Acta Cryst. (1996). D**52**, 551–563

Structure of *Escherichia coli* Inorganic Pyrophosphatase at 2.2 Å Resolution

JUSSI KANKARE,[a]† TIINA SALMINEN,[a,b]† REIJO LAHTI,[b] BARRY S. COOPERMAN,[c] ALEXANDER A. BAYKOV[d] AND ADRIAN GOLDMAN[a]*

[a]*Centre for Biotechnology, FIN-20521 Turku, Finland,* [b]*Department of Biochemistry, University of Turku, FIN-20500 Turku, Finland,* [c]*Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania, USA, and* [d]*A. N. Belozersky Institute of Physico-Chemical Biology, Moscow State University, Moscow 119899, Russia. E-mail: goldman@btk.utu.fi*

(*Received 20 October 1995; accepted 5 January 1996*)

Abstract

The refined crystal structures of hexameric soluble inorganic pyrophosphatase from *E. coli* (E-PPase) are reported to *R* factors of 18.7 and 18.3% at 2.15 and 2.2 Å, respectively. The first contains one independent monomer; the other, two independent monomers, in an *R*32 unit cell. Because the E-PPase monomer is small with a large open active site, there are relatively few hydrophobic interactions that connect the active-site loops to the five-stranded twisted β-barrel that is the hydrophobic core of the molecule. The active-site loops are, however, held in place by interactions between monomers around the threefold and twofold symmetry axes of the D_3 hexamer. Consequently, mutations of active-site residues (such as Glu20 and Lys104) often affect protein stability and oligomeric structure. Conversely, mutations of residues in the interface between monomers (such as His136 and His140) not only affect oligomeric structure but also affect active-site function. The effects of the H136Q and H140Q variants can be explained by the extended ionic interaction between H140, D143 and H136′ of the neighbouring monomer. This interaction is further buttressed by an extensive hydrogen-bonding network that appears to explain why the E-PPase hexamer is so stable and also why the H136Q and H140Q variant proteins are less stable as hexamers.

1. Introduction

The ubiquitous enzyme soluble inorganic pyrophosphatase (E.C. 3.6.1.1) (PPase) plays an important role in energy metabolism. PPi is the byproduct of many biosynthetic reactions such as protein, RNA and DNA synthesis. The cleavage of PPi thus shifts the equilibrium of these reactions, which have equilibrium constants around unity, towards biosynthesis (Kornberg, 1962). PPase may also have an important role in evolution by affecting the accuracy with which DNA molecules are copied during chromosome duplication (Kukko-Kalske & Heinonen, 1985; Lahti, 1983). PPase has been shown to be an essential enzyme both in bacteria (Chen *et al.*, 1990) and yeast (Lundin, Baltscheffsky & Ronne, 1991).

All known PPases require divalent metal ion, with Mg^{2+} conferring the highest activity. Magnesium has a dual role: free magnesium ion activates the enzyme, while magnesium pyrophosphate (Mg_2PPi) is the substrate for PPases (Cooperman, 1982; Lahti, 1983). PPi also complexes with polyamines that can, at least for *Streptococcus faecalis* PPase, carry out both of the functions that magnesium usually performs (Labadi, Jenei, Lahti & Lönnberg, 1991; Lahti, Hannukainen & Lönnberg, 1989).

The two best studied PPases are those from the yeast, *Saccharomyces cerevisiae* and *Escherichia coli*. *E. coli* PPase (E-PPase) is a homohexameric protein (Wong, Burton & Josse, 1970) containing 175 amino-acid residues per subunit (Lahti *et al.*, 1988). The catalytic mechanisms of E-PPase and yeast PPase (Y-PPase) are very similar (Cooperman, Baykov & Lahti, 1992). Catalysis in the presence of Mg^{2+} proceeds *via* single-step direct phosphoryl transfer to water without a phosphoryl enzyme intermediate (Gonzalez, Webb, Welsh & Cooperman, 1984). The first Pi released from the enzyme contains oxygen from the solvent. From studies on small molecules (Herschlag & Jencks, 1990), it is known that Mg^{2+} can activate water molecules for attack on phosphoester bonds; PPases seem to function in much the same way (Baykov *et al.*, 1996; Salminen *et al.*, 1995). A minimal kinetic scheme, fully accounting for the overall catalysis of PPi:Pi equilibration by E-PPase was proposed (Baykov, Shestakov, Kasho, Vener & Ivanov, 1990) and has been recently improved to explain the dependence of the functional properties of the enzyme both on pH and on [Mg^{2+}] (Baykov *et al.*, 1996). Catalysis by E-PPase proceeds with a total of 3–5 magnesium ions per active site; the exact number required for efficient catalysis is related to achieving the correct protonation state of the enzyme (Baykov *et al.*, 1996).

† These authors have contributed equally to this work.

©1996 International Union of Crystallography
Printed in Great Britain – all rights reserved

The *E. coli ppa* gene encoding for E-PPase has been cloned and sequenced (Lahti *et al.*, 1988), making it possible to study the relationship between structure and function by site-directed mutagenesis. For all 17 conserved polar and charged residues in the active site, a conservative (D–E, K–R *etc.*) replacement was made (Lahti *et al.*, 1990, 1991; Salminen *et al.*, 1995) and the kinetics of the variant enzymes were characterized (Käpylä *et al.*, 1995; Velichko *et al.*, 1995; Volk *et al.*, 1996). Some of the results from studies of E-PPase variants with altered oligomeric structures could be interpreted to suggest that only one trimer per E-PPase hexamer is catalytically active.

The only PPase structure available until recently was an unrefined 3.0 Å yeast apoenzyme structure (Terzyan *et al.*, 1984) [Brookhaven Protein Data Bank (Bernstein *et al.*, 1977) ID code 1PYP]. An initial report of a Y-PPase 2.35 Å X-ray structure with bound cations and substrate has been published but coordinates have yet to be released (Chirgadze *et al.*, 1991). In 1994, three independent X-ray studies of bacterial PPases were published: a refined crystal structure of *Thermus thermophilus* PPase (T-PPase) at 2.0 Å resolution (Teplyakov *et al.*, 1994) (2PRD) and two E-PPase crystal structures at resolutions of 2.5 Å (Oganessyan *et al.*, 1994) (1IGP) and 2.7 Å (Kankare *et al.*, 1994) (1EIP). The monomeric fold of E-PPase, T-PPase and Y-PPase is very similar (Kankare *et al.*, 1994; Teplyakov *et al.*, 1994), but Y-PPase is a homodimer (Terzyan *et al.*, 1984) unlike the bacterial PPases, which are homohexamers (Kankare *et al.*, 1994; Teplyakov *et al.*, 1994; Wong *et al.*, 1970).

Using E-PPase crystals that diffract to high resolution (Heikinheimo, Salminen, Cooperman, Lahti & Goldman, 1995), we can now describe the E-PPase structure in more detail and thus understand better how such a small protein monomer stabilizes a large open active site. We also hoped to find out what degree of conformational variability E-PPase exhibits in the two different crystal forms and thus have a more thorough explanation for some of the functional analyses of variant E-PPases.

2. Materials and methods

2.1. *Crystallization*

Overexpressed, purified E-PPase (Salminen *et al.*, 1995) crystals were grown by hanging-drop vapour diffusion using variants of conditions previously described (Heikinheimo *et al.*, 1995). E-PPase crystallized with cell dimensions of $a = b = 110.2$, $c = 156.0$ Å when the protein, at 15 mg ml^{-1} in 0.1 *M* Tris buffer, pH 8.0, 2 m*M* $MnCl_2$, was mixed 1:1 with and then equilibrated against a well solution containing 0.1 *M* HEPES, pH 7.0, 0.75 *M* sodium citrate, 4 m*M* KH_2PO_4 (final pH of solution $\simeq$ 8). E-PPase crystallized with cell dimensions of $a = b = 111.5$, $c = 76.5$ Å, when the protein at 30 mg ml^{-1} in 25 m*M* Tris buffer, pH 8.0, 4 m*M* $MnCl_2$, was mixed 1:1 with and then equilibrated against a well solution containing 2% PEG 6000, 0.85 *M* Li_2SO_4 and 50 m*M* KH_2PO_4, pH $\simeq$ 6. The former crystal form, the 'long *c*-axis form', contains two monomers of E-PPase hexamer per asymmetric unit; the latter crystal form, the 'short *c*-axis form', contains one monomer of the E-PPase hexamer per asymmetric unit (Heikinheimo *et al.*, 1995).

Table 1. *Data-collection statistics*

Crystal form	Short *c* axis	Long *c* axis
No. of crystals	1	2
Total no. of observations	106709	140425
Independent reflections	12454	16178
$R_{symm}(I)$(%)	7.3	8.2
Average $I/\sigma(I)$		
All data	16.5	23.6
Highest resolution shell	1.74 (2.1–2.0 Å)	3.0 (2.3–2.2 Å)
Completeness (%)		
All data	99.9	86.5
Highest resolution shell	99.7 (2.1–2.0 Å)	72.0 (2.3–2.2 Å)

2.2. *X-ray data*

Data were collected using an R-AXIS IIc image-plate system (Sato, Yamamoto, Imada & Katsube, 1992) and Cu $K\alpha$ radiation from a Rigaku RU200 X-ray generator run at 50 kV, 180 mA made monochromatic with a graphite monochromator. The diffraction data was processed and the intensities were reduced either using the commercial *MSC* software (Sato *et al.*, 1992) or using the program *DENZO* (Otwinowski, 1993). The diffraction data were generally weak beyond 2.2 Å, particularly for the long *c*-axis form (Table 1).

2.3. *Programs*

General crystallographic calculations and data-set scaling were either carried out with the program *PROTEIN* version 3.1 (Steigemann, 1991) or *XTALVIEW* (McRee, 1992). We used the program *X-PLOR* (Brünger, 1992; Brünger, Kuriyan & Karplus, 1987) for refinement and molecular replacement, and the program *O* (Jones, Zou, Cowan & Kjeldgaard, 1991) for fitting the model into the electron density. The secondary-structure elements were assigned with *PROCHECK* (Laskowski, MacArthur, Moss & Thornton, 1993) and the hydrogen bonds with *HBLUS* (McDonald & Thornton, 1994).

2.4. *Structure solution*

Having solved the structure of the long *c*-axis form at 2.7 Å, we decided to solve the structure of the short *c*-axis form at high resolution first because calculation and electron-density fitting would be faster, as there was only one molecule in the asymmetric unit. We, therefore, solved its structure by molecular replacement with *X-PLOR* using a monomer of the 2.7 Å model of

Table 2. *Refinement statistics*

Crystal form	Short *c* axis	Long *c* axis
Resolution range (Å)	8.0–2.15	8.0–2.2
Reflections with $\|F\| > 2\sigma(F)$	8942	15787
R factor on *F* (%)	18.7	18.3
R_{free} on *F** (%)	—	23.4 (5)
No. of non-H atoms	1409	2841
No. of solvent molecules	81	141
R.m.s. deviations†		
Bonds (Å)	0.02	0.01
Angles (°)	2.0	1.4
Planes (°)	25.7	23.4
Residues in allowed regions‡	93%	91.8%
Average *B* factor (Å^2)		
Protein	33.7	30.5
Solvent	45.6	39.9

* No data were excluded at the start of the short *c*-axis refinement, so no free *R* factor was calculated. † The deviations from ideality were calculated using the parameter set of Engh & Huber (1991). ‡ This is the percentage of residues falling in the most favourable regions of the Ramachandran plot, as assigned by *PROCHECK* (Laskowski *et al.*, 1993). No non-glycine residues were in the disallowed regions.

the long *c*-axis dimer E-PPase from our earlier work (Kankare *et al.*, 1994) as a search model. The long *c*-axis form has a local twofold axis parallel to the crystallographic twofold axis and intersecting the crystallographic threefold axis. Consequently, the rotation-function solution corresponded to no rotation at all and the translation-function solution, at 3 Å, produced a clear peak at (0.02, 0.0, 0.059) in fractional coordinates, corresponding to a slight change in *x* due to the change in cell dimensions, and a translation in *z* so that the formerly non-crystallographic twofold axis lay on a crystallographic twofold axis in the short *c*-axis cell. The *R* factor after molecular replacement was 37.2% (data from 15 to 5 Å).

We then refined the structure to diffraction limit in about 30 cycles starting from 3 Å. We used the following refinement protocols as described in the *X-PLOR* manual (Brünger, 1992): conventional positional refinement, simulated-annealing refinement, grouped *B*-factor refinement, and individual *B*-factor refinement. In general, a cycle consisted of two steps: inspection of the electron-density maps and manual model building with *O* (Jones *et al.*, 1991) and refinement with *X-PLOR*.

Refinement was started with a simulated-annealing (slow-cool) refinement followed by positional and grouped *B*-factor refinement at 3 Å. Our protocol for increasing the resolution of the data in the refinement was as follows. After convergence, the resolution was increased so that the number of reflections being refined increased by about 10% at a time. After each stepwise increase in resolution, a positional and grouped *B*-factor refinement was run. Having reached the desired resolution, we ran a slow-cool refinement, and calculated a systematic series of simulated-annealing omit maps by omitting approximately 5% of the residues at a time (Hodel, Kim & Brünger, 1992). This was done to reduce model bias and increase the radius of convergence of the refinement, and we repeated the above protocol after every increase in resolution (*i.e.* at 2.7, 2.5, 2.3 and 2.15 Å). We started adding water molecules to peaks in the $(F_o - F_c)$ difference electron-density maps once the resolution reached 2.5 Å. Our criteria were: difference electron-density peak with height greater than 3.5σ and reasonable hydrogen-bonding geometry. We removed water molecules from the model if their *B*-factors refined to values greater than 60 Å^2. At 2.3 Å, we switched from grouped *B*-factor refinement to restrained individual *B*-factor refinement, with the restraint set to the default *X-PLOR* value. The *R*-factor converged to a value of 18.7% at 2.15 Å resolution; the model has acceptable geometry (Table 2, Fig. 1) and the fit of the model to the map is good (Fig. 2).

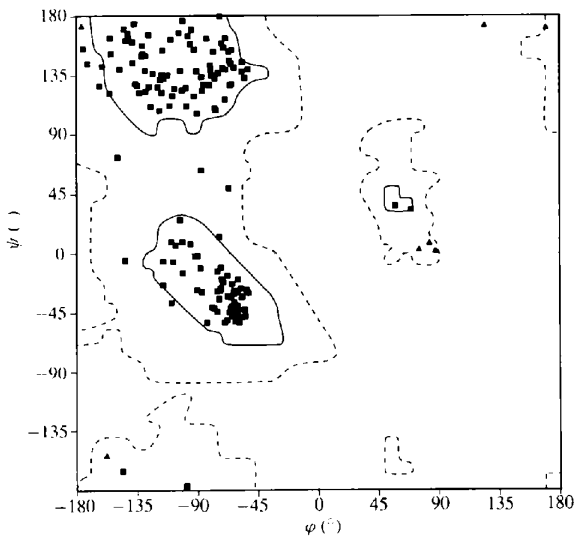


Fig. 1. A Ramachandran plot of the refined short *c*-axis model, drawn using the program *PROCHECK* (Laskowski *et al.*, 1993) with hand-smoothed boundaries based on those of *PROCHECK*. The solid line indicates allowed regions of the Ramachandran plot and the dotted line, additional allowed regions. Glycines are shown as triangles; all other residues, as squares.

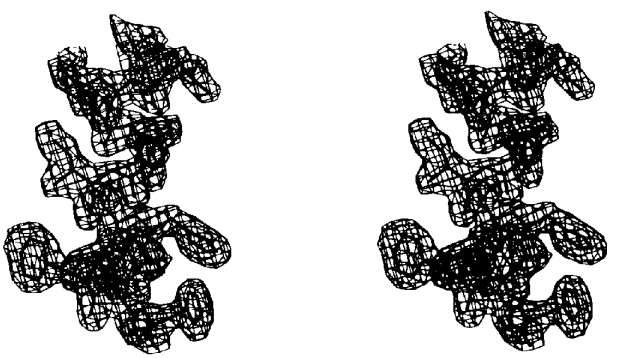


Fig. 2. A stereo drawing of a typical part of the $(2F_o - F_c)$ electron-density map of the short-axis crystal form at 2.15 Å contoured at 1σ showing the fit of the refined model (in black) of E-PPase to the electron density. The region shown is helix *A*.

The short c-axis refined model was used as a starting point for the refinement of the long c-axis form at high resolution rather than continuing the refinement of our original long c-axis crystal form model (Kankare *et al.*, 1994) because the short c-axis model had fewer errors. To minimize the differences between the two models, we constructed a search model by individually superimposing short c-axis monomers on the partly refined 2.7 Å non-crystallographic dimer, one on the 'top' monomer and one on the 'bottom' monomer. This produced a hybrid model: the spacing of the monomers in the dimer came from the 2.7 Å structure but the positions of the atoms in a monomer were from the refined short c-axis model. Because of the slight change in cell dimensions between the original long c-axis crystals and the ones that diffracted to 2.2 Å, we decided to use molecular replacement. This would, we hoped, reduce any bias in our hybrid model towards the monomer arrangement of the original 2.7 Å model. The molecular replacement, performed with *X-PLOR*, yielded a clear solution (R factor 44.3%; data from 15 to 5 Å), after which the individual monomers were refined as rigid bodies. This lowered the R factor to 33.0%.

The refinement of the long c-axis crystal form then proceeded like the refinement of the short form, except that very little manual intervention was needed. During the refinement the non-crystallographic monomers were restrained with an *X-PLOR* n.c.s. weight of 300 but the monomers were rebuilt individually. Waters were added from 2.5 Å on; grouped B-factor refinement was carried out from 2.7 to 2.3 Å; and restrained individual B-factor refinement thereafter. The non-crystallographic symmetry restraints on the monomers were released at 2.2 Å resolution. During the refinement the free R factor (Brünger, 1992) was monitored by omitting 5% of the reflection data. The conventional R factor converged to a value of 18.3% and the free R factor dropped to 23.4%. This model, too, has acceptable geometry (Table 2).

3. Results

3.1. *Secondary structure*

Of the 175 residues in E-PPase, we could model 167 residues of the single monomer per asymmetric unit in the short c-axis form (see *Materials and methods*), 173 residues of the type I monomers of the long c-axis form and 167 of the type II monomers in the long c-axis form (Fig. 3). In all cases we were unable to model the last two residues. In the type I monomers, we were able to model residues 98–100 and 146–148 due to a crystal packing contact which stabilizes them in a specific conformation (Fig. 3). They are in loop regions and otherwise disordered. Apart from this, the overall architecture of the three independent monomer structures determined from the two different crystal forms is very similar. The differences between the three independent monomers which, ignoring the loops, are in the range 0.2–0.25 Å (see below) give a measure of the coordinate error. This is consistent with the nominal resolution of the structures.

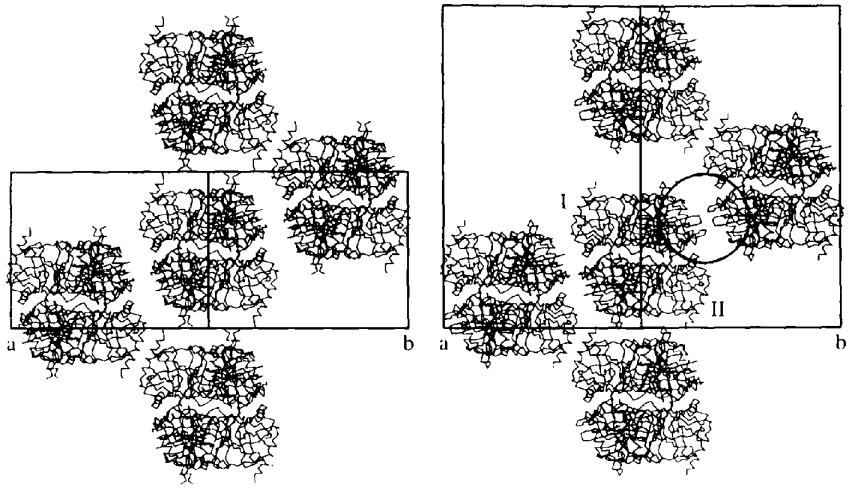


Fig. 3. A diagram showing the packing of E-PPase hexamers in two different crystal forms. Both views are along the $a + b$ dyad axis from the $a + b$ direction towards the origin. The unit cells are represented as solid lines. Five hexamers are drawn on each image and can be identified as the 'hamburger-without-beef' shaped molecules. On the left-hand side the packing of the short c-axis crystal form is shown: the twofold is crystallographic and all trimers are in identical environments. On the right-hand side the packing of the long c-axis form crystals is shown. Here the trimers I and II are in different environments and are related by a non-crystallographic twofold axis at (0, 0, 0.219) (Kankare *et al.*, 1994). The effect of the different packing of the two crystal forms can be seen in the loop regions 97–101 and 148–150 which appeared to be very disordered in the short c-axis crystal form as well as in the trimer II in the long c-axis crystals. On the contrary, this loop region could be traced for the trimer I in the long crystal form. This loop region is circled in the image.

Because the three independent monomers are so similar, the detailed description below applies equally to all, though it has been based on the type I monomers because there the model is most complete. The differences between the different models are discussed in §3.4 below.

The secondary-structure assignments (Table 3, Fig. 4) have changed somewhat from our earlier description (Kankare *et al.*, 1994). Previously, residues 8–10 were assigned as strand β1. Even though residues 8–10 show some β-sheet hydrogen-bonding character (Fig. 4*b*), *PROCHECK* does not assign them as a β-strand because the geometry is not quite good enough. Consequently we assign residues 15–21 as strand β1. There are then altogether eight β-strands and two α-helices, consistent with the T-PPase nomenclature (Teplyakov *et al.*, 1994). There were other changes in the assignments (Table 3); for instance strand β4 (earlier work, strand 5) is now six residues shorter. *PROCHECK* (Laskowski *et al.*, 1993) also identified three single-turn 3_{10}-helical segments: residues 2–4, 115–117 and 123–125. These additional elements, however, did not show characteristic helical hydrogen-bonding patterns and so do not form part of our secondary-structure description (Table 3; Fig. 4*b*).

Table 3. *Secondary structure assignments*

Secondary assignments from *PROCHECK* (Laskowski *et al.*, 1993).

Secondary structure	Residue number	Structural function
Loop	1–14	Excursion I
β1	15–21	β-barrel
Loop	22–27	Excursion II
β2	28–32	
Loop	33–38	
β3	39–44	
Loop	45–53	
β4	54–57	β-barrel
Loop	58–69	Excursion III
β5	70–73	β-barrel
Loop	74–83	β-barrel cap
β6	84–96	β-barrel
Loop	97–101	Short loop
β7	102–109	β-barrel
Loop	110–127	Excursion IV
αA	128–140	
Loop	141–150	
β8	151–156	
αB	158–172	β-barrel cap

3.2. *Monomeric fold*

The topology of the molecule (Kankare *et al.*, 1994) is still best described as a distorted, highly twisted five-stranded β-barrel with excursions (Fig. 4*a*). Strands β1, β4, β5, β7 and β6 form the barrel, which is capped on the top by α-helix *B*, and on the bottom with a loop between strands five and six (Fig. 4*a*). The active site, identified by sequence similarity (Cooperman *et al.*, 1992), structural homology with Y-PPase (Kankare *et al.*, 1994) and solution studies of variant proteins (Salminen *et al.*, 1995) lies in the bowl formed by the extensive excursions. Understanding how they are anchored to the barrel is, therefore, extremely important because 12 of the 17 putative catalytically important active-site residues (Cooperman *et al.*, 1992; Salminen *et al.*, 1995) reside there.

The first excursion, residues 1–14, from the core barrel participates in oligomeric contacts through Ser1 and also makes important hydrophobic contacts. Ala7 and Leu11, in contact with Ile58 (excursion III) and Phe169 (β-barrel cap), form part of the hydrophobic core of the β-barrel. A hydrophobic cluster that Leu2 forms with Val18 and Ile32 (excursion II) docks the second excursion to the barrel.

Lys29 and Arg43, two of the four catalytically important basic residues (Salminen *et al.*, 1995), are in the second long excursion (residues 22–53) between barrel strands β1 and β4. The excursion includes β2 and β3 (Table 3). To maintain correct active-site geometry, this large excursion must somehow be bound to the barrel. This is accomplished by hydrophobic contacts between excursions I and II (see above) and by a hydrogen bond between the carboxylate group of the barrel residue Glu20 and the backbone N atom of the residue Ile32 (Kankare *et al.*, 1996; Salminen *et al.*, 1995). Interactions between helix *A* (part of excursion IV) and the loop between β3–β4 loop that is also part of excursion II stabilize both elements of the active-site framework (Fig. 4*a*). Ser46 is hydrogen bonded to the carboxylate group of Glu145 (data not shown), and Tyr51 and Tyr55 form a hydrophobic cluster with helix *A* residues Phe137, Phe138 and Tyr141 in the classic herringbone manner (Burley & Petsko, 1985) (Fig. 5). Residues Leu39, Val41 and Phe44 participate in oligomeric contacts (see below).

The third excursion (59–68) also includes residues that are extremely important for catalytic activity. Residues 65–72 form a conserved structure/sequence motif diagnostic for all known pyrophosphatases (Kankare *et al.*, 1994). Furthermore, Asp65 and Asp70 bind an essential cation (Kankare *et al.*, 1996). Strand β6 (Table 3) is long, highly twisted and bends by about 90° at Pro89, which does not participate in the hydrogen-bonding pattern of the β-sheet. As a result, the amide N atoms of Arg86 and Val90 also do not form hydrogen bonds. The strand has two functions: its N-terminal part is an essential part of the β-barrel and contributes to the hydrophobic core, while its C-terminal part is hydrogen bonded to β7. Residues Leu93 and Met95 are also in hydrophobic contact with Ile134 and Phe138 of α-helix *A* and thus buttress the active site.

It is difficult to assign any specific structural role to the loop region 97–101 based on the crystal structures alone because it appears very disordered. However, Asp102 and excursion III together contribute to metal binding (Kankare *et al.*, 1996). The linker region (110–127) between strand 7 and helix *A* has no obvious role and, indeed, this region appears not to be very well conserved

in the structural alignment between the E-PPase and Y-PPase structures (Kankare *et al.*, 1994).

α-helix *A* forms an essential part of the active-site cavity wall and makes very important contributions to oligomeric contacts (see below). Strand β7 does not contribute any residues to the active-site cavity but is hydrogen bonded to the C-terminal part of β6 and is in hydrophobic contact with αA. It thus bridges the space between αA and β6 completing the large bowl-shaped active-site cavity (Fig. 4*a*). The polar residues of strand 8 are exposed to solvent and the residues in the region 145–155 (average main-chain *B* factors 56 Å^2) appear to be disordered.

3.3. *Oligomeric contacts*

Three lines of evidence suggest that the D_3 hexameric E-PPase should be considered as a dimer of trimers. Firstly, the accessible surface area buried per monomer on forming the trimer is 1340 Å^2 while the accessible surface area buried per monomer on forming a dimer is 660 Å^2 [calculated using the program *GRASP* (Nicholls, Sharp & Honig, 1991)]. Secondly, the single-amino-acid substitutions E20D, Y55F, K104R, H136Q and H140Q yield partially trimeric variant proteins, whereas no single amino-acid substitution partly dissociates E-PPase into dimers. Thirdly, the interactions around the threefold axis mostly involve backbone hydrogen bonds and hydrophobic interactions; the interactions around the twofold axis mostly electrophilic and involve side chains. The residues participating in contacts are coloured yellow in Fig. 4.

3.3.1. *Intertrimeric contacts.* The principal intertrimeric contact region involves the last two turns of helix *A* and the residues immediately C-terminal to the helix in excursion IV (Baykov *et al.*, 1995). These interact in an antiparallel manner with helix *A*′ and its C-terminal extension from a twofold-related monomer (Fig. 6). Other parts of the interface include the residues Ser46 to Phe50, part of the loop between β3 and β4; and Asn24 to Asp26. The latter region normally interacts *via* a water-mediated hydrogen-bond network but can bind a magnesium ion (Kankare *et al.*, 1996).

Helix *A* and its C-terminal extension contribute three residues from each monomer to form a three-centre ionic interaction that appears to be one of the major determinants of hexamer stability in E-PPase (Baykov *et al.*, 1995). As we have previously described (Baykov *et al.*, 1995; Kankare *et al.*, 1994), His136 of monomer I (*i.e.* His136′) interacts with Asp143 of monomer II, which in turn interacts with His140 from the same monomer. Assuming that one of the two histidines is positively charged (Baykov *et al.*, 1995), we now believe, based on the structure, that the positive charge is primarily on His136. Firstly, His136 and His136′ are more likely to be positively charged than His140 and His140′ because of the electrostatic repulsion if the charge is on the His140's (Fig. 6). Secondly, the rings of the more buried His140's are approximately parallel and 4 Å apart, so that a good hydrophobic interaction can form between them. Thirdly, the structure is consistent with His136 donating a hydrogen bond both to Asp143′ and to a water molecule (Fig. 6). The His136–water hydrogen bond would, of course, help stabilize the three-centre ionic interaction, as would the hydrogen bonds that the water molecule makes in turn to the backbone carbonyl group of Gln133 and the backbone amide group of Ala48′. Two hydrogen bonds, arranged as in an antiparallel sheet, form between the carbonyl O atom of Ala48 and the amide N atom of Phe50′ (and between Phe50 and Ala48′). There is also a hydrogen bond between the side-chain amide of Gln133 and the main-chain

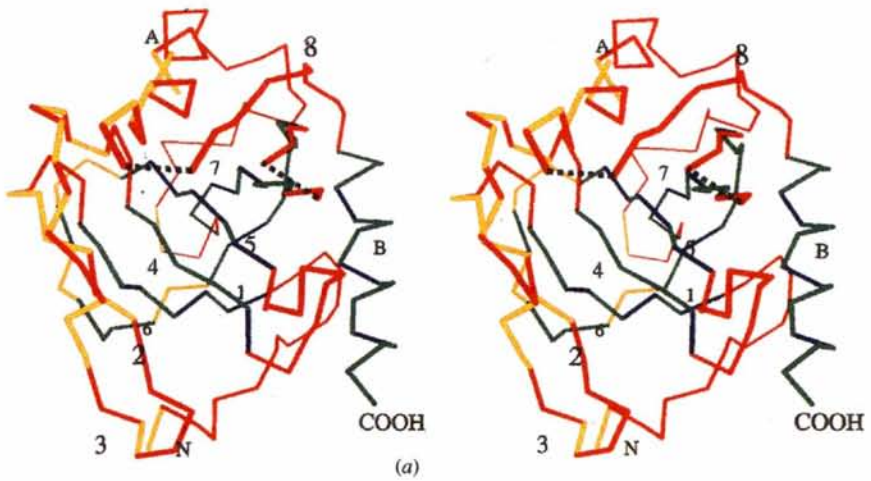
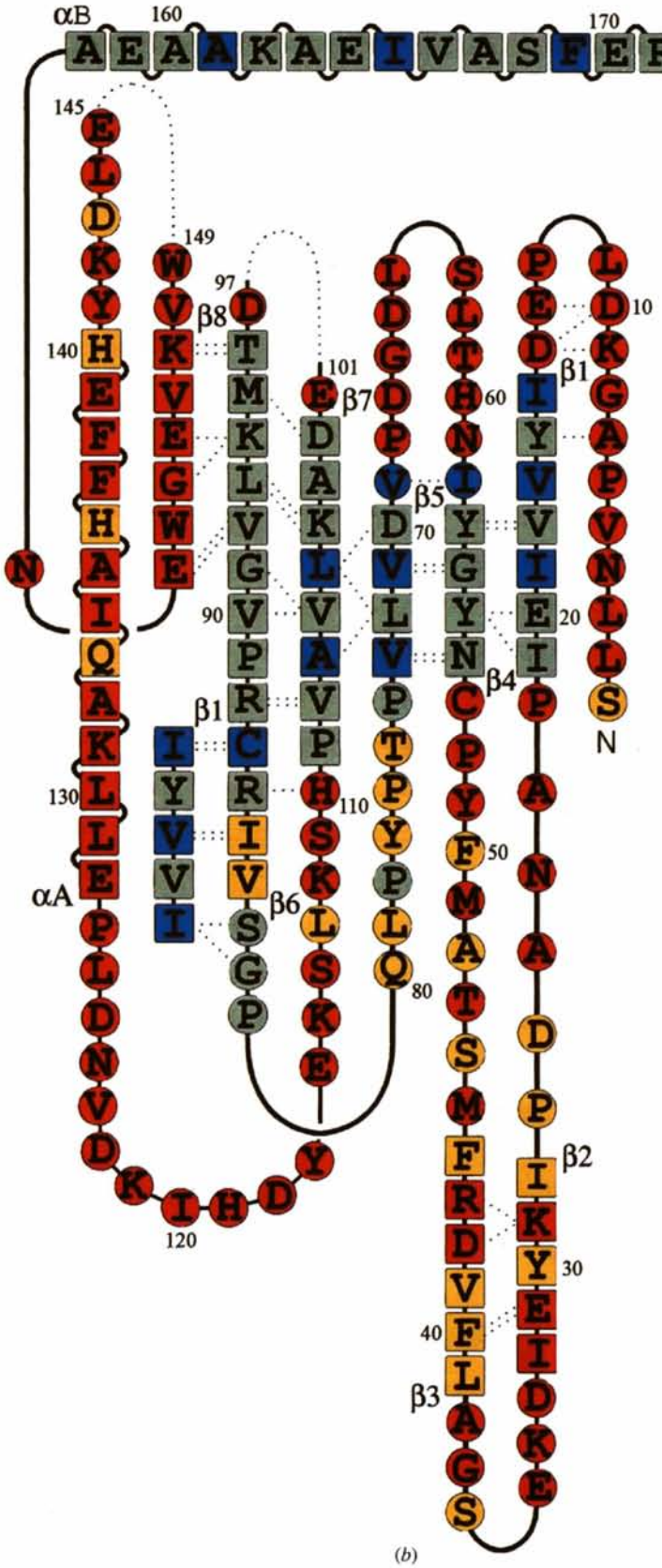


Fig. 4. (*a*). A stereo diagram of the C_α trace of *E. coli* PPase showing the overall topology of the enzyme. The colouring scheme follows this definition: the barrel and the parts belonging to it are coloured green and the excursions red. The residues belonging to the hydrophobic core of the barrel are blue and the residues which participate in intermonomeric contacts are yellow. The missing loops in the short *c*-axis form and the long *c*-axis form monomer II are shown as dotted lines. This figure was drawn using *MOLSCRIPT* (Kraulis, 1991).



(*b*)

Fig. 4 (cont.) (*b*) A diagram showing the hydrogen-bonding pattern between the main-chain atoms of *E. coli* PPase. The colouring scheme follows that of (*a*). The secondary-structure assignment was performed with *PROCHECK* (Laskowski *et al.*, 1993). All the residues which belong to some secondary-structure element are drawn as squares and the residues which were not assigned to any regular secondary structure are drawn as circles. The two long α-helices are drawn as boxes connected with arches.

carbonyl O atom of Ser46′. The intertrimeric interface is thus an intricate network of hydrogen bonds (Fig. 6).

3.3.2. *Intratrimeric contacts.* Unlike the interactions between trimers, which mostly involve side-chain hydrogen bonds, the intratrimeric contacts mostly involve main-chain hydrogen bonds and hydrophobic side-chain interactions (Fig. 7). The N-terminus is hydrogen bonded to the backbone O atom of Ser36 in the symmetry-related molecule (Ser36′) and there is a parallel β-sheet-like arrangement of hydrogen bonds where Val84 interacts with both Leu39′ and Val41′. In addition, the backbone carbonyl O atom of Leu113 accepts a hydrogen bond from the backbone N atom of Phe44′ (Fig. 7*a*). The barrel residues Leu79, Val84 and Ile85, which form a part of the monomeric hydrophobic core also contribute to the hydrophobic intratrimer contacts by forming a hydrophobic cluster with Ile28′, Val41′ and Phe44′ (Fig. 7*b*).

3.4. *The two crystal forms, crystal packing*

As mentioned above, E-PPase crystallizes in two different crystal forms in space group *R*32. Why the enzyme crystallizes in two different crystal forms is difficult to answer based on the crystallization setups;

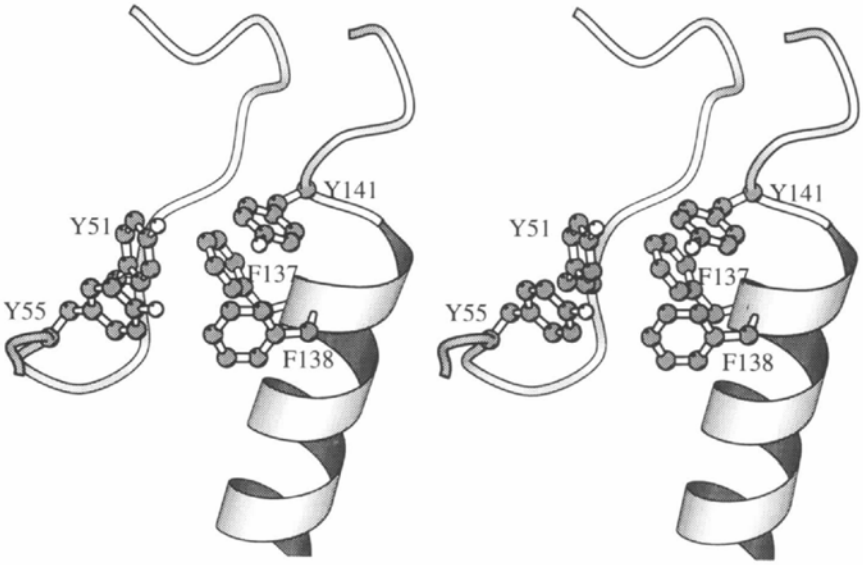


Fig. 5. A stereo diagram showing the interaction of three aromatic residues on helix *A* (the thin spiral) with two aromatic residues in the β3–β4 loop drawing with *MOLSCRIPT* (Kraulis, 1991). The aromatic residues do not interact in a stacking manner; rather, they interact edge-on.

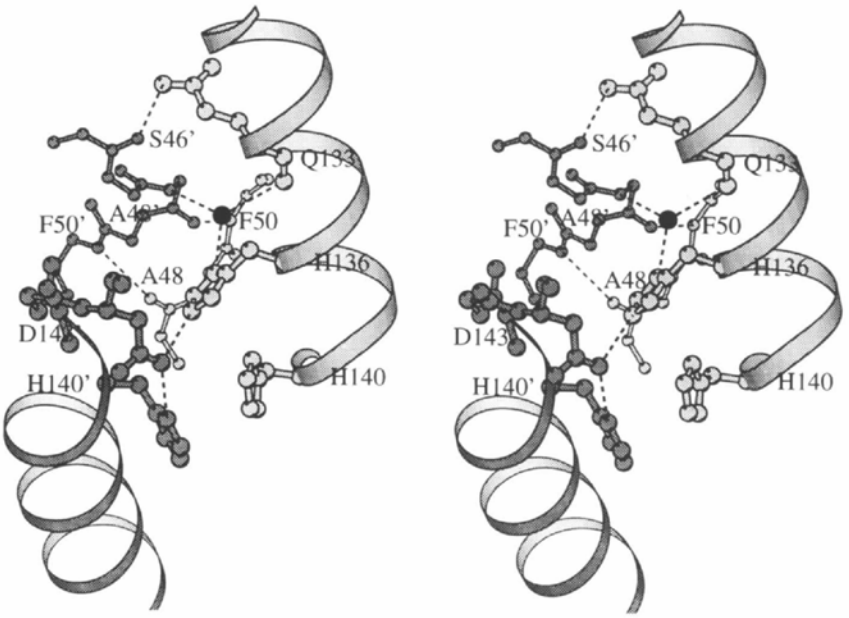
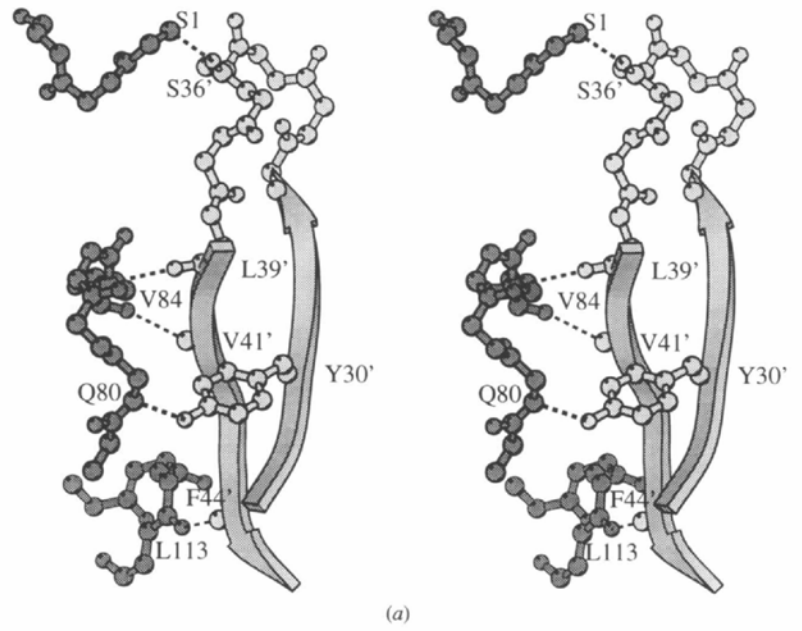


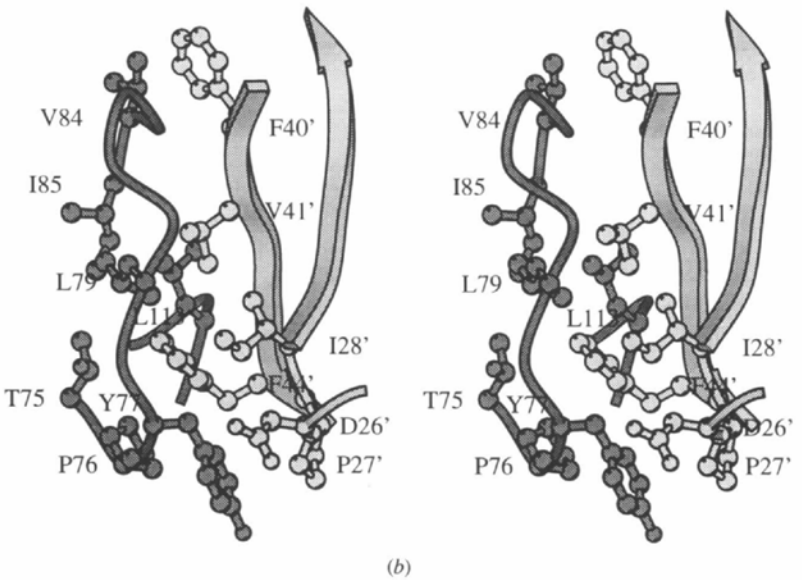
Fig. 6. A *MOLSCRIPT* (Kraulis, 1991) stereo diagram showing the major intertrimeric contacts between two dyad-related α-helices, shown as thin spirals. For clarity, approximately half an interface is shown in detail, hence the His140's seem more exposed than the His136's. In fact they are not. The two symmetry-related monomers are dark and light grey and the symmetry mates are distinguished with primed and unprimed labels. Hydrogen bonds are drawn with dashed lines. Water 8, which is part of the hydrogen-bond network around the His140′–Asp143′–HisH$^+$136 ionic interaction, is shown as a black sphere.

E-PPase (Heikinheimo *et al.*, 1995) readily crystallizes in both short *c*-axis and long *c*-axis forms with different salts and over a range of pH's. Whether a crystal grows in the short or the long *c*-axis form can be very sensitive to experimental conditions: addition of 0.1 m*M* of the inhibitor methane-1-hydroxyl-1,1-diphosphonate (MHDP: $O_3PCHOHPO_3$) to crystallization set ups which normally produced short *c*-axis crystals, produced long *c*-axis crystals. No trace of the inhibitor was, however, visible in difference electron-density maps (T. Salminen, unpublished results), presumably because of the complete insolubility of Mg_2MHDP and Mn_2MHDP (A. Baykov, unpublished results). The monomers in the short *c*-axis form (single monomer/asymmetric unit) are by crystal symmetry identical because the centre of the 32 symmetry of the molecule lies at the origin of the unit cell and thus coincides with the symmetry of the crystal. However, in the long *c*-axis crystal form, the trimers of the relevant E-PPase homohexamer are related only by non-crystallographic symmetry and are not necessarily identical (Fig. 3).

Monomers I and II from the long *c*-axis model could each be superimposed on the short *c*-axis model with an average root-mean-square deviation (r.m.s.d.) of 0.47 Å per C_α for 167 residues and upon each other with an r.m.s.d. of 0.20 Å per C_α (Fig. 8). The difference plot (Fig. 8) correlates very clearly with the *B*-factor plot, so it is very difficult to know whether the differences between



(*a*)



(*b*)

Fig. 7. (*a*) Hydrophilic intratrimeric contacts between monomers related by the threefold axis shown in a *MOLSCRIPT* (Kraulis, 1991) stereo diagram. Two adjacent symmetry mates around the threefold axis are dark and light grey and hydrogen bonds are shown with dashed lines. The curved arrows pointing upwards is β2, the one pointing down is β3. (*b*) Hydrophobic intratrimeric contacts between monomers related by the threefold axis shown in a *MOLSCRIPT* (Kraulis, 1991) stereo diagram. Two adjacent symmetry mates around the threefold axis are dark and light grey. The residues involved in hydrophobic interactions are marked. The view is approximately the same as in (*a*), but 'lower down' on the long β2–β3 loop; in (*a*), the top of the loop can be seen.

the models reflect errors in building or are real. Most of the differences are located in the loop areas in the models and are much larger for the side chains than main chain. The monomer I–short *c*-axis and monomer II–short *c*-axis differences are clustered in approximately the same areas of the structure (Fig. 8).

We, therefore, used the program *RgbSup* (P. Fitzgerald, personal communication) to increase the sensitivity of our superpositions. Unlike a normal alignment program, which maximizes the number of residues in the superposition, *RgbSup* minimizes the number of atoms in a superposition, using an adjustable distance cutoff. This allows one to identify possibly concerted movements in a molecule against an unchanged core. Using *RgpSup*, we found that all except one of the residues whose position changes more than 0.5 Å between the short *c*-axis form and monomer I of the long *c*-axis form are in the regions Asp33–Gly37, Leu62–Pro68, Met95–Ala103, Lys144–Val152 and the C-terminus from Glu170 on. A total of 29 C_α, not in the alignment, have an r.m.s.d. per C_α of 0.96 Å; the other 138 (in the alignment) superimpose with an r.m.s.d. of 0.27 Å, indicating that there is no overall distortion of the molecule. The results when the short *c*-axis form and monomer II of the long *c*-axis form are superimposed are similar: 34 atoms exceed the 0.5 Å cutoff with an r.m.s.d. per C_α of 0.88 Å, while the other 133 superimpose with an r.m.s.d. of 0.27 per C_α. The difference between the two superpositions is that the short *c*-axis form model and the monomer II model differ more at the C-terminus, where a stretch of seven residues, from Ala167 onwards, do not match.

These regions (Asp33–Gly37, Leu62–Pro68, Met95–Ala103, Lys144–Val152 and Glu70–end) correspond to areas with relatively high main-chain temperature factors (Fig. 8, average C_α temperature factor = 40.0 Å^2). Consequently, the motions observed in these regions in different crystal forms are presumably some indication of the range of possible conformational changes during catalysis (Kankare *et al.*, 1996), especially as two important catalytic loops (62–68 and 95–103) are somewhat disordered and are different in monomer I, monomer II and the short *c*-axis model.

Some of the differences between the monomers must be because of crystal packing. The electron density is very weak in the loop regions 97–101 and 145–149 for

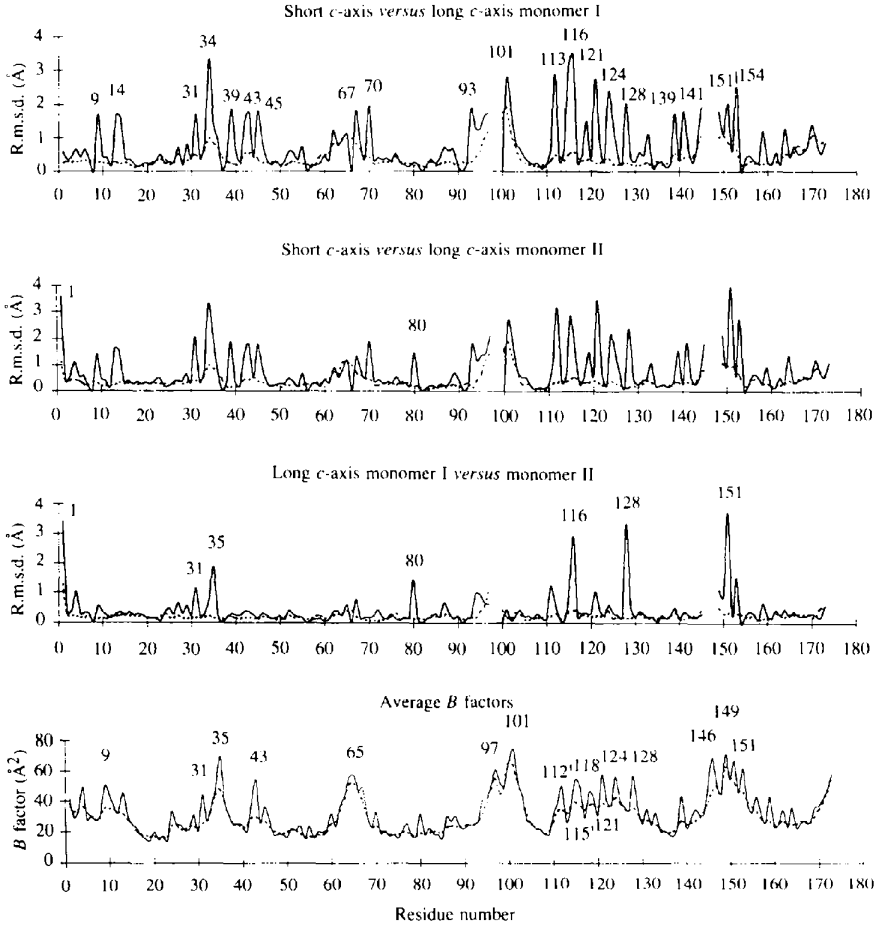


Fig. 8. A plot showing the differences between different models and an average *B*-value plot. The root-mean-square differences are shown for all residues. The differences for backbone atoms are drawn with dotted lines and with solid for side chains. The *B*-factor plot is an average of *B* values calculated over three models except for residues 98–100 and 146–148 which could only be traced in monomer I in the long *c*-axis crystal form.

the short crystal form and for the long *c*-axis monomer II. The situation for corresponding loops in trimer I monomers (Fig. 3) of the long *c*-axis form is somewhat different. The loops are restrained by two additional hydrogen bonds to the hexamer related by crystallographic symmetry: the backbone O atom of Glu98 is hydrogen bonded to Nε1 of Trp149 (data not shown). This serves to stabilize the loop sufficiently for an ion pair between Glu98 and Lys142 of the same monomer to form.

4. Discussion

4.1. *Effect of mutations on subunit structure*

Based on chemical-modification experiments, histidines had been suggested to be important for catalytic activity (Hirano, Ichiba & Hachimori, 1991; Samejima *et al.*, 1988) and so each histidine was substituted with glutamine (Baykov *et al.*, 1995). There are no conserved histidines (Cooperman *et al.*, 1992); there are no histidines in the active site (Kankare *et al.*, 1994); and no histidine is important for catalytic activity. However, two histidine variants, H136Q and H140Q, dissociate into trimers readily (Baykov *et al.*, 1995). The H136Q variant is more stable than the H140Q variant: the H136Q variant has a bell-shaped pH-stability profile and is maximally stable at pH 7.2; the H140Q variant becomes more stable at higher pH's. We have previously explained this as being because of one or more of the following: loss of the salt bridge, loss of precise hydrogen-bonding geometry, more entropy lost on immobilizing Gln as opposed to His, and loss of good hydrophobic interactions between His across the twofold (Baykov *et al.*, 1995).

The effects can now perhaps be explained more thoroughly. Experiments and electrostatic calculations both lead to the conclusion that forming an isolated ion pair does not have a negative ΔG because the enthalpy of desolvation is positive (reviewed in Honig & Nicholls, 1995). However, forming an ion pair can be made energetically favourable if additional residues are desolvated to form tight hydrogen bonds or further ionic interactions to the ion pair because the cost of desolvation is only half of the cost for the first ionic interaction. The geometry of the intertrimeric interface (Fig. 6) makes such an interpretation of the stability of wild-type PPase hexamers very appealing. How, then, are the variant proteins destabilized?

In the H136Q variant, the principal effect is loss of the ion-pair interaction between Asp143 and $HisH^+136'$ in the symmetry-related molecule (see *Intertrimeric contacts*). We assume that the ion pair then becomes Asp143–$HisH^+140$. However, Gln136′ should be able to form similar, though not as good, hydrogen bonds to Asp143 as His136′; especially as the interactions between His136′ and the backbone of Gln133′ and Ala48 are mediated by a water molecule allowing more flexibility in the positioning of Gln136′ in the variant (Fig. 4*b*). Thus, the free-energy difference between wild-type protein and the H136Q variant may not be as great as one might at first expect, particularly as the ΔG of formation of the Asp143–His136 ion pair by itself may not be negative. The two pK_a's seen when hexamer stability is plotted against pH for this variant (Baykov *et al.*, 1995) would then correspond to deprotonating first one His140, then the other.

Conversely, in the H140Q variant the ion pair (Asp143–$HisH^+136'$) is still intact. However, His140 is a more buried residue than His136 without any water molecule to act as an adapter. More hydrogen bonds are lost (potentially as many as four half-hydrogen bonds per Gln) and the hydrophobic interaction between the two His across the twofold is lost. This weakens the essential buttressing interaction to Asp143 that stabilizes the ion pair and the ionic interaction, instead of being stabilizing, may become destabilizing (see above and Waldburger, Schildbach & Sauer, 1995). Consequently, the H140Q variant could be less stable as a hexamer than the H136Q variant, as is in fact observed (Baykov *et al.*, 1995). We have no cogent rationale for the increase in stability of H140Q variant hexamers at high pH, however.

4.2. *Conformational flexibility and the effect of mutations on catalytic activity*

One of our original motivations for refining both E-PPase crystal forms was to try to find a structural explanation for the kinetic properties of the H136Q and H140Q variant proteins (Baykov *et al.*, 1995). Two major effects are seen: the trimeric variant PPases bind metal ions and Mg_2PPi much more weakly than the hexameric variant PPases do, and hexameric H136Q-PPase is more active than wild-type PPase.

The worsening of binding upon trimerization can be relatively easily rationalized. His136 and His140 are part of helix *A*, which is involved in hydrophobic interactions that help stabilize the active site (Fig. 5). Destabilizing helix *A* by exposing it to solvent (in the trimeric H136Q and H140Q variant PPases) will affect the organization of the active site and thus worsen substrate and metal binding. Conversely, mutations at active-site residues (Velichko *et al.*, 1995; Volk *et al.*, 1996) affect oligomeric structure. Finally, other bacterial PPases, when trimeric, are inactive (Hachimori, Shiroya, Hirato, Miyahara & Samejima, 1979; Schreier, 1980). To extend the argument still further, we expect that E-PPase dimers, when found, may also bind substrate and metal ions

poorly, because intratrimeric hydrophobic and hydrophilic interactions (see *Results*) (Fig. 7) stabilize the long excursion II and in particular the β-hairpin loop formed by strands 2 and 3 that carries three conserved active-site residues: Lys29, Glu31 and Arg43 (Cooperman *et al.*, 1992; Kankare *et al.*, 1994).

Explaining the increase in activity seen in hexameric H136Q-PPase is more difficult. We (Baykov *et al.*, 1995) earlier suggested that the weakened intersubunit contacts might allow for greater conformational flexibility and thus faster catalysis, if some of the rate-determining steps involve conformational change. Alternatively, the perfect hexameric symmetry might break down so that only one of the two sets of trimers could be active at any one time. The current three structural models, though only of apoenzymes, support increased conformational flexibility as the more likely explanation for increased enzyme activity. In the long *c*-axis form, the two monomers are not constrained by crystallographic symmetry to be identical, yet their mutual r.m.s.d. per C_α upon superposition is only 0.2 Å. The r.m.s.d. when either of them is superimposed on the short *c*-axis model is more than double: 0.47 Å. Thus, E-PPase maintains tight hexameric symmetry, even while catalytic loops (see *Results*) show evidence of conformational flexibility.

Our results thus explain why, although E-PPase is not an allosteric enzyme in the conventional sense, mutations at its interfaces do affect catalytic activity and, *mutatis mutandis*, mutations in the active site affect oligomerization. This appears to be a direct consequence of having to construct a large open active site with a relatively small polypeptide chain. One could, then, describe the residues as overdetermined in the literary sense: many have more than one 'meaning', which meaning depends on the question asked.

We thank Raija Anderson for technical assistance, and Nisse Kullberg and Dante Still for computing support. We also thank Paula Fitzgerald (Merck Research Laboratories, USA) for the program *RgbSup*. We acknowledge the support of the Finnish Academy of Sciences (grants 1444 to AG and 3875 to RL), of the Suomen Tiedeakatemia (for JK), of the NIH (DK13212 and TW00407), of the International Science Foundation and the Russian Government (M2J000 and M2J300 to AAB) and of the Russian Foundation for Basic Research (94-04-12658-a to AAB).*

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 1FAJ, R1FAJSF and 2EIP, R2EIPSF). Free copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Reference: GR0492). At the request of the authors, the atomic coordinates will remain privileged until 30 August 1996 and the structure factors will remain privileged until 30 January 2000.

Note added in proof: After this manuscript was accepted a paper that also describes the short-axis E-PPase structure at 2.2 Å appeared (Arutyunyan *et al.*, 1996).

References

Arutyunyan, E. G., Oganesyan, V. Yu., Oganesyan, N. N., Terzyan, S. S., Popov, A. N., Rubinskiy, S. V., Vainshtein, B. K., Nazarova, T. I., Kurilova, S. A., Vorobyova, N. N., Avaeva, S. M. (1996). *Kristallografia*, **41**, 84–96.

Baykov, A. A., Dudarenkov, V. Y., Kasho, V. N., Käpylä, J., Salminen, T., Hyytiä, T., Cooperman, B. S, 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*. In the press.

Baykov, A. A., Shestakov, A. S., Kasho, V. N., Vener, A. V. & Ivanov, A. H. (1990). *Eur. J. Biochem.* **194**, 879–887.

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. (1992). *X-PLOR Version 3.1. A system for X-ray Crystallography and NMR*. New Haven: Yale University Press.

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

Burley, S. K. & Petsko, G. A. (1985). *Science*, **229**, 23–28.

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. S., Strokopytov, B. V., Arutyunyan, É. 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 Biol. Sci.* **17**, 262–266.

Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.

Gonzalez, M. A., Webb, M. R., Welsh, K. M. & Cooperman, B. S. (1984). *Biochemistry*, **23**, 797–801.

Hachimori, A., Shiroya, Y., Hirato, A., Miyahara, T. & Samejima, T. (1979). *J. Biochem.* **86**, 121–130.

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

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

Hirano, N., Ichiba, T. & Hachimori, A. (1991). *Biochem. J.* **278**, 595–599.

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

Honig, B. & Nicholls, A. (1995). *Science*, **268**, 1144–1149.

Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* 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). *Biochemistry*. Submitted.

Käpylä, J., Hyytiä, T., Lahti, R., Goldman, A., Baykov, A. A. & Cooperman, B. S. (1995). *Biochemistry*, **34**, 792–800.

Kornberg, A. (1962). *Horizons in Biochemustry*, edited by M. Kasha & B. Pullman, pp. 251. New York: Academic Press.
Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.
Kukko-Kalske, E. & Heinonen, J. (1985). *Int. J. of Biochem.* **17**, 575–580.
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., Hannukainen, R. & Lönnberg, H. (1989). *Biochem. J.* **259**, 55–59.
Lahti, R., Pitkäranta, T., Valve, E., Ilta, I., Kukko-Kalske, E. & Heinonen, J. (1988). *J. Bacteriol.* **170**, 5901–5907.
Lahti, R., Pohjanoksa, K., Pitkäranta, T., Heikinheimo, P., Salminen, T., Meyer, P. & Heinonen, J. (1990). *Biochemistry*, **29**, 5761–5766.
Lahti, R., Salminen, T., Latonen, S., Heikinheimo, P., Pohjanoksa, K. & Heinonen, J. (1991). *Eur. J. Biochem.* **198**, 293–297.
Laskowski, R. A., MacArthur, M. V., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **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.
McRee, D. E. (1992). *J. Mol. Graphics*, **10**, 44–47.
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). *DENZO: an Oscillation Data Processing Program for Protein Crystallography*. New Haven, CT: Yale University.
Salminen, T., Käpylä, J., Heikinheimo, P., Goldman, A., Heinonen, J., Baykov, A. A., Cooperman, B. S. & Lahti, R. (1995). *Biochemistry*, **34**, 782–791.
Samejima, T., Tamagawa, Y., Kondo, Y., Hachimori, A., Kaji, H., Takeda, A. & Shiroya, Y. (1988). *J. Biochem.* **103**, 766–772.
Sato, M., Yamamoto, M., Imada, K. & Katsube, Y. (1992). *J. Appl. Cryst.* **25**, 348–357.
Schreier, E. (1980). *FEBS Lett.* **109**, 67–70.
Steigemann, W. (1991). *Crystallographic Computing*, edited by D. Moras, A. D. Podjarny & J.-C. Thierry, pp. 115–125. Oxford University Press.
Teplyakov, A., Obmolova, G., Wilson, K. S., Ishii, K., Kaji, H., Samejima, T. & Kuranova, I. (1994). *Protein Sci.* **3**, 1098–1107.
Terzyan, S. S., Voronova, A. A., Smirnova, E. A., Kuranova, I. P., Nekrasov, Y. V., Arutyunyan, É. G., Vainshtein, B. K., Höhne, W. & Hansen, G. (1984). *Bioorg. Khim.* **10**, 1469–1482.
Velichko, I., Volk, S., Dudarenkov, V. Y., Magretova, N., Chernyak, V., Goldman, A., Cooperman, B., Lahti, R. & Baykov, A. 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*. In the press.
Waldburger, C. D., Schildbach, J. F. & Sauer, R. T. (1995). *Nature Struct. Biol.* **2**, 122–128.
Wong, S. C. K., Burton, P. M. & Josse, J. (1970). *J. Biol. Chem.* **245**, 4353–4357.