An Unusual, His-dependent Family I Pyrophosphatase from Mycobacterium tuberculosis*

Marko Tammenkoski†, Stefano Benini‡, Natalia N. Magretova*, Alexander A. Baykov†, and Reijo Lahti†,‡

From the †Department of Biochemistry, University of Turku, FIN-20014 Turku, Finland, the ‡University of York, Heslington, York YO10 5YW, United Kingdom, and the *A. N. Belozersky Institute of Physico-Chemical Biology, Moscow State University, Moscow 119899, Russia

Received for publication, August 29, 2005, and in revised form, October 7, 2005 Published, JBC Papers in Press, October 20, 2005, DOI 10.1074/jbc.M509489200

Soluble inorganic pyrophosphatases (PPases) comprise two evolutionarily unrelated families (I and II). These two families have different specificities for metal cofactors, which is thought to be because of the fact that family II PPases have three active site histidines, whereas family I PPases have none. Here, we report the structural and functional characterization of a unique family I PPase from Mycobacterium tuberculosis (mtPPase) that has two histidine residues (His21 and His86) in the active site. The 1.3-Å three-dimensional and functional characterization of a unique family I PPase from Escherichia coli, although mtPPase lacks the intersubunit metal binding site found in E. coli PPase. The cofactor specificity of mtPPase resembles that of E. coli PPase in that it has high activity in the presence of Mg2+, but it differs from the E. coli enzyme and family II PPases because it has much lower activity in the presence of Mn2+ or Zn2+. Replacements of His21 and His86 in mtPPase with the residues found in the corresponding positions of E. coli PPase had either no effect on the Mg2+- and Mn2+-supported reactions (H186K) or reduced Mg2+-supported activity (H21K). However, both replacements markedly increased the Zn2+-supported activity of mtPPase (up to 11-fold). In the double mutant, Zn2+ was a 2.5-fold better cofactor than Mg2+. These results show that the His residues in mtPPase are not essential for catalysis, although they determine cofactor specificity.

Soluble inorganic pyrophosphatase (PPase)3 (EC 3.6.1.1), is an essential metal-dependent enzyme that converts pyrophosphate into orthophosphate. This simple reaction provides a thermodynamic pull for many biosynthetic reactions that yield pyrophosphate as a byproduct (1). Soluble PPases belong to two nonhomologous families (2, 3): family I PPases, which are fairly widespread in all types of organisms, and family II PPases, which are exclusive to bacteria. In eubacteria and archaebacteria, the family I PPases are usually homodimeric, whereas in eukaryotes, they are homodimers. In contrast, all family II PPases are homodimers. The subunit size is generally 19–22 kDa in hexameric PPases and 31–34 kDa in dimeric forms. Despite the variability of the subunit size in family I PPases, their active site cavities are formed by the same 13 functionally important polar residues, which is reflected in the conservation of the catalytic mechanism (4, 5). Although no sequence or overall structural similarity is observed between the two families, there is a striking similarity in the spatial arrangement of six key active site residues, a remarkable example of convergent enzyme evolution (6, 7).

Functionally, family II differs from family I in its preference for Mn2+ over Mg2+ as a cofactor. Mg2+ is a better cofactor for family I PPases, and it also activates family II PPases to the same extent. However, Mn2+ confers a 20-fold higher activity to family II PPases than Mg2+ (8, 9). The difference in cofactor specificity is thought to arise from the presence of histidine residues in the active site of family II PPases. There are a total of three such histidine residues, two act as ligands that bind to metal ions and one is believed to interact with bound PPi (6, 7). Histidine residues were not found in the active sites of any of the six family I PPases whose structures were solved by x-ray crystallography (4, 10–13).

Here, we present the structure of a family I PPase from Mycobacterium tuberculosis (mtPPase), which closely resembles the structure of other bacterial family I PPases but is unique in having two histidine residues (His21 and His86) in the active site. To elucidate the role of these histidine residues in mtPPase, we replaced them with the residues (Lys and Ala, respectively) found in the corresponding positions in the best studied bacterial family I PPase from Escherichia coli (ecPPase) and measured the effects of these replacements on catalysis and cofactor specificity.

EXPERIMENTAL PROCEDURES

Cloning of mtPPase—The Rv3628 gene encoding mtPPase was kindly provided by Dr. Florence Proux (Institut Pasteur, Paris) on a pDEST17 vector and was amplified by PCR using: 5′-CAACCGGCGCATATGC-3′ as forward primer and 5′-GA-AGAGAAGGCCTTTATACGTCAGGCTTGAAGCGC-3′ as reverse primer and 5′-GA-AGAGAAGGCCTTTATACGTCAGGCTTGAAGCGC-3′ as reverse primer. The PCR was performed using KOD Hot start DNA polymerase (Thermococcus kodakaraensis; Novagen) with an initial denaturation for 5 min at 95 °C followed by 35 cycles of denaturation for 30 s at 95 °C, annealing for 30 s at 60 °C, and extension for 2 min at 72 °C. A final extension was carried out for 3 min at 72 °C. The

References

1. W. Zhou, W. Tempel, Z.-J. Liu, L. Chen, L.-C. Clancy Kelley, B. D. Dillard, R. C. Hopkins, W. B. Arendall, III, J. P. Rose, J. C. Eneh, R. C. Hopkins, F. E. Jenney, Jr., H. S. Lee, T. Li, F. L. Poole, II, C. Shah, F. J. Sugar, M. W. W. Adams, J. S. Richardson, D. C. Richardson, and B.-C. Wang, Protein Data Bank entry 1TWL.

2. This work was supported by Academy of Finland Grant 201611, a grant for the National Graduate School in Informational and Structural Biology from the Ministry of Education and the Academy of Finland, European Community contract QLKCT2001-201018 (Structural and functional genomics of M. tuberculosis), Russian Foundation for Basic Research Grant 03-04-48798, and Ministry of Industry, Science, and Technologies of the Russian Federation Grant 1706-2003-4. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked ‘advertisement’ in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

3. The abbreviations used are: PPase, inorganic pyrophosphatase; ecPPase, Escherichia coli inorganic pyrophosphatase; mtPPase, Mycobacterium tuberculosis inorganic pyrophosphatase; tPPase, Thermus thermophilus inorganic pyrophosphatase; bsPPase, Bacillus subtilis inorganic pyrophosphatase; r.m.s., root mean square; MES, 4-morpholineethanesulfonic acid; CAPS, 3-(cyclohexylamino)propanesulfonic acid; CAPSO, 3-(3-cholamidopropyl)dimethylammonio)-2-hydroxy-1-propanesulfonic acid; TES, 2-(2-hydroxy-1,1-bis(hydroxymethyl)ethyl)amino)ethanesulfonic acid.

4. I. Pick, C. Tran, and B. Havenith, Protein Data Bank entry 1SC1.
purified PCR product was inserted via ligation-independent cloning using a modified pET28a vector (pET-YSBLIC) with a noncleavable His tag, giving a total overhang of 10 residues (MGSSHHHGGH). Mutagenesis was performed by an overlapping PCR technique using the Strategene QuikChange kit. All mutations were verified by DNA sequencing.

Expression and Purification of mtPPase—The mtPPase-pET-YSBLIC plasmid was first transformed into NovaBlue single plasmid production and then into E. coli BL21(DE3) cells for protein expression. A single colony was used to inoculate 10 ml of LB medium (15) containing 30 μg/ml kanamycin, which was then grown overnight at 30 °C with vigorous shaking (160 rpm in a Gallenkamp orbital shaker). The overnight culture (10 ml) was used to inoculate 1 liter of Terrific Broth (TB) medium (15), which also contained 30 μg/ml kanamycin, and was grown at 37 °C with vigorous shaking until the A600 reached 0.6. The temperature of the culture was decreased to 30 °C, and protein production was induced with 1 mM isopropyl-β-D-thiogalactopyranoside. After 5 h, the cells were harvested by centrifugation, washed twice with phosphate-imidazole buffer (50 mM sodium phosphate, 300 mM NaCl and 20 mM imidazole, pH 7.6) supplemented with 20% glycerol and 0.5% Triton X-100, and resuspended in 50 ml of the same medium supplemented with a tablet of Complete EDTA-free protease inhibitor mixture (Roche Diagnostics).

Cells were lysed for 30 min with 1 mg/ml of hen egg white lysozyme (Sigma). The cell suspension, kept on ice, was then sonicated using a Soniprep 150 MSE (with a 19-mm probe tuned to 23 kHz) with 10 pulses of 60 s with 30-s intervals between pulses at half-power to prevent foaming. After sonication, the lysed cell suspension was centrifuged for 30 min at 30,000 × g to remove cell debris, and the crude extract was loaded at a flow rate of 2 ml/min onto a 10-ml column containing nickel-charged chelating Sepharose (Amersham Biosciences) previously equilibrated with the phosphate-imidazole buffer. The column was washed at a flow rate of 2 ml/min with phosphate-imidazole buffer until the A280 baseline was stable, and a linear gradient of 20 column volumes from 20 to 500 mM imidazole was applied. Fractions containing mtPPase were pooled (50 ml), dialyzed overnight against 40 volumes of 20 mM Tris-HCl buffer (pH 8.0) containing 150 mM NaCl, concentrated to 5 ml by ultrafiltration using Vivaspin concentrators (Vivascience, Sartorius group), and loaded onto a Superdex 200 16/60 gel filtration column that had been washed with Tris-HCl buffer. The protein eluted at a flow rate of 2 ml/min onto a 10-ml column containing 50% glycerol and 0.5% Triton X-100, and resuspended in 50 ml of the same medium supplemented with a tablet of Complete EDTA-free protease inhibitor mixture (Roche Diagnostics).

To produce wild-type and variant mtPPase for functional studies, we employed a modified procedure for protein expression and isolation. E. coli C41(DE3) ril (16) cells transformed with corresponding mtPPase-pET28a constructs were used for both plasmid production and protein expression. A single colony was used to inoculate 1 ml of a 2× YT medium (15) supplemented with 30 μg/ml kanamycin (2× YT-kan). The cells were grown for 3 h at 37 °C and transferred to 50 ml of 2× YT-kan. The cells were grown for another 4 h and then stored overnight at 4 °C. One liter of 2× YT-kan was inoculated with 50 ml of bacterial culture, and after 1.5 h of incubation at 37 °C, the cells were induced with 1 mM isopropyl-β-D-thiogalactopyranoside for 5 h. The harvested cell pellets were kept frozen at −20 °C. After thawing on ice, 5 g of the cell paste were resuspended in 25 ml of ice-cold Tris/imidazole buffer (50 mM Tris-HCl, 20 mM imidazole-HCl, and 500 mM NaCl, pH 7.2) and homogenized twice in a French press (SLM Instruments) at 800 p.s.i. Cell debris was removed by centrifugation at 48,000 × g for 30 min at 4 °C. Nickel-chelating chromatography was performed as described above but using Tris/imidazole buffer instead of phosphate imidazole buffer. Superdex 200 gel filtration was performed in 20 mM Tris-HCl buffer (pH 7.2) containing 150 mM NaCl.

The purity of the eluted enzyme was checked by electrophoresis on 8–25% gradient polyacrylamide gels in the presence of 0.55% sodium dodecyl sulfate using the Phast System (Amersham Biosciences). Concentrations of mtPPase solutions were determined on the basis of a subunit molecular mass of 19.5 kDa and an extinction coefficient ε195 of 9.8 (as estimated from the amino acid composition using ProtParam).

Metal ions were removed from the enzyme stocks as follows. The enzyme solution was diluted to 1–5 mg/ml with 83 mM TES/KOH buffer (pH 7.2) containing 2 mM EDTA and 17 mM KCl, incubated for 1 day at 4 °C, and subjected to two 40-fold concentration/dilution cycles in a Centricon YM-30 centrifugal filter device (Amicon) using 83 mM TES/KOH buffer (pH 7.2) containing 50 mM EGTA and 17 mM KCl.

Finally, the enzyme solutions were concentrated to ~100 mg/ml and stored frozen at −70 °C.

Expression and Purification of ecPPase—Wild-type ecPPase was over-expressed in MCI061 E. coli cells and purified as described previously (17).

Crystallization of mtPPase and Structure Determination—Protein crystallization was performed at 20 °C using the hanging drop method. Precipitant and protein solutions (1 μl each) were mixed in 24-well Falcon plates equilibrated by vapor diffusion against 1 ml of precipitant. Crystals grew over 4 days from a solution of 1.7 M ammonium sulfate and 100 mM sodium acetate (pH 5.0) up to a dimension of ~0.3 × 0.3 × 0.8 mm.

A single crystal was soaked in 4 μl of a cryosolution containing 20% glycerol, 2 M ammonium sulfate, and 100 mM sodium acetate (pH 5.0). The crystal was then collected with a cryoloop and transferred to the goniometer head in a stream of cold (120 K) nitrogen (Oxford Cryosystems Cryostream). Data to a maximum resolution of 1.8 Å were collected on a rotating anode source with CuKα radiation using a MAR345 image plate detector. The crystal was then stored in liquid nitrogen for further data collection at the European Synchrotron Radiation Facility (Grenoble, France). Data to a maximum resolution of 1.3 Å were collected at the ID14-1 beamline (TABLE ONE). Images were autoindexed and integrated with DENZO and scaled with SCALEPACK (18). The space group was determined to be P612121 with one molecule in the asymmetric unit, giving a Matthews coefficient of 3.5 with 64.5% solvent content. The structure was solved by molecular replacement using the PPrase from Thermus thermophilus (ttPPase, Protein Data Bank code 2PRD) as a model and the program AmoRe (19) as implemented in the CCP4 interface. The model was chosen because, based on a BLAST search against the PDB, it had the most similar sequence (51% sequence identity). Residues 1 to 13 were removed from the model because the mtPPase sequence starts at position 14 of the ttPPase sequence. The best solution after rigid body fitting had a correlation coefficient of 46.5 and an Rfactor of 49.4%. The solution from AMoRe (19) was initially refined with REFMAC (20) to an Rfactor and Rfree of 41.0 and 46.4%, respectively. The model was then automatically updated using Arp/wArp (21) replacing the PDB 2PRD sequence with the mtPPase sequence. The Rfactor and Rfree dropped to 17.1 and 21.6%, respectively, with the model being almost completely autotraced. Model building was continued using the program COOT (22), and further cycles of refinement were carried out with REFMAC coupled with Arp/wArp to update the water molecules. Visual inspection of the waters was carried out with COOT (22). The model was then refined again without an automatic water update to a final Rfactor of 15.4% and Rfree of 16.9%. The final model statistics are summarized in TABLE ONE.
**TABLE ONE**

X-ray data collection and structure refinement statistics for mtPPase crystals

| Data                                      |        |
|-------------------------------------------|--------|
| Wavelength (Å)                            | 0.93   |
| Resolution range (Å)                      | 84.5–1.3 |
| Rsym                                      | 0.060  |
| Raw measurements                          | 558,826 |
| Unique reflections                        | 68,086 |
| Redundancy                                | 8.2    |
| High resolution bin (Å)                   | 1.32–1.3 |
| Completeness (%)                          | 99.2   |
| Completeness in high resolution bin (%)   | 98.2   |
| Greater than 3σ                           | 77.9   |
| Greater than 2σ in high resolution bin (%)| 53     |
| Redundancy in high resolution bin         | 7.2    |
| I/σ in high resolution bin               | 32     |
| Space group                               | P6_22  |
| a (Å)                                     | 98.46  |
| c (Å)                                     | 97.35  |
| Protein atoms                             | 1,358  |
| Solvent atoms                             | 238    |
| Temperature factor for protein atoms (Å²) | 16.53  |
| Temperature factor for solvent (Å²)        | 31.39  |
| r.m.s. deviation bond length deviation (Å) | 0.013  |
| r.m.s. deviation bond angle deviation (degrees) | 1.67  |
| Ramachandran most favored region (%)      | 94.0   |
| Ramachandran additional allowed region (%) | 6.0    |
| Ramachandran generously allowed region (%)| 0      |
| R<sub>f</sub>cat, R<sub>f</sub>free (%)   | 15.4, 16.9 |

**RESULTS**

**Expression and Purification of mtPPase**—SDS-PAGE analysis of crude extracts obtained from recombinant *E. coli* cells revealed an intense ~20-kDa band that was absent in the same cells transformed with empty pET28a plasmid (data not shown). Based on these results, both the wild-type and variant mtPPases were estimated to represent 20–30% of the total protein, and the recombinant enzymes were easily purified to homogeneity. The His-tagged PPases eluted as a single peak from the nickel-affinity column as the imidazole concentration was increased to ~200 mM. The pass-through fraction exhibited low but measurable PPase activity, presumably because of the chromosomal-encoded E-PPase that contains no His tag and, hence, does not bind to the nickel-affinity column. Finally, traces of contaminating proteins detected in the eluted mtPPase by SDS-PAGE were removed by Superdex 200 chromatography. This procedure yielded 60 to 120 mg of pure mtPPase per liter of cell culture (~10 g of cell paste).

**Three-dimensional Structure of mtPPase**—The quality of the data allowed the mtPPase (1SXV) polypeptide chain to be almost automatically traced into the electron density map. The mtPPase structure is composed of 162 amino acid residues of which only the last three (Gly<sup>160</sup>, Thr<sup>161</sup>, and His<sup>162</sup>) are not visible in the electron density map because they are disordered or in multiple conformations. Bound to the model are 1 sulfate ion, 3 molecules of glycerol (present in the cryoprotectant), and 238 molecules of water. Bound metal atoms were not found in the active site, as previously observed for other PPases crystallized in the absence of metal salts (e.g. 1JFD; Ref. 26).

The overall fold of the mtPPase subunit closely resembles the structure of the five other bacterial PPase structures that have been reported, including PPases from *E. coli* (10), *T. thermophilus* (11), *Staphylococcus acidocaldarius* (12), *Pyrococcus horikoshii* (13), and *Pyrococcus furiosus*. A structural superposition carried out for mtPPase on these structures using the Secondary Structure Matching (SSM) program (27) at the EBI server reveals that the core structure of all PPases are very similar. The lowest root mean square deviation (r.m.s. deviation = 0.93 Å) was with *E. coli* 1OBW, followed by the PPase from *P. furiosus* (1TWL; r.m.s. deviation = 1.07 Å). The main difference is in the N-terminal region, where mtPPase is shorter than any other reported PPase structure.

The active site of mtPPase contains the same 13 conserved polar residues found in other family 1 PPases (28). These residues, which participate in cofactor and substrate binding and activation of nucleophilic water (4, 17, 29), are arranged in a similar way in mtPPase and ecPPase (Fig. 1). However, a striking difference is that the mtPPase active site contains two additional His residues, His<sup>21</sup> and His<sup>46</sup>, in place of Lys<sup>34</sup> and Ala<sup>49</sup> in ecPPase. His<sup>46</sup> forms a hydrogen bond with a single

**His-dependent Pyrophosphatase of M. tuberculosis**

Values of the catalytic constant, $k_{cat}$, and the Michaelis constant, $K_{M}$, were obtained from the dependence of the reaction rate on the concentration of dimagnesium pyrophosphate (Mg<sub>2</sub>PP<sub>i</sub>) measured at fixed Mg<sup>2+</sup> concentrations. The dependence of $k_{cat}$ on pH were fit to Equation 3, where $k_{cat,lim}$ is the pH-independent catalytic constant, and $K_{a1}$ and $K_{a2}$ are the ionization constants for the essential basic and acidic groups, respectively. The dependence of $k_{cat}$ on metal cofactor concentration were fit to Equation 4, where [M] is the cofactor concentration, $K_{M}$ is the dissociation constant of the enzyme-metal complex, and $k_{cat,lim}$ is the limiting value of $k_{cat}$ at saturating [M] as shown in the following equations.

$$k_{cat} = \frac{k_{cat,lim}}{(1 + [H^+]/K_{a1} + [H^2]/K_{a2})} \quad \text{(Eq. 3)}$$

$$k_{cat} = \frac{k_{cat,lim}[M]}{(K_{M} + [M])} \quad \text{(Eq. 4)}$$

*Sedimentation*—Analytical ultracentrifugation was carried out at 2 °C in a Spinco E instrument (Beckman Instruments, Palo Alto, CA) with scanning at 280 nm. The samples contained 10 µM enzyme and 1 mM MgCl<sub>2</sub>. Before each run, the samples were incubated for 2 to 3 h at 2 °C. The sedimentation velocity was measured at 48,000 rpm, and the sedimentation coefficient, $s_{20,w}$, was calculated as described by Chervenka (23).

*Activity Measurements*—Rates of PP<sub>i</sub> hydrolysis were determined from continuous recordings of $P_i$ liberation using an automatic $P_i$ analyzer (24). Reactions were initiated by adding a suitable aliquot of enzyme solution and carried out for 3–4 min at 25 °C.

*Analysis of Kinetic Data*—Equations 1 and 2 describe the equilibrium activity of a mixture of hexamer (with a specific activity $A_{D}$) and its dissociated form (with a specific activity $A_{H}$) as a function of the H<sup>+</sup> concentration. Equation 2 was obtained by combining Equations 2 and 3 from Efimova et al. (25). $[E]_{t}$ is total enzyme concentration, $\alpha_{H}$ is the fraction of enzyme in hexamer form expressed in monomers, $K_{a}$ is the pH-independent dissociation constant for the hexamer, $K_{d}$ is the microscopic dissociation constant for H<sup>+</sup> binding to the dissociated form, and $m$ is the number of protons controlling hexamer stability. Equations 1 and 2 were fit to data simultaneously using SCIENTIST (MicroMath) as follows.

$$A = A_{D} + (A_{H} - A_{D})\alpha_{H} \quad \text{(Eq. 1)}$$

$$\frac{2}{3}[E]_{t}(1 - \alpha_{H})^2 - K_{d}\left(1 + \frac{[H^+]}{K_{a}}\right)^m \alpha_{H} = 0 \quad \text{(Eq. 2)}$$
bound sulfate ion, which mimics the product phosphate and occupies the same position as the sulfate in the ecPPase-sulfate complex (PDB code 1JFD; Ref. 26). The sulfate in mtPPase has an occupancy of 1, whereas the ecPPase sulfate has an occupancy of only 0.5, which may indicate stronger binding in mtPPase because of the bond with His186. Formation of this bond causes a shift of up to 3.9 Å in the main chain segment 85–91. Another difference between the two enzymes is that sulfate binding to mtPPase does not cause the molecular asymmetry observed in ecPPase (26); the biological hexamer is formed by identical subunits in the sulfate complex of mtPPase.

Superposition of the mtPPase biological hexamer and the ecPPase (1OBW) hexamer with SSM showed a very close match with a Q-score of 0.78 and a r.m.s. deviation of 1.13 Å. The number of intratrimer hydrogen bonds, detected by the CONTACT program in the CCP4 suite, is greater in mtPPase than in ecPPase (24 versus 12 per subunit), whereas intertrimeric bonds occur less frequently in mtPPase than in ecPPase (4 and 8, respectively). This highlights the prevalence of hydrophobic contacts between trimers in mtPPase.

**Catalytic Characteristics of Magnesium-activated mtPPase and Its Variants**—With 20 mM Mg2+, as the cofactor, mtPPase and ecPPase exhibited similar $k_{cat}$ values at pH 7.2, whereas the $K_m$ value was 4.4 times higher for mtPPase under these conditions (TABLE TWO). Replacement of the His residues, especially His21, causes an appreciable decrease in $k_{cat}$, although this effect mainly arose from shifts in the pH dependence of $k_{cat}$ in the variant proteins. The pH-independent $k_{cat}$ values, $k_{cat,lim}$ estimated from the bell-shaped curves shown in Fig. 2, decreased by a factor of only 1.8 and only for the H21K substitution (TABLE TWO). This substitution also caused an appreciable decrease in the $K_m$ value at pH 7.2, possibly because of a decrease in the forward rate constant (reflected in $k_{cat}$); by lowering $k_{cat}$, the variants more closely approach equilibrium kinetics.

Both His substitutions increased pK$_{cat,in}$, the acid dissociation constant for the group responsible for the decline in $k_{cat}$ in acidic medium (TABLE TWO). As a result, the mtPPase variants more closely resembled ecPPase. The effect of the substitutions on pK$_{cat,lim}$ was less apparent.

**Comparison of Mg2+, Mn2+, and Zn2+ as Cofactors**—The metal cofactor specificity of mtPPase was characterized by measuring $k_{cat}$ as a function of metal ion concentration (Fig. 3). The resulting profiles yielded two important parameters: the metal binding constant ($K_{M}$) and the limiting value of $k_{cat}$ at saturating cofactor concentrations. The results of this analysis (TABLE THREE) show that the cofactor specificities of mtPPase and ecPPase differ significantly. Like many other characterized family I PPases, ecPPase is relatively indiscriminate with respect to the metal cofactor in terms of $k_{cat,lim}$, although it exhibits different metal binding affinities ($K_{M}$). In contrast, wild-type mtPPase shows much greater $k_{cat,lim}$ values with Mg2+ than with Mn2+ or Zn2+ (TABLE THREE).

The substitutions had only a small effect on $K_{M}$ but, in terms of $k_{cat,lim}$, they markedly increased the potency of Zn2+ as the cofactor in comparison with Mg2+. Thus, the H21K substitution decreased the Mg2+-supported activity 4-fold and increased Zn2+-supported activity 3-fold. Substitution of H86A had a negligible effect on activity with Mg2+ but caused a 10-fold increase in the activity in the presence of Zn2+. The effects of the substitutions were clearly additive, and, like

---

**TABLE TWO**

| Enzyme                  | $k_{cat}$ pH 7.2 | $K_m$ pH 7.2 | $k_{cat,lim}$ pH 5–10 | pK$_{cat,ind}$ | pK$_{cat,lim}$ |
|------------------------|-----------------|--------------|-----------------------|----------------|---------------|
| mtPPase                | 240 ± 3 s$^{-1}$ | 8.3 ± 0.3 μM | 280 ± 9 s$^{-1}$      | 6.3 ± 0.1      | 9.2 ± 0.1     |
| H21K-mtPPase           | 81 ± 2 s$^{-1}$  | 4.3 ± 0.3 μM | 155 ± 14 s$^{-1}$     | 6.9 ± 0.2      | 9.7 ± 0.2     |
| H86A-mtPPase           | 200 ± 10 s$^{-1}$| 9.6 ± 0.7 μM | 270 ± 20 s$^{-1}$     | 6.6 ± 0.1      | 9.0 ± 0.1     |
| H21KH86A-mtPPase       | 47 ± 1 s$^{-1}$  | 2.6 ± 0.2 μM | 150 ± 20 s$^{-1}$     | 7.5 ± 0.3      | 9.4 ± 0.2     |
| ecPPase                | 182 ± 3 s$^{-1}$ | 1.9 ± 0.1 μM | 270 ± 10 s$^{-1}$     | 7.0 ± 0.1      | 9.3 ± 0.1     |
ecPPase, the double mutant was a much better catalyst with Zn$^{2+}$ than with Mg$^{2+}$. Interestingly, none of the substitutions affected catalysis with Mn$^{2+}$ as the cofactor.

Quaternary Structure Stability in Acidic Medium—Hexameric ecPPase was previously shown to dissociate into active trimers in a slightly acidic medium (25, 33). The trimer-trimer contact is formed by two groups of residues in ecPPase: His$^{336}$-His$^{440}$, Asp$^{142}$ triads of two subunits, which establish hydrogen-bonding and hydrophobic interactions; and Asn$^{25}$-Ala$^{25}$, Asp$^{25}$ triads, which interact through a Mg$^{2+}$ unit (34). Based on the results of site-directed mutagenesis studies, we previously suggested that ecPPase hexamer instability in acidic medium results from protonation of Asp$^{26}$ (25). The mtPPase provides a good model to test this proposal because the His-His-Asp triad is completely conserved, whereas the metal-binding triad is replaced by Lys-Gly-Glu, which lacks the ionizable Asp and is no longer involved in intersubunit interactions.

Fig. 4 compares the pH stability profiles for mtPPase and ecPPase. In these experiments, the enzymes were pre-equilibrated at the indicated pH, and their activities were measured at pH 7.2. Low oligomeric forms have lower activity and reform the hexamer slowly on the time scale of enzyme assays (25, 35, 36). Three differences between the two enzymes are apparent. First, the curve for mtPPase was markedly shifted to lower pH values, consistent with the proposed role for Asp$^{26}$ in ecPPase. Second, the final activity level reached in the acidic medium was lower for mtPPase. It should be noted that, as in ecPPase (25), the effect of low pH incubation on mtPPase activity was at least 60% reversible upon readjustment of the pH from 4.0 to 7.2 (data not shown). Third, the fitted value of $m$ in Equation 2 describing the observed profiles was 6 for ecPPase, in agreement with earlier estimates (25), but it was only 2 for mtPPase. Moreover, the sedimentation coefficient, $s_{20,w}$, measured for ecPPase under the conditions employed in Fig. 4 was 6.3 S at pH 7.2 and 4.0 S at pH 4, indicating a hexamer to trimer transition (25), whereas for mtPPase it was 7.1 ± 0.3 S and 2.0 ± 0.2 S, respectively, indicating a hexamer to monomer transition.

**DISCUSSION**

Active Site Structure and Cofactor Specificity in PPases—Despite the complete absence of similarity in the primary structures of family I and family II PPases, similarity is clearly seen in the structures of their active sites (6, 7), although they are not identical. In family II PPases, two His residues (His$^{9}$ and His$^{97}$ in Bacillus subtilis PPase (bsPPase)) are ligands for two metal ions (M1 and M2, respectively), and one His residue

---

**His-dependent Pyrophosphatase of M. tuberculosis**

---

**TABLE THREE**

Comparison of Mg$^{2+}$, Mn$^{2+}$, and Zn$^{2+}$ as cofactors of wild-type (WT) and variant mtPPases and wild-type ecPPase (pH 7.2)

| Enzyme | $K_{cat,lim}$ (Mg) | $K_{cat,lim}$ (Mn) | $K_{cat,lim}$ (Zn) | $K_M$ (Mg) | $K_M$ (Mn) | $K_M$ (Zn) |
|--------|-------------------|-------------------|-------------------|------------|------------|------------|
| mtWT   | 240 ± 10          | 13 ± 3            | 12 ± 5            | 1400 ± 200 | <10        | 50 ± 9     |
| mtH21K | 62 ± 2            | 14 ± 4            | 35 ± 2            | 1500 ± 100 | <10        | 50 ± 12    |
| mtH86A | 210 ± 10          | 14 ± 3            | 124 ± 5           | 900 ± 100  | <10        | 42 ± 7     |
| mtH21K/H86A | 53 ± 3   | 17 ± 4            | 130 ± 10          | 1300 ± 200 | <10        | 50 ± 10    |
| ecPPase| 170 ± 10          | 80 ± 8            | 310 ± 20          | 600 ± 100  | <20        | 40 ± 10    |
His-dependent Pyrophosphatase of M. tuberculosis

FIGURE 5. The evolution of His-containing family I PPases. A, partial sequences of 22 family I PPases having one or two histidines at the positions corresponding to residues 21 and 86 of mtPPase. The partial sequence of E. coli family I PPase is also shown. The entire enzyme sequences were aligned with ClustalW. Residues 21 and 86 are shown in bold, and the conserved polar active-site residues are shaded gray. The DHH sequences of the Frankia PPases are underlined. The % sequence identity with mtPPase is shown in parentheses. B, phylogenetic tree of selected prokaryotic family I PPases. The presence of active site histidines corresponding to positions 21/86 of PPases. The presence of active site histidines corresponding to positions 21/86 of PPases is indicated as H/H, H/X, or X/H. The numbers refer to various bacteria as follows: 1, M. tuberculosis; 2, M. leprae; 3, Mycobacterium avium; 4, Nocardia farcinica; 5, Brevibacterium linens; 6, Kineococcus radiotolerans; 7, Frankia sp. EAN1pec; 8, Frankia sp. Cc13; 9, Streptomyces coelicolor; 10, Streptomyces avermitilis; 11, Nocardoides sp.; 12, Thermobifida fusca; 13, Propionibacterium acnes; 14, Bifidobacterium longum; 15, Corynebacterium diphtheriae; 16, Corynebacterium efficiens; 17, Corynebacterium glutamicum; 18, Corynebacterium jeikeium; 19, Leifsonia xyli; 20, Arthrobacter sp.; 21, Tropheryma whippelii; 22, T. thermophilus; 23, Aquifex aeolicus; 24, E. litoralis; 25, N. aromaticivorans; 26, Variovorax paradoxus; 27, E. coli; 28, Salmonella typhimurium; 29, Legionella pneumophila; 30, Rickettsia typhi; 31, Rickettsia rickettsii; 32, Bartonella bacilliformis; 33, Helicobacter hepaticus; 34, Campylobacter jejuni; 35, Campylobacter upsaliensis; 36, Helicobacter pylori.

(His\(^{98}\)) interacts with the bound sulfate oxygen corresponding to the bridging oxygen of PP\(_1\) (6, 7, 37). In contrast, none of the family I PPases whose structures have been previously resolved have His residues in their active sites. Thus, mtPPase is the first family I PPase that is an exception to this rule. His\(^{98}\) of mtPPase, which is hydrogen bonded to sulfate, is clearly an analog of bsPPase His\(^{97}\). The analog of His\(^{21}\) in bsPPase cannot be unequivocally identified because the mtPPase structure lacks bound metal ions. It is closer to the M2 metal binding site and, hence, is more likely to correspond to His\(^{98}\) of bsPPase. This is also consistent with the fact that His\(^{21}\) in mtPPase is preceded by an Asp (D), similar to His\(^{97}\) of bsPPase, which lies in the so-called DHM motif (38) found in all family II PPases (39).

Based on our results, neither His residue is essential for catalysis by mtPPase. His substitutions only slightly decrease Mg\(^{2+}\)-supported activity, as characterized by \(k_{\text{cat,ind}}\) and the shift in the pH dependence of activity (TABLE TWO). On the other hand, these substitutions markedly changed the cofactor specificity, mainly by increasing Zn\(^{2+}\)-supported activity. Wild-type mtPPase is much more active in the presence of Mg\(^{2+}\) than Mn\(^{2+}\) or Zn\(^{2+}\), whereas its H21K/H86A variant exhibits the highest activity with Zn\(^{2+}\), similar to the His lacking ecPPase (TABLE THREE). The H86A substitution is the main cause for this because it activates Zn\(^{2+}\)-supported catalysis, whereas the H21K substitution both deactivates Mg\(^{2+}\)-supported catalysis and activates Zn\(^{2+}\)-supported catalysis, although to a lesser degree. Approaching this from the opposite perspective, restoration of His in the H21K/H86A variant, whose active site is virtually identical to that of ecPPase, somehow interferes with the ability of Zn\(^{2+}\) to act as a cofactor. Given the high effectiveness of the His imidazole group as a Zn\(^{2+}\) ligand, it appears that nonproductive Zn\(^{2+}\) binding must occur. This agrees with the recent finding (40) that family II PPases exhibit very low activity when Zn\(^{2+}\) is bound to the M2 site, which contains His\(^{97}\), the residue that we speculate mimics His\(^{21}\) of mtPPase. Based on the structure of family I, Saccharomyces cerevisiae PPase, which has a full complement of three essential bound metal ions (4, 5, 41), His\(^{86}\) should be close to the metal ion that comes with the substrate and may therefore also cause distortions in the enzyme transition state.

Interestingly, neither substitution markedly affected Mn\(^{2+}\)-supported activity. For family II PPases, Mn\(^{2+}\) is by far the best cofactor. This was initially attributed to the presence of His ligands in the metal-binding sites of family II PPase (8, 9). The fact that Mn\(^{2+}\) activation of...
His-dependent Pyrophosphatase of M. tuberculosis

mtPPase depends little on the presence of His ligands contradicts this explanation and is consistent with the low preference of Mn$^{2+}$ (and Mg$^{2+}$) for nitrogen ligands in comparison with oxygen ligands. Recent crystallographic studies, however, suggested an alternative explanation, that the metal cofactor in family II PPases requires a variable coordination geometry (37). Therefore, this alternative explanation is supported by the mtPPase data.

Oligomeric Structure Stability—The hexameric structure of mtPPase is maintained down to pH 5.5 and is converted into monomers at low pH values. This contrasts with ecPPase, whose hexamer starts dissociating at pH 6.5 (33, 36) but yields trimers that are stable down to pH 3.8 (36). Thus, in acidic medium, intertrimeric contacts are more stable than intratrimeric ones in mtPPase. These differences are easily explained by comparing the structures of the two enzymes. In the ecPPase trimer-trimer contact zone, there is a hydrophilic cavity formed by two pairs of Asn$^{24}$ and Asp$^{26}$ in the interacting subunits. Asp$^{26}$ substitutions greatly stabilize the ecPPase hexamer in acidic medium and reduce the value of $m$ in Equation 2, suggesting that protonation of this group in the wild-type enzyme is responsible for its dissociation into trimers (25). This explanation is strongly supported by the enhanced stability and decreased $m$ value for mtPPase, which lacks the polar cavity and Asp$^{26}$.

The reduced stability of the trimer in mtPPase may be a consequence of a smaller intratrimer contact area compared with ecPPase (1320 versus 1480 Å$^2$). Furthermore, the longer N-terminal extension found in mtPPase (from M. tuberculosis to the double-His PPases and possesses His21, also has a positively charged residue (Lys) in position 86. Of the 441 family I sequences, 252 have His in position 21 and 18 have His in position 86. Interestingly, another mycobacterium PPase (from Mycobacterium avium), which is the closest to the double-His PPases and possesses His$^{21}$, also has a positively charged residue (Lys) in position 86. Of the 441 family I sequences, 252 have Lys as residue 21, and 53 including ecPPase, have Ala as residue 86. Thus, strong conservation is not observed at these positions.

In the phylogenetic tree, all but two (from N. aromativorans and E. litoralis) of the His-PPases form a separate branch of eubacteria, including a sub-branch of two double-His PPases (Fig. 5B). It is likely that the appearance of His$^{21}$ in PPases during evolution stimulated the emergence of His$^{86}$ but not vice versa. All His-PPases, except those from N. aromativorans and E. litoralis, belong to the class of Actinobacteria (high G + C, Gram-positive bacteria) in the order Actinobacteria. Interestingly, Frankia PPases contain a DHH sequence (including the counterpart of mtPPase His$^{21}$) that may correspond to the functionally important DHH motif (38) found in all family II PPases (39) but not previously identified in family I PPases (Fig. 5A).

What advantage is gained by PPase upon acquiring His in its active site? In family II, the His residues allow for extraordinarily tight binding of cofactor in the M2 site (8, 39). Tight binding or changes in the binding affinity upon His mutation are not observed in mtPPase (TABLE THREE). However, the His residues allow mtPPase to function efficiently at lower pH values because of decreased $p$Ka$_{1}$ (TABLE TWO). Earlier studies of family I PPase from S. cerevisiae indicated that the decline in $p$Ka$_{1}$ in acidic media may result from protonation of the nucleophilic hydroxide ion held between the M1 and M2 metal ions, with concomitant deceleration of the PP$_{2}$ hydrolysis step (14). The His residues, especially His$^{86}$, may partially compensate for the effect of pH on this step by serving as general acids or electrophilic catalysts. The inhibition of His-containing PPases by transition metal ions may be used to regulate PPase activity in bacteria accumulating such ions.

Thus, the recruitment of His residues as catalytic aids occurred in both families of soluble PPases. This finding provides further support for the contention that the two families utilize the same catalytic mechanism and are thus a striking example of convergent evolution.

Acknowledgments—We thank ESRF for access to ID14-1 beamline and P. V. Kalnykov for help with analytical centrifugation.

REFERENCES
1. Kornberg, A. (1962) in Horizons in Biochemistry (Kasha, H., and Pullman, B., eds) pp. 251–264, Academic Press, Inc., New York
2. Young, T. W., Kuhn, N. J., Wadeson, A., Ward, S., Burgess, D., and Cooke, G. D. (1998) Microbiology 144, 2563–2571
3. Shintani, T., Uchiumi, T., Yonezawa, T., Salminen, A., Baykov, A. A., Lahti, R., and Hachimori, A. (1998) FEBS Lett. 439, 263–266
4. Heikinheimo, P., Lehtonen, J., Baykov, A., Lahti, R., Cooperman, B. S., and Goldman, A. (1996) Structure 4, 1491–1508
5. Harutyunyan, E. H., Kuranova, I. P., Vainshtein, B. K., Höhne, W. E., Lamzin, V. S., Dauter, Z., Tepleyakov, A. V., and Wilson, K. S. (1996) Eur. J. Biochem. 239, 220–228
6. Merkel, M. C., Fabrichnyi, I. P., Salminen, A., Kalkkinen, N., Baykov, A. A., Lahti, R., and Goldman, A. (2001) Structure 9, 289–297
7. Ahn, S., Miler, A. J., Füttetzer, K., Konopka, M., Illix, M., Young, T. W., and White, S. A. (2001) J. Mol. Biol. 313, 797–811
8. Parfenyev, A. N., Salminen, A., Halonen, P., Hachimori, A., Baykov, A. A., and Lahti, R. (2001) Biochem. J. 267, 24511–24518
9. Zvyaznov, A. B., Vener, A. V., Salminen, A., Goldman, A., Lahti, R., and Baykov, A. A. (2004) Biochemistry 43, 1065–1074
10. Kankare, J., Neale, G. S., Salminen, T., Ghumoff T., Cooperman, B. S., Lahti, R., and Goleman, A. (1994) Protein Eng. 7, 823–830
11. Tepleyakov, A., Obmolova, G., Wilson, K. S., Ishii, K., Kaji, H., Samejima, T., and Kuranova, I. (1994) Protein Sci. 3, 1098–1107
12. Leppänen, V. M., Nurmelm, H., Hansen, T., Lahti, R., Schafer, G., and Goldman, A. (1999) Protein Sci. 8, 1218–1231
13. Liu, B., Bartlam, M., Gao, R., Zhou, W., Pang, H., Liu, Y., Feng, Y., and Rao, Z. (2004) J. Biol. Chem. 279, 422–427
14. Belogurov, G. A., Fabrichnyi, I. P., Pohjanjoki, P., Kasvo, V. N., Lehtihuhta, E., Turkina, M. V., Cooperman, B. S., Goldman, A., Baykov, A. A., and Lahti, R. (2000) Biochemistry 39, 13931–13938
15. Sambrook, J., and Russell, D. W. (2001) Molecular Cloning A Laboratory Manual, 3rd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
16. Miroux, B., and Walker, J. E. (1996) J. Mol. Biol. 260, 289–298
17. Salminen, T., Käpplä, J., Heikinheimo, P., Kankare, J., Goldman, A., Heinonen, J., Baykov, A. A., Cooperman, B. S., and Lahti, R. (1995) Biochemistry 34, 782–791
18. Otsinowiski, Z., and Minor, W. (1997) Methods Enzymol. 276, 307–326
19. Navaza, J. (1994) Acta Crystallogr. Sect. D Biol. Crystallogr. 50, 157–163
20. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Acta Crystallogr. Sect. D Biol. Crystallogr. 53, 240–255
21. Morris, R. J., Zwart, P. H., Cohen, S., Fernandez, F. J., Kakaris, M., Kirillova, O., Vormheim, C., Perrakis, A., and Lamzin, V. S. (2004) J. Synchrotron Radiat. 11, 56–59
22. Emsley, P., and Cowtan, K. (2004) Acta Crystallogr. Sect. D Biol. Crystallogr. 60, 2126–2132
23. Csernok, A. (1992) Methods for the Analytical Ultracentrifuge, pp. 23–33, Spincso Division of Beckman Instruments, Inc., Palo Alto, CA
24. Baykov, A. A., and Avaeva, S. M. (1981) Anal. Biochem. 116, 1–4
25. Efimova, I. S., Salminen, A., Pohjanjoki, P., Lapinniemi, J., Magretova, N. N., Cooperman, B. S., Goldman, A., Lahti, R., and Baykov, A. A. (1999) J. Biol. Chem. 274, 3294–3299
26. Avaeva, S., Kirillova, S., Nazarova, T., Rodina, E., Vorobyeva, N., Skyvankina, V., Grigorjeva, O., Harutyunyan, E., Oganesyan, V., Wilson, K., Dauter, Z., Huber, R., and Mather, T. (1997) FEBS Lett. 410, 502–508
His-dependent Pyrophosphatase of M. tuberculosis

27. Krissinel, E., and Henrick, K. (2004) Acta Crystallogr. Sect. D Biol. Crystallogr. 60, 2256–2268
28. Baykov, A. A., Cooperman, B. S., Goldman, A., and Lahti, R. (1999) Prog. Mol. Subcell. Biol. 23, 127–150
29. Pohjanjoki, P., Lahti, R., Goldman, A., and Cooperman, B. S. (1998) Biochemistry 37, 1754–1761
30. DeLano, W. L. (2002) The PyMOL Molecular Graphics System, DeLano Scientific, San Carlos, CA
31. Salminen, T., Teplyakov, A., Kankare, J., Cooperman, B. S., Lahti, R., and Goldman, A. (1996) Protein Sci. 5, 1014–1025
32. Lee, B., and Richards, F. M. (1971) J. Mol. Biol. 55, 379–400
33. Borshchik, I. B., Magretova, N. N., Chernyak, V. Y., Sklyankina, V. A., and Avaeva, S. M. (1986) Biokhimiya 51, 1484–1489
34. Velichko, I. S., Mikalahti, K., Kasho, V. N., Dudarenkov, V. Y., Hyytiä, T., Goldman, A., Cooperman, B. S., Lahti, R., and Baykov, A. A. (1998) Biochemistry 37, 734–740
35. Baykov, A. A., Dudarenkov, V. Y., Kapyla, J., Salminen, T., Hyytiä, T., Kasho, V. N., Hugafvel, S., Cooperman, B. S., Goldman, A., and Lahti, R. (1995) J. Biol. Chem. 270, 30804–30812
36. Salminen, A., Efimova, I. S., Parfenyev, A. N., Magretova, N. N., Mikalahti, K., Goldman, A., Baykov, A. A., and Lahti, R. (1999) J. Biol. Chem. 274, 33898–33904
37. Fabrichniy, I. P., Lehtis, L., Salminen, A., Zyryanov, A. B., Baykov, A. A., Lahti, R., and Goldman, A. (2004) Biochemistry 43, 14403–14411
38. Aravind, L., and Koonin, E. V. (1998) Trends Biochem. Sci. 23, 17–19
39. Halonen, P., Tammenkoski, M., Niiranen, L., Hauhalaiti, S., Parfenyev, A. N., Goldman, A., Baykov, A., and Lahti, R. (2005) Biochemistry 44, 4004–4010
40. Zyryanov, A. B., Tammenkoski, M., Salminen, A., Kolomitytseva, G. Y., Fabrichniy, I. P., Goldman, A., Lahti, R., and Baykov, A. A. (2004) Biochemistry 43, 14395–14402
41. Heikinheimo, P., Tuominen, V., Ahonen, A. K., Teplyakov, A., Cooperman, B. S., Baykov, A. A., Lahti, R., and Goldman, A. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 3121–3126
42. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) Nucleic Acids Res. 22, 4673–4680