_{EH_2} and pK_{EH} and the pH-independent value of k_{cat}/K_m, determined by fitting pH profiles for k_{cat}/K_m to eq 1, and of pK_{ESH_2} and pK_{ESH} and the

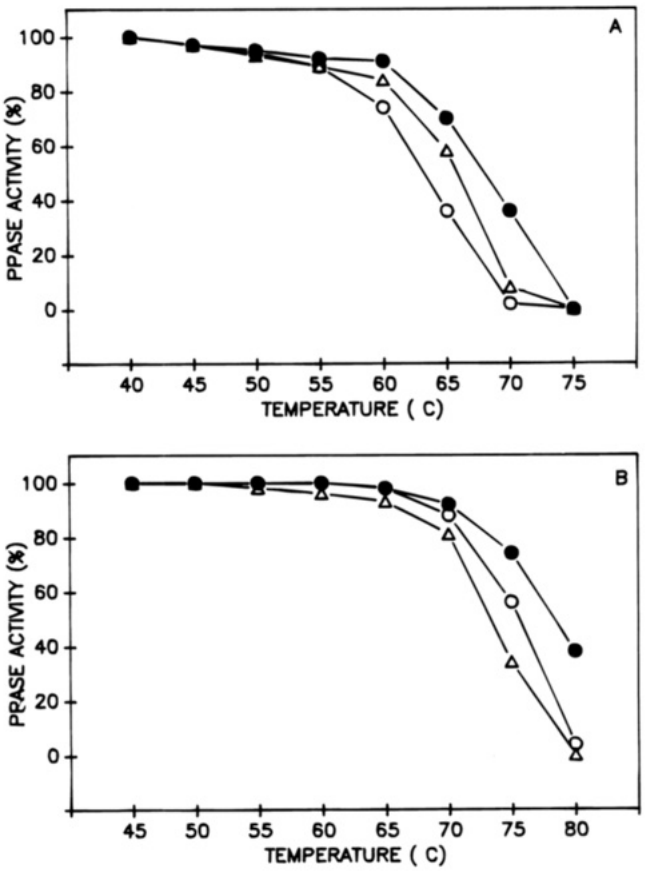

FIGURE 3: Thermal stability of the native *E. coli* PPase and its variants ED31 and DE65 in the absence (A) and presence (B) of magnesium (5 mM). Native and mutant PPases were incubated for 15 min at the temperatures indicated, and after that the residual PPase activity was measured. Symbols: ●, native PPase; ○, ED31; △, DE65.

pH-independent value of k_{cat}, determined by fitting pH profiles for k_{cat} to eq 2, are collected in Table 2. The values

$$k_{cat,app}/K_{m,app} = (k_{cat}/K_m)/(1 + [H^+]/K_{EH_2} + K_{EH}/[H^+]) \quad (1)$$

$$k_{cat,app} = k_{cat}/(1 + [H^+]/K_{ESH_2} + K_{ESH}/[H^+]) \quad (2)$$

shown for the D70E variant are only approximate, since measured activities were just above background. For this variant, the values for k_{cat}/K_m and for k_{cat} are upper limits. In addition, for the variants Y55F and K104R, pK_{EH_2} could not be evaluated separately from pK_{EH}; the pH profiles for k_{cat}/K_m were fit assuming pK_{EH_2} was equal to pK_{EH}, giving values for these parameters, as well as for k_{cat}/K_m, that are less reliable than is the case with the other variants. Similarly, for K104R, pK_{ESH_2} could not be distinguished from pK_{ESH}, and these parameters, as well as k_{cat}, are also less reliable.

Despite the uncertainties mentioned above, the parameter values collected in Table 2 do permit the following conclusions to be drawn. First, every substitution increased the pK_a of a basic group essential for k_{cat} (by 1.0−2.9 pH units) and for k_{cat}/K_m (by >1.2−2.9 pH units). In general, the relative size of the effect on pK_{ESH_2} mirrored that on pK_{EH_2} for a given substitution: largest for K29R and smallest for K34R. This result is consistent with the notion that pK_{ESH_2} and pK_{EH_2} measure the ionization constant of the same basic group in both the presence and absence of substrate. Second,

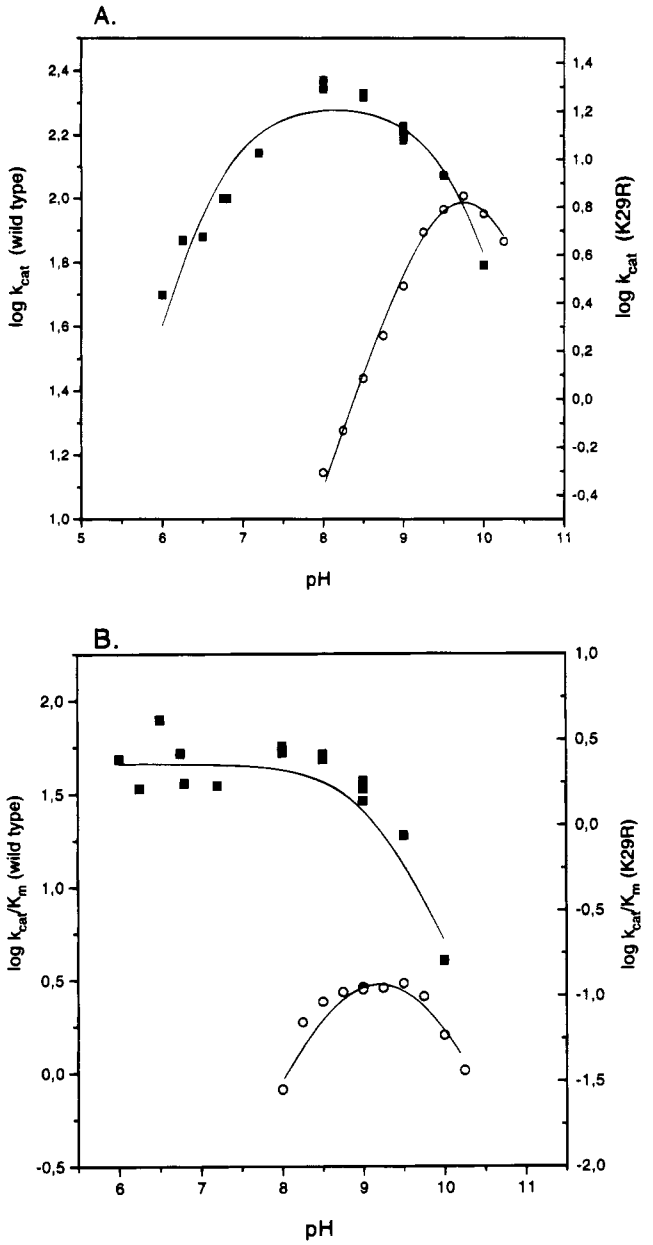

FIGURE 4: Dependence of $k_{cat,app}$ (A) and $k_{cat,app}/K_{m,app}$ (B) on pH for wild-type PPase (■) and for the K29R variant (○). Experimental conditions: 20 mM $[Mg]_{free}$ and $NaPP_i$ (2−350 μM) in 120 mM buffer (MES, HEPES, Tris, or 2-amino-2-methyl-1,3-propanediol) at desired pH. Substrate concentrations were calculated for Mg_2PP_i (Käpylä *et al.*, 1995). Lines are drawn to eqs 1 and 2 using parameters listed in Table 2.

no variant had as large a change in pK_{ESH} or pK_{EH}; the maximum change was 1.0 pH unit, but most changes were considerably smaller. Third, most of the variants had pH-independent values of k_{cat} and k_{cat}/K_m ≥20% of wild-type values. For such variants, the low relative activity at pH 8 (Table 1) can be attributed chiefly to the increase in the pK_a of the essential basic group.

Only two variants, K29R and D70E, had a very large (>10-fold: Table 2) reduction in pH-independent k_{cat}. K29R substitution, however, leads to increased Nile red binding (Table 1) so that the change in k_{cat} may not be directly caused by the substitution but by induced structural changes. With that caveat, Lys29 and Asp70 appear to be the most important residues for determining the speed of enzyme-catalyzed reaction. In addition, K29R, R43K, and K142R showed

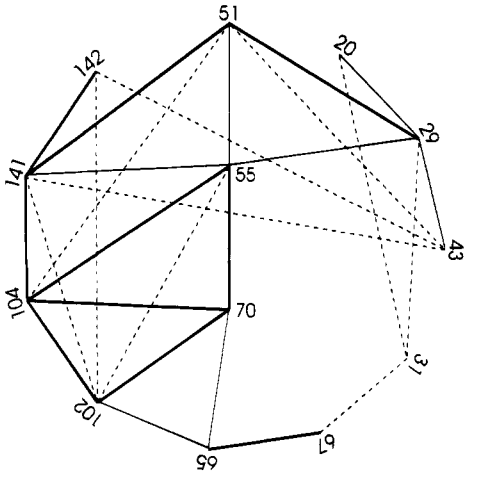

FIGURE 5: Schematic drawing showing the closest distances between side chains in the active site. The other residues appear to be arranged around Asp70 and Tyr55. Thick solid lines join residues with side-chain atoms less than 5 Å apart; thin solid lines, residues with side-chain atoms 5−6 Å apart; dashed lines, residues with side-chain atoms 6−7 Å apart. The distances used are the closest ones between any pair of atoms in side chain *j* and side chain *i*.

large increases (>5-fold) in K_m (Table 2). This is consistent with the proposal that they are involved in PP_i and P_i binding (Terzyan *et al.*, 1984; Kankare *et al.*, 1994).

DISCUSSION

What Is the Identity of the Essential Basic Group? As demonstrated in Table 2, with the possible exception of the D70E variant (see above) every conservative substitution we made at the active site increased the pK_a of an essential basic group, whether substrate was bound or not. Making the reasonable assumption, based on a conservation of catalytic mechanism among the variants, that each substitution affects the same essential group leads to the conclusion that the pK_a of this group is highly sensitive to the precise placement of at least 11 different charged or polar side chains in the active site.

What could this group be? In the wild-type enzyme, its pK_a is below 6 without substrate bound and 6.6 with substrate bound. The plausible candidates therefore are a histidine, a lysine, a tyrosine oxyanion or hydroxide having an unusually high conjugate acidity (low pK_a), or a high pK_a aspartate or glutamate. In deciding among these candidates, we have considered results obtained with both *E. coli* and *S. cerevisiae* PPases, as the active sites of the two enzymes appear to be quite similar (Kankare *et al.*, 1994).

We rule out histidine as a possible candidate on the basis of the absence of histidines in the active site (data not shown), the lack of a conserved His residue in soluble PPases, and the substantial retention of enzyme activity when His residues in *E. coli* PPase were substituted by site-specific mutagenesis (Cooperman *et al.*, 1992). Tyrosine also seems unlikely since even the least active YF variant, Y55F, retains some activity (Table 1) despite removal of the putative "essential" hydroxyl group. The deleted hydroxyl might be replaced by an adventitious water molecule, but as argued by Toney and Kirsch (1991) in considering the formally similar Y70F

Table 2: Kinetic Parameters of PP_i Hydrolysis as a Function of pH[a]

	pK_{EH_2}	pK_{EH}	k_{cat}/K_m ($\times 10^6$ M^{-1} s^{-1})[d]	pK_{ESH_2}	pK_{ESH}	k_{cat} (s^{-1})[d]	K_m (μM)
WT	<6	9.1 ± 0.1	46 ± 4 (100)	6.6 ± 0.2	9.7 ± 0.1	200 ± 14 (100)	4.3 ± 0.7
K29R	8.5 ± 0.1	9.8 ± 0.1	**0.17 ± 0.02 (0.4)**	9.5 ± 0.1	10.0 ± 0.2	**14 ± 3 (7)**	**82 ± 23**
E31D[b]	7.9 ± 0.1	9.2 ± 0.1	20 ± 2 (43)	8.3 ± 0.05	10.4 ± 0.15	69 ± 6 (34)	3.5 ± 0.5
K34R[b]	7.2 ± 0.2	8.6 ± 0.15	33 ± 7 (72)	7.6 ± 0.1	>10	52 ± 4 (26)	1.6 ± 0.4
R43K[b]	8.0 ± 0.1	8.9 ± 0.1	**4.7 ± 0.6 (10)**	8.2 ± 0.1	9.4 ± 0.2	100 ± 11 (50)	**21 ± 4**
Y55F	8.6[c]	8.6[c]	44[c] (96)	8.0 ± 0.1	9.5 ± 0.1	48 ± 3 (24)	
D65E	8.9 ± 0.2	9.3 ± 0.2	43 ± 9 (93)	9.0 ± 0.1	10.2 ± 0.1	92 ± 7 (46)	2.1 ± 0.5
D67E	8.0 ± 0.1	9.2 ± 0.1	49 ± 6 (107)	9.0 ± 0.1	10.7 ± 0.6	45 ± 9 (23)	0.9 ± 0.2
D70E	<7[c]	9[c]	**≤0.008 (0.02)**	nd	>9.5[c]	**≤0.1 (0.05)**	
D97E	8.2 ± 0.2	8.6 ± 0.2	65 ± 19 (141)	8.1 ± 0.1	9.9 ± 0.1	123 ± 8 (62)	1.9 ± 0.6
D102E	8.1 ± 0.1	8.8 ± 0.1	5.9 ± 1.0 (13)	8.3 ± 0.1	10.0 ± 0.3	30 ± 5 (15)	5.1 ± 1.2
K104R	8.9[c]	8.9[c]	13[c] (28)	9.35[c]	9.35[c]	230[c] (115)	
K142R[b]	8.0 ± 0.1	9.2 ± 0.1	**3.4 ± 0.4 (7)**	8.7 ± 0.05	9.6 ± 0.15	283 ± 25 (142)	**83 ± 13**

[a] The pH dependencies of $k_{cat,app}$ and $k_{cat,app}/K_{m,app}$ allow calculation of apparent ionization constants for enzyme with substrate bound ($Mg_2EMg_2PP_i$) indicated as ESH and for enzyme lacking substrate (Mg_2E) indicated as EH, as well as pH-independent values for k_{cat} and k_{cat}/K_m, as described by Knight *et al.* (1981) and Käpylä *et al.* (1995). The most drastic effects of pH-independent k_{cat}, k_{cat}/K_m, and K_m are shown in boldface. D102E, *which is unshaded, also shows a much larger change in* k_{cat}/K_m than the other residues. [b] These samples were contaminated with a small amount (about 0.5% of total protein) of endogenous, chromosomal-encoded wild-type PPase (see Materials and Methods). [c] See text. [d] Values in parentheses are percentages.

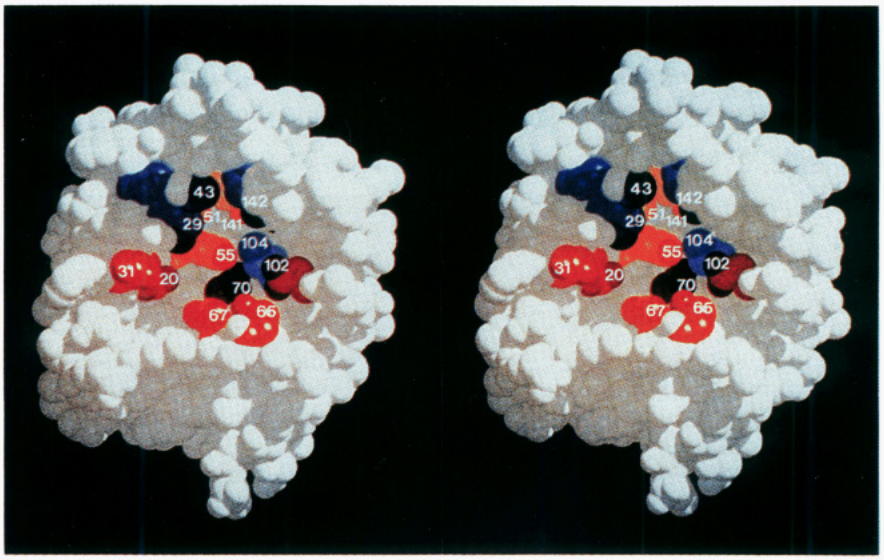

FIGURE 6: Space-filling stereoview of the *E. coli* PPase active site cavity, produced with MOLSCRIPT (Kraulis, 1991) and Raster3D. The residues are colored on the basis of their charge (Glu and Asp, red; Lys and Arg, blue; Tyr, yellow) and apparent importance for activity (Table 2; the more important residues are darker). From left to right, the negatively charged residues are Glu31, Glu20, Asp67, Asp70, Asp65, and Asp102, and the positively charged residues are Lys29, Arg43, Lys142, and Lys104. The structurally important Tyr55 is located deep in the active site cavity. Tyr51 and Tyr141 are behind Arg43 and Lys142, respectively. The most important active site residues are Asp70, Lys29, Arg43, and Lys142.

variant of aspartate aminotransferase, such replacement is unlikely on both entropic and steric grounds.

Alternatively, one of the seven conserved Asps or Glus (Table 1) could have an unusually high pK_a, presumably due to interaction with nearby negative charges; as an example, the second pK_a of cyclopropane-1,1−dicarboxylic acid is 7.43 (*Handbook of Chemistry and Physics*). Such a side chain would have high basicity and so could serve as an effective general base to activate a nucleophilic water to attack PP_i. It would probably not be directly coordinated to a metal ion, because that would lower its pK_a. On the basis of these considerations, one would conclude from the crystal structure of *S. cerevisiae* PPase complexed with manganese phosphate (Chirgadze *et al.*, 1991) that D97 is the most plausible candidate because, of the seven conserved carboxylate residues, yeast D147 (aligning with *E. coli* D97) is the only one *not* coordinated to a Mn^{2+}, yet it is close enough to other carboxylates to have a high pK_a. Nevertheless, neither D97 nor any other active site carboxylate side chain is convincing as the essential base, since an elevated carboxylate pK_a requires a particular active site geometry and it is highly unlikely that distorting the active site geometry by mutation would invariably raise the pK_a of the carboxylate still further. Once would naively expect exactly the reverse.

A lysine of low pK_a is also possible. The most plausible one is Lys29 because substituting it with the much more basic Arg leads, as expected, to a large increase (about 3 pH units) in both pK_{EH_2} and pK_{ESH_2}. However, substituting other residues increases pK_{EH_2} and pK_{ESH_2} by almost as much (Table 2), and we consider it unlikely that the indirect effects of such substitutions on the pK_a of Lys29 would have almost as much effect as directly replacing the lysine with arginine. The same argument would apply to Lys104.

The final, and the most likely, candidate is an hydroxide ion of unusually high conjugate acidity, stabilized by the interaction of its lone pairs with one or two Mg^{2+} ions and possibly Lys or Arg residues as well (coordination of water to one Mg^{2+} lowers its pK_a to 11.4; Baes & Mesmer, 1976).

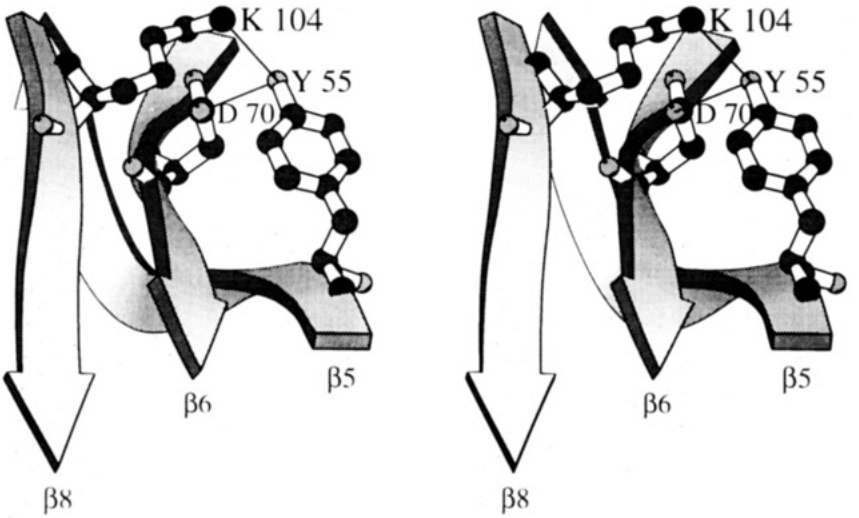

FIGURE 7: MOLSCRIPT (Kraulis, 1991) stereo diagram showing the Asp70−Tyr55−Lys104 hydrogen bonds. The three β-strands are represented as arrows and the hydrogen bonds as lines. Mutations that affect the hydrogen bonds between Asp70−Tyr55−Lys104 affect structure and thermostability.

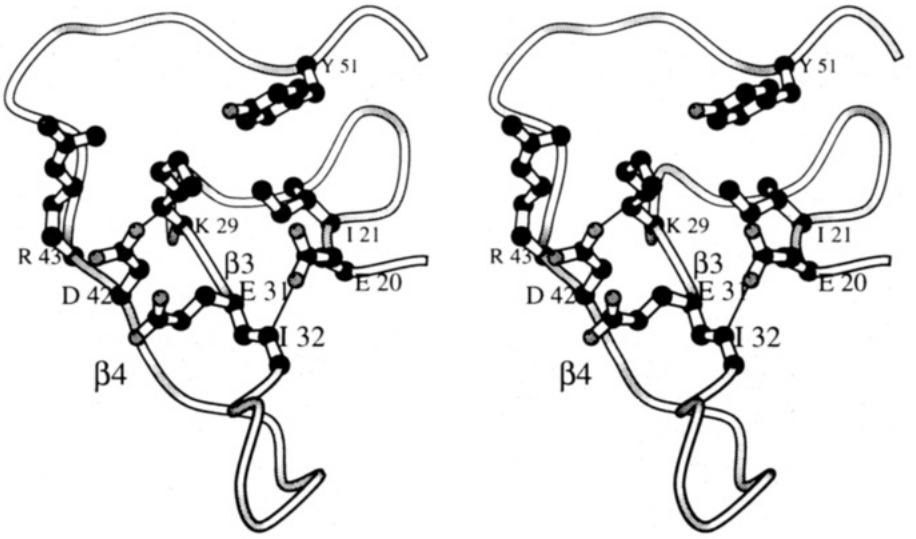

FIGURE 8: MOLSCRIPT (Kraulis, 1991) stereo diagram showing the key hydrogen bonds in the $\beta3-\beta4$ hairpin loop. A piece of secondary structure containing the $\beta3-\beta4$ hairpin loop is drawn as a coil, and the hydrogen bonds are drawn as lines. Four active site residues (Lys29, Glu31, Arg43, and Tyr51) are located in the $\beta3-\beta4$ hairpin loop. Mutations that affect the hydrogen bonds shown affect the structure and/or thermostability.

This possibility is attractive because it predicts that any distortion of the active site that disrupts the stabilizing interactions should lead to an increase in pK_a, as indeed was observed (Table 2). Thus a variant that replaces Arg with Lys or *vice versa* could increase the pK_a of the bound water because a direct interaction could have been disrupted. Similarly, replacing Glu with Asp or *vice versa*, or Tyr with Phe, could decrease the Mg^{2+}-OH_2 interaction, thereby raising the pK_a of the bound water. The hydroxide ion could be a specific base and activate a nucleophilic water molecule for attack on PP_i or it could be the attacking nucleophile itself. An essential Mg^{2+}-bound hydroxide ion in the active site of PPase is consistent with the strong inhibition of PPase by fluoride ion (Kurilova *et al.*, 1984; Baykov *et al.*, 1992), since F− could compete with OH^- for binding to Mg^{2+}.

Comparing the Mutant Data to the Structure. Of the 17 residues originally posited as putatively essential for PPase activity (Lahti *et al.*, 1990a; Cooperman *et al.*, 1992), our present results clearly show that 3 are unlikely to be so: certainly not Glu98 nor, most likely, Lys148 and Glu101. This agrees very well with the recently reported apoenzyme structure of *E. coli* PPase (Kankare *et al.*, 1994), in which these three residues are not structurally aligned with identical residues in the *S. cerevisiae* PPase apoenzyme structure. The remaining 14 residues are all quite close to one another, defining the active site cavity (Figure 5). In the apoenzyme structure of *E. coli* PPase (Kankare *et al.*, 1994), the active site cavity is no more than 7 Å across (Figure 6). A tight group of predominantly negatively charged residues (Tyr55, Asp65, Asp67, Asp102, and Lys104) forms a hydrogen bond network (Kankare *et al.*, 1994) centered around Asp70, which is perhaps the most crucial residue for enzymatic activity. The magnesium ions that activate the water molecule that we believe to be essential for catalysis (see above) may bind here. The other residues (Glu20, Lys29, Glu31, Arg43, Tyr51, Tyr141, and Lys142; Asp97 is, unfortunately, not well located in the apoenzyme structure) are more spread out and include more positively charged residues. They likely form the PP_i binding site, consistent with the large increase in K_m values observed for the K29R, R43K, and K142R variants

(Table 2). Although only Lys29, Arg43, Asp70, and Lys142 appear to be essential *chemically* (*i.e.*, raising the pH does not compensate for the effect of the substitution), Asp67, Asp102, and Lys104 are probably also essential *physiologically*; if they are mutated, the enzyme is inactive at physiological pH and the cell will die (compare the activities shown in Tables 1 and 2).

The current structural model (Kankare *et al.*, 1994) also provides explanations for some of the changes in dye binding and oligomeric structure. Tyr55 seems to be an essential part of the structural core of *E. coli* PPase. Its phenyl ring is in hydrophobic contact with the side chains of Tyr51 and Phe138 which, in turn, pack against Phe137 and Tyr141. Phe137, Phe138, and Tyr141 are also part of helix A that makes important contacts around the 2-fold axis (Kankare *et al.*, 1994). Furthermore, the hydroxyl group of Tyr55 makes hydrogen bonds to both the positively charged ϵ-amino group of Lys104 and the negatively charged γ-carboxylate group of Asp70 (Figure 7). Not surprisingly, therefore, the Y55F variant has altered hydrophobicity and oligomeric structure (Table 1). Conservative mutations in some of the residues that interact directly with Tyr55 *also* show changes (Table 1): the Y141F mutant is less thermostable; the K104R mutant shows changes in all the structural properties we measured.

Other conservative mutations affect the $\beta3-\beta4$ hairpin loop that carries four active site residues (Lys29, Glu31, Arg43, and Tyr51) and is involved in interactions across one of the monomer−monomer interfaces in the crystal structure (Kankare *et al.*, 1994). A hydrogen bond between the δ-carboxylate group of Glu20 and the backbone NH of Ile32 (Figure 8) appears to help position the hairpin loop correctly. In the E20D variant, the hydrogen bond probably forms poorly or not at all. This would make the $\beta3-\beta4$ loop [involved in monomer−monomer contacts (Kankare *et al.*, 1994)] more mobile, thus exposing hydrophobic surfaces within a single monomer and also affecting the oligomeric structure of the protein. The increased Nile red binding could be due to increased exposure of intermonomer surfaces or of intramonomer surfaces; we cannot distinguish which. The effect of the K29R substitution may be similar to that of the E20D substitution: Lys29 is on strand $\beta3$; its $\epsilon-$amino group hydrogen bonds Asp42 on $\beta4$ while the aliphatic portion of its side chain packs against Ile21 in the core of the molecule.

The results reported here provide an important step toward our overall goal of fully characterizing the structure and function of soluble inorganic pyrophosphatases. At present, we have narrowed to 12 the residues that could be catalytically essential in *E. coli* PPase—K34, R43, D67, D70, D97, D102, and K142 and E20, K29, E31, D65, and K104. For the latter five residues, even conservative substitution produces at least some change in PPase structure, thus making it difficult to interpret the functional changes we see purely in terms of alterations within the active site. On the other hand, such interpretations can be made with greater confidence for the former 7 residues. In the accompanying paper, we present a detailed analysis of the effects of D97E substitution on the functional properties of *E. coli* PPase.

ACKNOWLEDGMENT

We thank Mauno Vihinen for helpful discussions and Riitta Eerola, Teppo Hyytiä, Sari Kemppainen, Sari Latonen, Akiko Ohta, Taru Pitkäranta, and Katariina Pohjanoksa for skilled technical assistance.

REFERENCES

Baes, C. F., & Mesmer, R. E. (1976) *The hydrolysis of cations*, John Wiley, New York.
Baykov, A. A., & Avaeva, S. M. (1981) *Anal. Biochem. 116*, 1−4.
Baykov, A. A., Shestakov, A. S., Kasho, V. N., Vener, A. V., & Ivanov, A. H. (1990) *Eur. J. Biochem. 194*, 879−887.
Baykov, A. A., Alexandrov, A. P., & Smirnova, I. N. (1992) *Arch. Biochem. Biophys. 294*, 238−243.
Bond, M. W., Chiu, N. Y., & Cooperman, B. S. (1980) *Biochemistry 19*, 94−102.
Bradford, M. M. (1976) *Anal. Biochem. 72*, 248−254.
Chen, J., Brevet, A., Fromant, M., Leveque, 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 Biochem. Sci. 17*, 262−266.
Gonzalez, M. A., Webb, M. R., Welsh, K. M., & Cooperman, B. S. (1984) *Biochemistry 23*, 797−801.
Hamilton, C. M., Aldea, M., Washburn, B. K., Babitzke, P., & Kushner, S. R. (1989) *J. Bacteriol. 171*, 4617−4622.
Hamm, D. J., & Cooperman, B. S. (1978) *Biochemistry 17*, 4033−4040.
Heinonen, J., & Lahti, R. (1981) *Anal. Biochem. 113*, 313−317.
Jones, T. A. (1978), *J. Appl. Crystallogr. 11*, 268−272.
Jones, T. A., Zou, J. Y., Cowan, S. W., & Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
Kaneko, S., Ichiba, T., Hirano, N., & Hachimori, A. (1993) *Int. J. Biochem. 25*, 233−238.
Kankare, J., Neal, G. S., Salminen, T., Glumoff, T., Cooperman, B., Lahti, R., & Goldman, A. (1994) *Protein Eng. 7*, 823−830.
Käpylä, J., Hyytiä, T., Lahti, R., Goldman, A., Baykov, A. A., & Cooperman, B. S. (1995) *Biochemistry 34*, 792−800.
Knight, W. B., Fitts, S. W., & Dunaway-Mariano, D. (1981) *Biochemistry 20*, 4079−4086.
Kolakowski, F. L., Schlösser, M., & Cooperman, B. S. (1988) *Nucleic Acids Res. 22*, 10441−10452.
Komissarov, A. A., Sklyankina, V. A., & Avaeva, S. M. (1987) *Bioorg. Khim. 13*, 599−605.
Kornberg, A. (1962) in *Horizons in Biochemistry* (Kasha, H., & Pullman, P., Eds.) pp 251−264, Academic Press, New York.
Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
Kukko-Kalske, E., & Heinonen, J. (1985) *Int. J. Biochem. 17*, 575−580.
Kunkel, T. A., Roberts, J. D., & Zakour, R. A. (1987) *Methods Enzymol. 154*, 376−382.
Kuranova, I. P., Terzyan, S. S. Voronova, A. A., Smirnova, E. A., Vainstein, B. K., Höhne, W., & Hansen, G. (1983) *Biororg. Khim. 9*, 1611−1619.
Kurilova, S. A., Bogdanova, A. V., Nazarova, T. I., & Avaeva, S. M. (1984) *Bioorg. Khim. 10*, 1153−1160.
Labadi, I., Jenei, E., Lahti, R., & Lönnberg, H. (1991) *Acta Chem. Scand. 45*, 1055−1059.
Lahti, R. (1983) *Microbiol. Rev. 47*, 169−179.
Lahti, R., Pitkäranta, T., Valve, E., Ilta, I., Kukko-Kalske, E., & Heinonen, J. (1988) *J. Bacteriol. 170*, 5901−5907.
Lahti, R., Hannukainen, R., & Lönnberg, H. (1989) *Biochem. J. 259*, 55−59.

Lahti, R., Kolakowski, L. F., Heinonen, J., Vihinen, M., Pohjanoksa, K., & Cooperman, B. S. (1990a) *Biochim. Biophys. Acta 1038*, 338−345.

Lahti, R., Pohjanoksa, K., Pitkäranta, T., Heikinheimo, P., Salminen, T., Meyer, P., & Heinonen, J. (1990b) *Biochemistry 29*, 5761−5766.

Lahti, R., Salminen, T., Latonen, S., Heikinheimo, P., Pohjanoksa, K., & Heinonen, J. (1991) *Eur. J. Biochem. 198*, 293−297.

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

Messing, J. (1983) *Methods Enzymol. 101*, 20−78.

Norrander, J., Kempe, T., & Messing, J. (1983) *Gene 26*, 101−106.

Pflugrath, J. W., Saper, M. A., & Quiocho, F. A. (1985) in *Methods and Applications in Crystallographic Computing* (Hall, S., & Ashida, T., Eds.) pp 404−407, Oxford University Press, Oxford.

Raznikov, A., Egorov, T. A., Mirgorodskaya, O. A., Sklyankina, V. A., & Avaeva, S. M. (1992) *Biochemistry (Moscow) 57*, 1324−1332.

Rodriguez, R. L., & Tait., R. C. (1983) *Recombinant DNA techniques: an introduction*, p 149, Addison-Wesley, Reading, MA.

Sackett, D. L., & Wolff, J. (1987) *Anal. Biochem. 167*, 228−234.

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) *Molecular cloning. A laboratory manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Samejima, T., Tamagawa, Y., Kondo, Y., Hachimori, A., Kaji, H., Takeda, H., & Shiroya, Y. (1988) *J. Biochem. 103*, 766−772.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463−5467.

Terzyan, S. S., Voronova, A. A., Smirnova, I. P., Kuranova, I., Nekrasov, Y. V., Arutyunyan, E. K., Vainstein, B., Höhne, W., & Hansen, G. (1984) *Bioorg. Khim. 10*, 1469−1482.

Toney, M. D., & Kirsch, J. F. (1991) *Biochemistry 30*, 7456−7461.

Wong, S. C. K., Hall, D. C., & Josse, J. (1970) *J. Biol. Chem. 245*, 4335−4345.

BI941690K