Biochemical Characterization and Crystallographic Structure of an *Escherichia coli* Protein from the Phosphotriesterase Gene Family†

Jenny L. Buchbinder,‡ Robert C. Stephenson,‡ Mark J. Dresser, Jed W. Pitera, Thomas S. Scanlan,* and Robert J. Fletterick*

Departments of Biochemistry and Biophysics, Pharmaceutical Chemistry, and Cellular and Molecular Pharmacology, University of California, San Francisco, California 94143-0448

Received July 14, 1997; Revised Manuscript Received December 5, 1997

ABSTRACT: Phosphotriesterase homology protein (PHP) is a member of a recently discovered family of proteins related to phosphotriesterase, a hydrolytic, bacterial enzyme with an unusual substrate specificity for synthetic organophosphate triesters and phosphorofluoridates, which are common constituents of chemical warfare agents and agricultural pesticides. No natural substrate has been identified for phosphotriesterase, and it has been suggested that the enzyme may have evolved the ability to hydrolyze synthetic compounds in bacteria under selective pressure to meet nutritional needs. PHP, which has 28% sequence identity with phosphotriesterase, may belong to the family of proteins from which phosphotriesterase evolved. Here we report the cloning, expression, initial characterization, and high-resolution X-ray crystallographic structure of PHP. Biochemical analysis shows that PHP is monomeric and binds two zinc ions per monomer. Unlike phosphotriesterase, PHP does not catalyze the hydrolysis of nonspecific phosphotriesters. The structure, similar to that of phosphotriesterase, consists of a long, elliptical α/β barrel and has a binuclear zinc center in a cleft at the carboxy end of the barrel at the location of the presumptive active site.

Generally, evolution is presumed to require millions of years for selective adaptation of an organism to a particular ecological niche. Recent evidence suggests that evolution in microorganisms can occur over a shorter time period, on the order of decades or centuries rather than eons. Studies of artificial evolution of bacteria in a laboratory setting document the alteration of proteins on a short time scale in response to selective pressures (*1*). The evolution of novel activities may entail the adaptation of a substrate binding pocket in a protein to alter specificity for a given type of catalytic function or the evolution of a new catalytic activity in a pre-existing binding pocket which had no previous catalytic function. One such example is the artificial evolution of a bifunctional enzyme having 3-deoxy-D-arabinoheptulosonate-7-phosphate synthase and chorismate mutase activities in a strain of *Bacillus subtilis* deficient in chorismate mutase activity (*2*). The bifunctional enzyme apparently evolved from a monofunctional deoxy-D-arabinoheptulosonate-7-phosphate synthase by mutagenesis of an allosteric effector site for prephenate that previously had no catalytic activity. Another example is the evolution of β-galactosidase activity in a strain of *Escherichia coli* which had the gene for β-galactosidase deleted (*3*). The activity evolved under selective growth of the bacteria on media containing lactose as the carbon source.

The enzyme phosphotriesterase (PTE)[1] from *Pseudomonas diminuta* may be an example of natural evolution of a new enzymatic activity in response to changing environmental conditions. PTE is a metalloenzyme with a binuclear zinc center that catalyzes the hydrolysis of a variety of organophosphate triesters and phosphorofluoridates (*4*−*6*). All of the phosphate triesters found to be substrates of PTE are synthetic compounds, and the identity of any naturally occurring substrate for the enzyme is unknown. PTE has attracted interest because of its potential use in the detoxification of chemical waste and warfare agents and its ability to degrade agricultural pesticides (*4*). The gene for PTE is carried on a plasmid in *P. diminuta* and on a different plasmid in *Flavobacterium* sp. (*7*, *8*). The enzyme has a broad substrate specificity for phosphate triesters but no detectable activity with phosphate monoesters or diesters and no activity as an esterase or protease (*4*, *9*, *10*). PTE catalyzes the hydrolysis of the insecticide paraoxon at a rate approaching the diffusion limit and thus appears to be optimally evolved for utilizing this synthetic substrate (*4*, *11*). The synthesis of paraoxon was first reported in 1950 (*12*), suggesting that the phosphotriesterase activity has evolved in bacteria over the past several decades from an enzyme with a related catalytic function.

† This work was supported by a grant from the National Science Foundation (MB 9513300 to T.S.S.).

* To whom correspondence should be addressed. For T.S.S.: telephone (415) 476-3620, fax (415) 476-0688, e-mail scanlan@cgl.ucsf.edu. For R.J.F.: telephone (415) 476-5080, fax (415) 476-1902, e-mail flett@msg.ucsf.edu.

‡ These authors contributed equally to this work.

[1] Abbreviations: PHP, phosphotriesterase homology protein; PTE, phosphotriesterase; IPTG, isopropyl β-D-thiogalactopyranoside; ICP, inductively coupled plasma; Tris, tris(hydroxymethyl)aminomethane; Bis-Tris, 2-[bis(2-hydroxyethyl)imino]-2-(hydroxymethyl)-1,3-propanediol; CHES, 2-(*N*-cyclohexylamino)ethanesulfonic acid; SDS, sodium dodecyl sulfate; LB, Luria-Bertani; HPLC, high-performance liquid chromatography; CD, circular dichroism; MIR, multiple isomorphous replacement; NCS, noncrystallographic symmetry; rms, root-mean-square.

S0006-2960(97)01707-8 CCC: $15.00
Published on Web 03/25/1998

From a search of the *E. coli* genomic sequence database, an open reading frame (ORF) was identified coding for a protein with an amino acid sequence 28% identical to that of PTE (*13*). This protein was named phosphotriesterase homology protein (PHP) and is believed to belong to the family of proteins from which PTE evolved. Four other genes with homologous sequences have been subsequently identified in rat, mouse, mycobacterium, and mycoplasma (*14*−*16*; GenBank reference x99477). On the basis of sequence homology, the structure of PHP might be expected to be similar to that of PTE, and crystallographic structures are available for the apo form of PTE and its complexes with Zn and Cd (*6*, *17*, *18*). The structures of PTE show that it has an α/β barrel fold and possesses a binuclear metal center at its active site. In the complex of PTE with zinc, the two metal ions are separated by 3.3 Å. The zinc ions are bridged by a water molecule and the side chain of a carbamylated Lys169. One zinc ion exhibits tetrahedral coordination by ligands from carbamylated Lys169, His201, and His230 and a water molecule. The other zinc ion is coordinated with trigonal bipyramid geometry by ligands from His55, His57, carbamylated Lys169, Asp301, and a water molecule. The coordinated water molecule may act as the nucleophile in the hydrolysis reaction. Examination of an alignment of the primary sequences of PHP and PTE shows that the four histidine residues and the aspartate involved in coordination of zinc in PTE are conserved in PHP (*13*). An Ala is found in the sequence of PHP at the position of the carbamylated Lys in PTE; however, the next residue in the sequence of PHP is a glutamate. It seemed plausible that the carboxylate group of Glu125 might replace the carbamate of Lys169 in PTE and that the shift in position in the sequence might compensate for the longer length of the carbamylated Lys side chain. Here we report the cloning, purification, and initial biochemical characterization of the enzyme. In addition, we present the crystallographic structure of PHP complexed with two zinc ions, determined at 1.7 Å resolution, and offer a comparison with the structures of PTE and urease, two related, hydrolytic metalloenzymes that have the α/β barrel fold.

MATERIALS AND METHODS

Cloning of PHP. The *E. coli* phosphotriesterase homology protein (ePHP) sequence was identified previously using the BLAST searching algorithm (*19*) for sequences with homology to phosphotriesterase (*13*). A primer complementary to the 5′ end of the sequence that incorporated an *Nde*I restriction site and another primer complementary to the 3′ end of the sequence that incorporated a *Hin*dIII restriction site were synthesized. The primer sequences were 5′GGAATTCCATATGAGTTTTGATCCGAC for the N-terminal primer and 5′GCTCTCAAGCTTTTATTGGAAAAATTGAGA for the C-terminal primer. Both primers were prepared by the University of California−San Francisco (UCSF) Biomolecular Resources Center. The *php* gene was amplified out of an *E. coli* genomic DNA library obtained from the lab of C. Craik (*20*; originally from M. Obukowicz, Monsanto) using the polymerase chain reaction and then cloned into the pTactac vector (*21*). The resulting plasmid was termed pJWP1. The DNA sequence was confirmed by dideoxy sequencing performed by the UCSF Biomolecular Resource Center DNA Sequencing Facility.

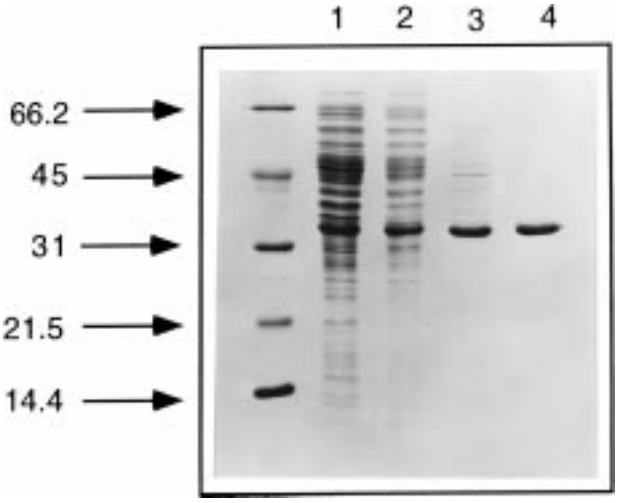

FIGURE 1: Purification of ePHP overexpressed in *E. coli*: lane 1, sample from lysed cells of *E. coli* overexpressing pJWP1; lane 2, dialyzed supernatant identical to the material loaded on the Q Sepharose column; lane 3, pooled fractions containing ePHP from Q Sepharose chromatography; and lane 4, pooled fractions containing ePHP from hydrophobic interaction chromatography. At the left, molecular size standards include serum albumin (66.2 kDa), ovalbumin (45 kDa), carbonic anhydrase (31 kDa), trypsin inhibitor (21.5 kDa), and lysozyme (14.4 kDa). Samples were separated by polyacrylamide gel electrophoresis on a 15% gel and visualized with Coomasie Brilliant Blue.

Protein Expression and Purification. The ePHP gene, cloned in plasmid PJWP1, was expressed in *E. coli* strain DH5α grown at 37 °C on LB media generally supplemented with 1 mM $ZnSO_4$ (except where otherwise specified for metal analysis experiments). One liter cultures of transformed bacteria were grown in the presence of 100 mg of ampicillin to an OD_{600nm} of 0.6 before induction of ePHP gene expression by addition of 2 mL of 0.5 M IPTG. Cultures were then grown for another 30 h at 37 °C before harvesting by centrifugation at 4000*g* for 15 min at 4 °C. Cell pellets were resuspended in 40 mL of cold 20 mM Bis-Tris·HCl at pH 6.6 (buffer A) and lysed by sonication. ePHP appeared to be the major protein in the crude lysate and soluble fraction as observed by polyacryamide gel electrophoresis (Figure 1, lanes 1 and 2). The cell homogenate was centrifuged at 18000*g* at 4 °C for 30 min to remove cellular debris, and the supernatant was dialyzed overnight at 4 °C against buffer A. After dialysis, the protein was loaded onto an anion exchange Q Sepharose (Pharmacia) column (2.5 cm diameter × 37 cm length) pre-equilibrated with buffer A. The protein was eluted with a 2 L linear gradient of 0 to 1 M NaCl in buffer A. Fractions containing ePHP, which eluted at approximately 0.25 M NaCl, were pooled, and ammonium sulfate was added to a final concentration of 1.2 M and the pH adjusted by the addition of Tris·HCl (pH 8.0) to a concentration of 100 mM before chromatography on a Rainin HPLC phenyl-5PW column (TosoHaas, 20 mm ID × 15 cm) pre-equilibrated in 1.5 M ammonium sulfate and 100 mM Tris·HCl at pH 8.0 (buffer C). The ePHP protein was eluted using a linear gradient from buffer C to buffer D (100 mM Tris·HCl at pH 8.0). Fractions containing ePHP protein were pooled and dialyzed to remove ammonium sulfate. Coomassie blue-stained SDS−PAGE gels showed that the protein was purified to 80% homogeneity after Q Sepharose chromatography (Figure

1, lane 3) and to better than 98% homogeneity after hydrophobic interaction chromatography (Figure 1, lane 4). A final yield of 10 mg of purified protein was obtained from 1 L of culture. The identity of the purified ePHP was confirmed by peptide sequencing of the N-terminal amino acids. The protein after polyacrylamide gel eletrophoresis was blotted onto nitrocellulose in preparation for sequencing, which was performed by the UCSF Biomolecular Resource Center. Metal analysis of the purified protein by inductively coupled plasma emission spectroscopy showed that 1.9 equiv of zinc was incorporated per monomer of protein.

Gel Filtration. The polymerization state of ePHP was determined by analytical gel filtration chromatography. ePHP (1 mg) was loaded onto a Pharmacia Sephacryl S-100 HR column (900 cm × 16 mm) along with molecular weight standards (1 mg each), ovalbumin, carbonic anhydrase, and bovine serum albumin. The column was equilibrated and eluted with 10 mM Tris and 150 mM NaCl (pH 7.5).

Metal Analysis of ePHP. ePHP was expressed in bacterial cultures grown on media supplemented with different metal ions, including $ZnSO_4$, $CoCl_2$, $MnCl_2$, $CuSO_4$, $FeSO_4$, $NiSO_4$, and $CdCO_3$. In each case, 1.0 mL of overnight culture of DH5α containing pJWP1 was used to inoculate 200 mL of LB media containing 20 mg of ampicillin. The cultures were grown to an OD_{600nm} of 0.4 (about 3 h), and then 400 μL of 0.5 M IPTG and 2 mL of a 100 mM solution of the appropriate metal were added to the culture. The $FeSO_4$ was dissolved in degassed water immediately before addition, and the $FeSO_4$, was added in two aliquots 90 min apart to avoid oxidation and precipitation of Fe^{2+}. Bacterial lysates were prepared as described above, and the protein was partially purified by chromatography on Q Sepharose using a step gradient. The Q Sepharose was first washed with 150 mM NaCl in buffer A followed by elution of ePHP with 200 mM NaCl in buffer A. Fractions containing ePHP were dialyzed against 50 mM NaCl and 20 mM Tris·HCl (pH 7.5) and submitted for metal analysis by inductively coupled plasma (ICP) emission spectroscopy performed by the Garratt-Callahan Co. (Millbrae, CA) and by XRAL Laboratories. Protein concentrations were determined by amino acid analysis at the University of Michigan, Protein and Carbohydrate Structure Facility (Ann Arbor, MI).

Enzyme Assays. ePHP was assayed for enzymatic activity against a panel of nonspecific substrates. Activity was determined by monitoring changes in absorbance at 405 nm for 5 min. The substrates *p*-nitrophenyl acetate, L-alanine nitroanilide, *p*-nitrophenyl sulfate, bis(*p*-nitrophenyl) phosphate, paraoxon, and *p*-nitrophenyl phosphate were assayed in the presence of 42 μg of purified Zn containing ePHP in 50 mM Tris·HCl and 100 mM NaCl (pH 7.8). All substrates were assayed at a concentration of 1 mM except for nitrophenyl phosphate which was assayed at a concentration of 0.2 mM. In addition, both *p*-nitrophenyl phosphate and paraoxon were assayed at pH 9.0 in 100 mM CHES buffer (*4*). The *p*-nitrophenyl phosphate was also assayed at pH 5.5 in 100 mM MES buffer. Carbonic anhydrase activity was assayed by monitoring the time required for a saturated carbon dioxide solution to lower the pH of 20 mM Tris·HCl from 8.3 to 6.3 at 0 °C (*22*).

Crystallization. ePHP was crystallized in hanging drops by the vapor diffusion method. The protein crystallized at a concentration of 15 mg/mL from a precipitant solution of 1.5 M ammonium sulfate, 10% 2-methyl-2,4-pentanediol, 100 μM $Zn(OAc)_2$, and 50 mM HEPES (pH 7.5). Three microliters of a solution containing the protein was mixed with 3 μL of the precipitant solution and the mixture allowed to equilibrate at room temperature. Crystals appeared after 3 days, but typically grew in clusters. The quality of crystals was improved by microseeding drops with crushed crystals. Single crystals obtained from seeding had dimensions of approximately 100 μm × 300 μm × 300 μm and diffracted to 1.7 Å resolution.

X-ray Diffraction Measurement and Data Processing. In preparation for diffraction measurements, crystals were transferred to an artificial mother liquor consisting of the precipitant solution with 20% glycerol added as a cryoprotectant. The crystals were mounted in rayon loops, 500 μm in diameter, and flash frozen under a stream of nitrogen at −180 °C. Diffraction data were measured with an Raxis II image plate detector using CuKα radiation generated by an 18 kW rotating anode operating at 50 kV and 300 mA. Data were processed with the programs Denzo and Scalepack (*23*). All data processing and subsequent refinement were performed on a Silicon Graphics workstation.

Structure Determination and Refinement. The unit cell dimensions of the crystal were determined with the program Denzo, and the crystallographic space group was assigned as $P2_1$ by analysis of pseudo precession images using the program Xprep. Calculation of a Matthew's coefficient (*24*), V_m, of 2.7 $Å^3$/Da indicated the presence of two monomers in the asymmetric unit. Inspection of peaks in a native Patterson map showed that the monomers of the asymmetric unit were related by a translation in fractional coordinates of 0.4 in *y* and 0.5 in *z*. Attempts to solve the structure by molecular replacement using the atomic coordinates of the apo form of PTE from *P. diminuta*, which has 28% sequence identity with ePHP, were initially unsuccessful. The multiple isomorphous replacement method was therefore used to determine experimental phases. Three heavy-atom derivatives were prepared by soaking crystals in a mother liquor [1 M ammonium sulfate, 10% 2-methyl-2,4-pentanediol, and 50 mM HEPES (pH 7.5)] containing heavy-atom compounds. Details concerning the concentrations of the heavy-atom compounds, mercuric chloride, Baker's dimercurial, and uranyl acetate and the lengths of the soaks are presented in Table 1. The binding sites of heavy-atom derivatives were determined from inspection of difference Patterson maps and difference Fourier maps, and MIR phases were calculated to 2.5 Å using the program Phases (*25*). The phases were improved by including anomalous data from the mercuric chloride and uranyl acetate derivatives in the phase calculation and by density modification using the program DM to perform histogram matching, solvent flattening, and noncrystallographic symmetry averaging (*26*). Once experimental phases were determined, the structure of the apo form of PTE from *P. diminuta* (*17*) could be roughly positioned in the electron density map. The coordinates of PTE were rotated according to a rotation search using the program Amore (*27*), and the experimental phases were used to perform a phased-translation search with the program X-PLOR (*26*). The initial *R* factor for the molecular replacement solution was 59% in the range of 50−3.5 Å resolution. No attempt at phase combination of the MIR and model phases was, therefore, attempted because of the good quality

Table 1: Crystallographic Data and Heavy-Atom Refinement Statistics

crystals (temperature of −180°)	space group $P2_1$	$a = 42.1$ Å, $b = 80.8$ Å, $c = 98.2$ Å, and $\beta = 97.1°$		
	native	baker's dimercurial	mercuric chloride	uranyl acetate
R_{merge}[a]	0.075	0.088	0.074	0.090
no. of reflections	286 468	122 476	236 370	262 809
no. of unique reflections	69 475	53 311	57 170	55 506
completeness (%)	98 (50−2.9 Å), 94 (2.1−2.0 Å), 72 (1.8−1.7 Å)	97 (50−2.7 Å), 86 (2.1−2.0 Å), 77 (1.9−1.8 Å)	99 (50−2.3 Å), 96 (2.1−2.0 Å), 79 (1.9−1.8 Å)	96 (50−2.6 Å), 91 (2.1−2.0 Å), 85 (1.9−1.8 Å
no. of sites		4	2	2
heavy-atom soaks		0.1 mM, 1 h	0.1 mM, 1 h	5 mM, 1 day
$R_{centric}$[b]		0.53 (4 Å)	0.62 (2.9 Å)	0.54 (2.9 Å)
R_{anom}[c]			0.048 (3.5 Å)	0.061 (2.9 Å)
phasing power[d]		2.4 (4 Å)	1.7 (2.9 Å), 1.0 (ano, 3.5 Å)	1.3 (2.9 Å), 1.5 (ano, 2.9 Å)
figure of merit		0.58 (10−2.9 Å)		

[a] $R_{merge} = \sum_i\sum_h|I(h) - I_i(h)|/\sum_i\sum_h I(h)$, where $I_i(h)$ and $I(h)$ are the ith and the mean measurement of the intensity of reflection h. [b] $R_{centric} = \sum||F_{PH} \pm F_P| - |F_H||/\sum|F_{PH} - F_P|$, where F_{PH} is the structure factor of the derivative, F_P is that of the native data, and F_H is the calculated heavy-atom structure factor. The upper resolution cutoffs used for the calculations are shown in parentheses. [c] $R_{anom} = \sum|F^+ - F^-|/\sum|F|$ for acentric anomalous scattering data. [d] Phasing power = rms($|F_H|/E$), where F_H is the calculated heavy-atom structure factor and E is the residual lack of closure error. ano indicates the calculated value based on the anomalous difference.

Table 2: Structure Refinement Statistics

R_{cryst}[a] (50−1.7 Å)	0.204
R_{free}[b] (50−1.7 Å)	0.241
rms deviation from ideal bond lengths (Å)	0.008
rms deviation from ideal bond angles (deg)	1.62

[a] $R_{cryst} = \sum_h|F_o - F_c|/\sum_h|F_o|$, where F_o and F_c are the observed and calculated structure factor amplitudes, respectively. [b] R_{free} is the cross-validation R factor determined for 10% of the total reflections, which were omitted from map calculations and refinement.

of the MIR map and the apparently substantial structural differences between PTE and ePHP. The PTE model was adjusted by visual inspection of the MIR map at 2.5 Å resolution. Model building was performed with O (*29*), and the structure was refined with X-PLOR (*28*) and Refmac (*30*). After one round of model fitting and refinement, the phases were extended to 1.7 Å resolution and all further fitting of the model was performed using $2F_o - F_c$ maps. An R_{free} was calculated before the start of refinement from 10% of the data, which was then excluded from all refinement and map calculations and monitored throughout the refinement. Multiple cycles of model building were alternated with positional refinement, isotropic B factor refinement, and simulated annealing. NCS restraints were applied to the two monomers composing the asymmetric unit during initial rounds of refinement. Removal of the NCS restraints at later stages of refinement resulted in a slightly lower R_{free} and improved geometry of the model. A composite omit map was calculated with X-PLOR over the entire structure to identify errors in the coordinates. Water molecules were added to the structure using the X-PLOR Waterpick subroutine. The water molecules were placed in the $2F_o - F_c$ map in peaks of electron density that exceeded 1.5σ and that were located a minimum of 2.5 Å and not more than 4 Å from nitrogen or oxygen atoms of the protein. The quality of the final model was assessed from Ramachandran plots and analysis of model geometry with the program Procheck (*31*). The structure of ePHP, refined at 1.7 Å resolution with good geometry, has an R of 20.4% and an R_{free} of 24.1% (Table 2). The excellent quality of the electron density map allowed the placement of all residues with the exception of the first residue at the N terminus, which is disordered. The average temperature factors for main chain atoms and side chain atoms are 11.0 and 13.6, respectively.

Superposition Analysis. The refined coordinates of ePHP were superimposed on the coordinates of the apo form of PTE using the program superimpose (*32*). The optimum superposition was obtained by removing residues where the structures showed the largest differences. The following portions of the ePHP structure were not included in the superposition: the region between β1 and α1, including antiparallel β strands β2 and β3 (residues 17−32); the loop between β7 and α5 (residues 160−164); the loop between β9 and α7 (residues 213−217); and the loop between β10 and α8 (residues 242−256). The overall rms deviation of the main chain atomic coordinates determined from the superposition of the two structures was 1.7 Å.

Electrostatic Calculations. Electrostatic surface potentials for ePHP and urease were calculated with the program GRASP (*33*). The accessible surface was determined with a probe radius of 1.4 Å. Calculations were performed using a solvent dielectric constant of 80 and a protein dielectric constant of 2.

RESULTS AND DISCUSSION

Sequence Alignment. The original alignment of ePHP and phosphotriesterase was reported by Scanlan and Reid (*13*). Since that time, four new family members have been identified (*14*−*16*; GenBank reference x99477). Figure 2 shows an alignment of all six known members of the phosphotriesterase family. The phosphotriesterase homology protein (ePHP) nomenclature adopted by Scanlan and Reid (*13*) is used with additional prefixes to indicate the organism from which the protein originated. The *E. coli* protein is labeled ePHP, the *Mycoplasma pneumoniae* protein mpPHP, the *Mycobacterium tuberculosis* protein mtPHP, the mouse protein muPHP, and the rat protein rPHP. The mouse and rat proteins are 92% identical at the amino acid level and may be species variants of the same enzyme. Phosphotriesterase, ePHP, mtPHP, and mammalian PHPs are 27−30% identical at the amino acid level. Although mpPHP is less similar to other members of the family, the level of sequence

```
PTE      1  MQTRRVVLKS  AAAAGTLLGG  LAGCASVAGS  IGTGDRINTV  RGPITISEAG
ePHP        ----------  ----------  ----------  ----------  ---MSFDPTG
mtPHP    1  ----------  ----------  ----------  ---MPELNTA  RGPIDTADLG
mpPHP    1  ----------  ----------  ----------  ------VRTV  LGDIDPKDLG
muPHP    1  ----------  ----------  ---------M  SSLSGKVQTV  LGLVEPSQLG
rPHP     1  ----------  ----------  ---------M  SSLSGKVQTV  LGPVEPSQLG

PTE     51  FTLTHEHICG  SSAGF-----  ----------  ------LRAW  PEFFGSRKAL
ePHP     8  YTLAHEHLHI  DLSGF-----  ----------  ---------K  NNVDCRLDQY
mtPHP   18  VTLMHEHVFI  MTTEI-----  ----------  ----------  -AQNYPEAWG
mpPHP   15  ICDCHDHLIK  NWGPE-----  ----------  ---------A  KEHPDFVMLS
muPHP   22  RTLTHEHLTM  TFDSFYCPPP  PCHEVTSKEP  IMMKNLFWIQ  KNPYSHRENL
rPHP    22  RTLTHEHLTM  AFDSFYCPPP  PCQEAASREP  IMMKNLFWIQ  KNPYSHQENL

PTE     80  A-----EKAV  RGLRRARAAG  VRTIVDVSTF  DIGRDVSLLA  EVSRAA--DV
ePHP    34  AF------IC  QEMNDLMTRG  VRNVIEMTNR  YMGRNAQFML  DVMRET--GI
mtPHP   42  DEDKRVAGAI  ARLGELKARG  VDTIVDLTVI  GLGRYIPRIA  RVAAAT--EL
mpPHP   41  N-----EAAI  KECL*FVHHG  GRSIVTMDPP  NVGRDVKRMV  AIAEQLKGKL
muPHP   72  QLNQEVGAIR  EELLYFKAKG  GGALVENTTT  GLSRDVHTLK  WLAEQT--GV
rPHP    72  QLNQEVEAVR  EELLYFKAKG  GGAVVENTTT  GLSRDVRTLK  WLAEQT--GV

PTE    123  HIVAATGL--  WFDPPLS---  ----MRLRSV  EELTQFFLRE  IQYGIE----
ePHP    76  NVVACTGY--  YQDAFFPEH-  ----VATRSV  QELAQEMVDE  IEQGID----
mtPHP   90  NIVVATGLYT  YNDVPFYFHY  LGPGAQLDGP  EIMTDMFVRD  IEHGIA----
mpPHP   85  NIIMATGF--  HKAAFYDKGS  SW--LAQVPV  NEIVPMLVAE  IEEGMDLYNY
muPHP  120  HIIAGAGF--  YVDATHSAA-  ----TRAMSV  EQLTDVLINE  ILHGAD----
rPHP   120  HIIAGAGF--  YVDATHFAA-  ----TRAMSV  EQLTDVLISE  ILHGAD----

PTE    160  ------DTGI  RAGIIKVATT  GKA--TPFQE  LVLKAAARAS  LATGVPVTTH
ePHP   115  ------GTEL  KAGIIAEIGT  SEGKITPLEE  KVFIAAALAH  NQTGRPISTH
mtPHP  136  ------DTGI  KAGILKCATD  EPG-LTPGVE  RVLRAVAQAH  KRTGAPISTH
mpPHP  131  SGPVVKRGKA  KAGIIKA-GT  GYAAIDRLEL  KALEAVAITS  ITTGAPVLVH
muPHP  159  ------GTSI  KCGVIGEIGC  SWP-LTDSER  KILEATAHAQ  AQLGCPVIIH
rPHP   159  ------GTSI  KCGVIGEIGC  SWP-LTDSER  KVLQATAHAQ  AQLGCPVIIH

PTE    202  TAASQRDGEQ  QAAIFESEGL  SPSRVCIGHS  DDTD-DLSYL  TALA-ARGYL
ePHP   159  TS-FSTMGLE  QLALLQAHGV  DLSRVTVGHC  DLKD-NLDNI  LKMI-DLGAY
mtPHP  179  THAGLRRGLD  QQRIFAEEGV  DLSRVVIGHC  GDST-DVGYL  EELI-AAGSY
mpPHP  180  TQ-LGTMAYE  AAQHLIDFGV  NPRKI*LSHL  NKNP-DEYYY  AKIIRELGVT
muPHP  202  PGRNPGAPFQ  IIRILQEAGA  DISKTVMSHL  DRTIFDKKEL  LEFA-QLGCY
rPHP   202  PGRNPGAPFQ  IIRVLQEAGA  DISKTVMSHL  DRSIFDKKEL  LEFA-QLGCY

PTE    250  IGLDHIPHSA  IGLEDNASAS  ALLGIRSWQT  RALLIKALID  QGYMKQILVS
ePHP   206  VQFDTIGKNS  Y---------  -----YPDEK  RIAMLHALRD  RGLLNRVMLS
mtPHP  227  LGMDRFGVDV  I---------  -----SPFQD  RVNIVARMCE  RGHADKMVLS
mpPHP  227  LCFDGPDRVK  Y---------  -----YPDCL  LAKHIKYLVD  LGFVKHITLA
muPHP  251  LEYDLFGTEL  L----NYQLS  PDIDMPDDNK  RIRRVHFLVD  EGYEDRILMA
rPHP   251  LEYDLFGTEL  L----NYQLS  PDIDLPDDNK  GLGGVRFLVN  EGYEDRILMA

PTE    300  NDWLFGFSSY  VTNIMDVMDR  VNPDGMAFIP  LRVIPFLR--  EKGVPQETLA
ePHP   242  MD--------  -ITRRSHLKA  NGGYGYDYLL  TTFIPQLR--  QSGFSQADVD
mtPHP  263  HDAC--CYFD  ALPEELVPVA  MPNWHYLHIH  NDVIPALK--  QHGVTDEQLH
mpPHP  263  LDAGRVLYQN  TMVKKGRSA-  LGLHICLNAL  FRSLKKWEFP  MRPLTPSWSK
muPHP  297  HD--------  -IHTKHRLMK  YGGHGYSHIL  TNIVPKML--  LRGLTERVLD
rPHP   297  HD--------  -IHTKHRLMK  YGVHGYSHIL  TNVVPKML--  LRGLTERVLD

PTE    348  GITVTNPARF  LSPTLRAS
ePHP   281  VMLRENPSQF  FQ
mtPHP  309  TMLVDNPRRI  FERQGGY
mpPHP  312  ISLKSWPLMH  RGSLIPRLCI  HGY
muPHP  336  KILIENPKQW  LTFK
rPHP   336  KILRENPKQW  LTFK
```

FIGURE 2: Alignment of the sequences of PTE and those of phosphotriesterase homology proteins from *E. coli* (ePHP), *M. tuberculosis* (mtPHP), *M. pneumoniae* (mpPHP), mouse (muPHP), and rat (rPHP). Sequences were aligned with the PILEUP and BESTFIT functions of the GCG Software package (Genetics Computer Group, Inc., Madison, WI). Residues that make up the Zn^{2+} binding site in phosphotriesterase and the analogous residues in the PHP proteins are boxed. Stars (*) in the *M. pneumoniae* sequence indicate positions where stop codons are found in the database sequence. The two loops in the PTE sequence, that between β9 and α7 and that between β10 and α8, which were found in the crystallographic structure of PTE to participate in the formation of the active site and to contain residues that form contacts with the substrate analogue diethyl 4-methylbenzylphosphonate are highlighted in gray. The residues within these loops that participate in binding the phosphonate inhibitor are shown in white.

identity increases to 29% if only residues 50−250 are considered. The aspartate and all four histidine residues that coordinate Zn^{2+} in phosphotriesterase are conserved across the six members of the phosphotriesterase family. Only the lysine at position 169 is not strictly conserved. This residue is replaced by a glutamate and is shifted by one position in the alignment for ePHP, muPHP, and rPHP. Both mtPHP and mpPHP have a lysine at the position corresponding to 169 in phosphotriesterase, and these lysines are likely to be similarly carbamylated.

Metal Ion Composition. The conservation in ePHP of the residues involved in zinc ligation in phosphotriesterase suggested that ePHP would also be a metalloenzyme. We therefore analyzed purified ePHP for metal content. Initial

Table 3: Analysis of Metal Content of ePHP Expressed in the Presence of Different Divalent Cations[a]

growth conditions	moles of metal ion per mole of ePHP							
	Mn	Fe	Co	Ni	Cu	Zn	Cd	total
1 mM Mn^{2+}	0.2			0.2		1		1.4
1 mM Fe^{2+}		0.8		0.1		1.5		2.4
1 mM Co^{2+}			0.4		0.1	0.8		1.3
1 mM Ni^{2+}		0.1				1		1.1
1 mM Cu^{2+}		0.1	0.1		0.2	1.1		1.5
1 mM Zn^{2+}		0.1				1.9		2.0
1 mM Cd^{2+}				0.1	0.1	0.3	0.7	1.2

[a] The protein concentration was determined by amino acid analysis, and the metal ion content was determined by ICP emission spectroscopy.

analysis of purified ePHP grown in simple LB media showed incorporation of approximately 0.7 mol of zinc, 0.2 mol of iron, and 0.2 mol of manganese per mole of protein (data not shown). The low metal stoichiometry suggested that the endogenous metal ion concentration of the LB growth media was insufficient to allow full metal incorporation in the overexpressed ePHP. The problem was overcome by supplementing the growth media with divalent cations. In the presence of 1 mM Zn^{2+}, ePHP incorporates two zinc atoms per protein molecule. The results of expression of ePHP in media containing different divalent transition metals, Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}, or Cd^{2+}, are shown in Table 3. *E. coli* PHP has a clear preference for binding Zn^{2+} over Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+}, and Cu^{2+}. In every case, Zn^{2+} in the LB growth media successfully competes for incorporation into ePHP, even in the presence of large excesses of alternative metal ions. Only Cd^{2+}, when present in large excess, displaces Zn^{2+}, and Cd^{2+} is unlikely to be associated with the protein under physiological conditions. Zinc and cadmium often display similar binding affinities for proteins, and cadmium toxicity in many animals is thought to arise from disruption of zinc metabolism (*34*). *E. coli* PHP also shows affinity for Fe^{2+}, and curiously, the presence of Fe^{2+} increases Zn^{2+} uptake, possibly by activation of an Fe^{2+} ionophore that transports Zn^{2+} to some extent as well.

Enzyme Activity Assays. ePHP was tested for general esterase, aminopeptidase, sulfatase, phosphatase, carbonic anhydrase, phosphodiesterase, and phosphotriesterase activities with the following substrates: *p*-nitrophenyl acetate, L-alanine nitroanilide, *p*-nitrophenyl sulfate, bis(*p*-nitrophenyl) phosphate, paraoxon, and *p*-nitrophenyl phosphate. No enzymatic activity was detected with any of these nonspecific substrates. The crude lysate of bacteria overexpressing ePHP was also assayed for phosphotriesterase activity because of concern that enzymatic activity might be lost during purification. No nonspecific phosphotriesterase activity, however, was detected in crude lysates.

Gel Filtration of ePHP. The phosphotriesterase protein from *Pseudomomas* is thought to function as a dimer (*7*) on the basis of ultracentrifugation studies and its dimeric association in X-ray crystallographic structures (*17*). To establish whether ePHP exists as a monomer or multimer, we used size exclusion chromatography and compared the migration of ePHP with that of protein standards with known molecular weights. As shown in Figure 3, ePHP elutes close to carbonic anhydrase at 29 kDa, which agrees well with the molecular weight of 32915 predicted from its amino acid sequence, and indicates that purified ePHP exists in a

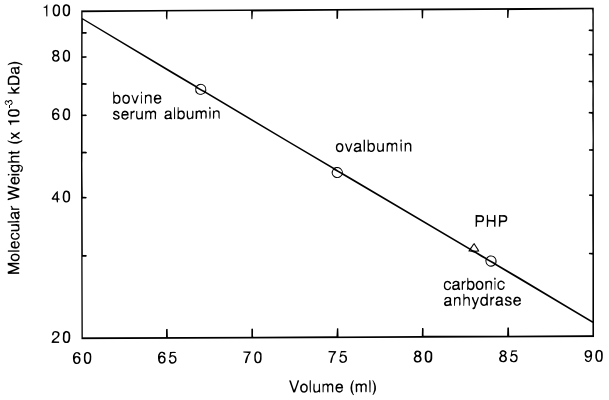

FIGURE 3: Determination of native molecular size using gel filtration chromatography. One milligram of each purified ePHP protein and three molecular mass standards (BSA, ovalbumin, and carbonic anhydrase) were loaded onto a 900 cm × 16 mm ID Sephacryl S-100 HR (Pharmacia) column pre-equilibrated with 10 mM Tris·HCl and 150 mM NaCl (pH 7.5). The flow rate was maintained at 0.5 mL/min using a peristaltic pump, and the eluant was monitored at 254 nm during collection of 6.4 mL fractions. Eluted proteins were identified using polyacrylamide gel electrophoresis.

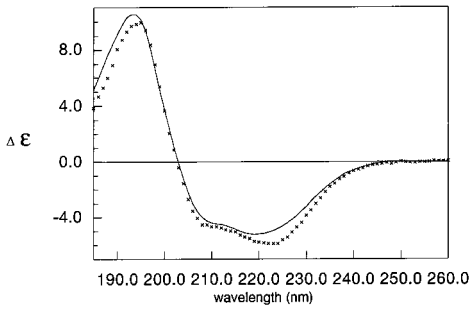

FIGURE 4: Comparison of the circular dichroism spectra of ePHP and triosephosphate isomerase. Circular dichroism spectra were recorded on an Aviv Circular Dichroism Spectrometer (model 62DS) instrument using a 1 mm path length cuvette. Purified ePHP protein at a concentration of 2.5 μM in buffer containing 50 mM KF and 10 mM $KHPO_4$ (pH 7.5) was used for data collection. The protein concentration was determined by amino acid analysis (University of Michigan). Data were recorded every 1 nm from 185 to 260 nm, and a buffer blank was subtracted. The solid line shows the circular dichroism spectrum for triosephosphate isomerase that was obtained from data provided in the Dicroprot V2.1 portion of the Antheprot program (Deleage and Geourjon, 1993). The crosses (×) show the CD spectrum of ePHP for comparison.

monomeric state. To rule out the possibility that hydrophobic interaction chromatography causes dimer dissociation, ePHP, only partially purified by ion exchange chromatography, was similarly subjected to analytical gel filtration chromatography. Again, ePHP was found to elute with a retention time consistent with monomer composition (data not shown). Although most known α/β barrels form dimers or larger multimers, ePHP appears to be monomeric at least in the absence of bound substrate.

Circular Dichroism. The crystallographic structure of PTE shows that the enzyme has an α/β barrel fold. ePHP shares 28% sequence identity with PTE, and structural similarities between the proteins were anticipated (*13*). The circular dichroism spectrum of ePHP and that of triosephosphate isomerase, an archetypical α/β barrel protein, are highly similar in shape and amplitude, consistent with ePHP also having the α/β barrel fold (Figure 4).

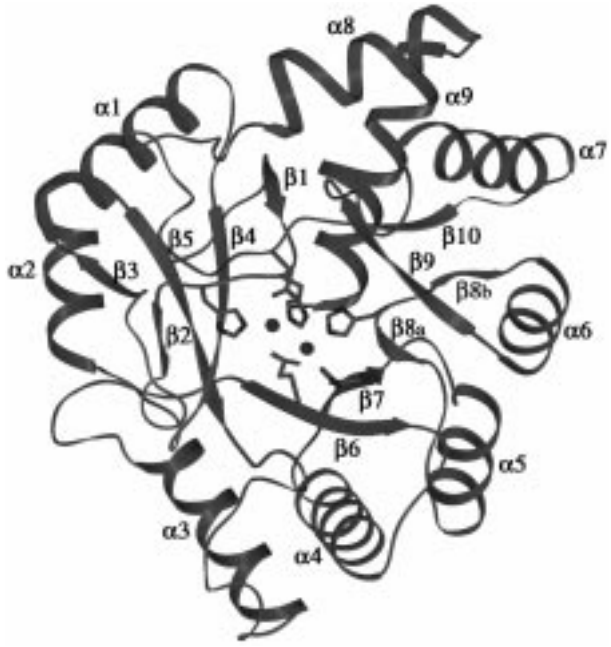

FIGURE 5: Ribbon representation of the structure of ePHP. The protein forms an α/β barrel and in addition has two antiparallel β strands, β2 and β3, inserted between the first β strand and α helix of the barrel. A helical excursion, α9, follows the last α helix of the barrel motif and extends inward toward the binuclear metal center. The α helices are colored purple, and the β strands are green. The two zinc ions are drawn as blue spheres located at the carboxyl end of the α/β barrel. The structure was rendered using the program Setor (*43*).

Description of the α/β Barrel Structure. The X-ray crystallographic structure of ePHP confirms predictions from the circular dichroism analysis. The structure shows that ePHP has an elliptical $(\alpha\beta)_8$ barrel fold with two antiparallel β strands inserted between strand β1 and helix α1 of the barrel (Figure 5). An additional helix, α9, follows α8, pointing inward from the C-terminal end of the barrel toward the binding pocket of the zinc center. The binuclear zinc center is located at the C-terminal end of the β barrel at perhaps the active site of this putative enzyme. The active sites of enzymes possessing the α/β barrel fold are commonly found at similar positions where catalytic residues are contributed either from the C-terminal ends of the β strands of the barrel or from the loops connecting the strands to the α helices that follow them. Figure 6 shows the coordination geometry at the binuclear zinc center. The two zinc ions are separated by 3.35 Å with one of the zinc ions coordinated by five ligands in a trigonal bipyramid arrangement and the other having tetrahedral geometry (Table 4). The ligands about the first zinc ion include the side chains of His14, His12, Asp243, and Glu125. The carboxylate groups of Asp243 and Glu125 occupy the apical positions of the trigonal bipyramid. The identity of the fifth ligand is unclear from the electron density (Figure 7) but may be some molecule from the *E. coli* extract from which the protein was purified or an unknown contaminant in the crystallization medium. The density does not appear to fit any buffer molecule used in crystallization. The carboxylate group of Glu125 and this molecule of unknown identity act as bridging ligands between the two zinc ions. The ligands around the second zinc ion include the side chains of Glu125, His158, and His186 and the molecule of unknown nature arranged with distorted tetrahedral geometry. In small molecule and protein complexes with metal ions, carboxylate groups generally show preferential coordination through the syn lone-pair electrons of the carboxylate oxygens (*35*). Such is the case here, where Glu125 and Asp243 coordinate the zinc ions through their syn lone pairs. Histidine generally binds to metal ions preferentially through the $N_{\epsilon 2}$ position (*36*), and in ePHP, three of the histidine ligands, His12, His14, and His186, are coordinated in this manner. His158 is coordinated with the less favorable $N_{\delta 1}$ interaction and has a strained geometry ($\chi_1 = 160°$ and $\chi_2 = 111°$). Binuclear metal centers of PTE and urease, also α/β barrel metalloproteins, show similar coordination schemes and

Table 4: Coordination Geometry at the Binuclear Zinc Center of ePHP

ligand−metal	distance (Å)	ligand−metal−ligand	angle (deg)
His12 $N_{\epsilon 2}$−Zn_1	2.3	His12 $N_{\epsilon 2}$−Zn_1−H_2O	117.5
His14 $N_{\epsilon 2}$−Zn_1	2.2	His14 $N_{\epsilon 2}$−Zn_1−His12 $N_{\epsilon 2}$	119.1
Asp243 $O_{\delta 2}$−Zn_1	2.4	His14 $N_{\epsilon 2}$−Zn_1−H_2O	123.4
Glu125 $O_{\epsilon 2}$−Zn_1	2.4	Asp243 $O_{\delta 2}$−Zn_1−Glu125 $O_{\epsilon 2}$	171.9
H_2O−Zn_1	2.0		
Glu125 $O_{\epsilon 1}$−Zn_2	2.2	Glu125 $O_{\epsilon 1}$−Zn_2−His158 $N_{\delta 1}$	94.8
His158 $N_{\delta 1}$−Zn_2	2.3	His186 $N_{\epsilon 2}$−Zn_2−His158 $N_{\delta 1}$	93.5
His186 $N_{\epsilon 2}$−Zn_2	2.2	Glu125 $O_{\epsilon 1}$−Zn_2−H_2O	111.8
H_2O−Zn_2	2.3	His186 $N_{\epsilon 2}$−Zn_2−H_2O	110.8

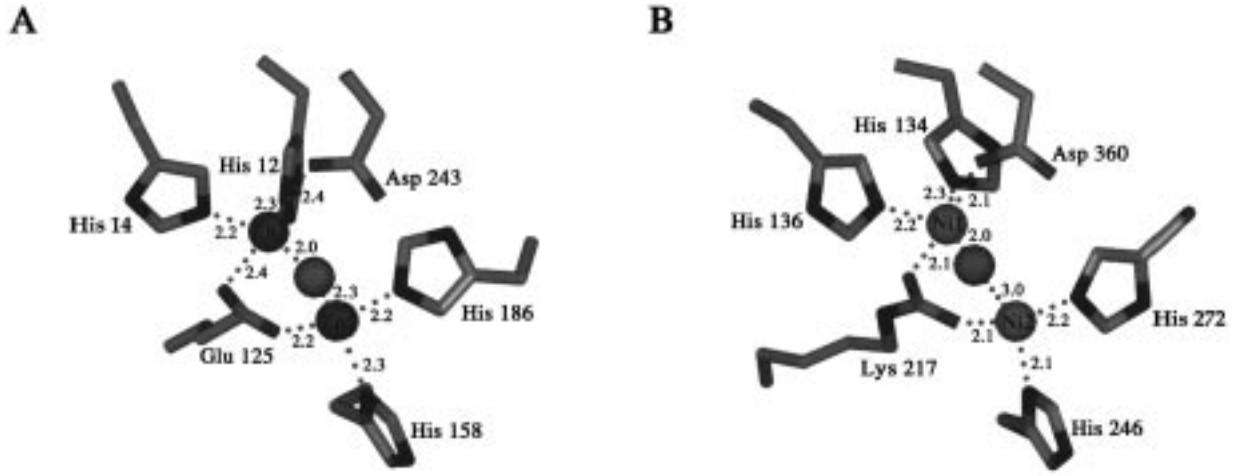

FIGURE 6: Comparison of the binuclear metal centers of (A) ePHP and (B) urease (*37*). Coordinated solvent molecules are drawn as gray spheres. The numbers shown indicate the distances in angstroms between metal ions and ligands.

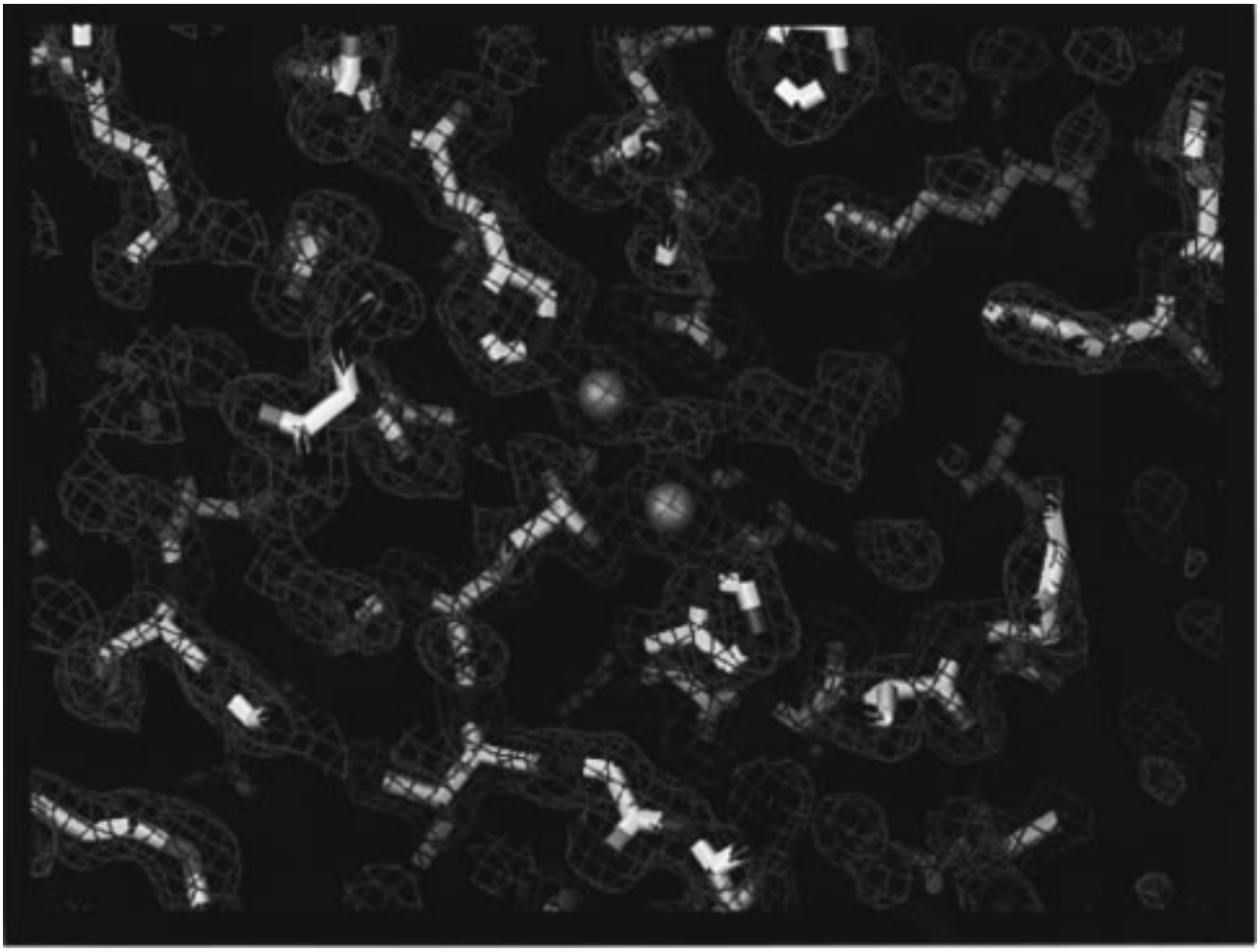

FIGURE 7: $2F_o - F_c$ electron density map contoured at σ in the vicinity of the binuclear zinc center of ePHP. The figure was produced using the program O (*29*).

structures (*18*, *37*). For comparison, the metal center of urease is shown in Figure 6. In both the structures of PTE and urease, a carbamylated lysine residue acts as a bridging ligand at the position of Glu125 in the ePHP structure. The side chain of Glu125 adopts an unusually strained conformation ($\varphi = -71°$, $\psi = 108°$, $\chi_1 = 59°$, $\chi_2 = 179°$, and $\chi_3 = -60°$) presumably because of distortion caused by coordination to the two zinc ions in ePHP. In the structures of PTE and urease, a water molecule serves as the second bridging ligand between the divalent cations at the analogous position of the mysterious solvent molecule shown in the structure of ePHP. It is possible that ePHP, which is similar in structure and perhaps function to these two hydrolytic metalloenzymes, may normally have a coordinated hydroxide at its zinc center, or alternatively, a substrate molecule may bind between the two zinc ions.

A long groove extends from the zinc center of ePHP along the protein surface between the β sheets (strands $\beta1-\beta5$ and strands $\beta6-\beta10$), which compose the barrel (Figure 8). The groove, approximately 17 Å long, 8 Å wide, and 4 Å deep, is lined by resides from loops at the carboxyl ends of the β strands. These loops include that between antiparallel β strands $\beta2$ and $\beta3$, and those connecting strands and helices within the barrel motif, including loops between $\beta5$ and $\alpha3$, $\beta6$ and $\alpha4$, $\beta7$ and $\alpha5$, $\beta8$ and $\alpha6$, $\beta9$ and $\alpha7$, and $\beta10$ and $\alpha8$. The putative active site of ePHP appears to be largely hydrophobic, made up of residues Leu15, Ile17, Leu19, Ile52, Tyr84, Ile126, Phe161, Val184, Leu189, Ile244, and Tyr216 (Figure 9). A calculation of the electrostatic surface potential of ePHP shows that the binding pocket containing the zinc ions lies in a region of positive potential produced by the divalent cations and positively charged residues in the vicinity, including Lys23, Lys213, and Arg246 (Figure 10). This concentration of positive charge contrasts with the overall negative charge of the protein, which has a p*I* of 5.2 and a charge of -12 at neutral pH as judged from the amino acid sequence. This finding suggests that, if ePHP is indeed an enzyme, it may bind an anionic substrate. Similarly, in purple acid phosphatase, which catalyzes the hydrolysis of negatively charged phosphate monoesters and which possesses a binuclear Fe(II)−Zn(II) center, the electrostatic potential is positive in the vicinity of the metal center at the active site (*38*). In contrast, for urease, which catalyzes the hydrolysis of urea, a neutral substrate, the electrostatic potential surrounding its binuclear nickel center appears to be slightly negative. The positive electrostatic potential at the binuclear zinc center of ePHP may serve to attract an anionic substrate into its presumed active site.

Superposition of ePHP and PTE. The refined structures of ePHP and the apo form of PTE were superimposed and found to have an overall rms deviation of 1.7 Å for the atomic coordinates of backbone atoms in aligned regions (Figure 11). PTE is larger than ePHP and has approximately 40 additional residues at the N terminus, including two

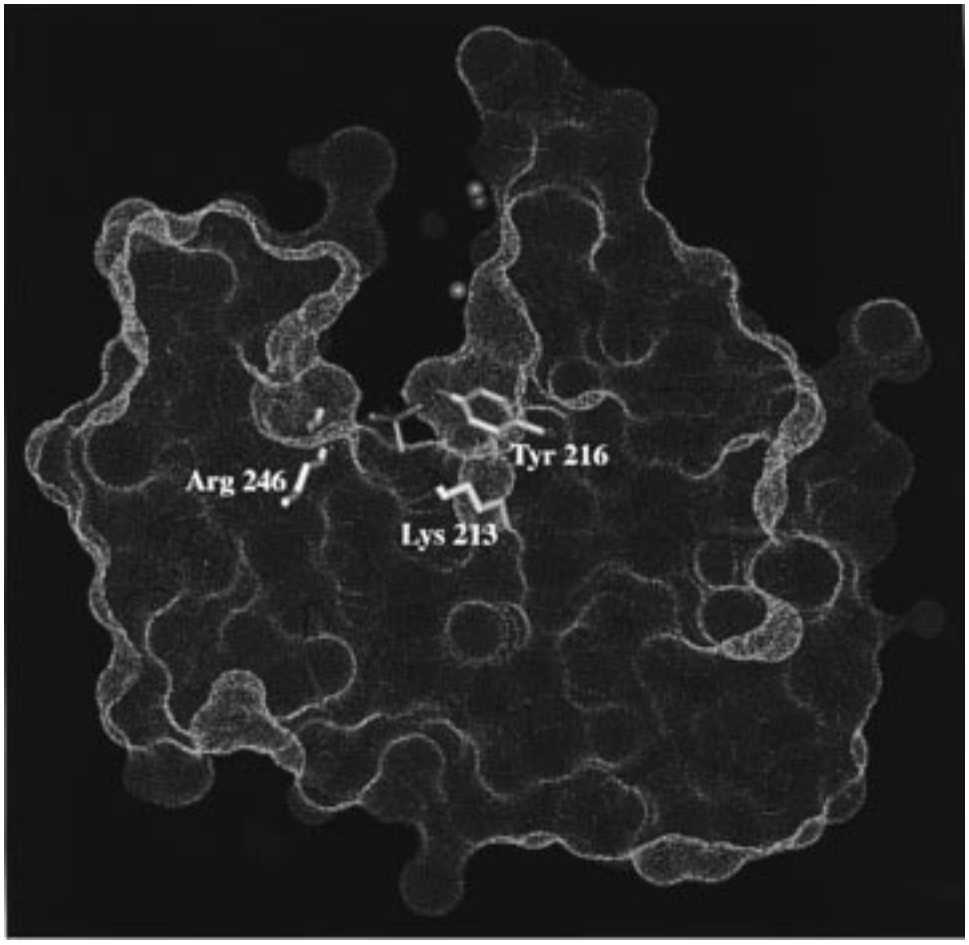

FIGURE 8: Conolly surface of ePHP. The surface was calculated with the program InsightII using a probe radius of 1.4 Å (Molecular Simulations, Inc., and ref *44*). The displayed slab thickness was adjusted to about 12.5 Å to show the long groove, which extends between the parallel β sheets of the α/β barrel. An ammonium sulfate ion from the crystallization medium binds at the base of the groove and forms charged hydrogen bonds with Lys213 and Arg246. Ordered water molecules are shown as blue spheres. The zinc center, though not visible here, is located approximately 7 Å from the sulfate ion and, in this view, would lie above the plane of the displayed section.

antiparallel β strands, which precede the first β strand of ePHP. Portions of the ePHP structure show large differences with respect to the PTE structure and were therefore not included in the superposition, including the region between β1 and α1 that encompasses antiparallel β strands β2 and β3 (residues 17−32), which are not present in PTE; the loop between β7 and α5 (residues 160−164); the loop between β9 and α7 (residues 213−217); and the loop between β10 and α8 (residues 242−256). In superimposed regions, relatively large structural differences are seen in the loop connecting β6 and α4, which has an rms deviation of 2.4 Å, the loop connecting β8 and α6 (residues 185−193), which has an rms deviation of 3.5 Å, helix α8, which has an rms deviation of 2.1 Å, and the loop connecting α8 and α9, which has an rms deviation of 2.3 Å.

Notably, significant differences between the two structures are found in the regions corresponding to the presumptive active site of ePHP. Aside from the residues involved in metal ligation, few of the residues in the active site are conserved. The positively charged residues in ePHP, Lys23, Lys213, and Arg246, are substituted in PTE with Trp, His, and Ile, respectively. In PTE, residues at the active site found to form contacts with the substrate analogue diethyl 4-methylbenzylphosphonate, namely, Phe306, Tyr309, Met317, Leu271, Trp131, and His257, are also not conserved. Met317 is the only relatively conservative substitution and is replaced by a hydrophobic residue, Leu250 in ePHP. Phe306, Tyr309, and Leu271, however, have no corresponding residues in ePHP and reside in regions of insertion in PTE relative to ePHP (Figure 2). In PTE, strand β9 is shorter than it is in ePHP and diverges into a loop that forms a flap over the active site that is not seen in the structure of ePHP. In PTE, the connecting region between β9 and α7 has an insertion of 14 residues with respect to the ePHP sequence and contains Leu271, found in the substrate binding pocket of PTE. The loop between β10 and α8 in PTE, which contains Phe306 and Tyr309, has an insertion of nine residues with respect to ePHP sequence. This loop, which differs substantially in the two structures, is also the location of the conserved aspartate (Asp243 in ePHP) involved in zinc ligation.

PTE lacks the long groove observed along the surface of ePHP, in part because of insertions in the loops and differences in the positions of the loops surrounding the active site. In particular, the loop between β9 and α7 occludes the active site in PTE. Even if the insertion in this loop is deleted from the PTE structure, however, the enzyme lacks a long groove because of structural differences between ePHP and PTE in the region connecting β1 and α1. In ePHP, a loop between antiparallel β strands, β2 and β3, contributes to a side wall of the groove. In PTE, the long loop replacing these strands is not in the proper position for

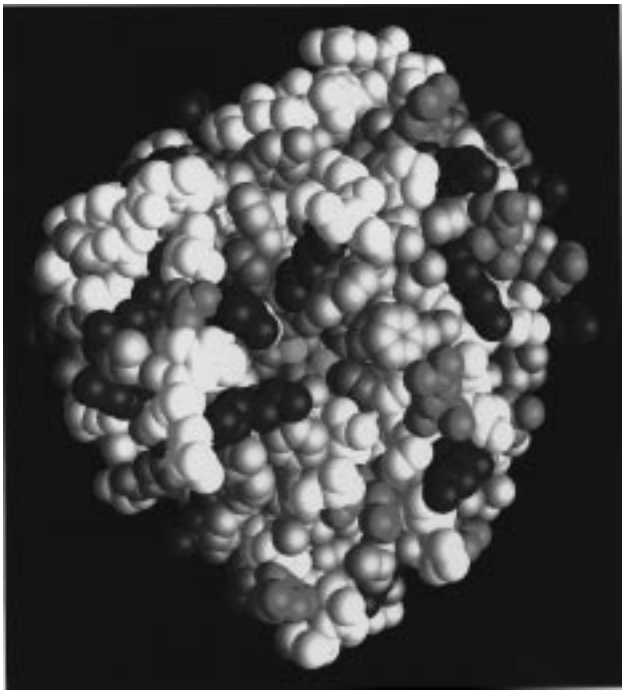

FIGURE 9: Space filling model of the structure of ePHP colored according to the nature of amino acid side chains. Basic residues are colored dark blue, acidic residues red, aromatic and other hydrophobic residues gold, histidines medium light blue, and zinc ions aqua.

formation of a long binding crevice. Several of the regions in the apo form of PTE that show large structural deviations from ePHP apparently undergo structural changes in the presence of bound divalent cations, including the region between β1 and α1, the loop between β7 and α5, the loop between β9 and α7, and the loop between β10 and α8 (*6*, *17*, *18*). Once the coordinates for the complex of PTE with divalent cations become available, we look forward to providing a detailed analysis and comparison of the structures.

If ePHP indeed belongs to the family of enzymes from which PTE evolved, the substantial sequence and structural differences indicate it is a distant ancestor. All α/β barrel proteins discovered so far have been found to be enzymes with active sites at similar positions at the carboxyl end of their α/β barrels. On the basis of the common location of active sites in these enzymes, it has been suggested that α/β barrel proteins are related to one another by divergent evolution from a common ancestor (*39*). ePHP is most similar in structure to a subclass of α/β barrel proteins that includes PTE, as well as urease and adenosine deaminase. These proteins all have long elliptical α/β barrels and bind metal ions with nearly identical ligands (*18*, *37*, *40*). In PTE and urease, Glu125 is replaced with a carbamylated lysine. In adenosine deaminase, Glu125 is replaced with an aspartate. Curiously, adenosine deaminase binds only one zinc ion despite having four histidine residues and an aspartate at positions analogous to those observed for the residues which bind zinc in ePHP. The three proteins in this subclass of α/β barrel proteins with known catalytic activities are hydrolytic enzymes, and this observation might lead one to speculate that ePHP may also catalyze some hydrolytic reaction.

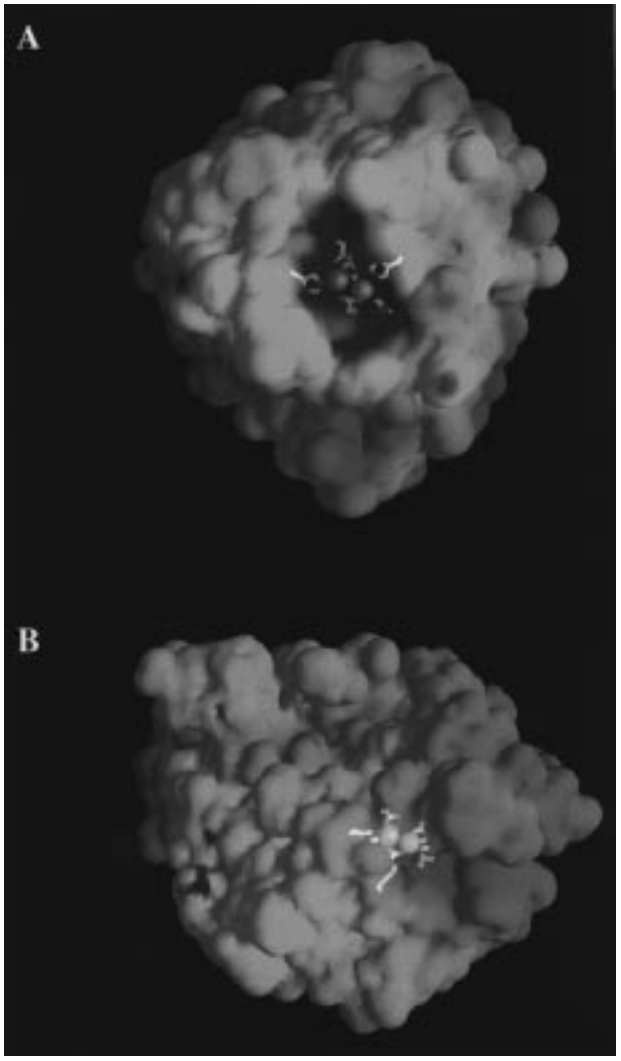

FIGURE 10: Representation of the electrostatic surface potentials of (A) ePHP and (B) urease. The accessible surfaces were determined for the structures with the metal ions and solvent molecules omitted. Calculations of electrostatic potential included the charge on the divalent metal ions. Histidines were assigned a neutral charge. Regions of positive potential are colored blue, and regions of negative potential are colored red. The orientation of the molecule is the same as that in Figure 5.

Zinc is found in a wide range of hydrolytic enzymes, including carbonic anhydrase, various carboxypeptidases, nucleases, and alkaline and acid phosphatases (*41*, *42*). Although assays so far have not detected enzymatic activity, the protein may require a specific substrate not yet identified. We propose that ePHP is most likely a hydrolytic enzyme, utilizing an anionic substrate. The long groove along the surface of ePHP suggests its substrate may be polymeric. Like in PTE, urease, adenosine deaminase, and carbonic anhydrase, at least one of the zinc ions may coordinate a water molecule, thus lowering its pK_a to enhance nucleophilicity. Alternatively, zinc may play a role analogous to that in carboxypepetidase and serve as an electrophilic catalyst, polarizing a carbonyl or phosphoryl oxygen bond to activate a substrate for nucleophilic attack. In addition, the metal ions may lower the pK_a of a substrate leaving group to facilitate bond cleavage. The numerous examples of metal ion catalysis in enzymes suggest many possibilities that will only be delineated once the true physiological function of ePHP is known.

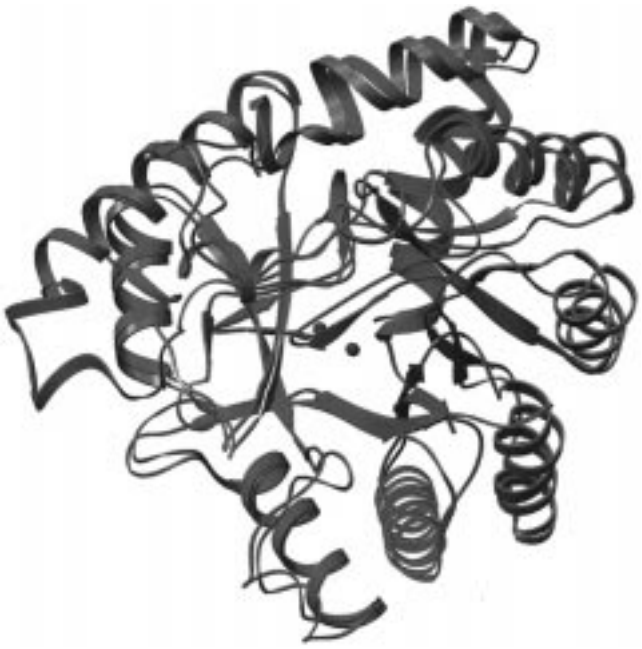

FIGURE 11: Superposition of the refined structures of ePHP and the apo form of phosphotriesterase (Brookhaven Protein Data Bank entry 1PTA). The structure of ePHP is displayed in purple and that of PTE in green.

ACKNOWLEDGMENT

We thank Ralph Reid for helpful discussions. We thank Dr. Alan Frankel for the use of his CD instrument and Dr. Kazuo Harada for help using the instrument. We also thank Dr. Charles Craik for the gift of the pTactac plasmid and the UCSF Biomolecular Resource Center for oligonucleotide synthesis and sequencing.

REFERENCES

1. Benner, S., and Ellington, A. D. (1988) *CRC Crit. Rev. Biochem. 23*, 369−426.
2. Llewellyn, D. J., Daday, A., and Geoffrey, D. S. (1980) *J. Biol. Chem. 255*, 2077−2084.
3. Krishnan, S., Hall, B. G., and Sinnott, M. L. (1995) *Biochem. J. 312*, 971−977.
4. Dumas, D. P., Caldwell, S. R., Wild, J. R., and Raushel, F. M. (1989) *J. Biol. Chem. 264*, 19659−19665.
5. Omburo, G. A., Kuo, J. M., Mullins, L. S., and Raushel, F. M. (1991) *Appl. Environ. Microbiol. 57*, 13278−13283.
6. Benning, M. M., Kuo, J. M., Raushel, F. M., and Holden, H. M. (1995) *Biochemistry 34*, 7973−7978.
7. McDaniel, C. S., Harper, L. L., and Wild, J. R. (1988) *J. Bacteriol. 170*, 2306−2311.
8. Mulbry, W. W., Karns, J. S., Kearney, P. C., Nelson, J. O., McDaniel, C. S., and Wild, J. R. (1986) *Appl. Environ. Microbiol. 51*, 926−930.
9. Donarski, W. J., Dumas, D. P., Heitmeyer, D. P., Lewis, V. E., and Raushel, F. M. (1989) *Biochemistry 28*, 4650−4655.
10. Dumas, D. P., Durst, H. D., Landis, W. G., Raushel, F. M., and Wild, J. R. (1990) *Arch. Biochem. Biophys. 277*, 155−159.
11. Caldwell, S. R., Newcomb, J. R., Schlecht, K. A., and Raushel, F. M. (1991) *Biochemistry 30*, 7438−7444.
12. Schrader, G. (1950) *Angew. Chem. 62*, 471−473.
13. Scanlan, T. S., and Reid, R. C. (1995) *Chem. Biol. 2*, 71−75.
14. Hou, X., Maser, R. L., Magenheimer, B. S., and Calvet, J. P. (1996) *Gene 168*, 157−163.
15. Himmelreich, R., Hilbert, H., Plagens, H., Li, B.-C., and Herrmann, R. (1996) *Nucleic Acids Res. 24*, 4420−4449.
16. Philipp, W. J., Poulet, S., Eiglmeier, K., Pascopella, L., Balasubramanian, V., Heym, B., Bergh, S., Bloom, B. R., Jacobs, W., Jr., and Cole, S. T. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 3132−3137.
17. Benning, M. M., Kuo, J. M., Raushel, F. M., and Holden, H. M. (1994) *Biochemistry 33*, 15001−15007.
18. Vanhooke, J. L., Benning, M. M., Raushel, F. M., and Holden, H. M. (1996) *Biochemistry 35*, 6020−6025.
19. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) *J. Mol. Biol. 215*, 403−410.
20. McGrath, M. E., Hines, W. M., Sakanari, J. A., Fletterick, R. J., and Craik, C. S. (1991) *J. Biol. Chem. 266*, 6620−6625.
21. McGrath, M. E., Erpel, T., Browner, M. F., and Fletterick, R. J. (1991) *J. Mol. Biol. 222*, 139−142.
22. Wilbur, K. M., and Anderson, N. G. (1948) *J. Biol. Chem. 176*, 147−154.
23. Otwinowski, Z. (1993) in *Data Collection and Processing, Proceedings of the CCP4 Weekend, 29−30 January 1993* (Sawyer, L., Isaacs, N., and Bailey, S., Eds.) pp 56−62, SERC, Daresbury Laboratory, Daresbury, U.K.
24. Matthews, B. W. (1968) *J. Mol. Biol. 33*, 491−497.
25. Furey, W., and Swaminathan, S. (1990) *American crystallographic association meeting abstracts, Series 2*, Vol. 18, p 73.
26. Zhang, K. Y. J., and Main, P. (1990) *Acta Crystallogr. A46*, 377−381.
27. Navaza, J. (1994) *Acta Crystallogr. A50*, 157−163.
28. Brunger, A. T. (1987) in *X-PLOR Manual version 3.1*, Yale University Press, New Haven, CT.
29. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard (1991) *Acta Crystallogr. A47*, 110−119.
30. Murshudov, G. N., Dodson, E. J., and Vagin, A. A. (1996) in *Proceedings of the CCP4 Study Weekend* (Dodson, E., Moore, M., Ralph, A., and Bailey, S., Eds.) pp 93−104, Council for the Central Laboratory of the Research Councils, Daresbury Laboratory, Daresbury, U.K.
31. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
32. Diederichs, K. (1995) *Proteins 23*, 187−195.
33. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins 11*, 281−296.
34. Martin, R. B. (1986) *Metal Ions in Biological Systems*, Vol. 20, pp 21−65, Marcel Dekker, Inc., New York.
35. Chakrabarti, P. (1990) *Protein Eng. 4*, 49−56.
36. Chakrabarti, P. (1990) *Protein Eng. 4*, 57−63.
37. Jabri, E., Carr, M. B., Hausinger, R. P., and Karplus, P. A. (1995) *Science 268*, 998−1004.
38. Klabunde, T., Strater, N., Frohlich, R., Witzel, H., and Krebs, B. (1996) *J. Mol. Biol. 259*, 737−748.
39. Farber, G. K., and Petsko, G. A. (1990) *Trends Biochem. Sci. 15*, 228−234.
40. Wilson, D. K., Rudolph, F. B., and Quiocho, F. A. (1991) *Science 252*, 1278−1284.
41. Wilcox, D. E. (1996) *Chem. Rev. 96*, 2435−2458.
42. Lipscomb, W. N., and Strater, N. (1996) *Chem. Rev. 96*, 2375−2433.
43. Evans, S. V. (1993) *J. Mol. Graphics 11*, 134−138.
44. Dayringer, H., Tramontano, A., Sprang, S., and Fletterick, R. J. (1986) *J. Mol. Graphics 4*, 82−91.

BI971707+