Mechanistic and Evolutionary Insights from Comparative Enzymology of Phosphomonoesterases and Phosphodiesterases across the Alkaline Phosphatase Superfamily

Fanny Sunden,† Ishraq AlSadhan,† Artem Y. Lyubimov, Susanne Ressl,§ Helen Wiersma-Koch,‡⊥ Jamar Borland,† Clayton L. Brown, Jr., † Tory A. Johnson, † Zorawar Singh, † and Daniel Herschlag*, †,

†Department of Biochemistry, Beckman Center, Stanford University, Stanford, California 94305, United States
‡Departments of Molecular and Cellular Physiology, Neurology and Neurological Science, Structural Biology, and Photon Science, Howard Hughes Medical Institute, Stanford University, Stanford, California 94305, United States
§Molecular and Cellular Biochemistry Department, Indiana University, Bloomington, Indiana 47405, United States
⊥Department of Biology, Indian River State College, Fort Pierce, Florida 34981, United States
∥Departments of Chemical Engineering and Chemistry, and Stanford ChEM-H (Chemistry, Engineering, and Medicine for Human Health), Stanford University, Stanford, California 94305, United States

Supporting Information

ABSTRACT: Naively one might have expected an early division between phosphate monoesterases and diesterases of the alkaline phosphatase (AP) superfamily. On the contrary, prior results and our structural and biochemical analyses of phosphate monoesterase PafA, from Chryseobacterium meningosepticum, indicate similarities to a superfamily phosphate diesterase [Xanthomonas citri nucleotide pyrophosphatase/phosphodiesterase (NPP)] and distinct differences from the three metal ion AP superfamily monoesterase, from Escherichia coli AP (EcAP). We carried out a series of experiments to map out and learn from the differences and similarities between these enzymes. First, we asked why there would be independent instances of monoesterases in the AP superfamily? PafA has a much weaker product inhibition and slightly higher activity relative to EcAP, suggesting that different metabolic evolutionary pressures favored distinct active-site architectures. Next, we addressed the preferential phosphate monoester and diester catalysis of PafA and NPP, respectively. We asked whether the >80% sequence differences throughout these scaffolds provide functional specialization for each enzyme’s cognate reaction. In contrast to expectations from this model, PafA and NPP mutants with the common subset of active-site groups embedded in each native scaffold had the same monoesterase:diesterase specificities; thus, the >107-fold difference in native specificities appears to arise from distinct interactions at a single phosphoryl substituent. We also uncovered striking mechanistic similarities between the PafA and EcAP monoesterases, including evidence for ground-state destabilization and functional active-site networks that involve different active-site groups but may play analogous catalytic roles. Discovering common network functions may reveal active-site architectural connections that are critical for function, and identifying regions of functional modularity may facilitate the design of new enzymes from existing promiscuous templates. More generally, comparative enzymology and analysis of catalytic promiscuity can provide mechanistic and evolutionary insights.

INTRODUCTION

Obtaining a fundamental understanding of how enzymes achieve their enormous rate enhancements and exquisite specificities and elucidating how new enzymes have evolved are central goals of biochemistry. Progress in these areas has biological implications and implications for the ultimate practical goal of effectively and efficiently designing enzymes with new and beneficial activities.

Site-directed mutagenesis, coupled with structural analysis, has been the prevailing approach in enzymology over the past decades. Nevertheless, approaches beyond subtractive site-directed mutagenesis are needed because enzyme energetics are non-additive.15–16 Our recent quantitative dissection of the functional interplay of five residues in the active site of Escherichia coli alkaline phosphatase (EcAP) provided unique mechanistic and evolutionary insights, as well as information that may help guide the design of new enzymes.17 Nevertheless, enzyme function is also dependent on its overall context, as the enzyme’s structure is, of course, needed for catalysis, but it is
not yet practical to comprehensively dissect the interconnections between active-site residues and the sea of surrounding residues and structural elements.

Given the natural complexities of enzymes and these practical limitations, the study of enzyme superfamilies has been enormously valuable. These studies have identified common and potentially critical structural and catalytic elements and, conversely, suggested residues and structural features that may allow individual enzymes to specialize in different reactions (e.g., refs 18−20). Of particular utility is the observation that enzymes across a superfamily often exhibit low but substantial activity for reactions of other superfamily members, a property referred to as ‘catalytic promiscuity’.21,22 Catalytic promiscuity can be used as a comparative tool in elucidating mechanism and as a probe of factors that may have influenced enzyme evolution (see refs 23 and 24).

The alkaline phosphatase (AP) superfamily has been particularly amenable to comparative studies, as members exhibit promiscuity for multiple reactions catalyzed by other superfamily members.24−35 We compare PafA, a phosphate monoesterase from Chryseobacterium meningosepticum, with two other enzymes within the AP superfamily that share its Zn2+ bimetallo core, one catalyzing the same reaction and the other catalyzing phosphate diester hydrolysis. Intriguingly, the PafA monoesterase has similarities to superfamily diesterases and differences from the three metal ion AP superfamily monoesterases such as EcaAP. These observations raise intriguing evolutionary questions and provide compelling opportunities to obtain mechanistic insights via comparative enzymology and analysis of catalytic promiscuity.32

■ EXPERIMENTAL SECTION

Plasmid Design. C. meningosepticum expresses the alkaline phosphatase gene pafA. This gene has been modified with a sequence expressing a C-terminal strepII tag with a Factor Xa cleavage site below and above the measured reaction. The cells were grown to an optical density of 0.6 for crystallization.

The alkaline phosphatase (AP) superfamily has been particularly amenable to comparative studies, as members exhibit promiscuity for multiple reactions catalyzed by other superfamily members.24−35 We compare PafA, a phosphate monoesterase from Chryseobacterium meningosepticum, with two other enzymes within the AP superfamily that share its Zn2+ bimetallo core, one catalyzing the same reaction and the other catalyzing phosphate diester hydrolysis. Intriguingly, the PafA monoesterase has similarities to superfamily diesterases and differences from the three metal ion AP superfamily monoesterases such as EcaAP. These observations raise intriguing evolutionary questions and provide compelling opportunities to obtain mechanistic insights via comparative enzymology and analysis of catalytic promiscuity.32

Scheme 1. Reaction Scheme for Phosphate Ester Hydrolysis

from initial rates, and the activity of the free enzyme, \( k_{cat}/K_M \), was determined. The kinetic parameters were shown to be first-order in both enzyme and substrate concentration, with concentrations varied over at least a 5-fold range. For pNPP, substrate concentrations high enough to obtain \( K_M \) and \( k_{cat} \) values were achievable, and these values were also reported; the concentration of pNPP was varied at least 5-fold below and above the measured \( K_M \) in all cases. The following pNPP concentrations were used: WT, 2.0 × 10−6−2.5 mM; T79S, 0.55−110 μM; K162A, 1.1 × 10−3−2.5 mM; R164A, 1.1 × 10−3−4.5 mM. Fits had R2 values of >0.98 in all cases. Reported errors were estimated from at least two independent kinetic measurements. Comparisons with independent enzyme preparations for each of the PafA mutants gave values within the reported error.

The \( K_M \) value for pNPP of the N100A mutant was too low to measure directly at pH 8.0. We therefore used an indirect method described previously.24 Brieﬂy, we used an inhibitor (tungstate), whose inhibition constant was determined using a substrate with a much higher \( K_M \) (me-pNPP), to raise the observed \( K_M \) value in these pNPP reactions. Apparent values for \( k_{cat} \) and \( K_M \) were obtained at 2.1, 6.3, and 12.5 mM tungstate. As expected for competitive inhibition, the \( k_{cat} \) values were constant (and the same as obtained in the absence of tungstate), and the observed \( K_M \) values increased with increasing tungstate concentration. From these observed values, \( K_M \) for the pNPP reaction was calculated using eq 1, the measured \( K_M \) apparent and the independently determined \( K_i \).
Kinetics. We followed the approach previously taken by O

limits in

Ma

π; T79S, 16 Information S3). As the observed activity arose from the same active site (Supporting WT PafA and the T79S, K162A, and R164A mutants to test if the reactive site of the least reactive PafA mutant, K162A, was too low to vary any need to correct values for pNPP or other more activated substrates. However, reactivity of the least reactive PafA mutant, K162A, was too low to measure with Me-P. Nevertheless, we were able to obtain an upper limit for this mutant with Me-P and \(k_{cat}/K_M\) values for all of the other PafA variants, and all of the conclusions herein could be made by these comparisons alone. Further, we used comparisons between substrates of differing intrinsic reactivity to obtain a \(k_{cat}/K_M\) estimate for K162A PafA relative to the other variants. The only underlying assumption is that each variant has the same ratio of chemical catalysis for the substrates, and this assumption is supported by comparisons of several variants and substrates (Supporting Information S4).

Occupancy of Metals and Phosphate in the PafA Active Site and Structure. To test for possible metal ion concentration dependent activation, such as observed for EcAP mutant E322Y,24 WT PafA was incubated with the following metal concentrations: 10 \(\mu\)M ZnCl\(_2\), 100 \(\mu\)M ZnCl\(_2\), 500 \(\mu\)M ZnCl\(_2\), 1.0 mM MgCl\(_2\), and 100 \(\mu\)M ZnCl\(_2\) in 10 mM sodium MOPS, pH 8.0, and 50 mM NaCl at 25 °C. No activation was observed after a week of incubation. The occupancy of metals and phosphorus in the WT PafA active site was determined with atomic emission spectroscopy as described previously.24 The following ratios were obtained: Zn\(^{2+}\)-protein = 3.1; Mg\(^{2+}\)-protein = 0.012; Ca\(^{2+}\)-protein = 0.53; P\(_2\)-protein = 2.9. The metal ion occupancies suggested that Mg\(^{2+}\) does not bind to the PafA active site, consistent with the lack of activation by 1 mM MgCl\(_2\). The Zn\(^{2+}\) value is consistent with occupancy of the Zn\(^{2+}\)-bimetallo site and partial occupancy of one or both of the remote Zn\(^{2+}\) sites observed in the crystal structure (see below). Although the results suggest half occupancy of a Ca\(^{2+}\) ion, addition of Ca\(^{2+}\) (1 mM) did not increase activity.

Crystallization and Crystallographic Data Collection for WT PafA. The affinity tag was cleaved from PafA by incubating WT PafA with factor Xa (New England BioLabs, Ipswich, MA) for 4 days in 50 mM Tris-HCl, pH 8.0, 200 mM NaCl, 10 mM maltose, 5 mM CaCl\(_2\). PafA was buffer exchanged into 10 mM Tris-HCl, pH 8.0, 50 mM NaCl, 100 \(\mu\)M ZnCl\(_2\) and concentrated to 2.5 mg/mL. Crystal growth was achieved via the hanging drop method using vapor diffusion where equal volumes of PafA and precipitant solution (22% polyethylene glycol (PEG) MW 20,000, 0.1 M sodium acetate, pH 4.5, 0.2 mM ammonium sulfate) were mixed on a cover slide and placed over a reservoir of 1 mL of precipitant solution at room temperature. Crystals were harvested using nylon loops (Hampton Research, Aliso Viejo, CA) or LithoLoops (Molecular Dimensions Inc., Almonte Springs, FL) and immediately flash frozen into liquid nitrogen; the high concentration of PEG MW 3550 in the precipitation solution was determined to be sufficient for cryopreservation. Crystallographic data were collected at the Stanford Synchrotron Radiation Laboratory (SSRL) Beamline 11-1. Radiation sensitivity of the crystals was immediately apparent during data collection, as the overall intensity of recorded reflections decreased with concurrent rapid increase of mosaicity after ~20–40 s-second exposures. To alleviate this problem, we collected data sets from seven PafA crystals and merged high-quality frames into a single data set.

Processing of Diffraction Data and Structure Determination. Raw diffraction images were indexed and integrated using oMOSlM,40 followed by a merging and scaling step using Pointless and Aimless.21 The final merged data set included 30–40 initial frames from each of the seven separate crystals and extended to 1.7 Å. The L-test42 detected significant twinning, with the twin fraction of 0.242; as a result, the merging R-values appear somewhat inflated. The final merging statistics are provided in Table 1. The structure was solved by molecular replacement (MR) using Phaser,43,44 with the modified structure of SPAP (PDB ID: 3Q3Q) used as a search model. MR was successful only when most of the divergent regions were omitted, with essentially the conserved alpha-beta core retained in the search model. The maps were steadily improved with multiple rounds of iterative manual modeling of the best-resolved portions in Coot45 alternating with automated refinement in phenix.refine.46 A set of two operators (−h−k−l) was utilized throughout the refinement. The final model included nearly the entire expressed PafA monomer, with the exception of 12 residues of the N-terminus and 1 residue of the C-terminus. At this stage, clear electron density could be seen for the two active-site Zn\(^{2+}\) ions as well as two additional Zn\(^{2+}\) sites at the

\[
K_{M,\text{apparent}} = \frac{K_M}{1 + [1]/K_i}
\]

(1)

Values of \(k_{cat}/K_M\) for Me-P (Chart 1) were obtained using a discontinuous kinetic assay, following the formation of free phosphate by a Malachite green assay (as described in ref 37). For the N100A and T79S mutants, the \(K_M\) was too low, given the sensitivity of this assay to obtain reliable \(k_{cat}/K_M\) values. Tungstate could not be used as above to raise the \(K_M\) and allow determination of \(k_{cat}/K_M\) because the Malachite Green assay is not compatible with tungstate. We therefore measured the activity of these mutants at pH 9.0, where the \(K_M\) is higher, and extrapolated to pH 8.0 using the ratio of pNPP \(k_{cat}/K_M\) values at pH 8.0 and pH 9.0 (eq 2; Supporting Information S2). As a control, the same measurements were carried out for WT PafA, and the measured and calculated values agreed within 2-fold.

\[
\frac{k_{cat}}{K_M} = \frac{(k_{cat}/K_M)_{pNPP,pi} + k_{cat}}{(k_{cat}/K_M)_{pi} + k_{cat}}
\]

(2)

Hydrolysis rates for phenyl phosphate (Ph-P) were measured by following production of free phenol. Production of free phenol was assayed as phenolate by taking aliquots of the reaction at specified time points and quenching in equal volume of 0.1 M sodium hydroxide. Absorbance was measured at 278 nm with 500 nm as a reference wavelength on a Tecan Infinite 200 PRO Microplate Reader (Tecan, Männedorf, Switzerland), against a standard curve of phenol (Sigma-Aldrich, St. Louis, MO).

The following buffers were used for pH dependences: MES (pH 6.0), MOPS (pH 7.0), CHES (pH 9.0), CAPS (pH 10.0) at 100 mM, each as a sodium salt in the presence of 500 mM NaCl and 100 \(\mu\)M ZnCl\(_2\) to ensure that there were no changes in pH-dependences or protonation states introduced by the mutations (Supporting Information S2).

Inhibition constants were determined for tungstate, vanadate, and inorganic phosphate (Pi) in 0.1 M sodium MOPS, pH 8.0, 0.5 M NaCl, 100 \(\mu\)M ZnCl\(_2\), using subsaturating concentrations of me-pNPP (Chart 1). The following inhibitor concentrations were used: WT, 0.16–28 \(\mu\)M vanadate, 1.6 × 10\(^{-3}\)–13 mM tungstate, 0.13–5.0 mM Pi; T79S, 16–450 \(\mu\)M vanadate, 6.3 × 10\(^{-3}\)–13 mM tungstate, 6.3 × 10\(^{-3}\)–12 mM Pi; N100A, 6.3 × 10\(^{-3}\)–13 mM vanadate, 6.3 × 10\(^{-3}\)–13 mM tungstate, 0.2–13 mM Pi; K162A, 16–450 \(\mu\)M vanadate, 6.3 × 10\(^{-3}\)–13 mM tungstate, 0.38–13 mM Pi; R164A, 1.6 × 10\(^{-3}\)–13 mM vanadate, 6.3 × 10\(^{-3}\)–13 mM tungstate, 0.38–13 mM Pi.

Inhibition constants for vanadate were also measured with pNPP for WT PafA and the T79S, K162A, and R164A mutants to test if the observed activity arose from the same active site (Supporting Information S3). As the \(K_M\) for pNPP hydrolysis by N100A PafA at pH 8.0 is very small, the \(K_M\) for tungstate was measured at pH 9.0 (0.1 M sodium CHES, 0.5 M NaCl, 100 \(\mu\)M ZnCl\(_2\)) for me-pNPP and pNPP. The inhibition constant was calculated with a non-linear least-squares fit to the data using eq 3 for competitive inhibition.

\[
\frac{k_{cat}}{K_M} = \frac{K_M}{1 + [1]/K_i}
\]

(3)

Accounting for Non-chemical Rate-Limiting Steps and Limits in \(k_{cat}/K_M\) Measurements. As previously observed for WT EcAP,38 the \(k_{cat}/K_M\) values for pNPP by WT PafA and a subset of its mutants are within the range observed when substrate binding, rather than a chemical step, is rate limiting. To allow comparisons of catalysis by PafA and its mutants that assess the ability of these enzymes to carry out the chemical step, we used information from intrinsically less reactive substrates so that comparisons reflect effects on substrate binding and the subsequent chemical step and not simply association kinetics. We followed the approach previously taken by O’Brien et al. for EcAP.34,38

In principle we could obtain and compare \(k_{cat}/K_M\) values for all PafA variants with a poorly reactive substrate, and thereby avoid any need to correct values for pNPP or other more activated substrates. However, reactivity of the least reactive PafA mutant, K162A, was too low to
The overall structure and active-site configuration of PafA closely match those of *Sphingomonas* sp. BSAR-1 (SPAP), an AP superfamily member closely related to PafA (Supporting Information S6). As noted by Bihani et al., SPAP shares specific active-site features with the AP superfamily phosphodiesterases such as NPP. These similarities are also seen with PafA (Figure 2). PafA, SPAP, and NPP share a nucleophilic threonine, while EcAP has a serine, and PafA, SPAP, and NPP contain a homologous active-site asparagine that appears to make identical interactions with oxygen atom (O₁) of the transferred phosphoryl group (Figure 2, Scheme 2, Supporting Information S5). EcAP lacks this asparagine residue, and, although the PafA, SPAP, and EcAP active sites all have arginine and lysine residues, the arginine and lysine residues in the EcAP active site originate from distinct helices and loops and make distinct interactions (Figures 1B and 2, Supporting Information S5 and S6, ref 51).

PafA was previously isolated and expressed by Berluti et al., and initial structure—function studies were carried out by Bihani et al. with SPAP. Our studies build upon this prior research and are consistent with their results.

**PafA’s Reaction Specificity for Phosphate Monoester versus Diester Hydrolysis.** Based on structural inspection, SPAP and PafA are expected to be phosphate monoesterases, as they lack a binding pocket for the second diester substituent and have positively charged residues situated to interact with both non-bridging oxygen atoms of phosphomonoester substrates (Figure 2, Supporting Information S6, ref 51). To formally test this expectation, we followed the PafA hydrolysis of phosphate monoester and diester substrates, pNPP and me-pNPP, respectively (Chart 1). We observed a preference of ~10⁵-fold for catalysis of the monoester reaction (Table 2, WT). This value represents a lower limit for the preferential chemical catalysis, as binding rather than the chemical step appears to be rate limiting for the pNPP reaction (see Experimental Section, Accounting for Non-chemical Rate-Limiting Steps and Limits in kcat/K₉ Measurements, and Supporting Information S4), and as the methyl substituent of the diester substrate does not make binding interactions made with substituents of physiological diesterase substrates.³⁴,³²

**Effect of Active-Site Mutations on PafA Catalysis.** We next assessed the effect of PafA active-site mutations on the phosphate monoesterase and diesterase activities. Each PafA active-site residue that appears to contact the phosphoryl group, N100, K162, and R164, was mutated to alanine, and the nucleophilic threonine residue, T79, was mutated to serine. Values of kcat/K₉ were first determined with a p-nitrophenoxide leaving group (Table 2; pNPP and me-pNPP, Chart 1). Whereas the diester reactions were uniformly slow, and thus very likely limited by the chemical step, the high kcat/K₉ values of ~10⁶ M⁻¹ s⁻¹ for pNPP, the monoester substrate, were within the range seen for rate constants for substrate binding to enzymes. Furthermore, mutations that gave 20–200-fold effects on the non-cognate diesterase reaction had essentially no effect on the cognate monoesterase reaction for a subset of the mutations (Table 2, i.e., T79S and N100A). This result is most simply accounted for by a non-chemical rate-limiting step that masks effects on the chemical step for the pNPP monoesterase reaction for WT PafA (Supporting Information S4).

To further investigate the effects of the mutations on phosphomonoester hydrolysis we therefore carried out reactions for each PafA variant with less reactive phosphate monoester substrates, phenyl phosphate (Ph-P) and methyl phosphate monoester substrates, phenyl phosphate (Ph-P) and methyl

---

**Table 1. X-ray Crystallographic Data Collection and Refinement Statistics**

<table>
<thead>
<tr>
<th>Data Collection</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>space group</td>
<td>I4</td>
</tr>
<tr>
<td>unit cell axes</td>
<td>a, b, c (Å) 113.8, 113.8, 71.4</td>
</tr>
<tr>
<td></td>
<td>α, β, γ (deg) 90, 90, 90</td>
</tr>
<tr>
<td>resolution range (Å)</td>
<td>36.0–1.7</td>
</tr>
<tr>
<td>Rwork/Rfree (%)</td>
<td>28.7 (50.0)</td>
</tr>
<tr>
<td>completeness (%)</td>
<td>96.6 (85.3)</td>
</tr>
<tr>
<td>multiplicity</td>
<td>5.4 (1.9)</td>
</tr>
<tr>
<td>CC1/2</td>
<td>95.8% (41.6%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Refinement</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>resolution range (Å)</td>
<td>36.0–1.7</td>
</tr>
<tr>
<td>no. unique reflections</td>
<td>48541 (4475)</td>
</tr>
<tr>
<td>Rwork/Rfree (%)</td>
<td>15.9/17.7</td>
</tr>
<tr>
<td>no. atoms</td>
<td>4771</td>
</tr>
<tr>
<td>average B-factors (Å²)</td>
<td>13.1, 29.4, 40.9</td>
</tr>
<tr>
<td>bond length (Å)</td>
<td>0.093</td>
</tr>
<tr>
<td>bond angle (deg)</td>
<td>0.62</td>
</tr>
<tr>
<td>Ramachandran statistics</td>
<td></td>
</tr>
<tr>
<td>favored regions (%)</td>
<td>97.2</td>
</tr>
<tr>
<td>allowed regions (%)</td>
<td>2.4</td>
</tr>
<tr>
<td>outliers (%)</td>
<td>0.4</td>
</tr>
<tr>
<td>PDBe code</td>
<td>5TJ3</td>
</tr>
</tbody>
</table>

---

**RESULTS**

**The PafA Structure.** Crystals of PafA were obtained, and its structure was determined to 1.7 Å resolution (Table 1). As expected based on sequence similarities, its core architecture matches that of EcAP, NPP, and other members of the AP superfamily (Figure 1AB, gray and white), and its Zn²⁺ bimetallo center is highly homologous to those of NPP and EcAP (Figure 2; see Supporting Information S5 for a structure-based sequence alignment).
phosphate (Me-P; Chart 1, Table 3). The larger mutational effects observed with the less reactive substrates supported the model of rate limiting binding for reaction of pNPP with the faster mutants. Hence, we used Me-P to compare the PafA

Figure 1. Global structure and active site of PafA, NPP, and EcAP. (A) Comparison of the global structure of PafA, NPP, and EcAP. The structurally conserved core and bimetallo Zn²⁺ (spheres) are shown in gray, conserved zinc ligands are shown in black, and auxiliary domains that are not conserved between all three enzymes are shown in color; PafA in green, NPP in magenta, and EcAP in blue. The active-site nucleophile (threonine or serine) is shown in blue, and other active-site residues and EcAP Mg²⁺ (sphere) are shown in orange. (B) Topographic diagrams for PafA, NPP, and EcAP, with arrows representing β-sheets and rods representing α-helices. Gray represents conserved sheets and white conserved helices, and the auxiliary regions that are not conserved between all three enzymes are colored: PafA, helices in olive, sheets in green; NPP, helices in purple, sheets in magenta; and EcAP, helices in light blue, sheets in blue. Structures are from 3tg0 (EcAP,96) and 2gso (NPP,32) with ligands removed for clarity.
mutants. As only an upper limit could be obtained for the Me-P reactivity of the most affected mutant (K162A), we used the observed constant ratio of reactivity of the different substrates to estimate this $k_{cat}/K_M$ value (see Experimental Section and Supporting Information S4); use of this value rather than the limit does not affect the conclusions drawn herein.

Given the uniformly low $k_{cat}/K_M$ values with the phosphate diester substrate, additional precautions and controls are required to ensure that a contaminant that is physically undetectable, but highly proficient at the alternative reaction, does not account for the observed promiscuous activity. We therefore determined inhibition constants for the phosphate monoester and diester reactions (Supporting Information S3). For each PafA variant, the same inhibition constant was therefore determined inhibition constants for the phosphate monoester and diester reactions (Supporting Information S3).

Table 2. PafA Phosphomonoesterase and Diesterase Reaction Kinetics

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>$k_{cat}/K_M$ (M$^{-1}$ s$^{-1}$)</th>
<th>$k_{cat}$</th>
<th>$K_M$ (M)</th>
<th>$k_{cat}$ (s$^{-1}$)</th>
<th>$k_{cat}$ (M$^{-1}$ s$^{-1}$)</th>
<th>$K_M$ (M)</th>
<th>$(k_{cat}/K_M)_{Me-pNPP}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>$1.6(0.4) \times 10^6$</td>
<td>(1)</td>
<td>$1.9(0.4) \times 10^{-4}$</td>
<td>$3.0 \times 10^2$</td>
<td>$1.7(0.5) \times 10^2$</td>
<td>(1)</td>
<td>$9.4 \times 10^6$</td>
</tr>
<tr>
<td>T79S</td>
<td>$6.3(1.5) \times 10^3$</td>
<td>2.5</td>
<td>$2.9(0.3) \times 10^{-6}$</td>
<td>1.8</td>
<td>$6.6(2.2) \times 10^{-1}$</td>
<td>26</td>
<td>$9.5 \times 10^6$</td>
</tr>
<tr>
<td>N100A</td>
<td>$7.6(2.0) \times 10^3$</td>
<td>2.1</td>
<td>$5.1(1.3) \times 10^{-7}$</td>
<td>$3.9 \times 10^{-1}$</td>
<td>$1.1(0.4) \times 10^{-1}$</td>
<td>$1.5 \times 10^1$</td>
<td>$6.9 \times 10^6$</td>
</tr>
<tr>
<td>K162A</td>
<td>$1.3(0.1) \times 10^3$</td>
<td>$1.2 \times 10^1$</td>
<td>$2.3(0.8) \times 10^{-4}$</td>
<td>$3.0 \times 10^{-2}$</td>
<td>$1.3(0.1) \times 10^1$</td>
<td>1.3</td>
<td>$1.0 \times 10^1$</td>
</tr>
<tr>
<td>R164A</td>
<td>$1.3(0.4) \times 10^3$</td>
<td>$1.2 \times 10^1$</td>
<td>$5.1(1.5) \times 10^{-5}$</td>
<td>6.6</td>
<td>$1.4(0.1) \times 10^1$</td>
<td>1.2</td>
<td>$9.3 \times 10^5$</td>
</tr>
</tbody>
</table>

*Assay conditions: 100 mM sodium MOPS, pH 8.0, 500 mM NaCl, 100 μM ZnCl$_2$. $b$ $k_{cat}$ values are $k_{cat}/K_M$ for WT divided by $k_{cat}/K_M$ for each mutant. By definition the value for WT PafA is one. $c$ $(K_M)$$_{Me-pNPP}$ $\gg$ 1.0 mM for WT and all mutants.
C. meningosepticum and salt and fresh water (coli complex regulation of Pi uptake and storage and roles of phosphate in stress and pathogenesis. 62,63 The di AP superfamily monoesterases such as E. coli phosphate diesterases and di ester reactions between all monoesterases and persistent distinctions between early division between phosphate monoesterase and diesterase of catalytic promiscuity.32

Table 3. $k_{\text{cat}}/K_i$ Values for Phosphomonoester Hydrolysis by WT and Mutant PafA Enzymes$^a$

<table>
<thead>
<tr>
<th>enzyme</th>
<th>$K_i$ (mM)</th>
<th>$K_{\text{cat}}$ (s$^{-1}$)</th>
<th>$K_i$ (mM)</th>
<th>$K_{\text{cat}}$ (s$^{-1}$)</th>
<th>$K_i$ (mM)</th>
<th>$K_{\text{cat}}$ (s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>$8.3(1.0) \times 10^{-1}$</td>
<td>(1)</td>
<td>$3.9(0.1) \times 10^{-1}$</td>
<td>(1)</td>
<td>$6.7(0.4) \times 10^{-4}$</td>
<td>(1)</td>
</tr>
<tr>
<td>T79S</td>
<td>$9.6(0.4) \times 10^{-3}$</td>
<td>$1.2 \times 10^{-2}$</td>
<td>$2.8(0.8) \times 10^{-1}$</td>
<td>$7.2 \times 10^{-1}$</td>
<td>$9.0(0.2) \times 10^{-2}$</td>
<td>$1.3 \times 10^{2}$</td>
</tr>
<tr>
<td>N100A</td>
<td>$1.3(0.2)$</td>
<td>$1.6$</td>
<td>$5.7(2.3) \times 10^{-1}$</td>
<td>$1.5 \times 10^{2}$</td>
<td>$1.9(0.1) \times 10^{-1}$</td>
<td>$2.8 \times 10^{2}$</td>
</tr>
<tr>
<td>K162A</td>
<td>$4.7(1.2)$</td>
<td>$5.7$</td>
<td>$&gt;13$</td>
<td>$&gt;3.3 \times 10^{3}$</td>
<td>$7.9(2.6) \times 10^{-2}$</td>
<td>$1.2 \times 10^{2}$</td>
</tr>
<tr>
<td>R164A</td>
<td>$&gt;13$</td>
<td>$&gt;16$</td>
<td>$1.3(0.4)$</td>
<td>$3.3 \times 10^{3}$</td>
<td>$3.0(0.7) \times 10^{-1}$</td>
<td>$4.5 \times 10^{2}$</td>
</tr>
</tbody>
</table>

Table 4. Inhibition of WT and Mutant PafA by Ground- and Transition-State Analogues$^a$

<table>
<thead>
<tr>
<th>enzyme</th>
<th>$K_i$ (mM)</th>
<th>$K_{\text{cat}}$ (s$^{-1}$)</th>
<th>$K_i$ (mM)</th>
<th>$K_{\text{cat}}$ (s$^{-1}$)</th>
<th>$K_i$ (mM)</th>
<th>$K_{\text{cat}}$ (s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>$8.3(1.0) \times 10^{-1}$</td>
<td>(1)</td>
<td>$3.9(0.1) \times 10^{-1}$</td>
<td>(1)</td>
<td>$6.7(0.4) \times 10^{-4}$</td>
<td>(1)</td>
</tr>
<tr>
<td>T79S</td>
<td>$9.6(0.4) \times 10^{-3}$</td>
<td>$1.2 \times 10^{-2}$</td>
<td>$2.8(0.8) \times 10^{-1}$</td>
<td>$7.2 \times 10^{-1}$</td>
<td>$9.0(0.2) \times 10^{-2}$</td>
<td>$1.3 \times 10^{2}$</td>
</tr>
<tr>
<td>N100A</td>
<td>$1.3(0.2)$</td>
<td>$1.6$</td>
<td>$5.7(2.3) \times 10^{-1}$</td>
<td>$1.5 \times 10^{2}$</td>
<td>$1.9(0.1) \times 10^{-1}$</td>
<td>$2.8 \times 10^{2}$</td>
</tr>
<tr>
<td>K162A</td>
<td>$4.7(1.2)$</td>
<td>$5.7$</td>
<td>$&gt;13$</td>
<td>$&gt;3.3 \times 10^{3}$</td>
<td>$7.9(2.6) \times 10^{-2}$</td>
<td>$1.2 \times 10^{2}$</td>
</tr>
<tr>
<td>R164A</td>
<td>$&gt;13$</td>
<td>$&gt;16$</td>
<td>$1.3(0.4)$</td>
<td>$3.3 \times 10^{3}$</td>
<td>$3.0(0.7) \times 10^{-1}$</td>
<td>$4.5 \times 10^{2}$</td>
</tr>
</tbody>
</table>

also exhibit extensive catalysis of $\sim 10^{12}$-fold for both reactions (Table 5).

**DISCUSSION**

Enzymes of the AP superfamily catalyze a range of reactions, with its bimetallo branch containing phosphate monoesterases and diesterases, along with phosphomutases and phosphonate esterase hydrolyses.24–27,32,33,58,59 Naively, one might expect an early division between phosphate monoesterase and diesterase superfamily members, such that there would be similarities between all monoesterases and persistent distinctions between these enzymes and the diesterases. However, PaFA and SPAP are phosphate monoesterases with similarities to superfamily phosphate diesterases and differences from the three metal ion AP superfamily monoesterases such as E. coli AP (EcAP) (Figures 1 and 2; Supporting Information S6; ref 51, comment 60). These observations raised intriguing evolutionary questions and provided us with compelling opportunities to obtain mechanistic insights via comparative enzymology and analysis of catalytic promiscuity.32

**Evolution of Phosphate Monoesterases with Distinct Functional Properties.** Why are there AP superfamily phosphate monoesterases with distinct active-site architectures, and why would one more resemble the AP superfamily phosphate diesterases? EcAP comes from a gut bacteria (E. coli), whereas PaFA is derived from a bacterium found in soil and salt and fresh water (C. meningosepticum).52,61 Further, phosphate is often a scarce resource for growth, and there is complex regulation of P$_i$ uptake and storage and roles of phosphate in stress and pathogenesis.52,65 The different catalytic constants we observe for PaFA and EcAP are consistent with different metabolic evolutionary pressures, arising from different lifestyles that favored distinct active-site architectures. EcAP exhibits very strong product inhibition by P$_i$ with a $K_i$ in the low micromolar and more than 3 orders of magnitude stronger than that for PaFA (Figure 4A). The strong binding could be functional—for example, limiting enzyme turnover until P$_i$ that has been produced can be cleared from the periplasm and utilized. Alternatively, the strong product inhibition may be a necessary trade-off to achieve the ~4-fold higher $k_{\text{cat}}/K_i$ for EcAP relative to PaFA (Figure 4B).66

While it is always possible that the functional differences between EcAP and PaFA arose from trapping of a suboptimal evolutionary solution as a local maximum on the enzyme’s fitness landscape for one of enzymes, there is also an intriguing possibility that E. coli and a subset of other bacteria utilize strong product inhibition for phosphate homeostasis and signaling in ways that remain to be elucidated. Our results open new biological questions that would be indiscernible without quantitative comparative enzymology.

**Modular Active-Site Architecture Facilitates Transitions between AP Superfamily Phosphate Monoesterases and Diesterases.** Within the EcAP active site, R166 interacts with both of the phosphoryl oxygen atoms that face away from the Zn$^{2+}$ bimetallo site (Figure 2). Consistent with this structural picture, mutation of R166 is deleterious for both phosphate monoesterase and diesterase reactions but is more deleterious for monoesterase reactions.24 Thus, if R166 of EcAP were mutated to allow the introduction of residues that interact with a diester R’ substituent, interactions with the other, non-esterified oxygen atom would be compromised, and the promiscuous diester reaction would be adversely affected. Based on these observations, the EcAP active-site architecture would be expected to impede evolutionary transitions between mono- and diesterases, and these observations may account for the absence of diesterase superfamily members that are highly related to EcAP. For analogous reasons, EcAP might be a suboptimal starting point for engineering a new phosphodies- terase.

In contrast, residues within PaFA’s active site interact with *either* of the two phosphoryl oxygen atoms facing away from the Zn$^{2+}$ bimetallo center (O$_1$ and O$_3$ in Scheme 2). Oxygen atom O$_1$ accepts hydrogen bonds from the side chain of N100 and the backbone amide of T79 (the active-site nucleophile), and
O₂ accepts hydrogen bonds from K162 and R164 (Figure 2). Correspondingly, mutation of N100 gives a uniform deleterious effect on monoester and diester hydrolysis, suggesting that this interaction facilitates both reactions, and mutations of K162 and R164 are deleterious only to phosphate monoester hydrolysis, with essentially no effect on the diester reaction (Figure 5A). Further, the opposite behavior is observed for NPP—i.e., mutations of the residues that constitute the R’ binding pocket are deleterious to NPP’s diesterase activity but not to its monoesterase activity.

The presence of distinct interactions with each of the two phosphoryl oxygen atoms or substituents in PafA and NPP allow interactions around one of the phosphoryl oxygen atoms to be altered and optimized for phosphate monoester or diester catalysis, while the interactions with the phosphoryl oxygen atom that contributes equally to both reactions can be maintained. We propose that the structural and functional modularity of the phosphoryl oxygen interactions facilitated evolutionary transitions between phosphate mono- and diesterases across the AP superfamily. Information of this type about the functional architecture of active sites may also be powerful in guiding efforts to reengineer enzymes for new functions.

**Beyond the O₂ and R’ Site in PafA and NPP: Extensive Sequence Variation in These Enzymes’ Scaffolds Does Not Contribute to Reaction Specificity.** The Zn²⁺ bimetallo site is a hallmark of a major branch of the AP superfamily and, as is generally the case for conserved features of superfamilies, is thought to play the same role across different reactions. For the AP Zn²⁺ bimetallo site, one Zn²⁺ presumably deprotonates and positions the nucleophilic serine or threonine, while the other Zn²⁺ stabilizes charge accumulation on the leaving group.

---

**Figure 3.** Comparative kinetics for wild-type and “common” PafA and NPP variants. (A) Schematics of the active residues in PafA (green) and NPP (magenta) that are removed to give mutants with common active-site residues. The residues removed interact with substrate oxygen O₂ in the PafA monoesterase reaction (Scheme 2) and the R’ group attached to this oxygen atom in the NPP diesterase reaction. (B) $k_{cat}/K_M$ values for WT PafA, WT NPP, and the common core mutants. Phosphomonooester reactions are shown by black bars (pNPP) and phosphodiester reactions (me-pNPP) by gray bars. Rate constants are from Table S. (C) The ratio of phosphomonooester to phosphodiester hydrolysis for WT PafA and WT NPP (gray) and for the common core mutants (black). The ratio of these ratios gives the relative specificity for the enzyme.
different scaffolds surrounding the conserved Zn$^{2+}$ ions and ligands could, in principle, adjust their positions or tune their charge densities to favor interactions with a phosphate monoester or diester transition state. There is no structural evidence to support such differences, but the differences could be small, transient, or only in electric field.

We therefore tested whether the PafA and NPP scaffolds, which differ in >80% of their residues (Supporting Information S5), provide functional specialization for each enzyme’s cognate reaction. We created “common” mutant forms of PafA and NPP with the disparate residues interacting with the O$_2$ atom in PafA and the R’ group in NPP removed but the rest of the surrounding PafA and NPP scaffolds maintained (Figure 3A). Remarkably, whereas WT PafA and NPP differ in their preference for reaction with phosphate mono- and diesters by >10$^{7}$-fold, the mutant versions catalyze both reactions equally well (Figure 3B,C; the 10$^{7}$-fold difference in preference was calculated as the ratio of the ratios of the specificity for pNPP over me-pNPP for each enzyme in Figure 3C; see also ref 67). Thus, despite having evolved to carry out different cognate reactions and despite very low sequence identity, PafA and NPP have essentially identical specificities after removal of only five characteristic local active-site residues—three from NPP and two from PafA (Figure 3A). While this represents the simplest and most likely model, the possibility of secondary conformational rearrangements of either enzyme that co-incidently lead to the same activity will require structural comparisons.

As the rate enhancements of these mutants are substantial (~10$^{31}$–10$^{32}$-fold) and rivaling enhancements observed for many fully evolved enzymes (Table 5), and as the >10$^{7}$-fold difference in native specificities appears to arise from interactions at a single phosphoryl position, these scaffolds seem to provide excellent jumping off points for evolving or engineering highly efficient and specific phosphate monoesters and diesterases.

Comparisons across the AP Superfamily Suggest Active-Site Roles for Specific Structural Elements. As emphasized in the Introduction, understanding enzyme function will require the identification of structural connections beyond the active site. Indeed, the notion that the enzyme scaffold positions binding and catalytic groups for catalysis is the fundamental precept of enzymology (e.g., refs 69–72). While the importance of elements beyond the active site is

---

### Table 5. Comparison of Rate Constants for WT and Mutants of PafA and NPP

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>$k_{cat}/K_{M}$ (M$^{-1}$ s$^{-1}$)</th>
<th>Rate Enhancement$^{ad}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>pNPP</td>
<td>me-pNPP</td>
</tr>
<tr>
<td>WT PafA</td>
<td>1.6(0.4) × 10$^9$</td>
<td>1.7(0.5) × 10$^9$</td>
</tr>
<tr>
<td>WT NPP$^{b}$</td>
<td>1.1</td>
<td>2.3</td>
</tr>
<tr>
<td>PafA K162A/R164A</td>
<td>16(6)</td>
<td>10(0.1)</td>
</tr>
<tr>
<td>NPP F91A/L123A/Y206A</td>
<td>1.6(0.1)</td>
<td>2.6(0.4)</td>
</tr>
</tbody>
</table>

$^{a}$Assay conditions: 100 mM sodium MOPS, pH 8.0, 500 mM NaCl, 100 μM ZnCl$_2$. $^{b}$Rate constants from ref 32. $^{c}$Calculated as the ratio of $(k_{cat}/K_{M})_{pNPP/(k_{cat}/K_{M})_{me-pNPP}}$. $^{d}$Rate enhancements are calculated as $(k_{cat}/K_{M})_{PafA}/(k_{cat}/K_{M})_{WT}$ using the water hydrolysis rate constants of $K_{W}^{pNPP}$ = 3.0 × 10$^{-11}$ M$^{-1}$ s$^{-1}$ and $K_{W}^{me-pNPP}$ = 3.3 × 10$^{-12}$ M$^{-1}$ s$^{-1}$. $^{e}$Calculated as the ratio of the ratios of the specificity for pNPP over me-pNPP for each enzyme in Figure 3C; see also ref 67. Table 5 shows the rate enhancements for the mutants compared to the WT.

---

**Figure 4.** Comparison of EcAP and PafA kinetics. (A) Values of $K_{M}$ for phosphonomonoester hydrolysis by WT PafA (black) and EcAP (gray) and inhibition constants ($K_i$) for binding the product, inorganic phosphate (P$_i$). Values are from Tables 2 and 4 and refs 38 and 56. (B) $k_{cat}/K_{M}$ values for phosphomoester (Me-P) hydrolysis by WT PafA (black) and EcAP (gray). Values are from Table 3 and ref 17.

---

**Figure 5.** Effect of PafA mutations on phosphate monoester and diester hydrolysis. (A) Ratio of $k_{cat}/K_{M}$ values for PafA mutants for phosphonomonoester (blue) and phosphodiester (yellow) reactions. Reactions are for Me-P and me-pNPP (Chart 1), and kinetic values are from Tables 2 and 3. (B) Ratio of the mutational effects on the phosphate monoester and diester reactions, with larger values representing larger effects on the monoester relative to diester reaction. The phosphoryl oxygen atom (Scheme 2) that each residue interacts with is noted and is based on the structural schematic of Figure 2.
clear, identifying their roles has been difficult. For example, most attempts to convert one enzyme to another start from highly related enzymes with a finite set of amino acid differences and have used trial-and-error strategies,\(^\text{73-77}\) and engineering efforts have generally not reproduced natural catalytic efficiencies.\(^\text{78-80}\) Here, comparisons across the AP superfamily allowed us to identify two structural elements that may make specific functional connections to the active site.

As emphasized above, PafA and NPP catalyze distinct reactions but have common interactions with phosphoryl oxygen atom O\(_1\) (Figures 2 and 3A). Strikingly, PafA and NPP share an inserted structural element that is not present in EcAP, which catalyzes the same reaction as PafA but has distinct O\(_2\) interactions. This insertion is situated between conserved sheets 6 and 7 in PafA and NPP and sits directly behind the common asparagine residue in their three-dimensional structures (Figure 1B and Supporting Information S7). Given this observation, we surveyed all 27 bimetallo AP superfamily members for which high-resolution structures are available. Of these, 10 contained a residue homologous to N100 of PafA and N111 of NPP, and in all cases these enzymes also contained an insertion at the position observed for PafA and NPP. Further, none of the enzymes lacking a corresponding asparagine contained this insertion (Supporting Information S8, Table 1). This correspondence suggests that the inserted structural elements play a common role in positioning the active-site asparagine for its interactions with an oxygen O\(_1\) of the transferred phosphoryl group (Figures 1 and 2).

There is also evidence for a helical element that may govern phosphate monoesterase activity. PafA and EcAP contain a helix (4a) that is inserted after \(\beta\) sheet 3 and is absent in NPP (Figure 1B; Supporting Information S9). Within or adjoining this helix are residues interacting with O\(_2\) in PafA (K162 and R164) and EcAP (D153) as well as the EcAP Mg\(^{2+}\) ligand T155 (Figures 1B and 2). Repeating the above structural analysis revealed that helix 4a is present in all structurally characterized AP superfamily members that catalyze reactions of monosubstituted phosphates and is absent in all members that catalyze reactions of phosphate diesters (Supporting Information S8, Table 1). Additional elements accompany the helix 4a insertion (e.g., helices 4a-e and \(\beta\) sheet 4a for PafA; helices 4a-g, and \(\beta\) sheet 4a-b for EcAP; Figure 1B) but are absent in the diesterases, perhaps because their role is to position helix 4a and the nearby active-site residues, akin to the structural elements behind the conserved asparagine of PafA and NPP described above. Thus, helix 4a and the remainder of this monoester-specific insert may help position active-site residues for interactions with O\(_2\) of monosubstituted phosphate substrates.

The absence of the 4a helix in diesterases could reflect the absence of selective pressure to maintain it, or it could be selected against to provide space for a bulky diester substituent on O\(_3\) and allow a more remote binding site to be sculpted for that substituent. Most generally, understanding whether active-site elements work cooperatively with adjacent structural elements or distinctly will help elucidate the evolutionary history of enzymes within and beyond the AP superfamily and may also suggest effective modular approaches to the problem of engineering new enzymes.

**Mechanistic Commonalities between EcAP and PafA Phosphate Monoesterases.** For catalysis to occur, an enzyme must stabilize a reaction’s transition state to a greater extent than the reaction’s ground state; otherwise, if both are stabilized to the same extent, the reaction barrier is the same on the enzyme as off of the enzyme so that there is no catalysis.\(^\text{81-85}\) In essence, such destabilization occurs in all enzymatic reactions, as fixation of reactants in the active site can be considered as entropically destabilizing, even though
quantification of this destabilization has proven elusive. Prior experiments with EcAP showed that P₂ binds orders of magnitude stronger to EcAP mutants when the anionic serine residue is removed. These and additional data provided strong evidence for electrostatic ground-state destabilization and a quantitative estimate of at least 10⁵-fold for this effect.⁵⁶,⁵⁷

Our results with PafA suggest that its active site harnesses analogous ground-state destabilization. Mutation of the nucleophilic threonine to serine decreases catalysis and weakens binding to vanadate and tungstate transition-state analogues⁶⁻⁶⁷ by ~100-fold (Figure 6B), similar to the ~20-fold decrease in activity (Figure 5) as expected for transition-state analogues (see also Discrimination between Transition-State Analogues in AP Superfamily Active Sites Suggests Precise Transition-State Recognition, below). But in contrast, this mutation increases inhibition by P₂ by >100-fold (Figure 6A, Table 4).

These effects are predicted from the ground-state electrostatic destabilization model shown in Figure 6C. The decreased activity and decreased transition-state analogue binding are predicted from loosening the positioning of the nucleophilic oxyanion, resulting from a loss of interactions with the methyl group of T79. The reaction and formation of covalent vanadate and tungstate adducts now requires overcoming an additional barrier for proper positioning and are correspondingly less favorable. In contrast, greater freedom of motion of the seryl oxyanion (Figure 6C, right) allows it to move away from anionic phosphate ligand, thereby lessening electrostatic repulsion and allowing the observed increase in P₂ binding. The 80-fold stronger binding corresponds to 2.2 kcal/mol of destabilization energy, providing a lower limit for the extent of ground-state destabilization, as some destabilization may remain in the mutant.⁸⁸

Interestingly, while EcAP and PafA both appear to utilize electrostatic ground-state destabilization to facilitate catalysis, the interactions that are responsible for positioning the oxyanion for nucleophilic attack and providing ground-state repulsion differ for the two enzymes. EcAP has a nucleophilic serine and so lacks the methyl group of PafA’s threonine residue that contributes to oxyanion positioning. Understanding the strategies used to position the threonyl and seryl nucleophiles will help uncover distinct evolutionary solutions to a common problem and may help define the types and extent of interactions needed to efficiently engineer new enzymes.

Another functional analogy between PafA and EcAP that involves distinct interactions was suggested from our mutagenesis results. Prior results revealed a ~10⁵-fold rate decrease from removal of the active-site Mg²⁺ of EcAP (via mutation of one of its ligands, E322; Figure 2)⁶⁸ and showed that this Mg²⁺ ion acts as part of a functional network with K328, D153, and a set of bound water molecules (Figure 7A, blue).¹⁷,²⁴ Such a large effect is rare for a group not directly involved in chemical catalysis (like the serine nucleophile) and implies important functional role of this interaction network. Thus, it was particularly noteworthy to observe a 10⁶-fold rate decrease in phosphate monoester catalysis upon mutation of K162 (Figure 5A), a residue that has no direct role in PafA’s reaction chemistry (Figure 2) and, similar to the observation of EcAP’s Mg²⁺ removal, had little effect on PafA’s diesterase activity (Figure 5A).

Intriguingly, although there is no structural homology or sequence identity in this region of PafA and EcAP, analogous interaction networks can be drawn for K162 of PafA and the Mg²⁺ ion of EcAP (Figure 7). In PafA (Figure 7B), K162 donates a hydrogen bond to the O₂ phosphoryl oxygen atom, like the Mg²⁺·OH₁ of EcAP (Figure 7A). K162 also hydrogen bonds to D38, the residue homologous to D51 of EcAP that is also a Mg²⁺ ligand. In addition, K162 hydrogen bonds to D305, the Zn₁ ligand that is homologous to D327 of EcAP.

Given the remarkable correspondence of active-site connections and similar large functional effects, the Mg²⁺ ion of EcAP and K162 of PafA may play analogous important functional roles and may represent distinct evolutionary solutions to analogous structural and functional challenges within AP superfamily active sites. Structural inspection suggests that these groups make connections within the active site that may simultaneously optimize the position of the two Zn²⁺ ions, the threonine or serine nucleophile, and the transferred phosphoryl group, via one of its non-bridging oxygen atoms (Figure 7). These observations may help define the type and degree of interconnections that are needed to engineer and re-engineer highly efficient enzymes.

Discrimination between Transition-State Analogues in AP Superfamily Active Sites Suggests Precise Transition-State Recognition. A recent high-resolution crystal structure of WT EcAP revealed that tungstate, like
vanadate, binds as a covalent pentavalent adduct, mimicking the reaction’s transition-state geometry. Extensive analysis of the catalytic properties of 20 EcAP variants relative to the binding of Pi, vanadate, and tungstate showed that tungstate binding mirrors activity much more closely than Pi binding and also more closely than vanadate binding. Thus, tungstate meets the most stringent criterion for acting as a transition-state analogue, although no stable analogue can perfectly mimic a transient, partially bonded transition state.

We also see a better correlation of activity and binding for tungstate for the five PafA variants studied herein (Figure 8). Tungstate gives a higher coefficient of determination (Pearson $R^2$) than vanadate and Pi and a best-fit slope closer to one (Table 6), as expected for a transition-state analogue, although no stable analogue can perfectly mimic a transient, partially bonded transition state.

We also see a better correlation of activity and binding for tungstate for the five PafA variants studied herein (Figure 8). Tungstate gives a higher coefficient of determination (Pearson $R^2$) than vanadate and Pi and a best-fit slope closer to one (Table 6), as expected for a transition-state analogue, although no stable analogue can perfectly mimic a transient, partially bonded transition state.

We noted previously that a subset of EcAP variants—mutants that removed the active-site Mg$^{2+}$ ion addressed in the preceding section—were predominantly responsible for the poorer correlation of binding and activity for vanadate than tungstate (Figure 8, red squares). Intriguingly, for PafA, the K162 mutant deviated substantially with vanadate but fell closer to the line with tungstate (Figure 8, red circles, the tungstate value is a limit). This result extends the analogy between K162 of PafA and the active-site Mg$^{2+}$ ion of EcAP (Figure 7) and suggests a high degree of precision to discriminate between tungstate and vanadate species within the active site of these extremely efficient enzymes.

<table>
<thead>
<tr>
<th>EcAP</th>
<th>PafA</th>
<th>EcAP</th>
<th>PafA</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>WT</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>D01A</td>
<td>D179S</td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td>D53A</td>
<td>D53A</td>
<td>0.97</td>
<td>0.97</td>
</tr>
<tr>
<td>R166S</td>
<td>R166S</td>
<td>1.3</td>
<td>1.3</td>
</tr>
<tr>
<td>D01A/D179S</td>
<td>1.1</td>
<td>1.1</td>
<td></td>
</tr>
<tr>
<td>D01A/R166S/D53A</td>
<td>1.2</td>
<td>1.2</td>
<td></td>
</tr>
<tr>
<td>E32Y</td>
<td>E32Y</td>
<td>0.35</td>
<td>0.35</td>
</tr>
<tr>
<td>K328A</td>
<td>K328A</td>
<td>0.41</td>
<td>0.41</td>
</tr>
<tr>
<td>D01A/D53A/R166S</td>
<td>0.75</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>D179S</td>
<td>D179S</td>
<td>0.66</td>
<td>0.66</td>
</tr>
</tbody>
</table>

Figure 8. Correlation of activity and affinity for ground- and transition-state analogues. Correlations of activity values ($k_{cat}/K_M$ measured with Me-P normalized to WT PafA, and $k_{cat}/K_M$ measured with pNPP normalized to WT EcAP) with inhibition ($K_i = 1/K_a$) by inorganic phosphate (Pi), tungstate, and vanadate. Data for PafA (circles) from Table 3 ($[k_{cat}/K_M]_{norm}$) and Table 4; data for EcAP (light squares) from ref 87. The solid lines are fits to the PafA data with fixed slopes of 1, as expected for a transition-state analogue; the rmsd given is for deviation from these lines; and Pearson’s $R^2$ values are also listed. Statistical tests were carried out as described by Peck et al., and additional statistical parameters for PafA and EcAP are given in Table 6. Values at the vertical dashed lines marked with arrows represent lower limits for affinity for AP variants and were not included in the fits. Similarly PafA points that are limits are denoted with an arrow and were excluded from fitting of the data. The PafA mutants are color-coded as noted in the attached legend. The EcAP mutants are specified and color-coded as follows: WT, white; mutants with the Mg$^{2+}$ ion removed (E322Y or E322A), red; all other mutants, blue; these data represent combinations of mutations at five active-site residues, and the individual data points can be found in ref 87.
GENERAL IMPLICATIONS

Over the past decades, there has been an enormous amount learned about how enzymes catalyze reactions, in terms of the underlying chemical mechanisms, the residues and cofactors that facilitate those reactions, and the overall structural context in which these reactions take place. Nevertheless, there is also an enormous amount that we do not yet understand, as exemplified by the limited ability to engineer new enzymes and the prevalence of trial-and-error approaches.

Given that we can identify the residues directly involved in binding and catalysis, the next stage of understanding will entail deciphering how protein scaffolds and interaction networks assemble and position active-site residues for catalysis. Enzyme superfamilies provide sets of enzymes with functional distinctions in similar overall structural contexts, and catalytic promiscuity provides a means to probe the functional consequences of the structural elements and sequence features that differ between superfamily members. Thus, superfamily members provide sufficient “contrast” to make meaningful comparisons, but sufficient similarities to allow inferences to be drawn from their juxtaposition, and catalytic promiscuity allows systematic and multi-dimensional functional analysis across superfamily members and properties.

Propitiously, enzymes of the AP superfamily have been highly amenable to comparative approaches (e.g., refs 32, 35, 93, and 94). We have taken advantage of the occurrence of distinct AP superfamily phosphate monoesterases, one of which has features similar to the superfamily diesterases that are absent in another monoesterase (Figures 1 and 2). Multi-faceted comparisons of structure, sequence, and cognate and promiscuous reactions of these superfamily members and their mutants in this work and previously have extended our systematic and multi-dimensional functional analysis across superfamilies of cognate and promiscuous reactions, the next stage of understanding will entail the robust engineering of novel, highly effective enzymes.

ASSOCIATED CONTENT

Supporting Information
The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.6b06186.

S1, PafA gene sequence; S2, pH dependencies for PafA WT and mutants; S3, inhibition curves for PafA WT and mutants; S4, determination of Me-P kinetics at pH 8; S5–S9, sequence alignments and structural comparisons in the AP superfamily (PDF)
Crystallographic data for PafA_refine_reflections (CIF)
Crystallographic data for PafA_refine_99 (CIF)

AUTHOR INFORMATION

*Corresponding Author

herschla@stanford.edu

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was funded by a grant from the U.S. National Institutes of Health to D.H. (GM49243); partial support for J.B., C.L.B., and Z.S. was provided by National Science Foundation grant (MCB-1121778). J.B. was an Amgen Scholar in the Stanford Summer Research Program, and C.B. was a Stanford ADVANCE Institute fellow. The SSRL Structural Molecular Biology Program is supported by the Department of Energy, Office of Biological and Environmental Research, and by the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program, and the National Institute of General Medical Sciences. We thank Emily Burnette for technical assistance, Axel Brunger for support, and Monarin Uervirojnangkoorn in the Brunger laboratory for discussion and advice on the data processing and structure refinement. We thank Patricia Babbitt, and Gemma Holliday and members of the Babbitt laboratory for discussions on phylogeny. We thank Jonathan Lassila, Craig Markin, and Ariana Peck for discussions about structural and energetic aspects of this work. We thank members of the Herschlag laboratory for discussions and comments on the manuscript.

REFERENCES

(3) Horovitz, A. Folding Des. 1996, 1, R121.
Closely resemble naturally occurring phosphate diesters and with evolution, and their extreme diversity prevents establishment of a high-confidence rooted evolutionary tree. Without discovery of additional relevant sequences and determination of more three-dimensional structures to guide the multiple alignments required, we cannot determine whether the PaFX- and NPP-like AP superfamily members originated from a monoesterase, diesterase, or generalist-like ancestor.

Me-P reactions are compared rather than reactions of pNPP, as the unactivated leaving group is more likely to mimic physiological substrates.

Members of the AP superfamily branched very early in evolution, and their extreme diversity prevents establishment of a high-confidence rooted evolutionary tree. Without discovery of additional relevant sequences and determination of more three-dimensional structures to guide the multiple alignments required, we cannot determine whether the PaFX- and NPP-like AP superfamily members originated from a monoesterase, diesterase, or generalist-like ancestor.

Members of the AP superfamily branched very early in evolution, and their extreme diversity prevents establishment of a high-confidence rooted evolutionary tree. Without discovery of additional relevant sequences and determination of more three-dimensional structures to guide the multiple alignments required, we cannot determine whether the PaFX- and NPP-like AP superfamily members originated from a monoesterase, diesterase, or generalist-like ancestor.

Members of the AP superfamily branched very early in evolution, and their extreme diversity prevents establishment of a high-confidence rooted evolutionary tree. Without discovery of additional relevant sequences and determination of more three-dimensional structures to guide the multiple alignments required, we cannot determine whether the PaFX- and NPP-like AP superfamily members originated from a monoesterase, diesterase, or generalist-like ancestor.

Members of the AP superfamily branched very early in evolution, and their extreme diversity prevents establishment of a high-confidence rooted evolutionary tree. Without discovery of additional relevant sequences and determination of more three-dimensional structures to guide the multiple alignments required, we cannot determine whether the PaFX- and NPP-like AP superfamily members originated from a monoesterase, diesterase, or generalist-like ancestor.
For EcAP, it has been shown that Pi binds in its fully deprotonated trianionic form, with the simultaneous uptake of a proton to neutralize the serine nucleophile (ref 56). This same binding presumably occurs with PafA. The stronger binding to the PafA T79S mutant is consistent with greater freedom of motion relieving remaining repulsion with the lone pair electrons of this oxygen atom at this position and/or a higher proton affinity from movement away from the Zn2 ion.

References:

(88) For EcAP, it has been shown that Pi binds in its fully deprotonated trianionic form, with the simultaneous uptake of a proton to neutralize the serine nucleophile (ref 56). This same binding presumably occurs with PafA. The stronger binding to the PafA T79S mutant is consistent with greater freedom of motion relieving remaining repulsion with the lone pair electrons of this oxygen atom at this position and/or a higher proton affinity from movement away from the Zn2 ion.