

Differential Catalytic Promiscuity of the Alkaline Phosphatase Superfamily Bimetallo Core Reveals Mechanistic Features Underlying Enzyme Evolution

Fanny Sunden¹, Ishraq AlSadhan¹, Artem Lyubimov^{2,3,4,5,6}, Tzanko Doukov⁷, Jeffrey Swan¹ and Daniel Herschlag^{1,8*}

¹Department of Biochemistry, Beckman Center, Stanford University, Stanford, CA 94305 USA; ²Departments of Molecular and Cellular Physiology; ³Neurology and Neurological Science; ⁴Structural Biology; ⁵Photon Science; ⁶Howard Hughes Medical Institute, Stanford University, Stanford, CA 94305, USA;

⁷Macromolecular Crystallographic Group, The Stanford Synchrotron Radiation Lightsource, National Accelerator Laboratory, Stanford University, Stanford, CA 94309; ⁸Departments of Chemical Engineering and Chemistry, Stanford University, Stanford, California 94305, USA; ⁹Stanford ChEM-H (Chemistry, Engineering, and Medicine for Human Health), Stanford University, Stanford, California 94305, USA

*Corresponding author: herschla@stanford.edu

ABSTRACT

Members of enzyme superfamilies specialize in different reactions but often exhibit catalytic promiscuity for one another's reactions, consistent with catalytic promiscuity as an important driver in the evolution of new enzymes. Wanting to understand how catalytic promiscuity and other factors may influence evolution across a superfamily, we turned to the well-studied Alkaline Phosphatase (AP) superfamily, comparing three of its members—two evolutionarily distinct phosphatases and a phosphodiesterase. We mutated distinguishing active-site residues to generate enzymes that had a common Zn^{2+} bimetallo core, but little sequence similarity and different auxiliary domains. We then tested the catalytic capabilities of these pruned enzymes with a series of substrates. A substantial rate enhancement of $\sim 10^{11}$ -fold for both phosphate mono- and diester hydrolysis by each enzyme indicated that the Zn^{2+} bimetallo core is an effective mono/di-esterase generalist and that the bimetallo cores were not evolutionarily tuned to prefer their cognate reactions. In contrast, our pruned enzymes were ineffective sulfatases, and this limited promiscuity may have provided a driving force for founding the distinct one-metal-ion branch that contains all known AP superfamily sulfatases. Finally, our pruned enzymes exhibited 10^7 - 10^8 fold phosphotriesterase rate enhancements, despite absence of such enzymes within the AP superfamily. We speculate that the superfamily active site architecture involved in nucleophile positioning prevents accommodation of the additional triester substituent. Overall, we suggest that catalytic promiscuity and the ease or difficulty of remodeling and building onto existing protein scaffolds have greatly influenced the course of enzyme evolution and will provide lessons for engineering new enzymes.

INTRODUCTION

Enzymes catalyze biological reactions with enormous rate enhancements—up to 10^{29} -fold—and do so with high specificity, thereby ensuring rapid and controlled flux of

substrates through metabolic pathways and permitting exquisite biological regulation (1,2). But how did these remarkably efficient enzymes arise? What are the mechanistic and structural properties that determine the catalytic proficiencies of a family of proteins and thus the catalytic landscape that evolution has traversed?

Catalytic promiscuity, the ability of an enzyme to carry out reactions other than its evolutionary optimized reaction, is expected to strongly bias the emergence of new enzymes, as it can provide an evolutionary head start following gene duplication (3,4). Correspondingly, the catalytic promiscuity of modern day enzymes and their conserved elements may be evolutionary vestiges and may provide clues to catalytic capabilities and mechanistic features that have influenced evolutionary history and insights for the engineering of new enzymes (4-13).

We can identify enzymes that are members of superfamilies—i.e., enzymes with similar overall folds and common underlying catalytic features—and these enzymes very likely emerged from a common evolutionary precursor (3,8,10,11,14,15). Indeed, extant superfamily members that catalyze different reactions often exhibit cross promiscuity, consistent with an evolutionary role for promiscuity (3,11,16). In some cases, for enzymes that have evolved subsequent to the last universal common ancestor (LUCA) of current cellular life, reconstructions can provide information about potential evolutionary pathways (17-21), whereas in other cases reconstructions are stymied as superfamily members were already present in LUCA or emerged very early after the common ancestor (22) (P.C. Babbitt, personal communication).

In addition to our ability to create models for the sequences of evolutionary precursors (17-21), we would like to know, from a mechanistic standpoint, why certain enzymes have evolved from one another, and conversely, why others did not. As noted

above, catalytic promiscuity is one factor that has likely played an important role, and we gain new insights into evolutionary pathways from the determinations of catalytic promiscuities herein. Nevertheless, catalytic promiscuity does not ensure the development and fixation of a new activity, even if it can provide a selective advantage; in addition to the inevitable role of chance in evolution, mutational paths, to be followed, need to avoid major sustained fitness dips that would prevent or greatly hinder optimization.

Thus, information about the physical and mechanistic features that determine catalytic efficiency and catalytic promiscuity will help us understand how enzymes have emerged throughout evolution and the evolutionary potential and evolvability of enzymes. Ultimately these insights may be harnessed to engineer new, beneficial enzymes ((4-13,23,24), and references herein (23) e.g. (25-27)).

We have taken advantage of the well-established functional, structural, and evolutionary connections within the Alkaline Phosphatase (AP) superfamily to relate the catalytic promiscuity of the superfamily's Zn^{2+} bimetallo core to the elaboration of enzyme classes across this superfamily (6,28-37). To ensure that our results would be representative rather than idiosyncratic we compared three Alkaline Phosphatase (AP) superfamily members, mutating the distinguishing active site residues to create mutant enzymes with a common Zn^{2+} bimetallo core but pervasive sequence differences and distinct auxiliary domains. Determination of the catalysis of these three 'pruned' enzymes across a series of reaction classes and consideration of catalytic and structural features allowed us to relate intrinsic catalytic prowess of the Zn^{2+} bimetallo core to evolutionary choices made across this superfamily.

RESULTS

The Bimetallo Branch of the AP Superfamily

The AP superfamily subdivides into two main subgroups, one with enzymes containing a single metal ion and

predominantly catalyzing sulfatase reactions and the other with enzymes containing the canonical alkaline phosphatase bimetallo center (38). We focused on the latter group, which possesses two metal ions coordinated by conserved residues in the active site and a universally conserved serine or threonine nucleophile (Fig. 1). This group includes phosphatases, phosphate diesterases, phosphomutases, among other enzymes (Chart 1)(30-34,39). The enzymes studied here typify a cross-section of the structural diversity seen across the AP superfamily bimetallo branch while maintaining Zn^{2+} ions in their bimetallo cores, which facilitates the comparisons herein (Fig. 2). *EcAP* and *Chryseobacterium meningosepticum* Alkaline Phosphatase PafA (PafA) are both phosphate monoesterases (Table 1A), but are evolutionarily distant as suggested by a low sequence identity (8.9%; Supplemental Fig. 1, (40)) and distinct inserted elements (Fig. 2A & B)(34,40). *Xanthomonas citri* Nucleotide Pyrophosphatase/Phosphodiesterase (NPP) has a subset of active site and scaffold features in common with PafA but is evolutionary distant from PafA (10% sequence identity; Supplemental Fig. 1, (40)) and is a phosphate diesterase rather than a monoesterase (Table 1A), related to a large set of AP superfamily diesterases (Fig. 2A & B)(34,40-42).

The active site residues common to these enzymes reside in analogous positions within conserved secondary structure elements (Fig. 2A, white & gray), and these secondary structure elements comprise highly similar tertiary structures (Fig. 2B, gray)(30,31,33,34,38-40,43). All three enzymes have a serine or threonine nucleophile, two Zn^{2+} ions, and identical Zn^{2+} ligands (Figs. 2A & C, black). Mechanistic similarities are also readily apparent. The conserved seryl/threonyl residue is a ligand of Zn_2 and serves as a nucleophile that attacks the phosphorus atom of the phosphate monoester or diester substrate; the leaving group oxygen atom, which develops negative charge in the transition state, is stabilized by Zn_1 ; and one of the oxygen atoms of the transferred phosphoryl group situates between the Zn^{2+}

ions, presumably making close electrostatic interactions that stabilize the trigonal bipyramidal transition state (Fig. 2D)(44).

These structural and mechanistic similarities contrast with local and global differences (Fig. 2). Figure 2C shows the residues that sit opposite to the bimetallo center (colored residues). The two monoesterases (*EcAP* and *PafA*) donate multiple hydrogen bonds to each of the two oxygen atoms on the transferred phosphoryl group that face away from the Zn^{2+} bimetallo site, interactions that are unique to each enzyme and suggestive of evolution along distinct pathways within the AP superfamily (40,41). The diesterase (NPP) makes interactions with the nonesterified outward-facing phosphoryl oxygen atom that are the same as for the *PafA* monoesterase, apparently reflecting a shared evolutionary origin (Fig. 2C: NPP, N111 (magenta); *PafA*, N100 (green), (37,40,41) whereas the other outward facing phosphoryl oxygen makes different interactions, in its esterified form with an NPP hydrophobic pocket and in its nonesterified form with a *PafA* hydrogen bond network (Fig. 2B & 2C: NPP, magenta; *PafA*, green) (34,37,45).

Mechanistic studies have shown, as expected, that mutations of groups making specialized interactions with a phosphoryl oxygen atom (in the monoesterases) or with the ester group attached to the transferred phosphoryl group (in the diesterase) preferentially lower activity of the cognate reaction and thereby decrease specificity (6,36,40,45). In addition to the local active site differences and very low sequence conservation within their structurally conserved cores (sequence identities: 6.0% for *EcAP*/NPP, 8.9% for *EcAP*/*PafA*, and 10% for NPP/*PafA* for the structurally conserved cores; Supplemental Fig. 1), these enzymes contain different sets of added helices and sheets (Fig. 2A & 2B, color coded regions)(40).

The Phosphate Monoesterase and Diesterase Activities of the Zn^{2+} Bimetallo Scaffold of AP

Superfamily Members. We first asked how effective the AP superfamily Zn^{2+} bimetallo center is in catalyzing phosphate monoester and diester hydrolysis when embedded in each of the three superfamily scaffolds after truncation of the active site side chains particular to each active site (Fig. 2C, colored residues interacting with the substrate non-bridging oxygen atoms facing away from the Zn^{2+} bimetallo center). It has been suggested that the scaffold tunes the Zn^{2+} bimetallo site to favor monoesterase over diesterase activity (34), and the ample sequence differences throughout these scaffolds could cause these or other catalytic differences. Further, while it is reasonable that the scaffolds make the same functional interactions in phosphate monoester and diester transition states, the charges of reactants and transition state structures for reactions vary, so there is no reason *a priori* to expect the same strength for each catalytic interaction and same catalytic contribution; in other words, we cannot predict the relative rate constants for reactions with different substrates, even ones making the same atom-to-atom interactions.

We created “pruned” versions of each enzyme by subtractive mutagenesis, denoted throughout with asterisks (i.e., *EcAP**, NPP*, *PafA**), by mutating the side chains directly contacting and within 4 Å of the substrate non-bridging oxygen atoms and, for NPP, side chains that interact with the methyl substituent of me-pNPP used herein (Fig. 2C, color-coded side chains; Chart 2)(45). Subtractive mutagenesis was used in all cases except for mutations of the nucleophile, T79S and T90S for *PafA* and NPP, respectively, so that all three pruned enzymes contained the same nucleophile. The pruned constructs had mutations as follows: *EcAP**: D101A, R166S, D153A, E322A, K328A; NPP*: T90S (nucleophile), F91A, N111A, L123A, Y205A; and *PafA**: T79S (nucleophile), N100A, K162A, R164A (Table 1B). Prior studies showed that the only non-alanine mutant, R166S, gives similar kinetics as the corresponding alanine mutation (36), and for each enzyme, we created at least one additional mutated version to control for

unintended consequences from the specific mutations made. For E322, mutation to tyrosine was employed as prior studies showed equivalent effects for it and alanine and as the homologous residue in NPP is a tyrosine (Y205)(6,46). The ratio between the rate constants for monoester and diester hydrolysis of each enzyme set fell within 3-fold of one another (Table 2), providing evidence against unintended or idiosyncratic effects.

We expected, most simply, that the bimetallo cores would not rearrange in the mutated enzymes, as numerous *EcAP* variants with mutations in these active site residues exhibit no measureable change in the bimetallo center or in more remote parts of the structure (6,36,47-50). Nevertheless, as we mutated four or five side chains simultaneously in each active site, we tested for structural rearrangements.

For each enzyme, its wild-type (WT) and pruned form gave CD spectra that were identical, within error (Supplemental Fig. 2). To obtain a more detailed view of possible structural rearrangements, crystals of PafA* and *EcAP** were obtained and their structures were determined to 2.03 and 2.45 Å resolution, respectively (Tables 3 and 4, Supplemental Fig. 3). The overall structures of WT *EcAP* and *EcAP** were superimposed with a backbone RMSD of 0.42 Å over 801 atoms (Fig. 3A), and WT PafA and PafA* were superimposed with a backbone RMSD of 0.18 Å over 449 atoms (Fig. 3B). The conserved active site Zn²⁺ ions and ligands were identical for the two sets of wild-type and pruned enzymes (Fig. 3C), with the only exception being the serine nucleophile of PafA* (T79S), which was not coordinated to Zn₂ and appeared to be interacting with another ion in the active site (Supplemental Fig. 3). This difference presumably arose because the PafA* crystals were grown at pH 4.3, a pH at which the enzyme is not active and the serine residue is likely neutral (Fig. 3C; data not shown; see also ref (51)).

Figure 4A compares the phosphate monoesterase activity for the WT and pruned versions of each enzyme (gray and black bars, respectively). As expected, the pruned monoesterases (*EcAP** & PafA*) are more diminished in monoester activity than the pruned diesterase (NPP*). Intriguingly, the monoesterase activities of the three pruned enzymes were within 5-fold.

Figure 4B shows the corresponding comparisons for phosphate diesterase activity. As expected, converse effects on the two reactions are observed—i.e., the diesterase pruned construct (NPP*) gave the largest effect. Remarkably, as for the monoesterase reaction, all three enzymes had very similar diesterase activity, again within a 5-fold range.

The phosphate monoesterase and diesterase results are summarized together in Figure 4C as the ratios of these activities. There is an enormous difference in the substrate preferences of the WT enzymes of ~12 orders of magnitude, and the intrinsic discrimination is greater still (i.e., the chemical step is not rate-limiting in all cases)(Fig. 4C, gray bars)(40,51). In contrast, for the pruned constructs the monoesterase:diesterase ratios are uniform, all within 2-fold (Fig. 4C, black bars (near axis)). Further, the pruned enzymes provide substantial rate enhancements of ~10¹⁰–10¹¹ fold rate for the two reactions, indicating that catalysis from the bimetallo scaffold rivals or exceeds that of many fully evolved enzymes (Fig. 5)(2-4).

The similarity of the catalytic ratios for pruned enzymes derived from two different monoesterases and a diesterase strongly suggests that the bimetallo site is *functionally* conserved across the AP superfamily, despite the very low sequence identity of the surrounding scaffold and despite the different additional structural elements (Fig. 2A & B)(33,39,40). Thus, specialization for the individual cognate reactions arises from the residues that have been pruned from each construct, and evolution has not independently tuned the

common Zn^{2+} bimetallo center to favor phosphate monoester or diester hydrolysis. Most pertinent for understanding catalysis and evolution in the AP superfamily, the Zn^{2+} bimetallo center appears to be a potent and general enzyme, as elaborated in the next section and in the Discussion.

Enzymatic Activity of the Zn^{2+} Bimetallo Scaffold as a Potential Evolutionary Driver Across the AP Superfamily. Interestingly, the monoesterase:diesterase ratios for the pruned enzymes were not just similar, but were also near unity (Fig. 4C; Tables 1B and 2). To test whether this similarity was deep-seated or coincidental due to the particular phosphate ester leaving group used, we carried out analogous measurements with phosphate mono- and diesters across a range of intrinsic reactivities, using phosphate mono- and diesters with a 3-nitrophenoxide or 3,4-dinitrophenoxide instead of *p*-nitrophenoxide leaving group and thereby spanning a range of leaving group pK_a values of three units and reactivity of $\sim 10^3$ fold (52-54) (& unpub results). Monoesterase:diesterase activity ratios near unity were observed for these substrates as well (Supplemental Table 1). We do not know if this identity represents an equivalence of individual catalytic contributions for the different reactions or differential individual effects that coincidentally sum to the same value.

The similarity of the monoesterase and diesterase activities of the AP superfamily bimetallo scaffold suggests a model in which this promiscuous scaffold provided a strong physical and functional foundation that Nature built upon to make distinct, more efficient, and specialized monoesterases and diesterases. Physically, two of the three phosphoryl oxygen atoms face away from the bimetallo center—these atoms are accessible to the local active site and serve as functional handles providing evolutionary opportunities to sculpt new interactions, or to co-opt existing residues for new purposes, along the pathway from a generalist to specialist or from a monoesterase to diesterase (or vice versa). Functionally, the substantial $\sim 10^{10}$ – 10^{11} fold rate enhancements

from the bimetallo scaffold would provide a considerable head start toward the evolution of fully optimized monoesterases and diesterases or allow significant catalysis while progressing from one activity to the other (Fig. 5)(2-4). The observation of different interactions in the *EcAP* and PafA monoesterase active sites (Fig. 2C, colored residues) indicates an evolutionary divergence, perhaps to provide distinct catalytic properties best suited to growth conditions of each parent organism (40,42,55).

Weak Sulfatase Activity of the Zn^{2+} Bimetallo Scaffold May Have Restricted Evolutionary Diversification within this AP Superfamily Branch. Highly promiscuous phosphate monoesterase and diesterase activities of the ancient Zn^{2+} bimetallo core may have promoted the evolution of AP-superfamily phosphate monoesterases and diesterases. Conversely, activities *not* found in extant AP superfamily bimetallo enzymes may be absent because they are not efficiently catalyzed by this motif. The AP superfamily sulfatases provide a particularly incisive test of this prediction, as the sulfatase and phosphatase substrates and transition states are highly analogous geometrically (56,57), but sulfatases are not found in the bimetallo branch, falling instead solely in the one metal-ion active site subgroup (38,43,58-60).

The sulfatase activities of our bimetallo scaffolds were extremely low, so low that only upper limits could be determined (Fig. 5, pNPS; Table 1B). Our most sensitive measurements revealed that the bimetallo scaffold is at least 10^5 fold less effective as a sulfatase than as a phosphate monoesterase or diesterase. Furthermore, the very low intrinsic reactivity of sulfate esters necessitates highly efficient catalysis (16,61). Thus, our results support a mechanistic mandate for the absence of bimetallo cores among sulfatases. Intriguingly, the AP superfamily appears to have overcome this limitation of bimetallo scaffold catalysis by utilizing only one metal ion for sulfatases, instead of two. While the ancient origin of the AP superfamily members prevents us from tracing possible ancestral

sulfatases and possible pathways for this conversion (P.C. Babbitt, personal communication), our results reveal a fundamental underlying mechanistic property that Nature faced as it navigated these evolutionary processes.

Phosphotriesterase Activity of the Zn²⁺ Bimetallo Scaffold? As all known cognate phosphate triesterases exist outside of the AP superfamily (62), we wondered if the AP superfamily bimetallo scaffold is also less efficient as a triesterase. Indeed, the triesterase rate enhancements were $\sim 10^3$ – 10^4 fold lower than those for reactions of phosphate diesters and monoesters (Fig. 5, paraoxon; Table 1B, Supplemental Fig. 4). The lower catalysis could arise from differences in substrate charge or transition state charge distribution, due to the tighter triester transition state structures, although such differences did not manifest for the mono- and di-esters despite their charge and transition state differences (44).

The bimetallo scaffold is at least 2–3 orders of magnitude more effective at catalyzing the phosphate triester reaction than sulfatase reaction, and the greater intrinsic reactivity of phosphate triesters than sulfate esters (63) would render it more probable for Nature to be able to take advantage of this catalytic promiscuity. Nevertheless, there is an apparent absence of an AP superfamily triesterase, and this may arise from the lower catalytic effectiveness of the bimetallo center for this reaction relative to phosphate mono- and diester reactions. However, it may also, or instead, arise from the rarity of phosphate triesters in biology and/or from factors intrinsic to the AP superfamily structural scaffold that render it difficult to sculpt binding sites for both substituents of the transferred phosphoryl group, as elaborated in the Discussion.

DISCUSSION

The probability of *de novo* evolution of proteins that adopt specific structures is extremely low, and the probability that such proteins exhibit substantial and specific

enzymatic activities is lower still. But once a structural motif evolves, gene duplication allows it to be put to multiple uses, and promiscuous catalytic activity greatly increases the probability that evolution will be able to find, utilize, and optimize a potentially beneficial activity (3-5,64). These factors have presumably led to the preponderance of enzymes with shared folds and shared catalytic features (3,10,11).

We have posed two fundamental questions at the interface of catalytic function and evolution, using the Zn²⁺ bimetallo scaffold of the AP superfamily: (1) Does the superfamily bimetallo core intrinsically favor certain reactions over others and, if so, are these preferences mirrored in evolutionary history? (2) Did evolution, subsequent to divergence in the AP superfamily, “tune” individual bimetallo cores and scaffolds to favor cognate over noncognate reactions?

We addressed the second question by determining the activity of three AP superfamily Zn²⁺ bimetallo scaffolds, absent of the specialized interactions specific to each subset of superfamily members. The AP superfamily members studied herein—*E. coli* Alkaline Phosphatase (*EcAP*), an evolutionarily distinct alkaline phosphatase from *Chryseobacterium meningosepticum* (*PafA*), and a phosphate diesterase from *Xanthomonas citri* (*NPP*)—have very low overall sequence identity and distinct architectural differences within their common bimetallo scaffolds (Fig. 2A & B, gray and white; Supplemental Fig. 1). Nevertheless, these enzymes, with a small set of divergent active site side chains pruned away (Fig. 2C, color-coded), have highly similar phosphate monoesterase and phosphate diesterase activities, each within a 5-fold range (Fig. 4A & 4B; Fig. 5). This observation is consistent with the common Zn²⁺–Zn²⁺ distance observed for AP superfamily members that catalyze hydrolysis of phosphate mono- and diesters (65), and these similarities suggest that electrostatic or geometric effects rooted in the Zn²⁺ bimetallo site did not specialize the active site Zn²⁺ bimetallo center for cognate over

noncognate reactions. Rather, active site features other than the bimetallo core are responsible for monoester versus diester specialization. As illustrated in Figs. 2 and 6, the bimetallo core provides a scaffold for these specializing auxiliary domains and side chains (see also Ref. (40,45,46)).

With this information in hand, we turn to the broad question of the interplay of catalytic promiscuity and the selection and optimization of enzymes through evolution. We found that the rate enhancements of our pruned enzymes for both phosphate monoester and diester hydrolysis are substantial, 10^{10} – 10^{11} fold (Fig. 5), rivaling or exceeding the catalytic power of many natural enzymes (2). Thus, an AP superfamily scaffold with the Zn^{2+} bimetallo center alone is a highly effective catalyst that could have provided an effective starting point for the evolution of the extant enzymes that specialize in each of these reactions, as depicted by the central generalist “bimetallo core” in Fig. 6 (see also simplified version of Fig. 6 in Supplemental Fig. 6). While the ancient origin of the Alkaline Phosphatase superfamily prevents tracing its evolutionary history, the generalist ability of the bimetallo core seems likely to have contributed to its use in both phosphate mono- and di-esterases, whether these activities arose from a common generalist or from one another.

In contrast to this phosphate mono-/di-esterase generalist activity, the pruned enzymes are many orders of magnitude less effective in catalyzing reactions of sulfate esters (Fig. 5). Intriguingly, sulfatase reactions are catalyzed exclusively by enzymes on a distinct one-metal ion branch of the AP superfamily (38,43,58-60). These observations suggest a mechanistic imperative for the absence of AP superfamily bimetallo sulfatasases and a functional driving force for the formation of a distinct sulfatase superfamily branch with just one divalent metal ion (Fig. 6, orange) (59). While the mechanistic origin of the inability of the bimetallo centers to catalyze sulfuryl transfer remains to be determined, linear free energy

comparisons of *EcAP* reactions reveal a strong dependence of catalytic power on substrate charge (56). Thus, to achieve sulfatase activity, the Alkaline Phosphatase superfamily may have had to undergo a more involved remodeling of its core, in addition to taking on auxiliary domains (38). We speculate that this extensive remodeling might have been evolutionary allowed if metal ligand mutations led to loss of the second metal ion and use of one or more of the freed ligands to aid sulfatase catalysis. Alternatively, the sulfatasases may have arisen from a highly evolved AP superfamily phosphatase, using the catalytic interactions beyond the bimetallo core to provide modest and selectable sulfate ester catalysis while the second metal site was repurpose; it is also possible that the phosphatasases arose from the addition of a second metal ion site to a sulfatase precursor.

We also determined the promiscuous phosphotriesterase activity of the AP superfamily bimetallo scaffold, a biological reaction that is catalyzed by members of the Amidohydrolase superfamily (66-68), but not known to be catalyzed within the AP superfamily. Catalysis by the pruned AP superfamily enzymes is substantial, and considerably greater than that for a sulfate ester, although still $\sim 10^3$ fold less than that for phosphate monoesters and diesters (Fig. 5).

So why did Nature chose structural and catalytic motifs other than those of the AP superfamily to catalyze phosphate triester reactions (62,66,69)? The simplest models are probabilistic, especially given the small number of phosphotriesterases found in Nature ((70)). When the selective advantage for a phosphotriesterase arose, random duplications may have occurred for members of the Amidohydrolase superfamily or duplications in that family may have had higher promiscuity and thus been favored early in the evolution of these new enzymes. However, there is another possibility, rooted in the structural architecture of the AP superfamily. Below we describe how the architecture has been used and remodeled through the evolution of phosphate mono- and diesterases,

and we then return to how architectural features may have limited the ability to evolve an AP superfamily triesterase.

Remarkably, the monoesterase PafA and diesterase NPP share a common auxiliary domain that is distinct from any in *EcAP*, despite PafA and *EcAP* both functioning as monoesterases (Fig. 6, Asn auxiliary domain, magenta) ((40,41)). This element is inserted in a common position in the bimetallo core and is structurally homologous (Fig. 2A). Its role appears to be to position the asparagine residue (Fig. 2C, N111 and N100 in NPP and PafA, respectively) that interacts with one of the phosphoryl oxygen atoms that faces away from the bimetallo core (Fig. 2D, O₁). This element provides an equivalent mono- and diesterase enhancement, consistent with interaction with an unesterified oxygen atom for mono- and diester substrates (40). Architecturally, there is enough room and there are sufficient interactions to build this auxiliary domain onto the bimetallo core to enhance catalysis.

How, then, are monoester and diester substrates distinguished? As highlighted in Fig. 6, PafA, *EcAP*, and all known Alkaline Phosphatase superfamily monoesterases contain a common active site alpha helix (Fig. 6, *EcAP* & PafA, “phosphatase helix,” light blue)(40). This helix appears to position active site residues that interact with the other outward-facing phosphoryl oxygen atom (Fig. 2D, O₂). However, the phosphatase helix sits in the area where the second phosphodiester substituent is situated, and it is absent in all known diesterases (40). Instead, residues that line the region exposed by the absence of the phosphatase helix are used to create a binding pocket for this substituent (Fig. 6, NPP, diester pocket, red), and different NPP subfamilies utilize different residues in this area and in nearby auxiliary domains to achieve distinct substituent specificities (34,71,72).

The different sets of auxiliary domains in *EcAP*, PafA, and NPP highlight the versatility of the bimetallo core as a scaffold onto which additional domains and elements

can be built (Fig. 6). What about the potential to build an Alkaline Phosphatase superfamily active site that is optimal for a phosphate triester substrate? Stereochemical studies indicate that the second diester substituent has a strong preference for the O₂ position (Fig. 2D) in diesterase reactions catalyzed by *EcAP*—i.e., in the same position it occupies in NPP reactions—despite the presence of the phosphatase helix and absence of a substituent binding pocket (37). This observation implies inhibition of positioning of the substituent in the opposite position. Indeed, inspection of active site interactions of *EcAP* and PafA reveals that the backbone amide of the serine or threonine nucleophile donates a hydrogen bond to this oxygen atom (O₁, Fig. 2D; Fig. 6).

Given that this backbone interaction is located right at the site of nucleophilic attack, we speculate that resculpting an AP superfamily member to replace this interaction and accommodate an additional triesterase substituent would be difficult, and this difficulty may account for the absence of known AP superfamily triesterases (Fig. 6, brown “X”). These observations underscore the role of idiosyncratic structural and chemical properties of each protein scaffold in determining evolvability for new enzymatic functions (73-77). Analogously, these features will determine which scaffolds can be efficiently engineered to effectively carry out new reactions.

Overall, our results suggest a profound connection between catalytic mechanism and evolution and support a prominent role for catalytic promiscuity in deciding which enzymes and scaffolds were adopted for particular catalytic roles and how they were adapted to take on new catalytic roles throughout evolution (Fig. 6). Furthermore, our results, combined with prior mechanistic and structural studies, provide perspective on how AP superfamily structural elements and domains may have been used to enhance and specialize catalysis (Fig. 6). Once a scaffold like the AP superfamily bimetallo core is in place, its further evolutionary elaboration will depend on what reactions are

“needed,” which reactions can be catalyzed by that scaffold, and how readily that scaffold can accommodate new structural and functional elements. The insights herein and elsewhere into the structural and mechanistic properties of AP and other superfamilies should help guide efforts to select, redesign existing, and design *de novo* enzymes for practical applications in chemistry, medicine, and engineering.

EXPERIMENTAL PROCEDURES

Protein Expression and Purification. Mutants of AP from *Escherichia coli* (*EcAP*), NPP from *Xanthomonas axonopodis* pv. *Citri*, and PafA from *Chryseobacterium meningosepticum* were expressed in *E. coli* SM547 λ (DE3) cells. Cells were grown to an optical density of 0.6–0.8 in rich medium and glucose (10 g tryptone, 5 g yeast extract, 5 g NaCl, and 2 g glucose per liter) with 50 μ g/mL carbenicillin at 37 °C. Protein expression was induced by adding 0.3 mM Isopropyl thiogalactopyranoside. Cultures were grown at 30 °C for 16–20 h and were harvested by centrifugation.

EcAP was purified from a fusion pMAL-p2X construct containing an N-terminal maltose binding protein (MBP) with a Factor Xa cleavage site between it and the natural N-terminal and a C-terminal StrepII tag end as previously described (6,47). Briefly, pelleted cells with *EcAP* were lysed with osmotic shock and purified over amylose resin. The resin was washed extensively and the protein was eluted with 10 mM maltose. The enzyme was buffer exchanged into storage buffer (10 mM sodium MOPS, pH 8.0, 50 mM NaCl, and 100 μ M ZnCl₂). Constructs with and without the MBP tag used for purification for NPP* and *EcAP** were shown to have the same activity, as did PafA* with and without the Strep tag.

NPP was purified from a similar pMAL-p2X fusion construct as *EcAP*, containing an N-terminal maltose binding protein (MBP) and C-terminal StrepII tags with a Factor Xa cleavage site between it and the natural N-terminal end, as previously

described (45). Briefly, cells containing NPP were lysed with osmotic shock and run over a Q-sepharose column. After extensive washing, the protein was eluted and protein-containing fractions were pooled. StrepTactin® resin (5 mL, IBA Life Sciences, Göttingen, Germany) was washed extensively with 0.5 M sodium hydroxide, and was packed onto a gravity column. Following neutralization of the resin with 1 M Tris-HCl, pH 8.0, and washing with several column volumes of column buffer (100 mM Tris-HCl, pH 8.0, and 150 mM NaCl), the pooled enzyme fractions were loaded over the resin. Following washing the resin with column buffer (100 mM Tris-HCl, pH 8.0, 150 mM NaCl) the enzyme was eluted with 5 mM desthiobiotin in the same buffer and buffer exchanged into storage buffer (100 mM Tris-HCl, pH 8.0, 150 mM NaCl, and 100 μ M ZnCl₂).

PafA was purified from a fusion pET-22b construct containing a C-terminal StrepII tag. The cell pellets were resuspended in column buffer (100 mM Tris-HCl, pH 8.0, and 150 mM NaCl) and either frozen for later purification or lysed by passing the suspension through an Emulsiflex (Avestin, Ottawa, ON) three times. The lysate was centrifuged to remove cell debris (20,000g for 20 min), and the supernatant was filtered through a 0.45 μ m filter. StrepTactin® resin (5 mL, IBA Life Sciences) was washed with sodium hydroxide as described above, and following neutralizing the resin, the filtrate was loaded over the resin. The resin was washed with 6 column volumes of column buffer and the protein was eluted with 5 mM desthiobiotin in column buffer. Fractions containing purified PafA were pooled and buffer exchanged into storage buffer (10 mM sodium MOPS, pH 8.0, 50 mM NaCl, and 100 μ M ZnCl₂).

Protein purity was assessed with sodium dodecyl sulfate–polyacrylamide gel electrophoresis and was estimated to be >95% for all enzymes by staining with Coomassie Blue.

Synthesis of phosphomonoester and diester substrates. Substrate 3,4-dinitrophenyl

and 3-nitrophenyl phosphate monoesters were synthesized according to previously described experimental procedures (3,4-diNPP and 3-NPP respectively)(6).

Aryl methyl diesters were prepared using the method originally reported by Ba-Saif et al. (78), using commercially available dimethyl chlorophosphate (me-3,4-diNPP and me-3-NPP respectively). After demethylation, the acid form of the methyl diester was isolated by resuspending the lithium salt in 5 M HCl and extracting into diethyl ether. This method resulted in approximately 1% contamination with aryl monoester. The material was further purified over a RediSep Rf Gold C18 (Teledyne Isco, Lincoln, NE) column equilibrated with 100 mM triethylammonium chloride, pH 5.4. Gradient chromatography was run with this buffer and a 0-40% acetonitrile gradient at 2% per min. Fractions containing the aryl methyl diester were pooled, dried by rotary evaporation, and the acid form of the aryl methyl diester isolated as described above. Using the isolated acid form, concentrations for kinetic assays were determined gravimetrically and by quantification by absorbance of the leaving group after total hydrolysis and values agreed within 20% in all cases.

Kinetic Assays. Activity measurements were performed at 25 °C in a Perkin-Elmer UV/bis Lambda 25 spectrophotometer (Perkin Elmer, Waltham, MA) in 0.1 M sodium MOPS, pH 8.0, 0.5 M NaCl, 100 μM ZnCl₂, unless otherwise noted. For substrates with a *p*-nitrophenolate (Chart 2) or 3,4-dinitrophenolate leaving group, the formation of free *p*-nitrophenolate or 3,4-dinitrophenolate from hydrolysis of the substrates was monitored continuously at 400 nm. For substrates with an 3-nitrophenolate leaving group, the formation of free *m*-nitrophenolate was monitored continuously at 394 nm. Rate constants were determined 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 of each varied over at least a 5-fold range. Linear 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 mutant gave values within error.

The following buffers were used in pH dependences that established that kinetics were followed in a pH-independent region for each of the pruned enzymes: sodium MES, pH 5.5 and 6.0; sodium MOPS, pH 7.0; sodium CHES, pH 9.0; sodium CAPS, pH 10, each at 100 mM and in the presence of 500 mM NaCl and 100 μM ZnCl₂ (Supplemental Fig. 4). The pH dependency for pNPP and me-pNPP is similar in each enzyme, suggesting that the dianionic form of pNPP is reacting and not the monoionic species. The pH-dependency for pNPP and me-pNPP for *EcAP** is also similar to the paraoxon pH dependency, suggesting that paraoxon is hydrolyzed in the same active site.

As the mutant activities are orders of magnitude lower than those for the wild type enzymes with their cognate substrates, we carried out inhibition experiments to ensure that the measured activities represent the pruned enzymes and not a trace contaminant and that phosphate monoester and diester hydrolysis reactions occurred at the same active site for each enzyme. Inhibition experiments with inorganic phosphate (P_i) were carried out with subsaturating amounts of substrate. The P_i concentration range was 0.1–100 mM for *EcAP*, 0.5–100 mM for NPP, and 0.13–100 mM for PafA (Supplemental Fig. 5). The observation of K_i values that are the same for the monoester and diester substrates and distinct from the K_i values for the wild type enzymes suggests that the same enzyme is catalyzing both reactions in each case and that this enzyme is not a contaminating wild-type enzyme. To further test the origin of the observed catalysis we purified pruned versions of each enzyme with the nucleophilic Ser or Thr residue mutated to Gly: *EcAP*: S102G/D153A/R166S/E322Y/K328A; NPP: T90G/F91A/N111A/L123A/Y205A; and PafA: T79G/N100A/K162A/R164A. In each

case, the observed reaction rate was decreased by one to two orders of magnitude relative to the pruned enzymes and no reaction was observed above the non-enzymatic background hydrolysis, strongly suggesting that the observed activities of the pruned constructs arose from the pruned enzymes studied herein.

Metal occupancy. To test for metal ion concentration dependent activation of expressed enzyme, PafA, NPP and *EcAP* were incubated with the following metal concentrations: 10 μM ZnCl_2 , 100 μM ZnCl_2 , 500 μM ZnCl_2 , 1.0 mM MgCl_2 and 100 μM ZnCl_2 in 10 mM sodium MOPS, pH 8.0, and 50 mM NaCl at 25 °C. No activation was observed for either PafA or NPP. The *EcAP* pruned variant containing the E322Y mutation displayed activation behavior, with a half time for activation of ~ 4 days, while the E322A-containing mutant maintained constant activity. These results mirrored the activation previously observed in the otherwise wild-type context (6). The E322Y and E322A AP* constructs had activities within 2-fold after activation. The occupancy of metal ions in the active site of all of the pruned enzymes and WT enzymes were determined with atomic emission spectroscopy (Supplemental Table 2), as described previously (6,47), following overnight incubation in storage buffer with 100 μM ZnCl_2 , at room temperature.

Circular dichroism (CD). CD spectra were collected for the WT and pruned enzymes, each with their MBP tags removed, in 10 mM sodium MOPS, pH 8.0, 50 mM NaCl 100 μM ZnCl_2 . The following enzyme concentrations were used: *EcAP* WT and pruned 1.4 μM ; NPP WT and pruned, 1.0 μM ; and PafA WT and pruned, 1.4 μM .

Crystallization and structure determination of PafA and EcAP*.* PafA* was buffer exchanged into 10 mM Tris-HCl, pH 8.0, 50 mM NaCl, 100 μM ZnCl_2 and concentrated to 2.5 mg/ml. Equal volumes of PafA and precipitant solution (22% PEG3350, 0.1 M sodium acetate, pH 4.4, 0.2 M ammonium sulfate) were mixed and placed

over a reservoir of 1 mL precipitant solution to crystallize by the hanging drop method. Crystals were harvested and frozen in liquid nitrogen without cryoprotectant, crystallographic data were collected at the Stanford Linear Accelerator at beamline 11-1, using the Pilatus 6M detector (Dectris) with thin slice oscillations (0.2°) for a total 240° rotation. Diffraction images were processed with iMosflm and scaled with AIMLESS and POINTLESS programs from the CCP4 suite of software (79). The dataset was complete (99.9%) and redundant (8.5-fold) to 2.03Å resolution, with good merging statistics, which are summarized in Table 3. Dataset resolution was kept at the limit of the detector (2.03Å), since strong merging statistics, high completeness and high redundancy persisted into the highest resolution bin (Table 3).

The structure was solved by molecular replacement (MR) using PHENIX(80) with the WT PafA structure (PDB ID: 5TJ3(40)) as a search model. Prior to MR, we modified the search model by removing all solvent and metal atoms, side-chains surrounding the two active site metal sites, and the side-chains of the mutated residues (T79S, N100A, K162A and R164A). The molecular replacement solution showed strong electron density peaks for the missing active site side-chains and the metal ions, a lack of density for the residues mutated to alanines, and a clear change in the position of the T79S loop. Additionally, a strong positive density peak was found coordinated between the two active site Zn^{2+} ions and S79, which we interpreted as Cl⁻. After several rounds of refinement with *phenix.refine* (80) and manual adjustments with *Coot* (81), the final model had excellent geometry and stereochemistry, with $R_{\text{work}} = 17.0\%$ and $R_{\text{free}} = 20.9\%$ (Table 3).

*EcAP** was buffer exchanged into 10 mM sodium MOPS, pH 7.0, 50 mM NaCl 100 μM ZnCl_2 and concentrated to 5 mg/ml. Equal volumes of *EcAP** and precipitant solution (23% PEG3350, 0.2 M NH_3F , 0.2 M sodium HEPES, pH 8.0) were mixed and placed over a reservoir of 1 mL precipitant solution to crystallize by the hanging drop method.

Crystals were harvested and frozen in liquid nitrogen with paratone-N oil as cryoprotectant. Crystallographic data were collected at the Stanford Linear Accelerator at beamline 12-2 using the Pilatus 6M PAD detector with thin slice oscillations and shutterless mode for 360° rotation, resulting in highly redundant 2.45 Å resolution dataset containing anomalous signal from the Zn, despite the energy offset compared to the Zn peak energy (12658 eV versus 9668 eV). Diffraction images were processed with the AUTOXDS script (Ana Gonzalez, SSRL) using the current XDS version, AIMLESS, and POINTLESS programs from the CCP4 package (79). Anomalous signal was kept during the processing. Dataset resolution cutoff was chosen based on the high completeness, >35 fold multiplicity, and >30% CC1/2 as suggested by the scaling program AIMLESS.

The structure was solved by molecular replacement with a single monomer (subunit A) from the 3TG0 PDB entry in the PHASER program(82) that produced a single solution in space group P6322 (#182), with two polypeptide chains in the unit cell. The molecular replacement solution showed negative peaks in the Fo-Fc map for the mutated side chains (D153A, E322A, K328A) or a lack of side chain density in the 2Fo-Fc map for D101A and R166S. The following features that were observed and incorporated into the structural model: (a) a well-populated metal site ligated by H162 and E164; (b) a less-occupied metal site ligated by H125 and D304; (c) the StrepII tag at the C-terminus providing a histidine ligand (H452) to the metal binding site (a). Anomalous difference maps using the refined model suggest that all metal binding sites have Zn²⁺ ions incorporated. The model was updated and checked in COOT (81), while being refined initially with maximum likelihood program REFMAC(83) and finally with BUSTER (version 2.10.2 (84)). The final model had excellent geometry and stereochemistry, in the top 1% percent distribution for this resolution range in MolProbity (85) and low residual factors values: R = 16.46% and R_{free} = 21.90%

(Table 4), justifying the inclusion of the weaker data in improving the resulting model.

ACKNOWLEDGEMENTS

We dedicate this paper to Patsy Babbitt on the occasion of her retirement in honor of her profound and creative insights into enzyme evolution and function. This work was funded by a grant from the US National Institutes of Health to D.H. (GM64798). 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 Filip Yabukarski for assistance with collecting crystallographic data for PafA*. We thank Patsy Babbitt and Jonathan Lassila for helpful discussions, advice, and communication of unpublished results. We thank Daniel Mokhtari and Ariana Peck and members of the Herschlag lab for discussions and comments on the manuscript.

AUTHOR CONTRIBUTION

F.S. and I.A. expressed and purified enzymes, performed all characterization of the enzymes, crystallized the enzymes, and mounted them for X-ray studies.

T.D. and A.L. solved the crystal structures for the PafA* and EcAP*.

J.S. Synthesized the substrates with alternative leaving groups.

D.H. and F.S. designed the study, analyzed the data and wrote the manuscript with contributions from all authors.

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest with the contents of this article.

REFERENCES:

1. Edwards, D. R., Lohman, D.C., Wolfenden, R. (2012) Catalytic proficiency: the extreme case of S-O cleaving sulfatases. *J. Am. Chem. Soc.* **134**, 525-531

2. Wolfenden, R., Snide, M.J. (2001) The depth of chemical time and the power of enzymes as catalysts. *Acc. Chem. Res.* **34**, 938-945
3. O'Brien, P. J., Herschlag, D. (1999) Catalytic promiscuity and the evolution of new enzymatic activities. *Chem. Biol.* **6**, R91-R105
4. Jensen, R. A. (1976) Enzyme recruitment in evolution of new function. *Annu. Rev. Microbiol.* **30**, 409-425
5. Khersonsky, O., Tawfik, D.S.. (2010) Enzyme promiscuity: a mechanistic and evolutionary perspective. *Annu. Rev. Biochem.* **79**, 471-505
6. Zalatan, J. G., Fenn, T.D., Herschlag, D. (2008) Comparative enzymology in the alkaline phosphatase superfamily to determine the catalytic role of an active-site metal ion. *J. Mol. Biol.* **384**, 1174-1189
7. O'Brien, P. J., Herschlag, D. (1999) Does the active site arginine change the nature of the transition state for alkaline phosphatase-catalyzed phosphoryl transfer? *J. Am. Chem. Soc.* **121**, 11022-11023
8. Gerlt, J. A., Babbitt, P.C. (2001) Divergent evolution of enzymatic function: mechanistically diverse superfamilies and functionally distinct suprefamilies. *Annu. Rev. Biochem.* **70**, 209-246
9. O'Brien, P. J. (2006) Catalytic Promiscuity and the Divergent Evolution of DNA Repair Enzymes. *Chem. Rev.* **106**, 720-752
10. Babbitt, P. C., Mrachko, G.T., Hasson, M.S., Huisman, G.W., Kolter, R., Ringe, D., Petsko, G.A., Kenyon, G.L., Gerlt, J.A. (1995) A functionally divers enzyme superfamily that abstracts the ac protons of carboxylic acids. *Science* **267**, 1159-1161
11. Glasner, M. E., Gerlt, J.A., Babbitt, P.C. (2006) Evolution of enzyme superfamilies. *Curr. Opin. Chem. Biol.* **10**, 492-497
12. Pandya, C., Farelli, J.D., Dunaway-Mariano, D., Allen, K.N. (2014) Enzyme Promiscuity: Engine of Evolutionary Innovation. *J. Biol. Chem.* **289**, 30229-30236
13. Allen, K. N., and Dunaway-Mariano, D. (2004) Phosphoryl group transfer: evolution of a catalytic scaffold. *Trends in biochemical sciences* **29**, 495-503
14. Meng, Q., Chen, K., Ma, L., Hu, S., Yu, J. (2011) A systematic identification of Kolobok superfamily transposons in *Trichomonas vaginalis* and sequence analysis on related transposases. *J. Genet. Genomics* **38**, 63-70
15. Gerlt, J. A., Allen, K.N., Almo, S.C., Armstrong, R.N., Babbitt, P.C., Cronan, J.E., Dunaway-Mariano, D., Imker, H.J., Jacobson, M.P., Minor, W., Poulter, C.D., Raushel, F.M., Sali, A., Shoichet, B.K., Sweedler, J.V. (2011) The Enzyme Function Initiative. *Biochemistry* **50**, 9950-9962
16. O'Brien, P. J., Herschlag, D. (1998) Sulfatase activity of E.coli alkaline phosphatase demonstrates a functional link to arylsulfatases, an evolutionary related enzymes family. *J. Am. Chem. Soc.* **120**, 12369-12370

17. Voordeckers, K., Brown, C.A., Vanneste, K., van der Zande, E., Voet, A., Maere, S., Verstrepen, K.J. (2012) Reconstruction of ancestral metabolic enzymes reveals molecular mechanisms underlying evolutionary innovations through gene duplication. *PLoS biology* **10**, 1-17
18. Thomson, J. M., Gaucher, E. A., Burgan, M. F., De Kee, D. W., Li, T., Aris, J. P., and Benner, S. A. (2005) Resurrecting ancestral alcohol dehydrogenases from yeast. *Nature genetics* **37**, 630-635
19. Bridgham, J. T., Carroll, S. M., and Thornton, J. W. (2006) Evolution of hormone-receptor complexity by molecular exploitation. *Science* **312**, 97-101
20. Ortlund, E. A., Bridgham, J. T., Redinbo, M. R., and Thornton, J. W. (2007) Crystal structure of an ancient protein: evolution by conformational epistasis. *Science* **317**, 1544-1548
21. Bar-Rogovsky, H., Stern, A., Penn, O., Kobl, I., Pupko, T., Tawfik, D.S. (2015) Assessing the prediction fidelity of ancestral reconstruction by a library approach. *Protein Eng. Des. Sel.* **28**, 1-12
22. Weiss, M. C., Sousa, F.L., Mrnjavac, N., Neukirchen, S., Roettger, M., Nelson-Sathi, S., Martin, F. (2016) The physiology and habitat of the last universal common ancestor. *Nat. Microbiol.* **1**, 1-8
23. Baier, F., Copp, J.N., Tokuriki N. (2016) Evolution of enzyme superfamilies: comprehensive exploration of sequence-function relationships. *Biochemistry* **55**, 6375-6388
24. Renata, H., Wang, Z.J., Arnold, F.H. (2015) Expanding the enzyme universe: accessing non-natural reactions by mechanism-guided directed evolution. *Angew. Chem. Int. Ed.* **54**, 3351-3367
25. Babbie, A., Tokuriki, N., Hollfelder, F. (2010) What makes an enzyme promiscuous? *Curr. Opin. Chem. Biol.* **14**, 200-207
26. Furnham, N., Dawson, N.L., Rahman, S.A., Thornton, J.M., Orengo, C.A. (2016) Large-Scale Analysis Exploring Evolution of Catalytic Machineries and Mechanisms in Enzyme Superfamilies. *J. Mol. Biol.* **428**, 253-267
27. Das, S., Dawson, N.L., Orengo, C.A. (2015) Diversity in protein domain superfamilies. *curr. opin. genet. dev.* **35**, 40-49
28. van Loo, B., et al. (2010) An efficient, multiply promiscuous hydrolase in the alkaline phosphatase superfamily. *Proc. Natl. Acad. Sci. USA* **107**, 2740-2745
29. Kim, A., et al. (2011) Divergence of chemical function in the alkaline phosphatase superfamily: structure and mechanism of the P-C bond cleaving enzyme phosphonoacetate hydrolase. *Biochemistry* **50**, 3481-3494
30. Jonas, S., van Loo, B., Hyvonen, M., Hollfelder, F. (2008) A new member of the alkaline phosphatase superfamily with a formylglycine nucleophile: structural and kinetic characterisation of a phosphonate monoester

- hydrolase/phosphodiesterase from *Rhizobium leguminosarum*. *J. Mol. Biol.* **384**, 120-136
31. Gijsbers, R., Ceulemans, H., Stalmans, W., Bollen, M. (2001) Structural and catalytic similarities between nucleotide pyrophosphatases/phosphodiesterases and alkaline phosphatases. *J. Biol. Chem.* **276**, 1361-1368
 32. Chapman, E., Best, M.D., Hanson, S.R., Wong, C.H. (2004) Sulfotransferases: structure, mechanism, biological activity, inhibition, and synthetic utility. *Angew. Chem. Int. Ed. Engl.* **43**, 3526-3548
 33. Galperin, M. Y., Jedrzejewski, M.J. (2001) Conserved core structure and active site residues in alkaline phosphatase superfamily enzymes. *Proteins* **45**, 318-324
 34. Zalatan, J. G., Fenn, T.D., Brunger, A.T., Herschlag, D. (2006) Structural and functional comparisons of nucleotide pyrophosphatase/phosphodiesterase and alkaline phosphatase: implications for mechanism and evolution. *Biochemistry* **45**, 9788-9803
 35. Olguin, L. F., Askew, S.E., O'Donoghue, A.C., Hollfelder, F. (2008) Efficient catalytic promiscuity in an enzyme superfamily: an arylsulfatase shows a rate acceleration of 10¹³ for phosphate monoester hydrolysis. *J. Am. Chem. Soc.* **130**, 16547-16555
 36. O'Brien, P. J., Lassila, J.K., Fenn, T.D., Zalatan, J. G., Herschlag, D. (2008) Arginine coordination in enzymatic phosphoryl transfer: evaluation of the effect of Arg166 mutations in *Escherichia coli* alkaline phosphatase. *Biochemistry* **47**, 7663-7672
 37. Lassila, J. K., Herschlag, D. (2008) Promiscuous sulfatase activity and thio-effects in a phosphodiesterase of the alkaline phosphatase superfamily. *Biochemistry* **47**, 12853-12859
 38. Jonas, S., Hollfelder, F. (2009) Mapping catalytic promiscuity in the alkaline phosphatase superfamily. *Pure Appl. Chem.* **81**, 731-742
 39. Galperin, M. Y., Bairoch, A., 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
 40. Sunden, F., AlSadhan, I., Lyubimov, A.Y., Ressler, S., Wiersma-Koch, H., Borland, J., Brown, C.L., Johnson, T.A., Singh, Z., Herschlag, D. (2016) Mechanistic and evolutionary insights from comparative enzymology of phosphomonoesterases and phosphodiesterases across the alkaline phosphatase superfamily. *J. Am. Chem. Soc.*
 41. Bihani, S. C., Das, A., Nilgiriwala, K. S., Prashar, V., Pirocchi, M., Apte, S. K., Ferrer, J., Hosur, M. V. (2011) X-ray structure reveals a new class and provides insight into evolution of alkaline phosphatases. *PloS one* **6**, 1-10
 42. Berlutti, F., Passariello, C., Selan, L., Thaller, M. C., Rossolini, G. M.

- (2001) The Chryseobacterium meningosepticum PafA enzyme: prototype of a new enzyme family of prokaryotic phosphate-irreversible alkaline phosphatases? *Microbiology* **147**, 2831-2839
43. Bond, C. S., Clements, P.R., Ashby, S.J., Collyer, C.A., Harrop, J.J., Guss, J.M. (1997) Structure of a human lysosomal sulfatase. *Structure*, 277-289
44. Lassila, J. K., Zalatan, J.G., Herschlag, D. (2011) Biological phosphoryl-transfer reactions: understanding mechanism and catalysis. *Annu. Rev. Biochem.* **80**, 669-702
45. Wiersma-Koch, H. I., Sunden, F., Herschlag, D. (2013) Site-directed mutagenesis maps interactions that enhance cognate and limit promiscuous catalysis by an alkaline phosphatase superfamily diesterase. *Biochemistry* **52**, 9167-9176
46. Zalatan, J. G., Fenn, T. D., Herschlag, D. (2008) Comparative enzymology in the alkaline phosphatase superfamily to determine the catalytic role of an active-site metal ion. *J. Mol. Biol.* **384**, 1174-1189
47. Sunden, F., Peck, A., Salzman, J., Ressler, S., Herschlag, D. (2015) Extensive site-directed mutagenesis reveals interconnected functional units in the alkaline phosphatase active site. *eLife* **4**, e06492
48. Le Du, M. H., Lamoure, C., Muller, B.H., Bulgakov, O.V., Lajeunesse, E., Ménez, A., Boulain, J.C. (2002) Artificial evolution of an enzyme active site: structural studies of three highly active mutants of Escherichia coli alkaline phosphatase. *J. Mol. Biol.* **316**, 941-953
49. Wang, J., Stieglitz, K., Kantrowitz, E.R. (2005) Metal Specificity Is Correlated with Two Crucial Active Site Residues in Escherichia coli Alkaline Phosphatase. *Biochemistry* **44**, 8378-8386
50. Murphy, J. E., Tibbits, T.T., Kantrowitz, E.R. (1995) Mutations at positions 153 and 328 in Escherichia coli alkaline phosphatase provide insight towards the structure and function of mammalian and yeast alkaline phosphatases. *J. Mol. Biol.* **253**, 604-617
51. O'Brien, P. J., Herschlag, D. (2002) Alkaline phosphatase revisited: hydrolysis of alkyl phosphates. *Biochemistry* **41**, 3207-3225
52. Jencks, W. P., Regenstein, J. (1976) *Handbook of biochemistry and molecular biology*, Cleveland, OH
53. Tull, D., Withers, S.G. (1994) Mechanisms of cellulases and xylanases: A detailed kinetic study of the exp- b -1,4-glycanase from Cellulomonas fimi. *Biochemistry* **33**, 6363-6370
54. Kirby, A. J., Varvoglis, A.G. (1966) The reactivity of phosphate esters. Monoester hydrolysis. *J. Am. Chem. Soc.* **89**, 415-423
55. Ratner, H. (1984) Flavobacterium Meningosepticum. *Infection Control* **5**, 237-239
56. Nikolic-Hughes, I., Rees, D.C., Herschlag, D. (2004) Do

- electrostatic interactions with positively charged active site groups tighten the transition state for enzymatic phosphoryl transfer? *J. Am. Chem. Soc.* **126**, 11814-11819
57. Cleland, W. W., Hengge, A. (2006) Enzymatic Mechanisms of Phosphate and Sulfate Transfer. *Chem. Rev.* **106**, 3252-3278
 58. Boltes, I., Czapinska, H., Kahnert, A., von Bülow, R., Dierks, T., Schmidt, B., von Figura, K., Kertesz, M.A., Usón, I. (2001) 1.3 Å Structure of Arylsulfatase from *Pseudomonas aeruginosa* Establishes the Catalytic Mechanism of Sulfate Ester Cleavage in the Sulfatase Family. *Structure* **9**, 483-491
 59. Lukatela, G., Krauss, N., Theis, K., Selmer, T., Gieselmann, V., von Figura, K., Saenger, W. (1998) Crystal structure of human arylsulfatase A: the aldehyde function and the metal ion at the active site suggest a novel mechanism for sulfate ester hydrolysis. *Biochemistry* **37**, 3654-3664
 60. Hanson, S. R., Best, M.D., Wong, C.H. (2004) Sulfatases: structure, mechanism, biological activity, inhibition, and synthetic utility. *Angew. Chem. Int. Ed. Engl.* **43**, 5736-5763
 61. Benkovic, S. J., Benkovic, P.A. (1966) Studies on sulfate esters. I. Nucleophilic reactions of amines with p-nitrophenyl sulfate. *J. Am. Chem. Soc.* **88**, 5504-5511
 62. Khersonsky, O., Tawfik, D.S. (2005) Structure–reactivity studies of serum paraoxonase PON1 suggest that its native activity is lactonase. *Biochemistry*
 63. Babbie, A. C., Bandyopadhyay, S., Olguin, L.F., Hollfelder, F. (2009) Efficient catalytic promiscuity for chemically distinct reactions. *Angew. Chem. Int. Ed. Engl.* **48**, 3692-3694
 64. Ohno, S. (1970) *Evolution by gene duplication*, Springer-Verlag, London
 65. Bobyr, E., Lassila, J. K., Wiersma-Koch, H. I., Fenn, T. D., Lee, J. J., Nikolic-Hughes, I., Hodgson, K. O., Rees, D. C., Hedman, B., and Herschlag, D. (2012) High-resolution analysis of Zn(2+) coordination in the alkaline phosphatase superfamily by EXAFS and x-ray crystallography. *J Mol Biol* **415**, 102-117
 66. Xiang, D. F., Kolb, P., Fedorov, A.A., Meier, M.M., Fedorov, L.V., Nguyen, T.T., Sterner, R., Almo, S.C., Shoichet, B.K., Raushel, F.M. (2009) Functional annotation and three-dimensional structure of Dr0930 from *Deinococcus radiodurans*, a close relative of phosphotriesterase in the amidohydrolase superfamily. *Biochemistry* **48**, 2237-2247
 67. Seibert, C. M., Raushel, F.M. (2005) Structural and Catalytic Diversity within the Amidohydrolase Superfamily. *Biochemistry* **44**, 6383-6391
 68. Holm, L., and Sander, C. (1997) An Evolutionary Treasure: Unification of a Broad Set of Amidohydrolases Related to Urease. *Proteins: Struct., Funct., Genet.* **28**, 72-82
 69. Elias, M., Dupuy, J., Merone, L., Mandrich, L., Porzio, E., Moniot, S., Rochu, D., Lecomte, C., Rossi,

- M., Masson, P., Manco, G., Chabriere, E. (2008) Structural basis for natural lactonase and promiscuous phosphotriesterase activities. *J. Mol. Biol.* **379**, 1017-1028
70. Bigley, A. N., Raushel, F.M. (2013) Catalytic Mechanisms for Phosphotriesterases. *Biochim. Biophys. Acta.* **1834**, 443-453
71. Nishimasu, H., Okudaira, S., Hama, K., Mihara, E., Dohmae, N., Inoue, A., Ishitani, R., Takagi, J., Aoki, J., Nureki, O. (2011) Crystal structure of autotaxin and insight into GPCR activation by lipid mediators. *Nat. Struct. Biol.* **18**, 205-212
72. Gorelik, A., Liu, F., Illes, K., Nagar, B. (2017) Crystal structure of the human alkaline sphingomyelinase provides insights into substrate recognition. *J. Biol. Chem.* **292**, 7087-7094
73. Kirschner, M. W., Gerhardt, J.C. (2005) *The Plausibility of Life: Resolving Darwin's Dilemma*, Yale University Press, Princeton
74. Feder, M. E. (2007) Evolvability of physiological and biochemical traits: evolutionary mechanisms including and beyond single-nucleotide mutation. *J. Exp. Biol.* **210**, 1653-1660
75. Kirschner, M., Gerhardt, J. (1998) Evolvability. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 8420-8427
76. Aharoni, A., Gaidukov, L., Khersonsky, O., Mc, Q. G. S., Roodveldt, C., and Tawfik, D. S. (2005) The 'evolvability' of promiscuous protein functions. *Nature genetics* **37**, 73-76
77. Poelwijk, F. J., Kiviet, D.J., Weinreich, D.M., Sander, T. (2007) Empirical fitness landscapes reveal accessible evolutionary paths. *Nature* **445**, 383-386
78. Basaif, S. A., Davis, A.M., Williams, A. (1989) Effective Charge-Distribution for Attack of Phenoxide Ion on Aryl Methyl Phosphate Monoanion - Studies Related to the Action of Ribonuclease. *J. Org. Chem.* **54**, 5483-5486
79. Evans, P. R., Murshudov, G.N. (2013) How good are my data and what is the resolution? *Acta Crystallogr., Sect D: Biol. Crystallogr.* **69**, 1204-1214
80. Adams, P. D. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr., Sect D: Biol. Crystallogr.* **66**, 213-221
81. Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr., Sect D: Biol. Crystallogr.* **66**, 486-501
82. McCoy, A. J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., Read, R.J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658-674
83. Murshudov, G. N., Vagin, A.A., Dodson, E.J. (1997) Refinement of Macromolecular Structures by the Maximum-Likelihood method. *Acta Crystallogr., Sect D: Biol. crystallogr.* **53**, 240-255
84. Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W., Roversi, P, Sharff, A., Smart, O.S., Vonrhein, C., Womack, T.O. (2016) BUSTER version 2.10.12.

- Global Phasing Ltd., Cambridge, United Kingdom
85. Chen, V. B., Arendall, W. B. 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr., Sect D: Biol. Crystallogr.* **66**, 12-21
86. Andrews, L. D., Zalatan, J.G., Herschlag, D. (2014) Probing the origins of catalytic discrimination between phosphate and sulfate monoester hydrolysis: comparative analysis of alkaline phosphatase and protein tyrosine phosphatases. *Biochemistry* **53**, 6811-6819
87. Bobyr, E., Lassila, J. K., Wiersma-Koch, H. I., Fenn, T. D., Lee, J. J., Nikolic-Hughes, I., Hodgson, K. O., Rees, D.C., Hedman, B., Herschlag, D. (2012) High-resolution analysis of Zn(2+) coordination in the alkaline phosphatase superfamily by EXAFS and x-ray crystallography. *J. Mol. Biol.* **415**, 102-117
88. Mohamed, M. F., Hollfelder, F.E. (2013) Efficient, crosswise catalytic promiscuity among enzymes that catalyze phosphoryl transfer. *Biochem. Biophys. Acta.* **1834**, 417-424

Tables

Table 1. Rate Constants for Reactions of Wild-type and Pruned *EcAP*, NPP, and PafA.

A

Enzyme	k_{cat}/K_M ($M^{-1}s^{-1}$)				
	pNPP	me-pNPP	dT5'-pNPP	pNPS	paraoxon
WT <i>EcAP</i>	3.3×10^7 ^b [6.3×10^8]	1.8×10^{1c}	$<2.0 \times 10^{-2}$	2.8×10^{-3d}	N.D.
WT NPP	1.0 ^e	2.3×10^{2e}	1.6×10^{6e}	2.0×10^{-5f}	$<5.9 \times 10^{-4}$
WT PafA	$1.6(0.5) \times 10^{6b}$ [4.3×10^{8j}]	$1.7(0.5) \times 10^{1b}$	$1.6(0.3) \times 10^{-3}$	$1.6(0.1) \times 10^{-4}$	N.D.

^a Assay conditions: 0.1 M sodium MOPS, pH 8.0, 0.5 M NaCl, 100 μ M ZnCl₂, and 500 μ M MgCl₂ at 25 °C, unless noted otherwise. Errors (in parenthesis) in activities are standard deviations from a duplicate with the same or independent enzyme preparations.

^b Data from ref. (51) and ref. (40). Diffusion rather than the chemical step is rate limiting for pNPP hydrolysis by WT *EcAP*. The estimated second-order rate constant with the chemical step rate limiting is given in the square brackets, calculated in ref. (40,47,86).

^c Data from ref. (6).

^d Data from ref. (86).

^e Data from ref. (34). Assay conditions: 0.1 M Tris-HCl, pH 8.0, 0.5 M NaCl, 100 μ M ZnCl₂, and at 25 °C.

^f Data from ref. (37). Assay conditions: 0.1 M sodium MES, pH 6.5, 0.5 M NaCl, 100 μ M ZnCl₂.

B

Enzyme	k_{cat}/K_M ($M^{-1}s^{-1}$)				
	pNPP ^b	Me-pNPP	dT5'-pNPP	pNPS	paraoxon
<i>EcAP</i>*	$1.7(0.4) \times 10^{-1}$	$1.3(0.3) \times 10^{-1}$	$6.8(0.3) \times 10^{-3}$	$<1 \times 10^{-6}$	$7.2(1.6) \times 10^{-3}$
D101A/D153A/R166S/E322A/K328A					
NPP*	$1.7(0.6) \times 10^{-1}$	$1.0(0.4) \times 10^{-1}$	$1.8(0.1)$	$<4 \times 10^{-6}$	$8.6(1.8) \times 10^{-3}$
T90S/F91A/N111A/L123A/Y205A					
PafA*	$7.7(0.8) \times 10^{-1}$	$4.7(0.2) \times 10^{-1}$	$7.4(0.1) \times 10^{-1}$	$<4 \times 10^{-6}$	$5.4(1.3) \times 10^{-2}$
T79S/N100A/K162A/R164A					

^a Assay conditions: 0.1 M sodium MOPS, pH 8.0, 0.5 M NaCl, 100 μ M ZnCl₂ and 500 μ M MgCl₂ at 25 °C. Errors (in parenthesis) in activities are standard deviations from a duplicate with the same or independent enzyme preparations.

Table 2. Rate Constants for Alternative Pruned Constructs of *EcAP*, NPP and PafA

Mutant					pNPP ²⁻		me-pNPP	
					k_{cat}/K_M (M ⁻¹ s ⁻¹)	k_{rel} ^c	k_{cat}/K_M (M ⁻¹ s ⁻¹)	k_{rel} ^d
<i>EcAP</i> pruned constructs								
D101	D153	R166	E322	K328				
A	A	S	A	A	0.17(0.04)	(1)	0.13(0.03)	(1)
G	A	S	Y	A	0.13(0.02) ^b	1.3	0.12(0.01)	1.1
N	A	S	Y	A	0.19(0.03) ^b	0.9	-	-
A	A	S	Y	A	0.17(0.03) ^b	1.0	0.13(0.01)	1.0
NPP pruned constructs								
T90	F91	N111	L123	Y205				
S	A	A	A	A	0.17(0.06)	(1)	0.10(0.04)	(1)
S	A	G	A	A	0.086(0.005)	1.9	0.080(0.004)	1.3
PafA pruned constructs								
T79	N100	K162	R164					
S	A	A	A		0.77(0.08)	(1)	0.47(0.02)	(1)
T	A	A	A		0.10(0.02)	7.7	0.14(0.01)	3.4

^a Assay conditions: 100 mM sodium MOPS, pH 8.0, 500 mM NaCl, 100 μM ZnCl₂ at 25 °C. Errors (in parenthesis) in activities are standard deviations from a duplicate with the same or independent enzyme preparations.

^b Rate constants from ref.(47).

^c Calculated as the ratio of $(k_{cat}/K_M)_{pNPP, pruned} / (k_{cat}/K_M)_{pNPP, alt. pruned}$.

^d Calculated as the ratio of $(k_{cat}/K_M)_{me-pNPP, pruned} / (k_{cat}/K_M)_{me-pNPP, alt. pruned}$.

Table 3. X-ray Crystallographic Data Collection and Refinement Statistics for PafA*

Data Collection	
Space group	I 4
Unit cell axes	
a, b, c (Å)	113.33 113.33 69.790
α, β, γ (°)	90 90 90
Resolution range (Å)	40.07–2.03 (2.10–2.03)
R_{merge} (%)	16.8 (151.8)
R_{pim} (%)	6.1 (54.1)
$\langle I \rangle / \langle \sigma I \rangle$	9.6 (2.7)
Completeness (%)	99.9 (99.9)
Multiplicity	8.5 (8.8)
CC_{1/2}	0.995 (0.660)
Refinement	
Resolution range (Å)	40.07 - 2.03
No. unique reflections	28559 (2831)
R_{work}/R_{free} (%)	17.04/20.85
Number of atoms	4241
Average B-factors (Å²)	
protein	37.1
water	37.7
ligands	48.3
R.m.s. deviation from ideality	
Bond length (Å)	0.013
Bond angles (°)	1.18
Ramachandran statistics	
Favored regions (%)	95
Allowed regions (%)	4.3
Outliers (%)	0.39
PDB code	5TOO

Table 4. X-ray Crystallographic Data Collection and Refinement Statistics for *EcAP**

Data Collection	
Space group	P 63 2 2
Unit cell axes	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	160.72 160.72 138.36
α , β , γ (°)	90 90 120
Resolution range (Å)	38.60–2.45 (2.54–2.45)
R _{merge} (%)	59.4 (560.0)
R _{pim} (%) ^c	10.3 (106.0)
$\langle I \rangle / \langle \sigma I \rangle$	7.8 (1.2)
Completeness (%)	99.5 (95.6)
Multiplicity	35.5 (30.4)
CC _{1/2}	0.995 (0.318)
Refinement	
Resolution range (Å)	38.60 – 2.45
No. unique reflections	39204
R _{work} /R _{free} (%)	16.46/21.90
Number of atoms	7042
Average B-factors (Å ²)	
protein	56.0
water	54.4
ligands	68.5
R.m.s. deviation from ideality	
Bond length (Å)	0.010
Bond angles (°)	1.13
Ramachandran statistics	
Favored regions (%)	97.75
Allowed regions (%)	2.25
Outliers (%)	0.00
PDB code	5TPQ

FIGURES

Chart 1.
Reactions of the AP Superfamily

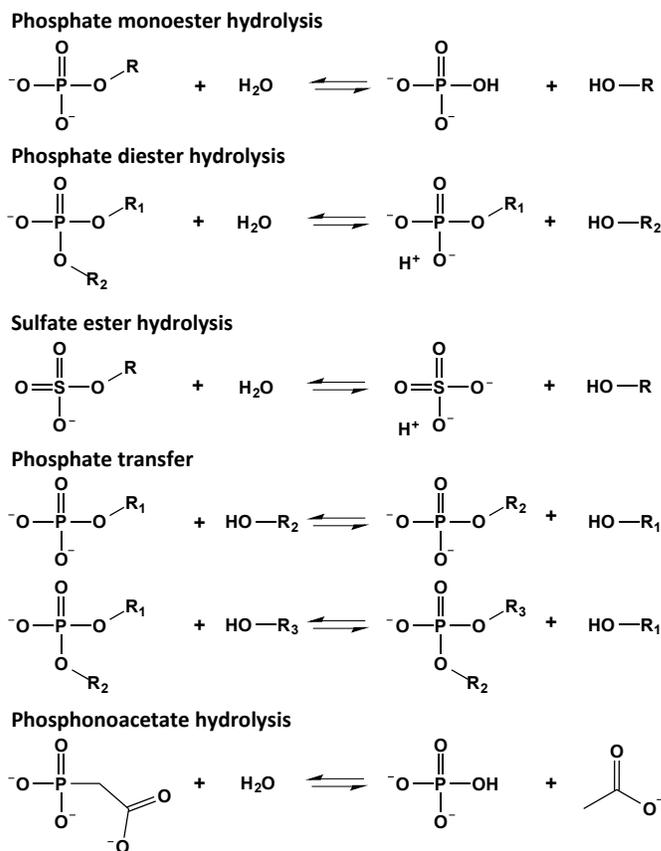


Chart 2.
Substrates Used in Study.

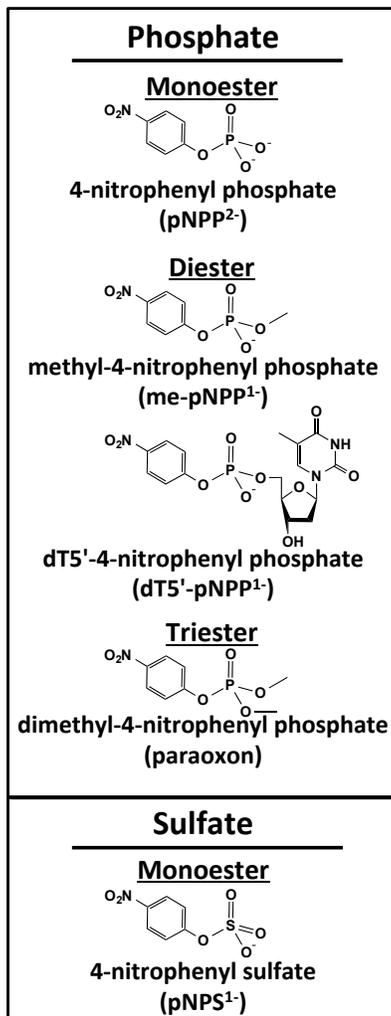


Figure 1. AP superfamily bimetallo active site and reaction pathway. Generic active site of AP bimetallo enzyme showing the alkoxide nucleophile (Ser or Thr), Zn²⁺ atoms, and Zn²⁺ ligands with a depiction of the reaction's transition state, and general reaction mechanism for AP superfamily enzymes via a covalent intermediate (E-P), shown for reaction of a phosphate monoester giving an inorganic phosphate (P_i) product.

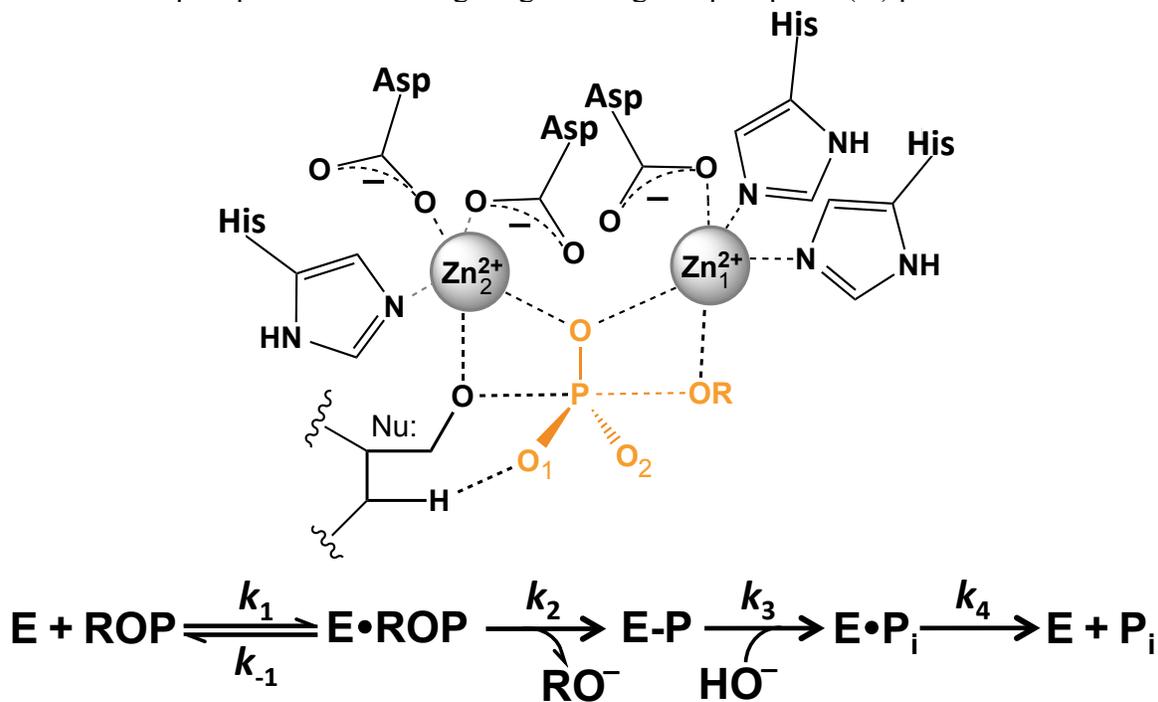


Figure 2. Overall structure and active sites of *EcAP*, *PafA*, and *NPP*. (A) Topology diagrams showing structurally conserved domains in gray and white and auxiliary domains that vary between enzymes in color, with blue for *EcAP*, magenta for *NPP*, and green for *PafA*. Residues that were mutated in the pruned constructs are colored in orange. (B) Structure of WT *EcAP* (3TG0(87)), *NPP* (2GSN(34)), and *PafA* (5TJ3(40)) with the conserved Zn^{2+} ligands and nucleophiles are shown in black, Zn^{2+} ions in gray, and the auxiliary domains and residues mutated in the pruned constructs colored as in part (A). (C) Active site schematics with a phosphoryl group and leaving group (red) represented as in the reaction's transition state and the residues mutated in each pruned enzyme shown in color; blue for *EcAP*, magenta for *NPP*, and green for *PafA*. (D) Schematic representation of the presumed ground and transition state interactions for reactions of the pruned enzymes.

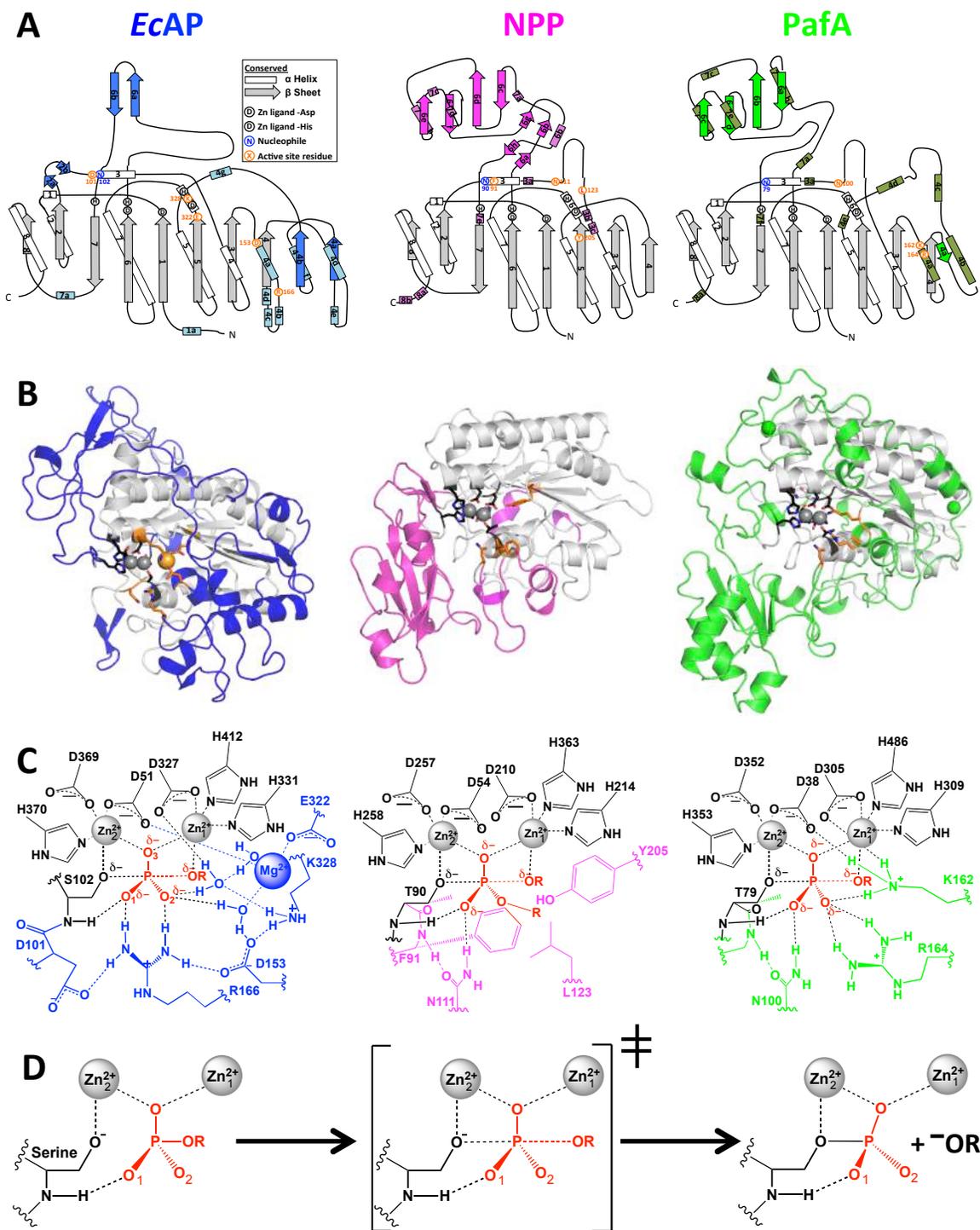


Figure 3. Structural comparisons of wild type and pruned PafA and EcAP. (A) Superimposition of the overall structure of WT *EcAP* (3TG0(87); blue) and *EcAP** (5TPQ, light blue). (B) Superimposition of active site of WT PafA (5TJ3(40); dark green) and PafA* (5TOO; pale green). (C) Overlay of active sites of WT and pruned *EcAP* and PafA. Structures and coloring as in parts (A) and (B). Blue residue numbers correspond to *EcAP* and green correspond to PafA; oxygen atoms are depicted in red.

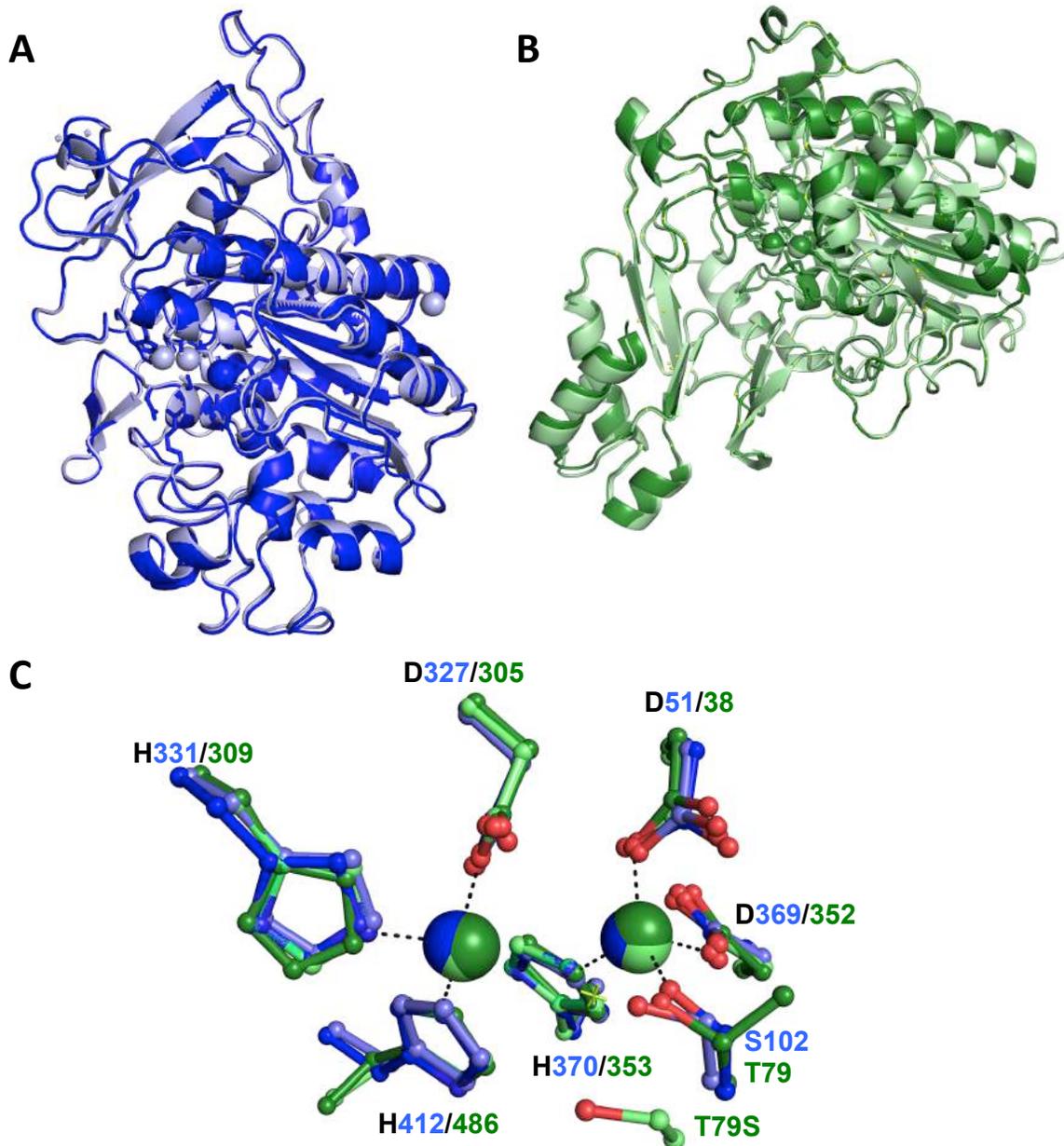


Figure 4. Comparison of phosphate monoester and diester reactions catalyzed by WT and pruned *EcAP*, *NPP*, and *PafA*. (A) Values of k_{cat}/K_M for phosphate monoester, pNPP, hydrolysis. For WT, values in dark gray are measured and values in light gray are calculated to represent reactions with the chemical step rate limiting (values are from Table 1A and ref. (40,47,51,86)). The values for the pruned enzymes are shown in black (Table 1B). (B) Values of k_{cat}/K_M for hydrolysis of minimal diester substrate me-pNPP. For WT, the value in light gray is for the full cognate diester substrate, dT5'-pNPP (Table 1A). The values for the pruned enzymes with the minimal diester are shown in black. (C) Specificity of the enzymes as ratio of reaction rates of monoester and diester hydrolysis, for WT (gray) and pruned enzymes (black). For the WT enzymes, the gray bars are obtained from the corresponding bars in parts (A) and (B) for mono- and di-esters, respectively. The values for the pruned enzymes are near unity and therefore difficult to see. Ratios for the reactions are given in Supplemental Table 3.

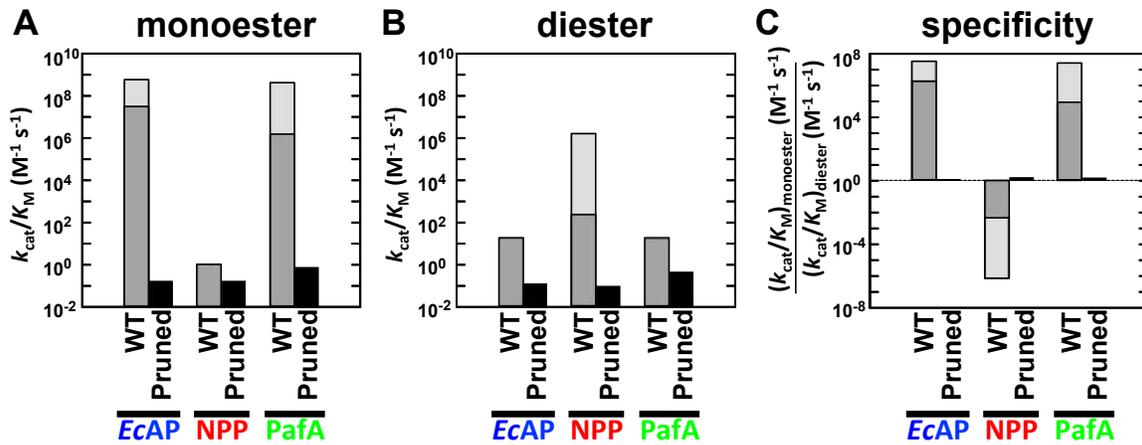


Figure 5. Rate enhancements by pruned enzymes. Rate enhancements for hydrolysis of substrates native to the AP superfamily and the phosphate triester paraoxon by the pruned enzymes, with values for *EcAP** in blue, *NPP** in red, and *PafA** in green. Arrows for sulfate ester hydrolysis signify upper limits. Values are from Supplemental Table 4.

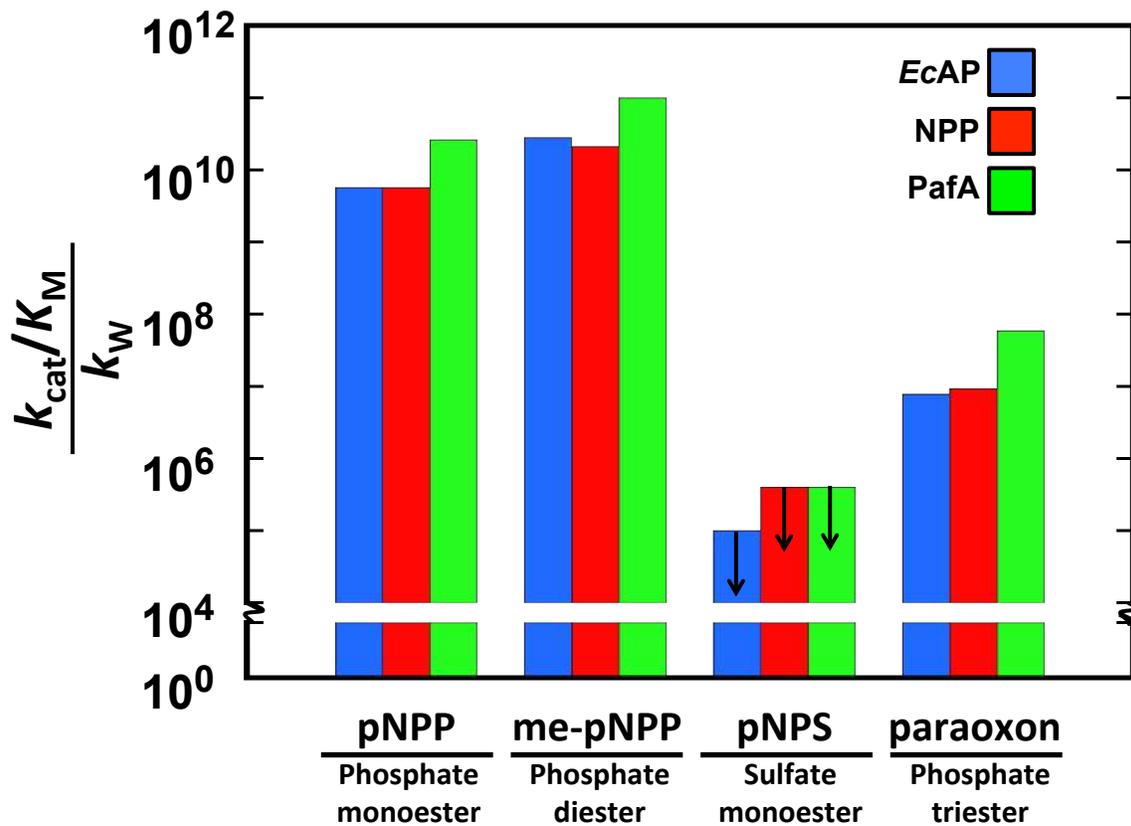
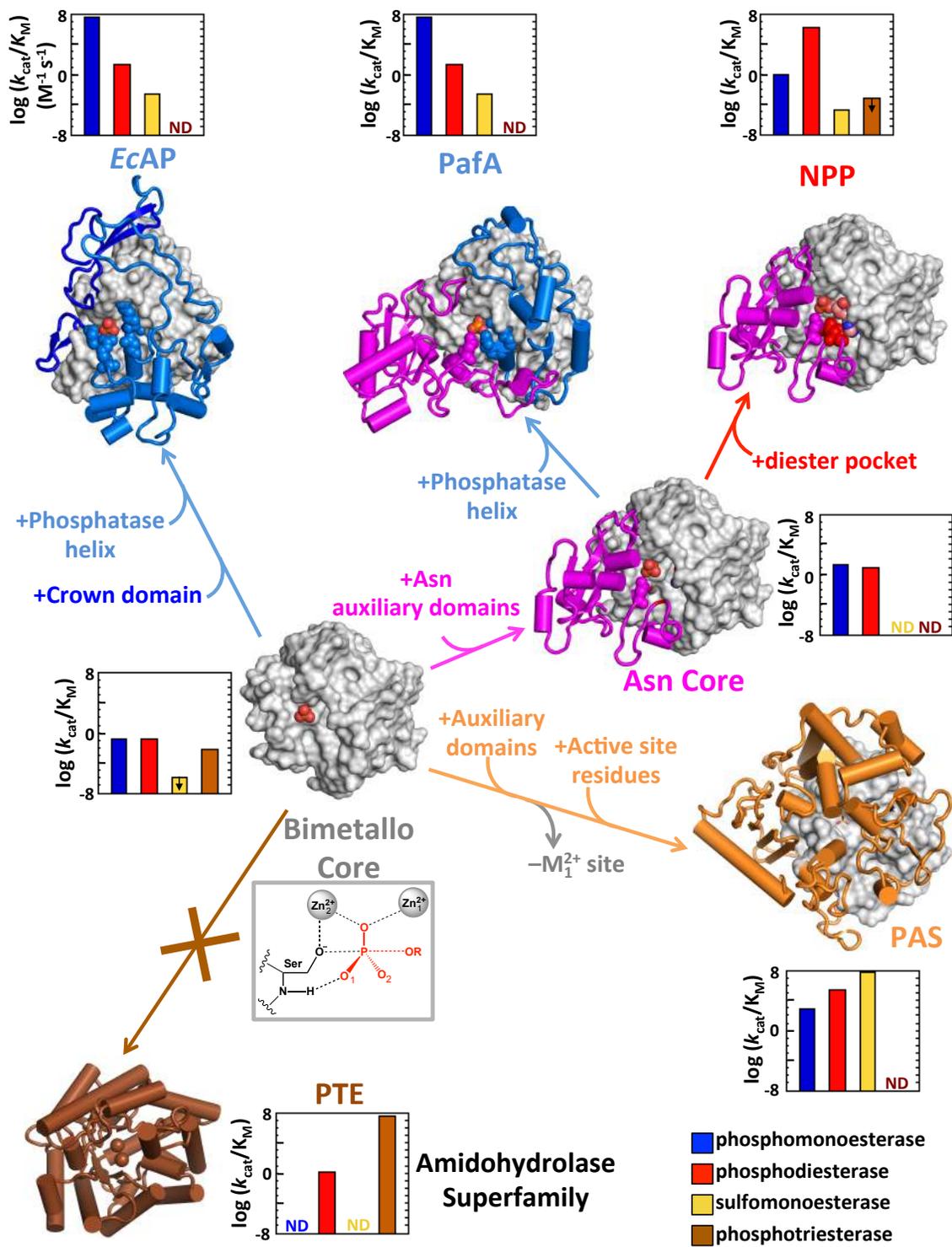


Figure 6. Building catalysis and specificity from the AP superfamily bimetallo core. The bimetallo core (gray, hypothetical core extracted from NPP (2GSN)(34)) is a phosphate monoesterase and diesterase, but a poor sulfatase and a less efficient triesterase. This low catalytic promiscuity may have led to the loss of one of the metal ion sites in proceeding to the AP superfamily sulfatases (orange). PafA and NPP, despite carrying out phosphate monoesterase and diesterase reactions, respectively, share common “Asn auxiliary domains” (magenta)(40). AP superfamily phosphate monoesterases, such as *EcAP* and PafA, contain a common “Phosphatase Helix” (light blue), whereas this helix is absent in diesterases and the resultant cavity contains residues involved in diester substrate binding (“Diester pocket”, red)(40). The AP superfamily lacks phosphate triesterases, and the interaction of the nucleophile backbone amide with one of the monoester or diester phosphoryl oxygen atoms (Fig. 2D) may render the AP superfamily scaffold difficult to evolve or reengineer as a triesterase. The rate for each enzyme construct is shown in graphically next to the structures, phosphomonoesterase in blue, phosphodiesterase in red, sulfatase in orange, and phosphotriesterase in brown. The values for WT *EcAP*, PafA, NPP and *EcAP** are from Tables 1 and 2. The values for the Asn core are from Sunden et al. (40). The PAS kinetic constants are from Babbitt et al. (63), and those for PTE from the amidohydrolase superfamily are from Mohamed et al. (88). The schematic lays out possible evolutionary relationships, but the ancient origin of the AP superfamily prevents evolutionary ordering.



**Differential Catalytic Promiscuity of the Alkaline Phosphatase Superfamily Bimetallo
Core Reveals Mechanistic Features Underlying Enzyme Evolution**

Fanny Sunden, Ishraq AlSadhan, Artem Lyubimov, Tzanko Doukov, Jeffrey Swan and
Daniel Herschlag

J. Biol. Chem. published online October 25, 2017

Access the most updated version of this article at doi: [10.1074/jbc.M117.788240](https://doi.org/10.1074/jbc.M117.788240)

Alerts:

- [When this article is cited](#)
- [When a correction for this article is posted](#)

[Click here](#) to choose from all of JBC's e-mail alerts

Supplemental material:

<http://www.jbc.org/content/suppl/2017/10/25/M117.788240.DC1>