Structural and Functional Comparisons of Nucleotide Pyrophosphatase/Phosphodiesterase and Alkaline Phosphatase: Implications for Mechanism and Evolution[†,‡]

Jesse G. Zalatan,[§] Timothy D. Fenn,[||] Axel T. Brunger,[||] and Daniel Herschlag*[,§,⊥]

Departments of Chemistry, Biochemistry, Molecular and Cellular Physiology, Neurology, and Neurological Sciences, Stanford Synchrotron Radiation Laboratory, and Howard Hughes Medical Institute, Stanford University, Stanford, California 94305

Received April 29, 2006; Revised Manuscript Received May 31, 2006

ABSTRACT: The rapid expansion of the amount of genomic and structural data has provided many examples of enzymes with evolutionarily related active sites that catalyze different reactions. Functional comparisons of these active sites can provide insight into the origins of the enormous catalytic proficiency of enzymes and the evolutionary changes that can lead to different enzyme activities. The alkaline phosphatase (AP) superfamily is an ideal system to use in making such comparisons given the extensive data available on both nonenzymatic and enzymatic phosphoryl transfer reactions. Some superfamily members, such as AP itself, preferentially hydrolyze phosphate monoesters, whereas others, such as nucleotide pyrophosphatase/phosphodiesterase (NPP), preferentially hydrolyze phosphate diesters. We have measured rate constants for NPP-catalyzed hydrolysis of phosphate diesters and monoesters. NPP preferentially catalyzes diester hydrolysis by factors of 10^2-10^6, depending on the identity of the diester substrate. To identify features of the NPP active site that could lead to preferential phosphate diester hydrolysis, we have determined the structure of NPP in the absence of ligands and in complexes with vanadate and AMP. Comparisons to existing structures of AP reveal bimetallo cores that are structurally indistinguishable, but there are several distinct structural features outside of the conserved bimetallo site. The structural and functional data together suggest that some of these distinct functional groups provide specific substrate binding interactions, whereas others tune the properties of the bimetallo active site itself to discriminate between phosphate diester and monoester substrates.

Enzymes in the alkaline phosphatase (AP)[1] superfamily catalyze a wide variety of phosphoryl and sulfuryl transfer reactions (*1*). Members of this superfamily include phosphatases, phosphodiesterases, phosphoglycerate mutases, phosphopentomutases, and sulfatases. AP itself preferentially catalyzes phosphate monoester hydrolysis, presumably to harvest phosphate for nucleic acids and metabolites, and its structural and functional properties have been extensively studied (*2*). The nucleotide pyrophosphatase/phosphodiesterase (NPP) enzymes are members of the AP superfamily that catalyze phosphate diester hydrolysis (*1*, *3*, *4*). In eukaryotes, these enzymes are found at cell surfaces, either as transmembrane proteins or secreted, and hydrolyze extracellular phosphate diesters to affect a variety of biological processes (*5*, *6*).

Although no experimental structural information was previously available for NPP, a model for the NPP active site that is strikingly similar to the AP active site was proposed on the basis of sequence comparisons and homology modeling. The model consists of two active site metal ions, six conserved metal ligands, and a Thr residue positioned in a manner analogous to that of a Ser residue in AP (Scheme 1) (*3*). In AP-catalyzed phosphate monoester hydrolysis reactions, the Ser alkoxide displaces the leaving group in the first step of a double-displacement mechanism to produce a covalent enzyme−phosphate intermediate, which is subsequently hydrolyzed by water in the second step of the reaction (*2*). An analogous mechanism was proposed for NPP-catalyzed phosphate diester hydrolysis on the basis of the predicted similarity of the bimetallo active site with AP and the observation that the active site Thr

[†] This work was supported by a grant from the NIH to D.H. (GM64798). J.G.Z. was supported in part by a Hertz Foundation Graduate Fellowship. T.D.F. was supported by the Universitywide AIDS Research Program of the University of California (F03-ST-216). Portions of this research were conducted at the Advanced Light Source, a national user facility operated by Lawrence Berkeley National Laboratory.

[‡] Coordinates and structure factors have been deposited in the RCSB Protein Data Bank as entries 2GSN (holoenzyme structure), 2GSO (vanadate-bound structure), and 2GSU (AMP-bound structure).

* To whom correspondence should be addressed: Department of Biochemistry, Beckman Center, B400, Stanford University, Stanford, CA 94305-5307. Phone: (650) 723-9442. Fax: (650) 723-6783. E-mail: herschla@stanford.edu.

[§] Department of Chemistry, Stanford University.

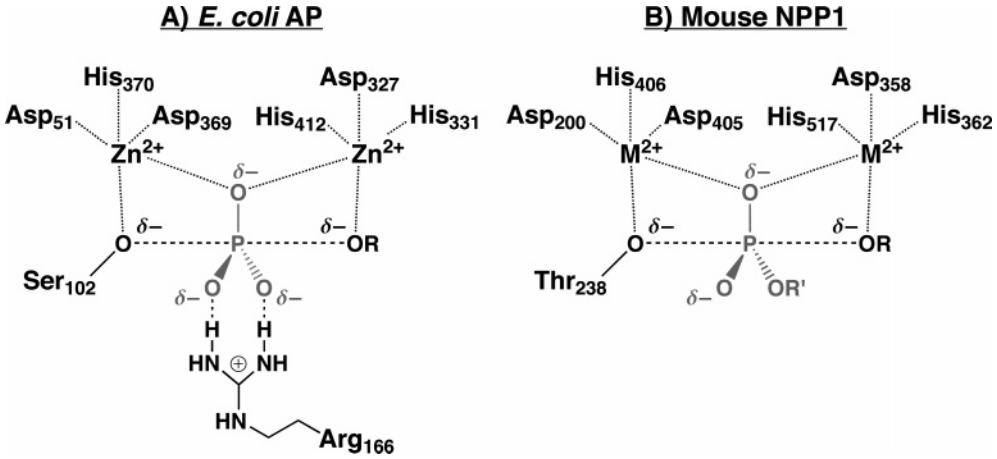
[||] Departments of Molecular and Cellular Physiology, Neurology, and Neurological Sciences, Stanford Synchrotron Radiation Laboratory, and Howard Hughes Medical Institute, Stanford University.

[⊥] Department of Biochemistry, Stanford University.

[1] Abbreviations: NPP, nucleotide pyrophosphatase/phosphodiesterase; AP, alkaline phosphatase; *Xac*, *X. axonopodis pv. citri*; $MpNPP^-$, methyl 4-nitrophenyl phosphate; pNPP, 4-nitrophenyl phosphate; bis-$pNPP^-$, bis-4-nitrophenyl phosphate; dT-5′-$pNPP^-$, thymidine 5′-monophosphate 4-nitrophenyl ester; MBP, maltose binding protein; PEG, polyethylene glycol; MPD, 2-methyl-2,4-pentanediol; ALS, Advanced Light Source; wt, wild type; PDB, Protein Data Bank.

10.1021/bi060847t CCC: $33.50 © 2006 American Chemical Society
Published on Web 07/25/2006

Scheme 1: Interactions of the Conserved Bimetallo Active Sites of (A) *E. coli* AP from Crystal Structures (*12*, *13*) and (B) Mouse NPP1 from a Homology Model (*3*) with a Transition State Model[a]



[a] Reactions are depicted in terms of a transition state representation with partial bond formation and bond cleavage. No information about bond orders and bond lengths is conveyed in this schematic.

residue could be covalently coupled to nucleoside monophosphates and inorganic phosphate (*3*, *7*−*11*). The two active site metal ions are proposed to activate the incoming nucleophile and stabilize charge buildup on the leaving group (*3*), again by analogy to the mechanism for AP (Scheme 1) (*12*, *13*).

In addition to catalyzing hydrolysis reactions of phosphate monoesters, AP also has a low level of activity for phosphate diester hydrolysis, and it was suggested that such promiscuous activity could have been important for the evolution of new enzymes following duplication of an ancestral phosphohydrolase gene (*14*). The existence of the NPP enzymes, which are evolutionarily related to AP and appear to preferentially hydrolyze phosphate diesters, provides an opportunity to investigate the evolutionary changes since the divergence from a common ancestor that left the NPP and AP active sites structurally similar yet highly specific for different reactions.

Active sites containing two metal ions have been suggested to be well-suited to catalyzing phosphoryl transfer reactions in general (*15*, *16*), and there are many examples of enzymes optimized for either phosphate monoester or phosphate diester hydrolysis (*17*, *18*). However, the features of bimetallo active sites that dictate specificity for either monoesters or diesters are not well understood. The most prominent distinguishing feature between phosphate monoesters and diesters is the presence of an additional functional group on the transferred phosphoryl group in diester reactions (R′ in Scheme 1). Binding interactions with this functional group could assist in specific binding and positioning of diester substrates and lead to preferential diester hydrolysis.

Another distinguishing feature between phosphate monoester and diester reactions is the nature of the transition state. In solution, phosphate monoesters react through loose transition states with nearly complete cleavage of the bond to the leaving group and little formation of the bond to the nucleophile (*19*−*23*), whereas phosphate diesters proceed through tighter transition states with less cleavage of the bond to the leaving group and more formation of the bond to the nucleophile (*23*−*27*). A growing body of experimental results suggests that enzymes catalyze reactions of phosphate monoesters and diesters through transition states that are similar to their respective solution transition states (see references in ref *28*). It is possible that an enzyme active site could be optimized for either the loose transition state for monoester hydrolysis or the tighter, synchronous transition state for diester hydrolysis. Potential mechanisms for accomplishing such optimization include adjusting the positions of the two metal ions, changing the nucleophile strength, or altering the local electrostatic properties of the active site to recognize transition states with different charge distributions.

A final distinguishing feature between phosphate monoesters and diesters is the total amount of negative charge on the phosphoryl group; dianionic phosphate monoesters bear more negative charge than monoanionic phosphate diesters. The possibility that the difference in total charge could be important for discrimination between substrates was suggested by the observation that the rate enhancements for a wide range of AP-catalyzed reactions correlate with the amount of negative charge on the nonbridging oxygen atom situated between the two Zn^{2+} ions (*29*). This preference for highly negatively charged substrates was attributed to the proximity of the two divalent cations. However, the existence of many binuclear metalloenzymes that preferentially react with phosphate diesters, including NPP, suggests that the presence of two closely spaced divalent cations alone does not necessitate preferential phosphate monoester hydrolysis. Instead, electrostatic properties of the bimetallo site could be tuned to be complementary to a particular substrate charge through alteration of the spacing between the metal ions, through variation of the properties of coordinating metal ion ligands, or by the presence or absence of other nearby charged groups.

To undertake a comparative analysis of the NPP and AP active sites aimed at determining how these similar active sites have such different substrate preferences, we have cloned, expressed, and purified a bacterial member of the NPP family. The rate constants measured for NPP-catalyzed phosphate diester and monoester hydrolysis suggest that preferential diester hydrolysis arises from both specific binding interactions for diester substrates and optimization of features of the bimetallo site. To identify active site features that contribute to preferential diester hydrolysis, we

have determined the structure of NPP by X-ray crystallography in the absence of ligands and in complexes with vanadate and AMP. Comparisons with the AP active site and kinetic results suggest structural features that may be important for discrimination between diester and monoester substrates.

MATERIALS AND METHODS

Materials. Genomic DNA from *Xanthomonas axonopodis pv. citri* and *Xanthomonas campestris pv. campestris* was a gift from J. Roden and M. B. Mudgett. Genomic DNA from *Porphyromonas gingivalis* and *Bacteroides fragilis* was a gift from D. Relman. Genomic DNA from *Caulobacter crescentus* was a gift from L. Shapiro. Methyl 4-nitrophenyl phosphate ($MpNPP^-$) was synthesized as described previously (*28*). All other reagents were obtained from commercial sources.

Identification of Candidate Genes. To identify candidate NPP genes amenable to expression in an *Escherichia coli* host, we performed a BLASTP search (*30*) of the bacterial database with the catalytic domain of mouse NPP1 (*3*). We identified six NPP homologues that contained the threonine nucleophile and the six Zn^{2+} ligands that are conserved in NPP and AP. The source organisms and the corresponding GI numbers for the NPP proteins are *Zymomonas mobilus* (3089614), *C. crescentus* (13422610), *X. axonopodis pv. citri* (21109117), *X. campestris pv. campestris* (21113820), *P. gingivalis* (34396301), and *B. fragilis* (52216418). Several more bacterial sequences satisfying these criteria have been deposited since the initial search was performed in 2004. Genomic DNA samples from five of the six organisms listed above were readily obtained (see Materials) and used to clone NPP genes into the pET21c vector (Novagen). After growth and induction of small-scale *E. coli* cultures harboring the NPP constructs, protein overexpression was assayed by SDS−PAGE and phosphodiesterase activity in the soluble fraction was assayed using the $MpNPP^-$ substrate. The NPP from *X. axonopodis pv. citri* displayed both protein overexpression and phosphodiesterase activity and was used for subsequent experiments. Details of cloning and protein expression for this construct are given below.

Cloning. NPP from *X. axonopodis pv. citri* (*Xac*) was PCR amplified from genomic DNA and cloned into pET21c using standard molecular biology procedures. DNA sequencing was performed on an ABI3000 capillary sequencer and confirmed that the sequence of the cloned construct was identical to the reported genomic sequence (GI number 21109106). To generate a C-terminal His-tagged construct (NPP−His), the stop codon was mutated to allow read-through to the His tag in the pET21c vector using the QuikChange site-directed mutagenesis kit (Stratagene). An N-terminal maltose binding protein (MBP) fusion construct (NPP−MBP) was obtained by cloning the DNA sequence corresponding to residues 41−432 (from the native *Xac* NPP signal peptide cleavage site to the stop codon; see below) into the pMAL-p2X vector (New England Biolabs). This vector includes coding regions for an N-terminal signal peptide for periplasmic export and a Factor Xa cleavage site that cleaves between NPP and MBP to release *Xac* NPP residues 40−432 (vector-introduced sequence fortuitously encoded *Xac* NPP residue 40). For protein expression, all constructs were cloned into *E. coli* SM547(DE3) (*29*), a strain that lacks the native *phoA* gene (*31*), to avoid potential contaminating phosphate monoesterase activity from AP.

Expression and Purification of C-Terminally His-Tagged Xac NPP. *E. coli* SM547(DE3) cells containing NPP−His were grown to an OD_{600} of 0.4 in LB medium with 50 μg/mL carbenicillin at 37 °C. Cultures were cooled to 18 °C, and IPTG was added to a final concentration of 0.3 mM to induce NPP expression. After 6 h, a second aliquot of carbenicillin was added. After 21 h, cells were harvested by centrifugation, resuspended in 10 mM sodium phosphate (pH 7.4), 0.5 M NaCl, and 5 mM imidazole, and lysed using a French pressure cell. The clarified cell lysate was passed over a HiTrap Chelating HP column charged with Ni^{2+} (Amersham) and eluted in a linear gradient of imidazole. Peak fractions were pooled, concentrated by centrifugation through a filter (10 kDa cutoff, Amicon), loaded onto a gel filtration column (Superose 12, Pharmacia), and run with 10 mM Tris-HCl (pH 8.0) and 50 mM NaCl. The elution volume of the peak fraction was consistent with a molecular mass of ~50 kDa, suggesting that NPP is a monomer under these conditions. Finally, peak fractions were pooled, diluted 1:5, and loaded onto an 8 mL Source Q column (Pharmacia). The column was washed with 3 column volumes of 10 mM Tris-HCl (pH 8.5) and 10 mM NaCl, and the protein was eluted with a linear gradient from 10 to 250 mM NaCl in 10 mM Tris-HCl (pH 8.5). Peak fractions were pooled and buffer exchanged into 10 mM Tris-HCl (pH 8.0) and 50 mM NaCl by centrifugation through a filter (10 kDa cutoff, Amicon). The final yield of the purified protein was low (0.1 mg from a 6 L culture) but sufficient for determination of relative rates of phosphate diester and monoester hydrolysis reactions, as well as for characterization of post-translational processing. The full-length *Xac* NPP gene is predicted to have an N-terminal signal peptide with a cleavage site after residue 25 or 27 (*32*). To determine if a signal peptide was cleaved from the purified protein, a sample was analyzed by mass spectrometry. The two major peaks that were observed corresponded to cleavage after residues 40 and 41.

Expression and Purification of Xac NPP from the MBP Fusion Construct. *E. coli* SM547(DE3) cells containing NPP−MBP were grown to an OD_{600} of 0.5 in rich medium and glucose (10 g of tryptone, 5 g of yeast extract, 5 g of NaCl, and 2 g of glucose per liter) with 50 μg/mL carbenicillin at 37 °C. Cultures were cooled to 30 °C, and IPTG was added to a final concentration of 0.3 mM to induce NPP expression. After 7−10 h, a second aliquot of carbenicillin was added. After 20 h, cells were harvested by centrifugation. Following osmotic shock and centrifugation, the supernatant was adjusted to 10 mM Tris-HCl (pH 8.0) and 10 μM $ZnCl_2$ and filtered through a 0.45 μm membrane (Nalgene). To concentrate the large volume obtained from the osmotic shock, the sample was passed over a 15 mL Q-Sepharose FF column (Amersham) and eluted with 0.5 M NaCl. Protein-containing fractions were pooled, concentrated to less than 5 mL by centrifugation through a filter (10 kDa cutoff, Amicon), diluted 1:10, adjusted to pH 6.0 by dropwise addition of 1 M MES (pH 2.8), and loaded onto an 8 mL Source S column (Pharmacia). The column was washed with 2 column volumes of buffer A [20 mM MES (pH 6.0) and 100 μM $ZnCl_2$], and the protein was eluted

with a linear gradient of 0 to 0.5 M NaCl in buffer A. Peak fractions were pooled and adjusted to pH 7.4 by dropwise addition of 1 M Tris-HCl (pH 7.4). The sample was then loaded onto a 50 mL amylose column (New England Biolabs), washed with 3 column volumes of amylose column buffer [20 mM Tris-HCl (pH 7.4) and 200 mM NaCl], and eluted with 10 mM maltose in amylose column buffer. Peak fractions were pooled, and the solution was adjusted to 50 mM Tris-HCl (pH 8.0), 200 mM NaCl, 10 mM maltose, and 5 mM $CaCl_2$ by addition of appropriate volumes of 1 M Tris-HCl (pH 8.0) and 1 M $CaCl_2$. To cleave the fusion protein, Factor Xa was added to the solution (~1 unit of Factor Xa per milligram of fusion protein, Novagen). After incubation for 16 h at room temperature, the cleavage reaction mixture was diluted 1:2 in water, adjusted to pH 6.0 by dropwise addition of 1 M MES (pH 2.8), and loaded onto an 8 mL Source S column (Pharmacia). The column was washed with 2 column volumes of 0.1 M NaCl in buffer A, and the protein was eluted with a linear gradient of 0.1 to 0.4 M NaCl in buffer A. Peak fractions were pooled and stored at 4 °C. Typical yields for a 6 L culture were 20−30 mg of pure protein. The purity was estimated to be >95% as judged by band intensities on Coomassie blue-stained SDS−polyacrylamide gels. For kinetic assays, the protein was diluted into 10 mM Tris-HCl (pH 8.0) and stored at 4 °C. For crystallization, the protein was concentrated to ~10 mg/mL in 10 mM MES (pH 6.0), 50 mM NaCl, and 100 μM $ZnCl_2$ by centrifugation through a filter (10 kDa cutoff, Amicon). The protein was stable at 4 °C for at least 2 months. The protein concentration was determined by absorbance at 280 nm in 6 M guanidine hydrochloride and 20 mM sodium phosphate (pH 6.5) using a calculated extinction coefficient of 79 920 M^{-1} cm^{-1} (*33*). Gel filtration chromatography (Shodex KW-803) followed by static light scattering (DAWN EOS, Wyatt Technologies) indicated that purified NPP was a monomer in solution at pH 6.0 and 8.0.

NPP Kinetics. All reported rate constants were obtained using NPP purified from the MBP fusion construct. Rate constants obtained from the NPP−His construct were identical within error. Reactions were performed in 0.1 M Tris-HCl (pH 8.0), 0.5 M NaCl, and 100 μM $ZnCl_2$ at 25 °C in quartz cuvettes. The absorbance of the 4-nitrophenolate product was monitored continuously at 400 nm using a Uvikon 9310 spectrophotometer, and rate constants were determined from initial rates. Values of k_{cat}/K_M were obtained under conditions where the reaction was shown to be first-order in both enzyme and substrate over at least 10-fold ranges in concentration.

Inhibition constants for vanadate and 5′-AMP were determined with subsaturating substrate concentrations under standard assay conditions at pH 8.0. Stock solutions of 100 mM sodium orthovanadate and 100 mM AMP were prepared for use in both inhibition assays and crystal soaks (see below).

The pH dependences of k_{cat}/K_M for $MpNPP^-$ and pNPP were obtained with 0.1 M buffer, 0.5 M NaCl, and 100 μM $ZnCl_2$ at 25 °C. The following buffers were used: NaAcetate (pH 4.5), NaMES (pH 5.5−6.5), NaMOPS (pH 7.5), Tris-HCl (pH 8.0−8.5), NaCHES (pH 9.5), and NaCAPS (pH 10.5−11.4). Control reactions with 0.02−0.1 M buffer showed that the different buffers did not exert any significant effects on the NPP reaction. At pH >6, the absorbance of the product was monitored continuously at 400 nm in quartz cuvettes using a Uvikon 9310 spectrophotometer. At pH <6, we monitored product formation discontinuously by removing aliquots from the reaction, quenching in an equal volume of 0.2 M NaOH, and immediately measuring the absorbance at 400 nm. In control reactions at pH 6.5 and 8.0, observed rate constants obtained from continuous and discontinuous assays agreed within 10%. For discontinuous assays, reactions were conducted in borosilicate glass tubes siliconized with *N*-(triethoxysilylpropyl)-*O*-polyethylene oxide urethane (United Chemical Technologies). Aside from the quartz cuvettes used for continuous assays, most other plastic and glass reaction vessels, both untreated and treated with various siliconizing agents, led to product formation time courses with pronounced downward curvature indicative of enzyme sticking to tube walls at enzyme concentrations below 1 μM. Observed second-order rate constants were obtained from initial rates, and both substrate and enzyme concentrations were varied over at least 10-fold ranges at each pH.

At low pH, both $MpNPP^-$ and pNPP exhibited saturation behavior. At pH 4.5, the apparent K_M for pNPP was too low to allow direct measurement of k_{cat}/K_M. Instead, k_{cat}/K_M was determined indirectly by measuring the apparent second-order rate constant in the presence of inhibitory concentrations of AMP for both the $MpNPP^-$ and pNPP reactions. The value of k_{cat}/K_M for the pNPP reaction could then be calculated from the observed value of k_{cat}/K_M with inhibition and the inhibition constant for AMP measured in the $MpNPP^-$ reaction at this pH.

Because the solubility of $ZnCl_2$ decreases as the pH increases, the observed rate decrease at pH >10 could be due to loss of metal ion from the active site (Zn^{2+} is the active site metal ion in purified NPP; see Results). Increasing the concentration of $ZnCl_2$ in the reaction at pH 11.4 from 100 to 500 μM did not increase the observed rate, suggesting that the metal ion sites in NPP were saturated and the observed rate decrease at high pH (see Results) was not due to the loss of metal ion.

Determination of the pK_a of pNPP. The second ionization constant of pNPP was determined by direct titration. Solutions containing 5 or 20 mM pNPP and 0.5 M NaCl were stirred at room temperature. A small quantity of NaOH was added to ensure that all pNPP was in the dianionic form. The solutions were titrated with small aliquots of 0.1 M HCl, and the pH was recorded with a pH meter. Titration of the same solutions in the reverse direction with 0.1 M NaOH gave the same pK_a values within 0.04 pH unit. The resulting pK_a value was 4.79 ± 0.03, in good agreement with a previous measurement of 4.95 in 1 M KCl (*34*).

Preparation of Metal-Free NPP. Metal-free NPP was prepared by treatment in batch with Chelex 100 resin (Bio-Rad). Approximately 0.1 g of Chelex 100 resin was added to a 0.5 mL solution of ~10 μM NPP and 0.1 M Tris-HCl (pH 8.0), and the mixture was gently agitated at room temperature. After 1−2 h, the resin was pelleted by centrifugation and the supernatant NPP solution was transferred to a tube containing fresh Chelex 100 resin. After being gently agitated for an additional 1−2 h, the resin was pelleted by centrifugation and the supernatant NPP solution was transferred to a fresh tube. The incubations were performed in glass tubes siliconized with *N*-(triethoxysilylpropyl)-*O*-polyethylene oxide urethane to prevent enzyme

Table 1: Kinetic Parameters for *Xac* NPP and *E. coli* AP[a]

	NPP		wt AP		R166S AP	
substrate	k_{cat}/K_M (M^{-1} s^{-1})	k_{rel}[b]	k_{cat}/K_M (M^{-1} s^{-1})	k_{rel}[b]	k_{cat}/K_M (M^{-1} s^{-1})	k_{rel}[b]
$pNPP^{2-}$	1.1	(1)	3.3×10^7	(1)	1.0×10^5	(1)
$pNPP^-$	79	72	−	−	−	−
$MpNPP^-$	2.3×10^2	2.1×10^2	−	−	0.48	4.8×10^{-6}
bis-$pNPP^-$	2.3×10^3	2.1×10^3	5.0×10^{-2}	1.5×10^{-9}	5.0×10^{-2}	5.0×10^{-7}
dT-5′-$pNPP^-$	1.6×10^6	1.5×10^6	−	−	9.7×10^{-3}	9.7×10^{-8}

[a] Values of k_{cat}/K_M for reactions catalyzed by NPP were obtained in 0.1 M Tris-HCl (pH 8.0), 0.5 M NaCl, and 100 μM $ZnCl_2$ at 25 °C except the values for $pNPP^{2-}$ and $pNPP^-$, which were obtained from the fit to the pH−rate profile (Scheme 3B and Figure 2B). The uncertainties for the values of k_{cat}/K_M, defined as the standard deviations for repeated measurements, are within ±20% except for that of the slow reaction of R166S AP with dT-5′-$pNPP^-$, which has a standard deviation of ±50%. The observed value of k_{cat}/K_M for pNPP at pH 8.0 of 0.9 M^{-1} s^{-1} is identical within error to the value of 1.1 M^{-1} s^{-1} obtained for $pNPP^{2-}$ from the fit. Values of k_{cat}/K_M for reactions catalyzed by wt and R166S AP are from previous work (*14*, *28*, *59*) except for that for the reaction of R166S AP with dT-5′-$pNPP^-$ which was obtained for this work. AP reactions were performed in 0.1 M NaMOPS (pH 8.0), 0.5 M NaCl, 1 mM $MgCl_2$, and 100 μM $ZnSO_4$ at 25 °C. [b] $k_{rel} = (k_{cat}/K_M)_{substrate}/(k_{cat}/K_M)_{pNPP^{2-}}$.

loss to tube walls (see NPP Kinetics). All kinetic assays with metal-free NPP were performed with buffers treated with Chelex 100.

R166S AP Purification and Kinetics. R166S AP was purified as previously described (*28*), and kinetic constants were also determined as previously described (*28*).

Crystallization. NPP [at 10 mg/mL, in 10 mM NaMES (pH 6.0), 50 mM NaCl, and 100 μM $ZnCl_2$] was crystallized at 20 °C by the hanging drop method against 0.1 M Bis-Tris-HCl (pH 6.0 or 6.5) and 17−21% PEG 3350. Crystals containing vanadate were obtained by soaking protein crystals in 0.1 M NaHEPES (pH 7.5), 19% PEG 3350, and 500 μM sodium orthovanadate for 10 min. Cocrystals with AMP were obtained by including 250 μM AMP in the crystallization solution. Crystals were quickly passed through a 30% MPD solution in mother liquor (or soaking condition) before direct immersion in liquid nitrogen.

Data Collection. NPP crystallized in space group $P2_12_12_1$ with two monomers in the asymmetric unit. Diffraction data were collected at ALS (Lawrence Berkeley National Laboratory) on beamlines 8.2.1 and 8.2.2. To obtain experimental phases, a three-wavelength MAD data set was collected at the Zn edge. The following wavelengths were used: 1.2826 (peak), 1.2831 (edge), and 1.2782 Å (remote). Data were collected in inverse beam mode in 10° wedges to collect Friedel mates as close in time as possible. All data were integrated and scaled using DENZO and SCALEPACK, respectively (*35*). Data statistics are summarized in Table 2.

Structure Determination and Refinement. Data integration and scaling indicated an orthorhombic Bravais lattice, but the correct space group assignment could not be made due to ambiguity in the systematic absences. To avoid the several phasing possibilities in the orthorhombic setting, Zn^{2+} positions were determined initially in space group $P2_1$ using SOLVE (*36*) and density modification was performed with RESOLVE (*37*). The resulting figures of merit for phasing from SOLVE and RESOLVE were 0.39 and 0.62, respectively. An experimentally phased map of the active site is shown in Figure S2 of the Supporting Information. Approximately 95% of the protein main chain was built using PHENIX (*38*), which provided a sufficient molecular replacement model for determination of the proper orthorhombic space group setting ($P2_12_12_1$) using Phaser (*39*). At this point, σ_A-weighted $2F_o - F_c$ and $F_o - F_c$ maps were inspected, and a complete model comprising residues 44−425 (the full-length construct comprises residues 40−432),

Table 2: Crystallographic Data and Model Statistics

	holoenzyme (2GSN)	vanadate-bound (2GSO)	AMP-bound (2GSU)
Data Collection			
beamline	ALS 8.2.1	ALS 8.2.2	ALS 8.2.2
wavelength (Å)	1.2826	0.9537	0.9537
space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
unit cell			
a (Å)	65.7	66.0	65.6
b (Å)	78.7	78.8	78.5
c (Å)	129.5	129.7	129.5
resolution range (Å)	50−1.75	50−1.30	50−2.00
effective resolution ($I/\sigma = 2$) (Å)	1.85	1.45	2.10
no. of unique reflections	67365	147720	41903
completeness	98.1	89.4	91.3
I/σ (highest-resolution shell)	16.8 (0.8)	20.1 (0.6)	12.7 (1.1)
R_{merge}[a] (%)	9.7	8.5	14.2
Refinement Statistics			
R_{factor} (R_{free})[b] (%)	17.1 (20.3)	17.2 (19.3)	23.2 (30.3)
no. of protein atoms	6006	6006	6006
no. of solvent atoms	739	847	504
no. of ligand atoms	4	14	50
average *B* factor	22.4	14.6	28.7
rmsd for bond lengths (Å)	0.01	0.01	0.01
rmsd for bond angles (deg)	1.1	1.2	1.4

[a] $R_{merge} = \sum|I_{obs} - I_{ave}|/\sum I_{obs}$. [b] $R_{factor} = \sum||F_{obs}| - |F_{calc}||/\sum|F_{obs}|$. See ref *81* for a description of R_{free}.

including two Zn^{2+} ions per protein monomer, was built using Coot (*40*). Five percent of the observed data was set aside for cross validation, and initial simulated annealing refinement was carried out using a maximum likelihood amplitude-based target function as implemented in CNS (*41*), resulting in an *R* factor of 28.7%. Further refinement was carried out using Refmac (*42*), and solvent was built using ARP/wARP (*43*). Each stage of refinement was interspersed with manual corrections and model adjustments using Coot. The final round of refinement in Refmac treated each monomer as an independent TLS group (*44*). The values of *R* and R_{free} for the final refined model were 17.1 and 20.3%, respectively. Estimated standard deviations for atomic positions were calculated by matrix inversion using SHELXL-97 (*45*).

For the vanadate-soaked crystal, phases from the holoenzyme model were used in rigid body refinement as implemented in Refmac, resulting in an initial *R* factor of 30.4%. Two Zn^{2+} ions per protein monomer were modeled in the

active site, followed by positional and *B* factor refinement in Refmac, resulting in a decrease in the *R* factor of 7%. The σ_A-weighted $F_o - F_c$ map clearly indicated the presence of vanadate in a trigonal bipyramidal geometry in the active site (Figure 6). An idealized trigonal bipyramidal vanadate molecule was modeled into the active site using O (*46*), with one axial oxygen replaced with Thr90 Oγ. Both bond length (2.0 ± 0.1 Å) and bond angle (±3° from the ideal trigonal bipyramidal geometry) restraints were applied to the vanadate molecule, based on Cambridge Structural Database entries YIQLEP and YIQLIT. Further rounds of refinement and solvent building were carried out as described above for the holoenzyme model. The final round of refinement modeled both the Zn^{2+} ions and the vanadate molecule using anisotropic displacement parameters, resulting in final values of *R* and R_{free} of 17.2 and 19.3%, respectively. Estimated standard deviations for atomic positions were calculated by matrix inversion using SHELXL-97 (*45*).

For the AMP cocrystal, phases from the holoenzyme model were used in rigid body refinement as implemented in Refmac, resulting in an initial *R* factor of 32.4%. Two Zn^{2+} ions per protein monomer were modeled in the active site, followed by positional and *B* factor refinement in Refmac, resulting in a decrease in the *R* factor of 5%. The σ_A-weighted $F_o - F_c$ map clearly indicated the presence of AMP in the active site (Figure 7). One AMP molecule per active site was modeled using O (*46*), and further rounds of refinement and solvent building were carried out as described above for the holoenzyme model. Because the *B* factors for the AMP molecules were refined to unreasonably large values relative to those of nearby protein atoms, the AMP occupancy was manually adjusted until the refined *B* factors were comparable to those of surrounding residues in the active site. The final AMP occupancies were 0.6 in the A site and 0.8 in the B site. The values of *R* and R_{free} for the final refined model were 23.2 and 30.3%, respectively.

The relatively high values of *R* and R_{free} for the AMP data set are most likely due to increased error in the measured structure factors (T. D. Fenn, E. Pozharski, and M. A. Wilson, manuscript in preparation). This error resulted in a value of R_{merge} higher than that observed for the holoenzyme and vanadate data sets (Table 2). Reprocessing the data and/or several rounds of simulated annealing refinement did not improve the *R* and R_{free} values, and aggressive rejection during scaling did not improve R_{merge}.

Analysis of Sequence Conservation in the NPP Family. Sequences from the first 100 hits in a WU-BLAST search (*30*) of the Uniprot database with the sequence of *Xac* NPP were obtained. Sequences that were replicated one or more times within the first 100 hits were identified, and the replicate sequences were removed from the set. Sequences were considered replicates if they were ≥97% identical in a pairwise Clustalw alignment (*47*), which allows for allelic differences and sequencing errors. It was not possible to remove replicates simply by allowing only one sequence for each species because at least seven NPP paralogs have been identified in higher eukaryotes (*5*). Sixty-eight unique sequences were aligned with Probcons (*48*), and the resulting multiple-sequence alignment was processed to identify sequences with conserved residues at all six Zn^{2+} ligands and the active site Thr. From these 57 sequences, additional sites of interest were located in the alignment and analyzed to determine if the residue present in *Xac* NPP was conserved. Perl scripts were used to process output from both the BLAST search and the Probcons alignment.

RESULTS

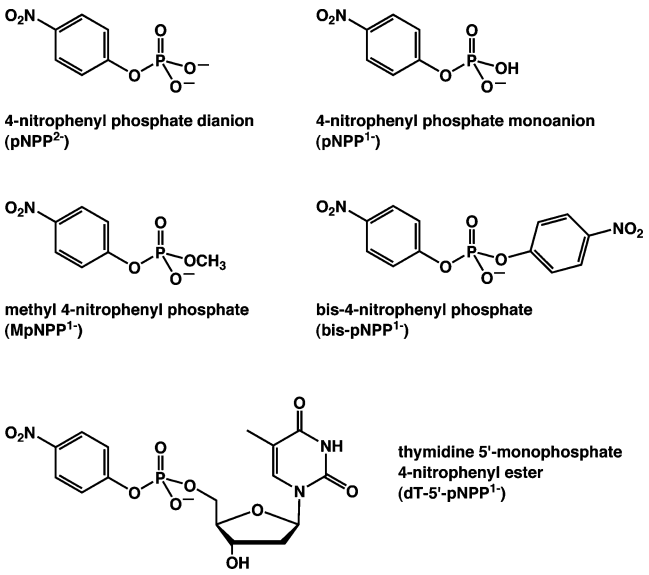
For detailed structural and functional comparisons of NPP and AP, a well-behaved, soluble protein from the NPP family was required. Previous studies of NPP, however, utilized eukaryotic family members, which are often extracellular membrane-associated proteins with multiple domains (*5*). Truncated versions of these proteins that contain the phosphodiesterase domain lack activity, possibly due to the loss of protein stability (*49*, *50*). We therefore used the sequence for the phosphodiesterase domain of mouse NPP1 in a BLAST search (*30*) of the bacterial database to identify NPP homologues that might be more amenable to structural and functional studies. We identified six candidate bacterial genes, obtained genomic DNA samples for five of the host species, and cloned NPP genes into the pET21c vector for expression tests in *E. coli* (see Materials and Methods). NPP from *X. axonopodis pv. citri* (*Xac*) had soluble phosphodiesterase activity and was used for subsequent experiments.

Expression and Characterization of Xac NPP. For the initial characterization of *Xac* NPP, a His-tagged construct was prepared in the pET21c vector, overexpressed in *E. coli*, and purified to homogeneity. The elution volume from a gel filtration column was consistent with a monomeric protein of ~50 kDa (see Materials and Methods). Because the eukaryotic NPP homologues are extracellular proteins and *E. coli* AP is exported to the periplasm (*51*, *52*), we expected that *Xac* NPP might also be exported to the periplasm. Sequence analysis using the SignalP server (*32*) suggested the presence of an N-terminal signal peptide, and this prediction was confirmed by mass spectrometry on the purified protein, which indicated cleavage of the signal peptide after residue 40 or 41 (data not shown).

Rate constants were measured for the reactions of NPP with phosphate diester and monoester substrates at pH 8.0 (see below). Because the yield from purification of the His-tagged expression construct was relatively low, all subsequent functional and structural experiments were performed with an NPP construct expressed as an MBP fusion and cleaved to release a truncated protein corresponding to *Xac* NPP residues 40−432 (see Materials and Methods). The kinetic constants reported below were obtained with NPP from the MBP fusion construct and are identical within error to those obtained with the His-tagged construct.

NPP Reactions with Phosphate Diesters. Values of k_{cat}/K_M for *Xac* NPP-catalyzed hydrolysis reactions of $MpNPP^-$, bis-$pNPP^-$, and dT-5′-$pNPP^-$ (Scheme 2) at pH 8.0 are reported in Table 1. There was no evidence for saturation up to 2 mM $MpNPP^-$, indicating that the K_M for this substrate is greater than 2 mM. For bis-$pNPP^-$, deviations from a linear dependence on substrate concentration could be detected at 1 mM, suggesting that K_M is greater than or equal to ~1 mM. The reaction of dT-5′-$pNPP^-$ exhibited saturation behavior and could be fit to the Michaelis−Menten equation with an apparent K_M of 110 μM. Notably, k_{cat}/K_M increases by a factor of $\sim 10^4$ as the size of R′ increases from a methyl group to a thymidine group, suggesting the presence of an R′ binding site in NPP that is important for recognition of phosphate diester substrates.

Scheme 2



Rate constants for hydrolysis reactions catalyzed by eukaryotic NPP family members have been previously measured with a variety of phosphate diester substrates (*5*, *53*, *54*). The values of k_{cat}/K_M obtained herein for reactions of *Xac* NPP with bis-pNPP$^-$ and dT-5′-pNPP$^-$ are within 10- and 3-fold, respectively, of the values reported for reactions of these substrates with NPP from bovine intestine (*53*).

NPP Reactions with Phosphate Monoesters. On the basis of the similarity of the NPP and AP active sites, it seemed reasonable that NPP could have phosphate monoesterase activity. Furthermore, although AP itself preferentially hydrolyzes phosphate monoesters, it has a low level of activity for phosphate diester hydrolysis (*14*). Previously, the activity of NPP with phosphate monoester substrates had not been quantitatively determined. It has been reported that no activity could be detected for the reaction of bovine intestinal NPP with 4-nitrophenyl phosphate (pNPP), but a detection limit was not reported (*53*). Several NPP enzymes can be radiolabeled with [γ-^{32}P]ATP (*3*, *9*−*11*), and amino acid analysis of the radiolabeled protein indicates the presence of phosphothreonine (*9*), suggesting that NPP has some phosphate monoesterase activity.

We observed a low level of activity for pNPP hydrolysis in *Xac* NPP. The value of k_{cat}/K_M at pH 8.0 of 0.9 M^{-1} s^{-1} is approximately 200-fold lower than the value of 2.3×10^2 M^{-1} s^{-1} obtained for the reaction of the diester MpNPP$^-$. As with MpNPP$^-$, no evidence for saturation was obtained at ≤2 mM pNPP. To confirm that the phosphate monoesterase activity arose from NPP and not from a small amount of a contaminating enzyme that is a proficient monoesterase, we measured inhibition constants for reactions of NPP with MpNPP$^-$ and pNPP. If the phosphate diester and monoester hydrolysis reactions occur in the same active site, the K_I for a given inhibitor should be identical for the two reactions.

Vanadate is a potent inhibitor of enzymes that catalyze phosphoryl transfer reactions and adopts a trigonal bipyramidal geometry analogous to the transition state for phosphate ester hydrolysis (*55*−*57*). Vanadate inhibited the reactions of NPP with MpNPP$^-$ and pNPP with K_I values of 52 and 47 μM, respectively (Figure 1A). 5′-AMP also inhibits the reactions of NPP with MpNPP$^-$ and pNPP, which is consistent with the presence of a nucleoside-binding pocket in NPP as suggested by the reactivity of dT-5′-pNPP$^-$. The values of K_I were 268 and 258 μM, respectively (Figure 1B). The observation that the inhibition constants obtained for reactions of phosphate diesters and monoesters are identical within error strongly suggests that these reactions occur in the same active site in NPP. Identical inhibition constants for the two reactions were also obtained with tungstate (K_I = 160 μM) and 5′-GMP (K_I = 280 μM) as inhibitors (Figure S1 of the Supporting Information).

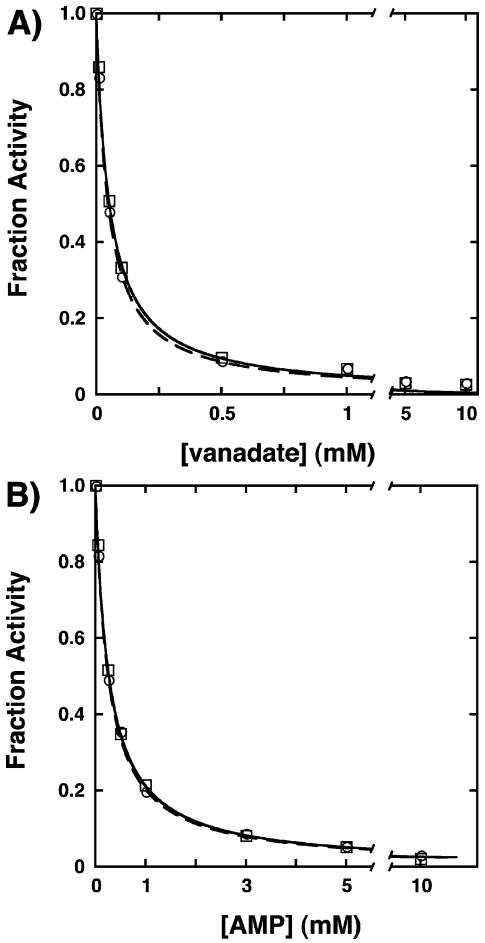


FIGURE 1: Identical inhibition of diesterase and monoesterase activities. Inhibition of MpNPP$^-$ hydrolysis (□) and pNPP^{2-} hydrolysis (○) by vanadate (A) and AMP (B). Activity was normalized by the observed rate constant in the absence of inhibitor. The lines are nonlinear least-squares fits for competitive inhibition and give K_I values of 52 ± 3 and 47 ± 3 μM for vanadate inhibition of diesterase and monoesterase activities, respectively, and K_I values of 268 ± 2 and 258 ± 10 μM for AMP inhibition of diesterase and monoesterase activities, respectively.

pH−*Rate Profiles.* The pH dependence of a reaction provides valuable information about the ionization states of the active substrate and enzyme species. Under subsaturating conditions, the pH dependence reflects ionizations of the free enzyme and substrate. For hydrolysis reactions of simple phosphate diesters, which have pK_a values of <2 (*58*), variations in the observed rate constant over the pH range from 4 to 12 are likely to reflect only ionizations of the free enzyme.

The pH dependence for NPP-catalyzed hydrolysis of MpNPP$^-$ follows a bell-shaped profile with a broad maxi-

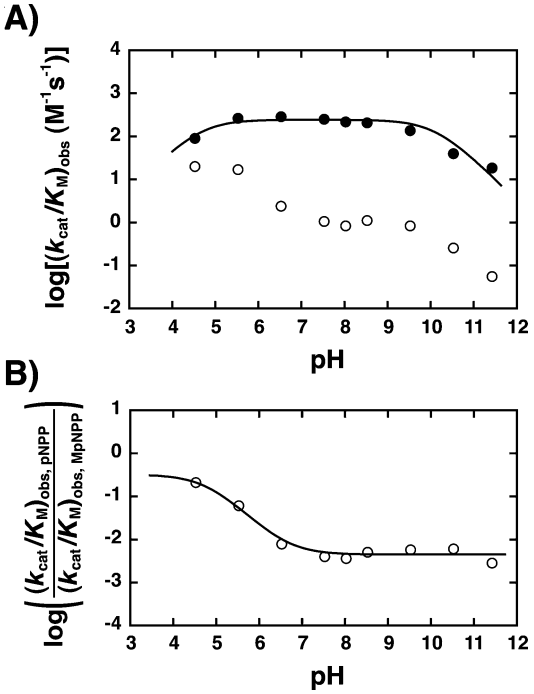


FIGURE 2: pH−rate profiles for diesterase and monoesterase reactions demonstrate that NPP reacts with both $pNPP^-$ and $pNPP^{2-}$. (A) pH−rate profiles for observed second-order rate constants (k_{cat}/K_M) for $MpNPP^-$ (●) and pNPP hydrolysis (○). The line through the $MpNPP^-$ data is a nonlinear least-squares fit to a model with two rate-controlling enzymatic ionizations with pK_a values of 4.6 and 10.1 (Scheme 3A). (B) NPP reacts with both $pNPP^-$ and $pNPP^{2-}$. The observed rate constants for pNPP hydrolysis were normalized by those for $MpNPP^-$ hydrolysis to eliminate effects arising from enzymatic ionizations (Scheme 3B and eq 5). The line is a nonlinear least-squares fit to a model in which NPP reacts with both $pNPP^-$ and $pNPP^{2-}$ with a fixed, experimentally determined pK_a of 4.79 and gives fitted values of $(k_{cat}/K_M)_{pNPP^-}$ and $(k_{cat}/K_M)_{pNPP^{2-}}$ of 79 and 1.1 M^{-1} s^{-1}, respectively. If the pK_a is not fixed, the fit gives indistinguishable results, with a fitted pK_a of 4.67 and values of $(k_{cat}/K_M)_{pNPP^-}$ and $(k_{cat}/K_M)_{pNPP^{2-}}$ of 93 and 1.1 M^{-1} s^{-1}, respectively.

mum over the pH range from 5 to 10 (Figure 2A, filled symbols). Within the pH range from 4.5 to 11.4, the observed loss of activity at extreme pH values was fully reversible upon dilution into buffer at pH 8.0. The pH−rate profile can be described by a model with two ionizations of the enzyme as depicted in Scheme 3A and described by eqs 1 and 2. A nonlinear least-squares fit to this model gives values of pK_{E1} and pK_{E2} of 4.6 and 10.1, respectively.

$$v_{obs} = (k_{cat}/K_M)_{obs}[E]_o[MpNPP]_o = (k_{cat}/K_M)_{MpNPP}[E_H][MpNPP] \quad (1)$$

$$(k_{cat}/K_M)_{obs} = \frac{(k_{cat}/K_M)_{MpNPP}}{1 + 10^{pK_{E1}-pH} + 10^{pH-pK_{E2}}} \quad (2)$$

Analysis of pH−rate profiles for reactions of AP suggests that the active site Ser is present as an alkoxide and the pK_a is less than 5.5 (*59*). On the basis of the similarity of the AP and NPP active sites and reactions that are catalyzed, it is likely that the NPP active site Thr is present as an alkoxide. If this model is correct, the acidic limb of the pH dependence for NPP-catalyzed $MpNPP^-$ hydrolysis could reflect protonation of the active site Thr with a pK_{E1} value of 4.6. There is some uncertainty in the fitted value of pK_{E1} because the data do not extend below pH 4.5 due to the irreversible loss of enzymatic activity. Furthermore, it is not possible to rule out models for the acidic limb involving a different enzyme residue that must be deprotonated for activity or correct folding.

There is no expectation from the structure of the active site (see below) that the hydrolysis mechanism requires a general acid that would account for the basic limb of the pH−rate profile. For AP, it was suggested that the basic limb of the pH dependence, with an observed pK_a of 8, reflected deprotonation of a Zn^{2+}-coordinated water molecule to form an inactive metal−hydroxide complex (*59*). If this model is correct, the shift to a higher pK_a in NPP of at least two units could reflect an active site that is optimized to bind less highly charged substrates. The observed pK_a of 10.1 could reflect deprotonation of an active site water molecule, the presence of an unidentified enzyme residue that must be protonated for activity, hydroxide ion-mediated protein unfolding, or loss of active site metal ions at high pH (see Materials and Methods).

Because phosphate monoesters have pK_a values within the experimentally accessible pH range (*58*), the pH dependence for NPP-catalyzed pNPP hydrolysis can be used to determine the ionization state of the substrate. Three mechanistic possibilities can be distinguished. First, the phosphate monoester dianion ($pNPP^{2-}$, Scheme 2) could be the reactive species, analogous to the situation observed for AP-catalyzed phosphate monoester hydrolysis (*59*). In this case, as the pH decreases below the pK_a of pNPP, the observed rate constant should decrease log linearly. Alternatively, the phosphate monoester monoanion ($pNPP^-$, Scheme 2) could be the reactive species, in which case the observed rate constant should decrease log linearly as the pH increases above the pK_a of pNPP. Finally, it is possible that NPP can react with both monoanionic and dianionic phosphate monoester substrates. In this case, the dependence of the observed rate constant on pH will have a complex dependence on the values of the bimolecular rate constants for each substrate.

The pH dependence for NPP-catalyzed hydrolysis of pNPP is shown in Figure 2A (empty symbols). From pH 9.5 to 7.5, the observed rate constant is unchanged, but as the pH is decreased below 7, an increase in the observed rate constant is observed. The data are consistent with a model in which NPP can react with both $pNPP^-$ and $pNPP^{2-}$, as depicted in Scheme 3B and described by eqs 3 and 4.

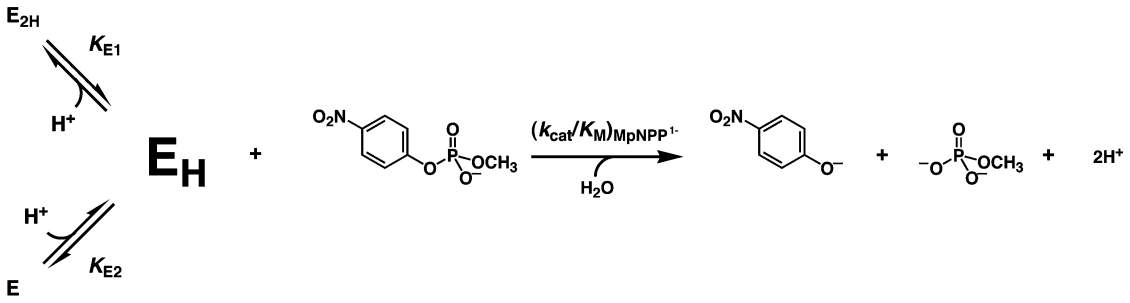
$$v_{obs} = (k_{cat}/K_M)_{obs}[E]_o[pNPP]_o = (k_{cat}/K_M)_{pNPP^-}[E_H] \times [pNPP^-] + (k_{cat}/K_M)_{pNPP^{2-}}[E_H][pNPP^{2-}] \quad (3)$$

$$(k_{cat}/K_M)_{obs} = \left[(k_{cat}/K_M)_{pNPP^{2-}} + \frac{(k_{cat}/K_M)_{pNPP^-} - (k_{cat}/K_M)_{pNPP^{2-}}}{1 + 10^{pH-pK_a}}\right] \times \left(\frac{1}{1 + 10^{pK_{E1}-pH} + 10^{pH-pK_{E2}}}\right) \quad (4)$$

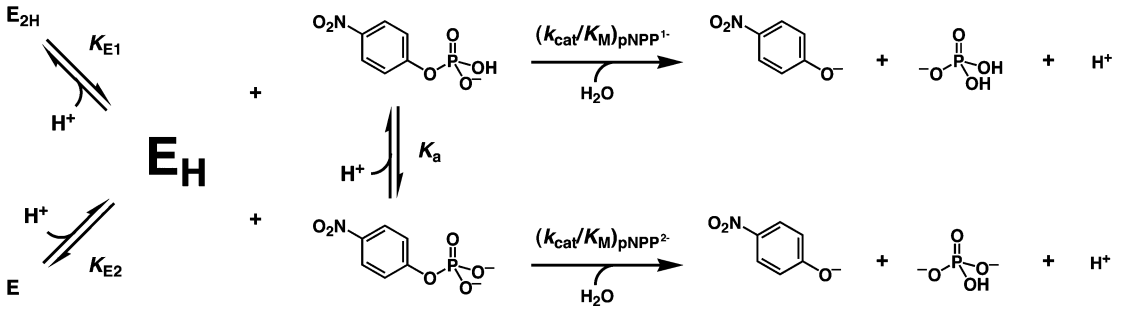
To eliminate effects arising from enzymatic ionizations, the ratio of observed rate constants for pNPP and $MpNPP^-$

Scheme 3

A) Kinetic Scheme for Reaction of MpNPP as a Function of pH



B) Kinetic Scheme for Reaction of pNPP as a Function of pH



hydrolysis is plotted as a function of pH in Figure 2B (see eq 5).

$$\frac{(k_{cat}/K_M)_{obs,pNPP}}{(k_{cat}/K_M)_{obs,MpNPP}} = \left[(k_{cat}/K_M)_{pNPP^{2-}} + \frac{(k_{cat}/K_M)_{pNPP^-} - (k_{cat}/K_M)_{pNPP^{2-}}}{1 + 10^{pH-pK_a}}\right]\left[\frac{1}{(k_{cat}/K_M)_{MpNPP}}\right] \quad (5)$$

A nonlinear least-squares fit of eq 5 to the data in Figure 2B using the experimentally determined pK_a of pNPP of 4.79 (see Materials and Methods) gives an excellent fit to the data and values of $(k_{cat}/K_M)_{pNPP^-}$ and $(k_{cat}/K_M)_{pNPP^{2-}}$ of 79 and 1.1 M^{-1} s^{-1}, respectively.

The reactivity of NPP with both $pNPP^-$ and $pNPP^{2-}$ is not surprising. $pNPP^-$ is analogous to a phosphate diester with a proton in place of the diester side chain (R′), and $pNPP^{2-}$ is the reactive species in the largely similar AP active site. In solution, however, phosphate monoester monoanions react via a mechanism distinct from that of both phosphate monoester dianions and phosphate diesters. The additional proton available to phosphate monoester monoanions leads to proton transfer from the nonbridging oxygen to the leaving group oxygen prior to leaving group bond cleavage (*20*, *22*, *60*). In a reaction with NPP with an active site metal ion present to stabilize charge buildup on the leaving group, however, $pNPP^-$ is likely to behave in a manner analogous to that of a simple phosphate diester. Consistent with this model, the value of k_{cat}/K_M for $pNPP^-$ is within 3-fold of the value for $MpNPP^-$, in which R′ is a methyl group (Table 1). The $\sim 10^2$-fold faster rate for $pNPP^-$ relative to that for $pNPP^{2-}$ is independent of contributions from interactions in an R′ binding site and thus reflects the intrinsic preference of the NPP active site for reactions of phosphate diesters over phosphate monoesters.

Phosphohydrolase Activity of NPP Is Zn^{2+}-Dependent. The model for the NPP active site (Scheme 1) is based on sequence comparisons to AP and strongly suggests the presence of two metal ions in the active site. Treatment of NPP with 10 mM EDTA resulted in inactivation (>300-fold loss in activity in reaction with $MpNPP^-$), confirming that the phosphohydrolase activity is metal ion-dependent. Using Chelex 100 resin (Bio-Rad) to remove metal ions (see Materials and Methods), we were able to both inactivate NPP (>10-fold loss in activity) and separate the protein from the chelator. Upon incubation of the demetalated enzyme with 100 μM $ZnCl_2$, the observed rate constant for hydrolysis of $MpNPP^-$ was completely restored to the value obtained with the protein prior to treatment with the chelator. Similar results were reported with mouse NPP1 (*3*).

The loss of activity upon treatment with metal ion chelators and the Zn^{2+}-mediated restoration of activity suggest that Zn^{2+} can bind in the NPP active site and support phosphohydrolase activity. These results do not establish the identity of the metal ion in vivo, however, because the *E. coli* periplasm is not the native environment for *Xac* NPP. Furthermore, the protein purification procedure is not necessarily expected to preserve the metal ions present in the active site in the cellular environment. Nevertheless, confirmation that Zn^{2+} supports phosphohydrolase activity in NPP is important for reactivity comparisons with AP, which contains two Zn^{2+} ions in the active site (Scheme 1).

Reactivity Comparisons of NPP and AP. Table 1 summarizes values of k_{cat}/K_M for reactions of *Xac* NPP and *E. coli* AP with phosphate diester and monoester substrates. Figure 3 shows the relative rates of diester and monoester hydrolysis for each enzyme, demonstrating that the preferences of NPP and AP for diester and monoester substrates are reversed. NPP reacts with diesters in preference to

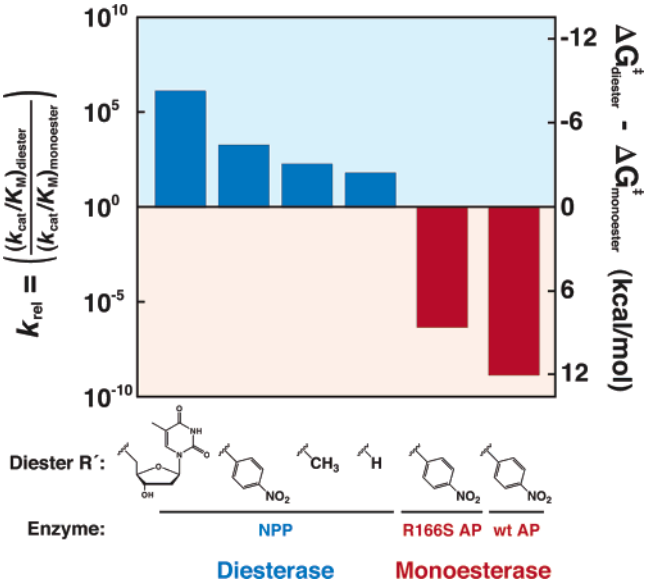


FIGURE 3: Relative reactivity of diester and monoester substrates with NPP and AP. The bar graph shows k_{rel} values for reactions of different diester substrates with NPP and AP. k_{rel} is the ratio of $(k_{cat}/K_M)_{diester}$ to $(k_{cat}/K_M)_{monoester}$. Several diester substrates with different R′ groups were used, and in all cases, $(k_{cat}/K_M)_{monoester}$ refers to the reaction with $pNPP^{2-}$ (Scheme 2). NPP reacts faster with diesters than monoesters (blue bars), and k_{rel} decreases as the size of the diester R′ group decreases. R166S AP and wt AP react faster with monoesters than diesters (red bars; k_{rel} for both AP variants is for bis-$pNPP^{-}$ relative to $pNPP^{2-}$). Values of k_{cat}/K_M are from Table 1.

monoesters by a factor of 10^2 for simple diesters and by a factor of 10^6 for a diester with a thymidine R′ group (Figure 3 and Scheme 2). In contrast, wild-type (wt) AP reacts with monoesters in preference to diesters by a factor of 10^9 (*14*).

In AP, Arg166 coordinates the nonbridging oxygen atoms of phosphate monoester substrates (Scheme 1) and makes a significant contribution to specificity for monoester substrates. Mutation of Arg166 to Ser weakens the preference of AP for monoesters relative to diesters from 10^9 to 10^6 (Table 1 and Figure 3) (*14*). This mutation exerts its effect entirely by decreasing the rate of phosphate monoester hydrolysis. Reactivity with phosphate diesters is unaffected, presumably because the arginine side chain can readily flip out of the active site to allow phosphate diester binding in wt AP (*14*). Nevertheless, the remaining 10^6-fold preference for monoesters relative to diesters in R166S AP represents the majority of the energetic discriminatory effect, and contributions from other active site features must be substantial.

According to a homology model (*3*) and the Zn^{2+} binding experiments described above, the NPP active site is expected to be largely similar to that of R166S AP. In this context, the intrinsic preference of the NPP active site for diesters by factors of 10^2 in the absence of contributions from an R′ binding site is remarkable in comparison with the 10^6-fold preference for monoesters in R166S AP. The differential discrimination of R166S AP versus NPP corresponds to ~10 kcal/mol. As discussed above, the bimetallo site itself could be optimized either for distinct features of the transition states for monoester and diester hydrolysis or for the difference in total charge on monoester and diester substrates. The physical origins of such effects in a largely conserved bimetallo site could involve repositioning of the active site Zn^{2+} ions or ligands, and homology models are not sufficient for such comparisons. We therefore determined the structure of NPP to allow for detailed comparisons of the NPP and AP active sites.

Structure of Xac NPP at 1.85 Å Resolution. The structure of *Xac* NPP was determined by X-ray diffraction, with an initial phase set obtained from the anomalous signal from the active site Zn^{2+} ions. Data collection and refinement statistics are reported in Table 1. Although the overall level of sequence homology of NPP and AP is remote and simple pairwise sequence alignments do not correctly align active site residues, there is substantial structural homology between NPP (Figure 4A) and AP (Figure 4B). Figure 4C shows an alignment of the structures of NPP and AP generated by minimizing the distances between corresponding active site residues (see Figure 5) and demonstrates that NPP and AP share a structurally conserved α/β core (colored red). The topology of the α/β core is also conserved, consistent with divergent evolution from a common ancestral sequence (Figure S3 of the Supporting Information). A sequence alignment generated from the structural alignment indicates that 199 of 382 NPP residues in noncontiguous stretches of sequence are homologous to residues in AP (Figure S4 of the Supporting Information). The overall level of sequence identity is 8%, and the level of identity within the homologous sequences is 16%, in both cases well below the 25% threshold at which sequence identity alone is a good predictor of homology (*61*).

There are several structural elements in NPP and AP that do not align with respect to each other. NPP residues 262−359 and AP residues 147−297, the regions colored green and silver on the left in Figure 4C, respectively, have no corresponding elements in the aligned structures except for four residues in AP from position 201 to 204 that align with residues in NPP (Figure S4 of the Supporting Information). There are also several noncontiguous stretches of sequence in AP, corresponding to the region colored silver on the right in Figure 4C, that make substantial contributions to the AP dimer interface and have no corresponding elements in NPP. AP is known to be active as a dimer in solution (*2*). *Xac* NPP crystallized as a dimer in the asymmetric unit; however, gel filtration chromotography and static light scattering suggest that NPP is a monomer in solution (see Materials and Methods), and the concentration dependence of activity measurements provides no indication that dimerization is required for function.

Inspection of the NPP active site region confirms the prediction of the homology model that the bimetallo sites of NPP and AP are conserved (*3*). NPP and AP both contain two active site Zn^{2+} ions, a set of six conserved metal ligands, and a Thr or Ser residue positioned to act as a nucleophile (Figure 5). In particular, the $Zn^{2+}-Zn^{2+}$ distances are almost identical. In NPP, the distances may vary slightly between the two monomers in the asymmetric unit from 4.26 ± 0.08 to 4.36 ± 0.08 Å. (Estimated standard devations were obtained as described in Materials and Methods.) For comparison, the $Zn^{2+}-Zn^{2+}$ distances in AP in the absence of bound ligand are reported to be 4.26 and 4.28 Å (*62*).

One notable difference between AP and NPP around the active site is the presence or absence of a third metal ion site. In AP, a Mg^{2+} ion is situated near the active site. Although this Mg^{2+} ion does not appear to directly contact

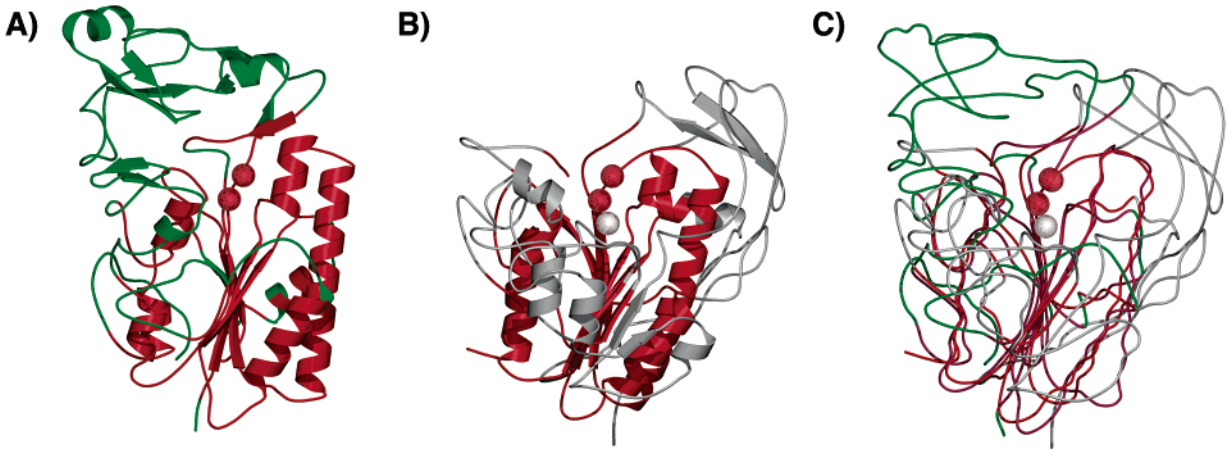


FIGURE 4: Structure of NPP and comparisons with AP. (A) Ribbon drawing of NPP. (B) Ribbon drawing of AP (PDB entry 1ED9). (C) Structural alignment of NPP and AP. The structures were aligned by minimizing the distances between corresponding residues in each active site (see Figure 5). The color scheme represents regions that align between the structures (red), regions of NPP that do not align with AP (green), and regions of AP that do not align with NPP (silver). The active site Zn^{2+} ions overlay and are colored red, and the Mg^{2+} ion of AP is colored silver. This figure and all subsequent structural figures were generated with POVScript+ (*82*).

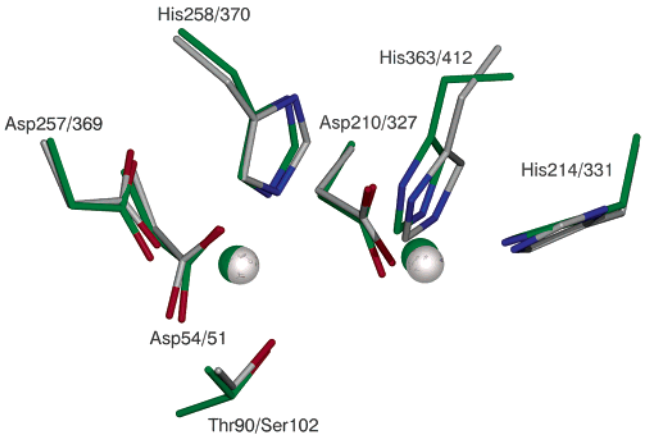


FIGURE 5: Overlay of the active sites of NPP (green) and AP (silver). The structures of NPP and AP (PDB entry 1ED9) in the absence of ligands were aligned in O (*46*) by minimizing the distances between corresponding Zn^{2+} ligands in each active site, using the Cα and Cβ atoms of these residues. Residue numbers are for NPP and AP, respectively.

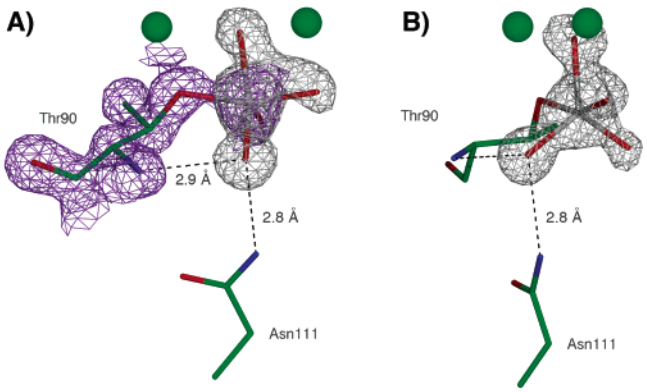


FIGURE 6: Vanadate bound in the active site of NPP. Vanadate adopts a trigonal bipyramidal geometry in the NPP active site, with one axial bond to the active site nucleophile (Oγ of Thr90). One equatorial oxygen is positioned between the two active site Zn^{2+} ions (green), and another is within hydrogen bonding distance of the backbone amide nitrogen of Thr90 and the side chain of Asn111. The remaining equatorial oxygen likely corresponds to the substrate oxygen esterified with R′ (Scheme 1) and is directed toward the R′ binding pocket (see Figure 8). Electron density from the σ_A-weighted $F_o - F_c$ map prior to inclusion of vanadate in the refined model is colored gray, and electron density from the σ_A-weighted $2F_o - F_c$ map surrounding Thr90 is colored purple (panel A only). (A) The $2F_o - F_c$ electron density indicates the presence of a covalent bond between Oγ of Thr90 and vanadium. (B) View of the active site along the axial bond of vanadate showing the trigonal plane. The $2F_o - F_c$ electron density is omitted for clarity.

the substrate, it has been suggested to be important for organizing the active site structure (*12*).[2] In NPP, no third metal site is observed (Figure S5 of the Supporting Information). Tyr205 occupies the region corresponding to the Mg^{2+} site in AP and is positioned to hydrogen bond with Asp54, which is also one of the Zn^{2+} ligands. In AP, the corresponding contact with Asp51 is made by the Mg^{2+} ion (Figure S5 of the Supporting Information).

Structure of NPP in Complex with Vanadate at 1.45 Å Resolution. The structure of NPP in complex with vanadate provides insight into potential interactions between the active site and the transition state. Vanadate binds in the NPP active site and adopts a trigonal bipyramidal geometry analogous to the transition state for phosphoryl transfer reactions (Figure 6). One of the axial positions is occupied by Oγ of Thr90, the active site nucleophile, and the opposing axial oxygen atom is positioned to interact with an active site Zn^{2+} in a manner likely to correspond to the leaving group interaction in a phosphate ester hydrolysis reaction. One equatorial oxygen atom is positioned between the two active site Zn^{2+} ions, and another is within hydrogen bonding distance of amide nitrogen atoms from the backbone of Thr90 and the side chain of Asn111. There are no protein functional groups within hydrogen bonding distance of the remaining equatorial oxygen, and the structure of NPP with AMP described below

[2] An alternative model was proposed for the role of Mg^{2+} in AP on the basis of the observation by X-ray crystallography that the Ser side chain is not positioned correctly for nucleophilic attack when the third metal site contains Zn^{2+} instead of Mg^{2+} (*62*). This observation was interpreted in terms of a model in which a Mg^{2+}-coordinated hydroxide ion acts as a general base to deprotonate the Ser nucleophile. However, the simplest model that accounts for the existing functional and structural data is that the Ser nucleophile is present as a Zn^{2+}-coordinated alkoxide (*59*). The Mg^{2+} site could help organize the AP active site (*12*) and may also contribute to a positively charged electrostatic environment, which could enhance reactivity with phosphate monoesters (see below).

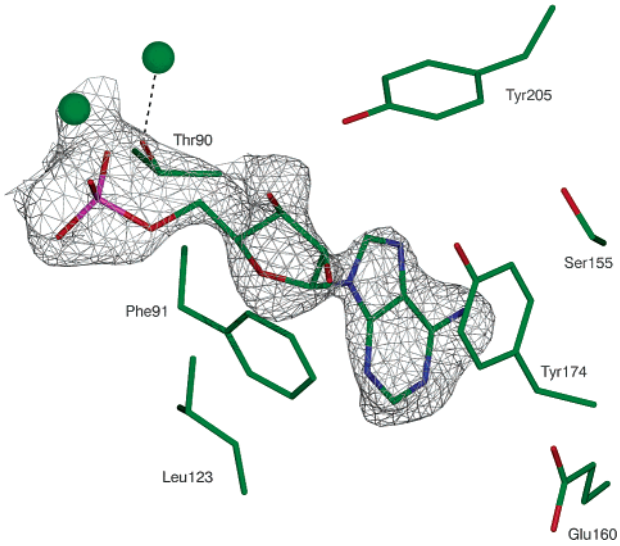


FIGURE 7: AMP bound in the active site of NPP. AMP binds in the active site of NPP with the 5′-phosphate near the Zn^{2+} ions (green) and the nucleoside in a binding pocket comprised of Phe91, Leu123, Ser155, Glu160, Tyr174, and Tyr205. The nucleoside corresponds to R′ in Scheme 1. Electron density from the σ_A-weighted $F_o - F_c$ map prior to inclusion of AMP in the refined model is colored gray.

confirms that this oxygen atom is likely to correspond to the substrate oxygen esterified with R′ (Scheme 1).

The active site $Zn^{2+}-Zn^{2+}$ distances in the vanadate-bound structure are 4.13 ± 0.01 and 4.17 ± 0.01 Å in each monomer, respectively, a decrease of approximately 0.1−0.2 Å relative to those in the holoenzyme structure. A similar 0.1−0.2 Å decrease was observed in the structure of AP with bound vanadate, in which the $Zn^{2+}-Zn^{2+}$ distances are reported to be 4.08 and 4.13 Å, respectively (*13*). The geometry of vanadate in the NPP active site is significantly distorted from an ideal trigonal bipyramid, with a bond angle along the axial bond axis of 157 ± 2°. It is not clear if this distortion is functionally significant. Comparable distortions of vanadate geometry have been observed in ribonuclease A (*63*). For ribonuclease A, the distortions were attributed to constraints introduced by binding interactions in the active site, although there are also likely to be intramolecular geometric constraints inherent in the 2′,3′-cyclic uridine vanadate ester species that was bound in the active site. The cause of vanadate distortion in the NPP active site is not readily apparent. The axial bond of vanadate bound in the active site of AP may also be distorted; the reported axial bond angle is 170° (*13*). The slightly larger distortion in NPP, if beyond error, could arise from differences in contacts to the equatorial oxygen atoms. In AP, Arg166 coordinates two equatorial oxygen atoms in a bidentate manner, while in NPP, an amide nitrogen atom of Asn111 coordinates only one of the equatorial oxygen atoms (Figure 8). The distortion in NPP could also arise because vanadate lacks any functionality analogous to R′ in a phosphate diester (Scheme 1), which could potentially make contacts that stabilize an idealized trigonal bipyramidal geometry. Finally, differences in the electrostatic properties of the bimetallo site, if present, could play a role.

Structure of NPP in Complex with AMP at 2.10 Å Resolution. Functional data with dT-5′-pNPP$^-$ suggest that an R′ binding pocket in NPP contributes at least 10^4-fold to catalysis of phosphate diester reactions (Figure 3). To identify the residues that comprise the R′ binding pocket, we determined the structure of NPP in complex with 5′-AMP, which is analogous to the nucleoside product of dT-5′-pNPP$^-$ hydrolysis. AMP was used instead of TMP because purines bind NPP roughly 10-fold more tightly than pyrimidines in inhibition assays (data not shown).

The conformation of AMP bound noncovalently in the active site of NPP is shown in Figure 7. The 5′-phosphate binds in the bimetallo site, and the adenosine binds in a pocket comprised of side chains from residues Phe91, Leu123, Ser155, Glu160, Tyr174, and Tyr205. The side chains of Val129 and Pro157 and the backbone of Trp156 also form part of the surface of the pocket but do not appear to be close enough to contact AMP directly. No contacts are observed that suggest specificity for a particular nucleoside. Glu211 is positioned to form hydrogen bonds with both the 2′- and 3′-hydroxyls, and Lys176 can form a hydrogen bond with the 2′-hydroxyl. There is some variability in the positions of these residues between the two monomers in the asymmetric unit, and in one of the monomers, only the contact between Glu211 and the 2′-hydroxyl is observed. 2′-Deoxy-AMP inhibits NPP with a K_I value of 110 μM (data not shown), which is similar to the inhibition constant for AMP of 260 μM, indicating that the observed contacts to the 2′-hydroxyl do not provide significant specificity for ribo- or deoxyribonucleosides.

Conservation Patterns in the NPP Family. The structures of NPP in complex with vanadate and AMP identified residues such as Asn111 and those comprising the R′ binding pocket that may be important for preferential phosphate diester hydrolysis by NPP. If these residues are indeed functionally significant, then they would likely be conserved in homologous NPP family members. To determine if these residues are conserved, a WU-BLAST search was performed with *Xac* NPP as the query sequence (*30*), and the hits were analyzed as described in Materials and Methods. In 57 of 68 unique sequences included in a multiple-sequence alignment, all six Zn^{2+} ligands and the active site Thr were conserved. These sequences are widely distributed across bacterial and eukaryotic species. Within the 57 sequences with all bimetallo site residues conserved, we examined sites corresponding to Asn111, Phe91, Leu123, Ser155, Glu160, Tyr174, and Tyr205 and observed widespread conservation of these residues across the NPP family. Asn111 and Tyr205 are the most highly conserved, occurring in 55 of 57 and 56 of 57 sequences, respectively. Phe91, Leu123, and Tyr174 are also conserved, but to a lesser extent, occurring in 46 of 57, 20 of 57, and 43 of 57 sequences, respectively. Glu160 is conserved in 23 of 57 sequences and in 21 of the remaining sequences is replaced with Asp. Finally, Ser155 is conserved in only 3 of 57 sequences and is most often replaced with Phe or Met. Eukaryotic NPPs are known to have distinct specificities for different diester substrates (*5*, *6*), and it is thus not surprising that some of the putative R′ binding site residues are not as well conserved as others. It is possible that the residues that are less conserved are involved in determining specificity for different types of diester substrates, and these residues could therefore be diagnostic of specificity for phosphate diester substrates within the NPP family.

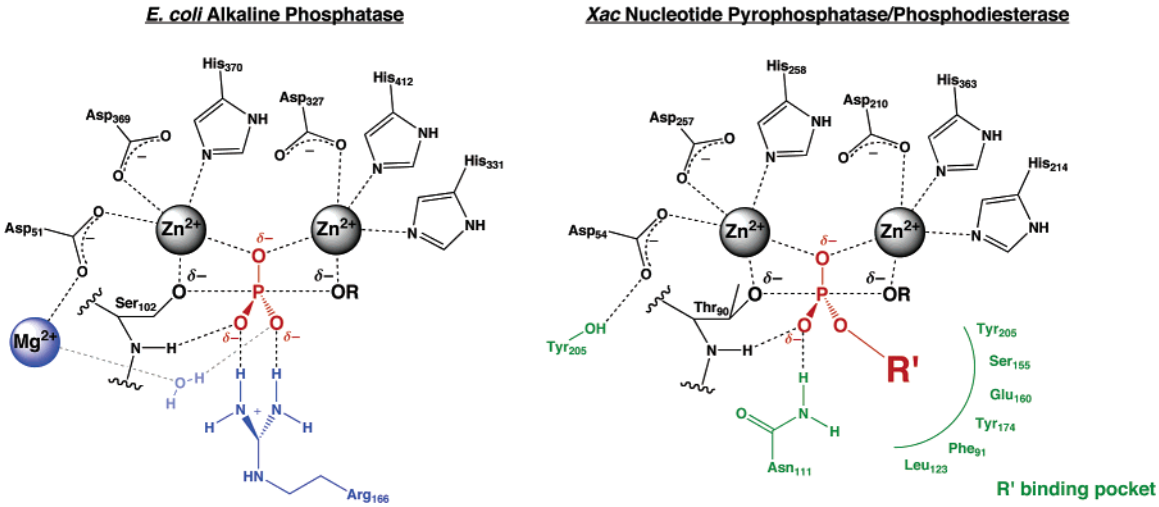


FIGURE 8: Revised active site schematics for AP and NPP highlight similarities and differences between these superfamily members. Functional groups unique to AP are colored blue, and functional groups unique to NPP are colored green. Conserved functional groups are colored black. The phosphate ester substrates are depicted in terms of a transition state representation with partial bond formation and bond cleavage. No information about bond orders and bond lengths is conveyed in this schematic.

DISCUSSION

Comparative Analysis of NPP and AP. *Xac* NPP and *E. coli* AP both catalyze hydrolysis reactions of phosphate diesters and monoesters. NPP preferentially catalyzes diester hydrolysis by factors ranging from 10^2 for simple diesters with hydrogen or methyl R′ groups to 10^6 for a diester with a thymidine R′ group, corresponding to an energetic discrimination ranging from 2.5 to 8.4 kcal/mol in favor of diesters (Figure 3). In contrast, AP preferentially catalyzes monoester hydrolysis by a factor of 10^9, corresponding to an energetic discrimination of 12 kcal/mol in favor of monoesters. The overall difference in specificity between NPP and AP is up to 10^{15}-fold, more than 20 kcal/mol (Figure 3). Despite the substantial difference in reactivity, the structure of NPP reveals that the bimetallo active sites of NPP and AP are nearly identical. How then do these similar active sites discriminate between phosphate monoester and diester substrates?

Similarities and differences between the AP and NPP active sites, taking into account features identified in the structural analysis of NPP, are presented in Figure 8. Several features were identified from the NPP structure as potential determinants of specificity for phosphate diester substrates. Phe91, Leu123, Ser155, Glu160, Tyr174, and Tyr205 contribute to an R′ binding site that increases reactivity with phosphate diesters by a factor of 10^4 for a substrate with a thymidine substituent on the transferred phosphoryl group (Figure 3). The R′ binding site could assist in binding the substrate and positioning the phosphoryl group relative to the active site nucleophile and Zn^{2+} ions.

Arg166 may perform an analogous function for phosphate monoester substrates in AP. Mutation of Arg166 to Ser weakens the preference of AP for monoesters relative to diesters from 10^9 to 10^6 (*14*), and the observation that Arg166 coordinates the nonbridging oxygen atoms suggests that Arg166 could assist with specific binding and positioning of phosphate monoester substrates (*12*, *13*). The remaining 10^6-fold preference for monoesters in the absence of Arg166 corresponds to an energetic discrimination of 7.3 kcal/mol (Figure 3). Possible origins for this remaining preference are discussed below.

NPP reacts ~10^2-fold faster with $pNPP^-$ than with $pNPP^{2-}$. This difference may reflect the intrinsic preference of the NPP active site for phosphate diesters in the absence of a contribution from the R′ binding site. As discussed in the introductory section, the bimetallo site itself could mediate specificity for phosphate diesters or monoesters by recognizing differences either in the nature of the transition state or in the total substrate charge. The structures of the bimetallo sites in NPP and AP are strikingly similar, however, and provide no obvious clues about the features that mediate such discrimination. Because reactions of phosphate monoesters proceed through loose transition states while phosphate diesters proceed through tighter, synchronous transition states, the simplest way to optimize a bimetallo site for diesters instead of monoesters might be to decrease the distance between the active site metal ions. The difference between axial bond O−P−O distances in loose and synchronous transition states has been suggested to be up to 1.4 Å (*64*), and a change in the Zn^{2+}−Zn^{2+} distance of this magnitude would have been readily detectable in the X-ray structures. However, no decrease in the Zn^{2+}−Zn^{2+} distances in NPP relative to AP was observed in comparisons of both the holoenzyme and vanadate-bound structures. Instead, the observed Zn^{2+}−Zn^{2+} distances in NPP were slightly longer than those reported for AP by 0.04−0.10 Å, although the differences are likely to be within the coordinate error of the AP structures. Nevertheless, it is possible that a decrease in the Zn^{2+}−Zn^{2+} distance in NPP relative to that in AP that is too small to be detected in the X-ray structures has a significant effect or that the active sites allow differential motion of the Zn^{2+} ions in NPP and AP that favors the monoesterase or diesterase reaction.

Discrimination between diester and monoester substrates could also arise from bimetallo sites that are optimized for differences in total substrate charge (*29*). The pH−rate profiles for AP and NPP provide functional evidence that supports this model. The basic limbs of the pH−rate profiles in AP (*59*) and NPP may result from deprotonation of Zn^{2+}-coordinated water molecules which forms inactive metal−hydroxide complexes, and the basic limb for NPP occurs 2 pH units higher than that for AP. The higher pK_a for deprotonation of an active site water in NPP could reflect

an active site that is optimized for the electrostatic properties of phosphate diester monoanions, which are less negatively charged than phosphate monoester dianions.

Residues that form the second sphere around the active site and make contacts with the ligands to the Zn^{2+} ions could modulate the electrostatic properties of the bimetallo sites by adjusting the Zn^{2+} positions or providing a different total charge density in the active site environment. The slightly longer $Zn^{2+}-Zn^{2+}$ distances in NPP relative to AP, if beyond error, could weaken the preference of the bimetallo site for negative charge between the two Zn^{2+} ions and favor reactions of phosphate diesters. Also, a Mg^{2+} ion contacts Asp51 in AP, while in NPP, the side chain hydroxyl of Tyr205 forms a hydrogen bond with Asp54 in NPP (Figure 8 and Figure S5). Thus, the Mg^{2+} ion in AP could provide extra positive charge density that assists in binding and reacting with phosphate monoesters in preference to diesters. This effect could be mediated by polarization of negative charge density on Asp51 away from the Zn^{2+} ion that this residue coordinates. In NPP, the absence of the Mg^{2+} site could lead to an electrostatic environment that prefers diester monoanions relative to monoester dianions.

In AP, a Mg^{2+}-bound water molecule is positioned to contact a nonbridging phosphate ester oxygen atom (Figure S5), and in NPP, this position corresponds to the oxygen esterified with the R′ group (Figure 8). Thus, instead of overall electrostatic differences between the AP and NPP active sites, the observed 10^2-fold faster rate for $pNPP^-$ relative to that for $pNPP^{2-}$ could arise from unfavorable solvation of a negatively charged oxygen atom in $pNPP^{2-}$ in a site that prefers to be occupied by an uncharged, esterified oxygen atom.

An additional possibility is that the bimetallo sites in NPP and AP are functionally equivalent in reactions of diesters and monoesters and the Mg^{2+} site has no significant effect. Because R166S AP preferentially reacts with phosphate monoesters, this model implies that specificity for phosphate monoesters is an intrinsic feature of the conserved bimetallo site. In this case, the specificity of NPP for phosphate diesters in the absence of contributions from the R′ binding site could arise from the interaction with Asn111, which is positioned to hydrogen bond with a nonbridging phosphoryl oxygen atom in NPP (Figure 8). Finally, it is also possible that substitution of Thr in NPP for Ser in AP is significant, although there is no reason to expect that these residues would give rise to intrinsic differences in reactivity as both are presumably Zn^{2+}-bound alkoxide nucleophiles in the active site.

Biological Function of Xac NPP. Although we do not have phenotypic or functional data for *Xac* NPP in its native cellular context, the biological role of bacterial NPP proteins could be similar to that of their eukaryotic counterparts, involving the processing of extracellular phosphate diester signaling molecules or precursors to signaling molecules. No experimental structural information is available for a eukaryotic NPP family member, but sequence comparisons with *Xac* NPP suggest that the eukaryotic homologues have a similar bimetallo active site as well as residues that contribute to an R′ binding site. Several eukaryotic NPP homologues have additional domains that are absent in bacterial NPP proteins and have an undetermined role in biological function (*5*). An alternative possibility for the biological role of bacterial NPP proteins might be harvesting nucleosides from the extracellular medium. *Xac* NPP does have detectable 3′−5′ nuclease activity with a value of k_{cat}/K_M for sequential cleavage of an RNA oligonucleotide of $\sim 10^3$ M^{-1} s^{-1} (J. Zalatan and M. Forconi, unpublished results).

Implications for Enzyme Evolution. NPP and AP are evolutionarily related members of the AP superfamily that have evolved to catalyze different reactions with remarkable specificity despite having very similar active sites. There are also members of the AP superfamily that appear to have nearly equivalent activities for phosphate monoester and diester hydrolysis (*65*). These enzymes may represent intermediate states along the evolutionary path from a specific phosphate monoesterase to a diesterase or vice versa.

Both NPP and AP have promiscuous activities for the reactions that are preferentially catalyzed by their homologous counterpart. Promiscuous enzyme activities have been suggested to support a model for enzyme evolution in which enzymes that can catalyze alternative reactions with a low level of activity will have a selective advantage for evolutionary optimization of that activity, after a gene duplication event (*66*−*71*). Additional support for this model comes from the existence of enzymes that, like NPP and AP, are optimized for reactions that are promiscuous activities of their evolutionarily related homologues (*67*, *72*−*74*).

There is a growing number of examples of promiscuous activities that can be substantially enhanced with one mutation or a few mutations (*72*−*78*). In some cases, these enhancements occur without significant detrimental effects on the native function, suggesting that optimization of a new enzyme activity prior to gene duplication could sometimes occur without negative effects on organism fitness (*74*). This type of evolutionary scenario could have been operative in the evolution of new activities in the AP superfamily. If NPP evolved from an ancestral AP-like precursor, mutations that introduced parts of an R′ binding site could have enhanced diesterase activity without decreasing monoesterase activity. The ancestral precursor could have even retained a residue analogous to Arg166 to maintain reactivity with phosphate monoesters. In AP, Arg166 provides specificity for phosphate monoesters and has no significant effect on reactions of AP with diesters, presumably because the arginine side chain can move out of the AP active site to allow for diester binding (*14*).

The structurally homologous regions of *Xac* NPP and *E. coli* AP comprise only ∼50% of each protein, and the overall level of sequence identity is 8% (Figure 4 and Figure S4). Similar examples of enzyme superfamilies with conserved active site functional elements and widely divergent sequences have been reported (*79*, *80*). Most of the divergence in sequence identity between NPP and AP is not likely to contribute to differential reactivity between phosphate monoester and diester substrates; some of the observed sequence changes could result from random drift at unconstrained or weakly constrained positions that still allow the same overall fold (*80*), whereas other changes could be important for protein−protein interactions or other in vivo functions not directly related to the chemical transformation. It is also possible that some changes contribute to recognition and hydrolysis of large phosphate ester substrates such as oligonucleotides but have little effect on reactivity with smaller substrates.

Conclusions. Our current structural and functional data provide a partial explanation for the remarkable discrimina-

tion between phosphate diesters and monoesters in the NPP and AP active sites. Several possibilities for the functional groups that provide the remaining discrimination are suggested and can be tested in future structure−function experiments. The ability to perform comparative experiments on the similar active sites of NPP and AP provides a powerful tool for improving our understanding of enzyme catalytic mechanisms and the evolution of new enzyme activities.

ACKNOWLEDGMENT

We thank Matthew Footer for advice on surface modification of glass, Daniel Ratner for assistance with X-ray data collection, Daniel Riordan for assistance with sequence analysis, and members of the Herschlag and Brunger labs for comments on the manuscript.

SUPPORTING INFORMATION AVAILABLE

Identical inhibition of diesterase and monoesterase activities by tungstate and GMP, experimentally phased map of the NPP active site, topology diagrams, structure-based sequence alignment of NPP and AP, and a comparison of the AP Mg^{2+} site with the corresponding region of NPP. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Galperin, M. Y., Bairoch, A., and Koonin, E. V. (1998) A superfamily of metalloenzymes unifies phosphopentomutase and cofactor-independent phosphoglycerate mutase with alkaline phosphatases and sulfatases, *Protein Sci. 7*, 1829−1835.
2. Coleman, J. E. (1992) Structure and mechanism of alkaline phosphatase, *Annu. Rev. Biophys. Biomol. Struct. 21*, 441−483.
3. Gijsbers, R., Ceulemans, H., Stalmans, W., and Bollen, M. (2001) Structural and catalytic similarities between nucleotide pyrophosphatases/phosphodiesterases and alkaline phosphatases, *J. Biol. Chem. 276*, 1361−1368.
4. Galperin, M. Y., and Jedrzejas, M. J. (2001) Conserved core structure and active site residues in alkaline phosphatase superfamily enzymes, *Proteins 45*, 318−324.
5. Stefan, C., Jansen, S., and Bollen, M. (2005) NPP-type ectophosphodiesterases: Unity in diversity, *Trends Biochem. Sci. 30*, 542−550.
6. Goding, J. W., Grobben, B., and Slegers, H. (2003) Physiological and pathophysiological functions of the ecto-nucleotide pyrophosphatase/phosphodiesterase family, *Biochim. Biophys. Acta 1638*, 1−19.
7. Landt, M., and Butler, L. G. (1978) 5′-Nucleotide phosphodiesterase: Isolation of covalently bound 5′-adenosine monophosphate, an intermediate in the catalytic mechanism, *Biochemistry 17*, 4130−4135.
8. Culp, J. S., Blytt, H. J., Hermodson, M., and Butler, L. G. (1985) Amino acid sequence of the active site peptide of bovine intestinal 5′-nucleotide phosphodiesterase and identification of the active site residue as threonine, *J. Biol. Chem. 260*, 8320−8324.
9. Oda, Y., Kuo, M. D., Huang, S. S., and Huang, J. S. (1993) The major acidic fibroblast growth factor (aFGF)-stimulated phosphoprotein from bovine liver plasma membranes has aFGF-stimulated kinase, autoadenylylation, and alkaline nucleotide phosphodiesterase activities, *J. Biol. Chem. 268*, 27318−27326.
10. Belli, S. I., Mercuri, F. A., Sali, A., and Goding, J. W. (1995) Autophosphorylation of PC-1 (alkaline phosphodiesterase I/nucleotide pyrophosphatase) and analysis of the active site, *Eur. J. Biochem. 228*, 669−676.
11. Clair, T., Lee, H. Y., Liotta, L. A., and Stracke, M. L. (1997) Autotaxin is an exoenzyme possessing 5′-nucleotide phosphodiesterase/ATP pyrophosphatase and ATPase activities, *J. Biol. Chem. 272*, 996−1001.
12. Kim, E. E., and Wyckoff, H. W. (1991) Reaction mechanism of alkaline phosphatase based on crystal structures. Two-metal ion catalysis, *J. Mol. Biol. 218*, 449−464.
13. Holtz, K. M., Stec, B., and Kantrowitz, E. R. (1999) A model of the transition state in the alkaline phosphatase reaction, *J. Biol. Chem. 274*, 8351−8354.
14. O'Brien, P. J., and Herschlag, D. (2001) Functional interrelationships in the alkaline phosphatase superfamily: Phosphodiesterase activity of *Escherichia coli* alkaline phosphatase, *Biochemistry 40*, 5691−5699.
15. Beese, L. S., and Steitz, T. A. (1991) Structural basis for the 3′−5′ exonuclease activity of *Escherichia coli* DNA polymerase I: A two metal ion mechanism, *EMBO J. 10*, 25−33.
16. Steitz, T. A. (1999) DNA polymerases: Structural diversity and common mechanisms, *J. Biol. Chem. 274*, 17395−17398.
17. Strater, N., Lipscomb, W. N., Klabunde, T., and Krebs, B. (1996) Two-metal ion catalysis in enzymatic acyl- and phosphoryl-transfer reactions, *Angew. Chem., Int. Ed. 35*, 2024−2055.
18. Wilcox, D. E. (1996) Binuclear metallohydrolases, *Chem. Rev. 96*, 2435−2458.
19. Kirby, A. J., and Jencks, W. P. (1965) Reactivity of nucleophilic reagents toward *p*-nitrophenyl phosphate dianion, *J. Am. Chem. Soc. 87*, 3209−3216.
20. Kirby, A. J., and Varvoglis, A. G. (1967) Reactivity of phosphate esters. Monoester hydrolysis, *J. Am. Chem. Soc. 89*, 415−423.
21. Thatcher, G. R. J., and Kluger, R. (1989) Mechanism and catalysis of nucleophilic substitution in phosphate esters, *Adv. Phys. Org. Chem. 25*, 99−265.
22. Hengge, A. C., Edens, W. A., and Elsing, H. (1994) Transition-state structures for phosphoryl-transfer reactions of *p*-nitrophenyl phosphate, *J. Am. Chem. Soc. 116*, 5045−5049.
23. Hengge, A. C., and Onyido, I. (2005) Physical organic perspectives on phospho group transfer from phosphates and phosphinates, *Curr. Org. Chem. 9*, 61−74.
24. Kirby, A. J., and Younas, M. (1970) The reactivity of phosphate esters. Reactions of diesters with nucleophiles, *J. Chem. Soc. B*, 1165−1172.
25. Kirby, A. J., and Younas, M. (1970) The reactivity of phosphate esters. Diester hydrolysis, *J. Chem. Soc. B*, 510−513.
26. Hengge, A. C., and Cleland, W. W. (1991) Phosphoryl-transfer reactions of phosphodiesters: Characterization of transition states by heavy-atom isotope effects, *J. Am. Chem. Soc. 113*, 5835−5841.
27. Hengge, A. C., Tobin, A. E., and Cleland, W. W. (1995) Studies of transition-state structures in phosphoryl transfer teactions of phosphodiesters of *p*-nitrophenol, *J. Am. Chem. Soc. 117*, 5919−5926.
28. Zalatan, J. G., and Herschlag, D. (2006) Alkaline phosphatase mono- and diesterase reactions: Comparative transition state analysis, *J. Am. Chem. Soc. 128*, 1293−1303.
29. Nikolic-Hughes, I., O'Brien, P. J., and Herschlag, D. (2005) Alkaline phosphatase catalysis is ultrasensitive to charge sequestered between the active site zinc ions, *J. Am. Chem. Soc. 127*, 9314−9315.
30. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J. H., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs, *Nucleic Acids Res. 25*, 3389−3402.
31. Chaidaroglou, A., Brezinski, D. J., Middleton, S. A., and Kantrowitz, E. R. (1988) Function of arginine-166 in the active site of *Escherichia coli* alkaline phosphatase, *Biochemistry 27*, 8338−8343.
32. Bendtsen, J. D., Nielsen, H., von Heijne, G., and Brunak, S. (2004) Improved prediction of signal peptides: SignalP 3.0, *J. Mol. Biol. 340*, 783−795.
33. Gill, S. C., and von Hippel, P. H. (1989) Calculation of protein extinction coefficients from amino acid sequence data, *Anal. Biochem. 182*, 319−326.
34. Bourne, N., and Williams, A. (1984) Effective charge on oxygen in phosphoryl ($-PO_3^{2-}$) group transfer from an oxygen donor, *J. Org. Chem. 49*, 1200−1204.
35. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol. 276*, 307−326.
36. Terwilliger, T. C., and Berendzen, J. (1999) Automated MAD and MIR structure solution, *Acta Crystallogr. D55*, 849−861.
37. Terwilliger, T. C. (2000) Maximum-likelihood density modification, *Acta Crystallogr. D56*, 965−972.
38. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., and Terwilliger, T. C. (2002) PHENIX: Building

new software for automated crystallographic structure determination, *Acta Crystallogr. D58*, 1948−1954.
39. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C., and Read, R. J. (2005) Likelihood-enhanced fast translation functions, *Acta Crystallogr. D61*, 458−464.
40. Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics, *Acta Crystallogr. D60*, 2126−2132.
41. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr. D54*, 905−921.
42. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr. D53*, 240−255.
43. Lamzin, V. S., and Wilson, K. S. (1993) Automated refinement of protein models, *Acta Crystallogr. D49*, 129−147.
44. Winn, M. D., Isupov, M. N., and Murshudov, G. N. (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement, *Acta Crystallogr. D57*, 122−133.
45. Sheldrick, G. M., and Schneider, T. R. (1997) SHELXL: High-resolution refinement, *Methods Enzymol. 277*, 319−343.
46. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models, *Acta Crystallogr. A47*, 110−119.
47. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) Clustal W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, *Nucleic Acids Res. 22*, 4673−4680.
48. Do, C. B., Mahabhashyam, M. S. P., Brudno, M., and Batzoglou, S. (2005) ProbCons: Probabilistic consistency-based multiple sequence alignment, *Genome Res. 15*, 330−340.
49. Gijsbers, R., Ceulemans, H., and Bollen, M. (2003) Functional characterization of the non-catalytic ectodomains of the nucleotide pyrophosphatase/phosphodiesterase NPP1, *Biochem. J. 371*, 321−330.
50. Cimpean, A., Stefan, C., Gijsbers, R., Stalmans, W., and Bollen, M. (2004) Substrate-specifying determinants of the nucleotide pyrophosphatases/phosphodiesterases NPP1 and NPP2, *Biochem. J. 381*, 71−77.
51. Malamy, M., and Horecker, B. L. (1961) Localization of alkaline phosphatase in *E. coli* K12, *Biochem. Biophys. Res. Commun. 5*, 104−108.
52. Inouye, H., Barnes, W., and Beckwith, J. (1982) Signal sequence of alkaline phosphatase of *Escherichia coli*, *J. Bacteriol. 149*, 434−439.
53. Kelly, S. J., and Butler, L. G. (1977) Enzymic hydrolysis of phosphonate esters. Reaction mechanism of intestinal 5′-nucleotide phosphodiesterase, *Biochemistry 16*, 1102−1104.
54. Moe, O. A., and Butler, L. G. (1983) The catalytic mechanism of bovine intestinal 5′-nucleotide phosphodiesterase, *J. Biol. Chem. 258*, 6941−6946.
55. Lindquist, R. N., Lynn, J. L., and Lienhard, G. E. (1973) Possible transition-state analogs for ribonuclease. The complexes of uridine with oxovanadium(IV) ion and vanadium(V) ion, *J. Am. Chem. Soc. 95*, 8762−8768.
56. Van Etten, R. L., Waymack, P. P., and Rehkop, D. M. (1974) Transition metal ion inhibition of enzyme-catalyzed phosphate ester displacement reactions, *J. Am. Chem. Soc. 96*, 6782−6785.
57. Davies, D. R., and Hol, W. G. J. (2004) The power of vanadate in crystallographic investigations of phosphoryl transfer enzymes, *FEBS Lett. 577*, 315−321.
58. Jencks, W. P., and Regenstein, J. (1976) Ionization Constants of Acids and Bases, in *Handbook of Biochemistry and Molecular Biology* (Fasman, G. D., Ed.) CRC, Cleveland, OH.
59. O'Brien, P. J., and Herschlag, D. (2002) Alkaline phosphatase revisited: Hydrolysis of alkyl phosphates, *Biochemistry 41*, 3207−3225.
60. Grzyska, P. K., Czyryca, P. G., Purcell, J., and Hengge, A. C. (2003) Transition state differences in hydrolysis reactions of alkyl versus aryl phosphate monoester monoanions, *J. Am. Chem. Soc. 125*, 13106−13111.
61. Sander, C., and Schneider, R. (1991) Database of homology-derived protein structures and the structural meaning of sequence alignment, *Proteins: Struct., Funct., Genet. 9*, 56−68.
62. Stec, B., Holtz, K. M., and Kantrowitz, E. R. (2000) A revised mechanism for the alkaline phosphatase reaction involving three metal ions, *J. Mol. Biol. 299*, 1303−1311.
63. Wladkowski, B. D., Svensson, L. A., Sjolin, L., Ladner, J. E., and Gilliland, G. L. (1998) Structure (1.3 Å) and charge states of a ribonuclease A-uridine vanadate complex: Implications for the phosphate ester hydrolysis mechanism, *J. Am. Chem. Soc. 120*, 5488−5498.
64. Hengge, A. C. (1998) Transfer of the PO_3^{2-} group, in *Comprehensive Biological Catalysis* (Sinnott, M. L., Ed.) pp 517−542, Academic Press, London.
65. Zhang, L., Balcerzak, M., Radisson, J., Thouverey, C., Pikula, S., Azzar, G., and Buchet, R. (2005) Phosphodiesterase activity of alkaline phosphatase in ATP-initiated Ca^{2+} and phosphate deposition in isolated chicken matrix vesicles, *J. Biol. Chem. 280*, 37289−37296.
66. Jensen, R. A. (1976) Enzyme recruitment in evolution of new function, *Annu. Rev. Microbiol. 30*, 409−425.
67. O'Brien, P. J., and Herschlag, D. (1999) Catalytic promiscuity and the evolution of new enzymatic activities, *Chem. Biol. 6*, R91−R105.
68. Palmer, D. R. J., Garrett, J. B., Sharma, V., Meganathan, R., Babbitt, P. C., and Gerlt, J. A. (1999) Unexpected divergence of enzyme function and sequence: "*N*-Acylamino acid racemase" is *o*-succinylbenzoate synthase, *Biochemistry 38*, 4252−4258.
69. Penning, T. M., and Jez, J. M. (2001) Enzyme redesign, *Chem. Rev. 101*, 3027−3046.
70. Copley, S. D. (2003) Enzymes with extra talents: Moonlighting functions and catalytic promiscuity, *Curr. Opin. Chem. Biol. 7*, 265−272.
71. James, L. C., and Tawfik, D. S. (2003) Conformational diversity and protein evolution: A 60-year-old hypothesis revisited, *Trends Biochem. Sci. 28*, 361−368.
72. Wise, E. L., Yew, W. S., Akana, J., Gerlt, J. A., and Rayment, I. (2005) Evolution of enzymatic activities in the orotidine 5′-monophosphate decarboxylase suprafamily: Structural basis for catalytic promiscuity in wild-type and designed mutants of 3-keto-L-gulonate 6-phosphate decarboxylase, *Biochemistry 44*, 1816−1823.
73. Yew, W. S., Akana, J., Wise, E. L., Rayment, I., and Gerlt, J. A. (2005) Evolution of enzymatic activities in the orotidine 5′-monophosphate decarboxylase suprafamily: Enhancing the promiscuous D-arabino-hex-3-ulose 6-phosphate synthase reaction catalyzed by 3-keto-L-gulonate 6-phosphate decarboxylase, *Biochemistry 44*, 1807−1815.
74. Aharoni, A., Gaidukov, L., Khersonsky, O., Gould, S. M., Roodveldt, C., and Tawfik, D. S. (2005) The 'evolvability' of promiscuous protein functions, *Nat. Genet. 37*, 73−76.
75. Schmidt, D. M. Z., Mundorff, E. C., Dojka, M., Bermudez, E., Ness, J. E., Govindarajan, S., Babbitt, P. C., Minshull, J., and Gerlt, J. A. (2003) Evolutionary potential of $(\beta/\alpha)_8$-barrels: Functional promiscuity produced by single substitutions in the enolase superfamily, *Biochemistry 42*, 8387−8393.
76. Gould, S. M., and Tawfik, D. S. (2005) Directed evolution of the promiscuous esterase activity of carbonic anhydrase II, *Biochemistry 44*, 5444−5452.
77. Norrgard, M. A., Ivarsson, Y., Tars, K., and Mannervik, B. (2006) Alternative mutations of a positively selected residue elicit gain or loss of functionalities in enzyme evolution, *Proc. Natl. Acad. Sci. U.S.A. 103*, 4876−4881.
78. Yoshikuni, Y., Ferrin, T. E., and Keasling, J. D. (2006) Designed divergent evolution of enzyme function, *Nature 440*, 1078−1082.
79. Gerlt, J. A., and Babbitt, P. C. (2001) Divergent evolution of enzymatic function: Mechanistically diverse superfamilies and functionally distinct suprafamilies, *Annu. Rev. Biochem. 70*, 209−246.
80. Orengo, C. A., and Thornton, J. M. (2005) Protein families and their evolution: A structural perspective, *Annu. Rev. Biochem. 74*, 867−900.
81. Brunger, A. T. (1992) Free *R* value: A novel statistical quantity for assessing the accuracy of crystal structures, *Nature 355*, 472−475.
82. Fenn, T. D., Ringe, D., and Petsko, G. A. (2003) POVScript+: A program for model and data visualization using persistence of vision ray-tracing, *J. Appl. Crystallogr. 36*, 944−947.

BI060847T