

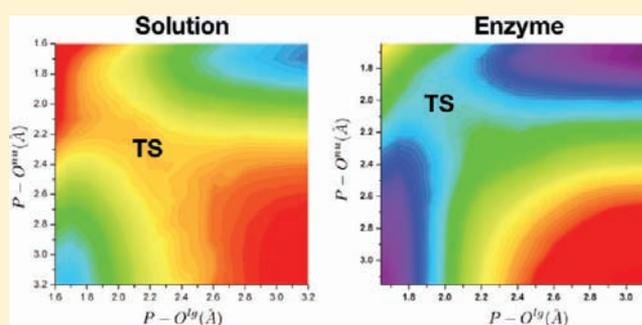
QM/MM Analysis Suggests That Alkaline Phosphatase (AP) and Nucleotide Pyrophosphatase/Phosphodiesterase Slightly Tighten the Transition State for Phosphate Diester Hydrolysis Relative to Solution: Implication for Catalytic Promiscuity in the AP Superfamily

Guanhua Hou[†] and Qiang Cui^{*,†}

[†]Department of Chemistry and Theoretical Chemistry Institute, University of Wisconsin, Madison, 1101 University Avenue, Madison, Wisconsin 53706, United States

S Supporting Information

ABSTRACT: Several members of the Alkaline Phosphatase (AP) superfamily exhibit a high level of catalytic proficiency and promiscuity in structurally similar active sites. A thorough characterization of the nature of transition state for different substrates in these enzymes is crucial for understanding the molecular mechanisms that govern those remarkable catalytic properties. In this work, we study the hydrolysis of a phosphate diester, MpNPP⁻, in solution, two experimentally well-characterized variants of AP (R166S AP, R166S/E322Y AP) and wild type Nucleotide pyrophosphatase/phosphodiesterase (NPP) by QM/MM calculations in which the QM method is an approximate density functional theory previously parametrized for phosphate hydrolysis (SCC-DFTBPR). The general agreements found between these calculations and available experimental data for both solution and enzymes support the use of SCC-DFTBPR/MM for a semiquantitative analysis of the catalytic mechanism and nature of transition state in AP and NPP. Although phosphate diesters are cognate substrates for NPP but promiscuous substrates for AP, the calculations suggest that their hydrolysis reactions catalyzed by AP and NPP feature similar synchronous transition states that are slightly tighter in nature compared to that in solution, due in part to the geometry of the bimetallic zinc motif. Therefore, this study provides the first direct computational support to the hypothesis that enzymes in the AP superfamily catalyze cognate and promiscuous substrates via similar transition states to those in solution. Our calculations do not support the finding of recent QM/MM studies by López-Canut and co-workers, who suggested that the same diester substrate goes through a much looser transition state in NPP/AP than in solution, a result likely biased by the large structural distortion of the bimetallic zinc site in their simulations. Finally, our calculations for different phosphate diester orientations and phosphorothioate diesters highlight that the interpretation of thio-substitution experiments is not always straightforward.



I. INTRODUCTION

Although a high-level of catalytic specificity has been regarded as an important hallmark of enzymes, it is increasingly recognized that many enzymes have promiscuous catalytic activities.^{1–5} Moreover, it has been proposed that catalytic promiscuity plays an important role in enzyme evolution since it can give an enzyme an evolution “head start”, providing a modest rate enhancement that is sufficient as a selective advantage.^{1,6–9} Therefore, identifying factors that dictate the level of catalytic promiscuity in enzymes can help better understand enzyme evolution and improve design strategies for evolving new catalytic functions.

In this context, members in the Alkaline Phosphatase (AP) superfamily present striking examples of catalytic specificity and promiscuity.^{10,11} They have been demonstrated to catalyze the hydrolysis of a broad range of substrates that differ in charge, size, intrinsic reactivity, and transition state (TS) nature.¹² For

example, *E. coli* AP not only catalyzes the hydrolytic reaction of phosphate monoesters for its physiological function but also exhibits promiscuous activity for the hydrolysis of phosphate diesters and sulfate esters of diverse structural/chemical features. Similarly, although the main function of Nucleotide pyrophosphatase/phosphodiesterase (NPP) is to hydrolyze phosphate diesters, it can also cleave phosphate monoesters and sulfate esters with considerable acceleration over solution reactions. The catalytic proficiencies (defined by the ratio of k_{cat}/K_M and rate of the uncatalyzed reaction in solution, k_w) vary greatly, ranging from $>10^{27}$ for the cognate activity^{13,14} to $\sim 10^6$ for the promiscuous activity.¹⁵ The reaction specificities of AP and NPP (characterized by ratios of k_{cat}/K_M for cognate and promiscuous substrates in the two enzymes) for phosphate

Received: June 7, 2011

Published: November 18, 2011

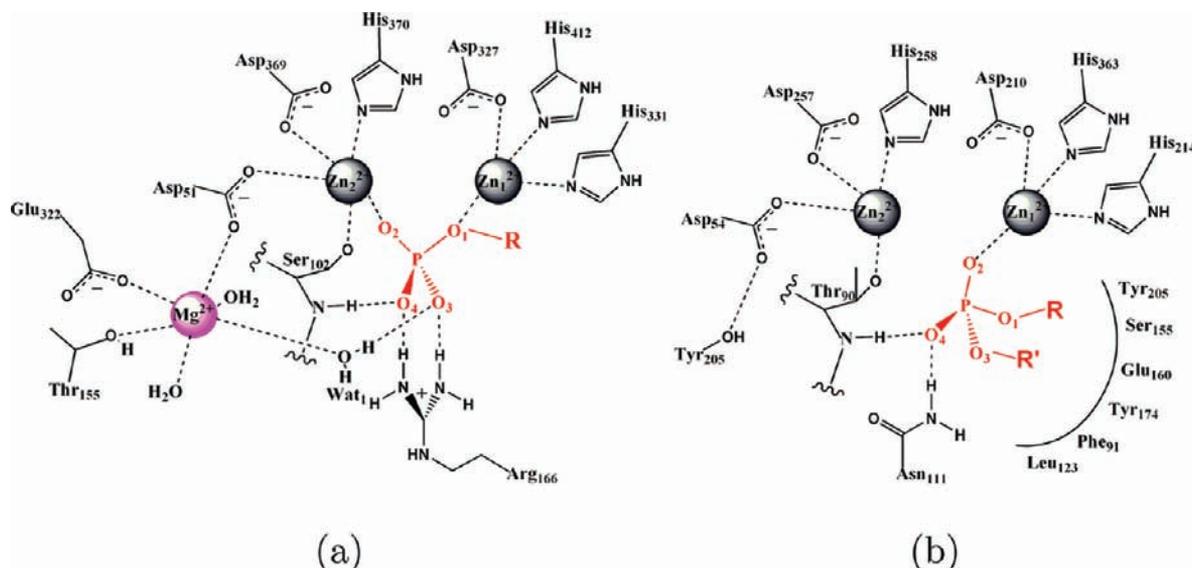


Figure 1. The active sites of Alkaline Phosphatase (AP) and Nucleotide pyrophosphatase/phosphodiesterase (NPP) are generally similar, with a few distinct differences. (a) *E. coli* AP active site. (b) *Xac* NPP active site. The cognate substrates for AP and NPP are phosphate monoesters and diesters, respectively. The labeling scheme of substrate atoms is used throughout the paper. We propose that diesters and monoesters have different binding modes in the active site (see Section III.B for discussions).

mono- and diesters differ by up to a remarkable level of 10^{15} fold.^{14,16} These significant levels of catalytic specificity and promiscuity are particularly striking in light of the fact that AP and NPP are very similar in their active site features yet have limited sequence identity (8%): as illustrated in Figure 1, both AP and NPP feature a highly conserved bimetallic zinc active site with the same set of metal ligands (three Asp and three His residues). These characteristics make this pair of enzymes ideal for in-depth comparative analyses, i.e., to understand how they combine the high levels of catalytic proficiency and promiscuity by making use of similar active sites.

Extensive work has been carried out to characterize the structure and function of AP and NPP. Crystal structures^{16,17} show that, in spite of the similarities, several differences can be noted between these enzymes. First, the AP active site has additional positively charged motifs, in particular a magnesium ion and Arg166; these are replaced in NPP by charge-neutral residues, Thr205 and Asn111, respectively. The extra positive charges in AP likely help stabilize phosphate monoesters over diesters due to difference in the charge states of these substrates (dianionic vs monoanionic). Second, the NPP active site is featured with additional hydrophobic residues (e.g., Leu123, Phe91 and nearby residues, see Figure 1), which are expected to help bind diesters more tightly than monoesters. Motivated by these observed differences, systematic analyses over the past few years have helped quantitatively account for a significant fraction of the 10^{15} fold differential catalytic specificity in AP and NPP.^{14,16} (1) Arg166 in AP interacts favorably with two negatively charged nonbridging phosphoryl oxygen atoms present in phosphate monoesters but not diesters, giving a preference to monoesters of $\sim 10^4$ fold; (2) the hydrophobic R' binding pocket of NPP provides $\sim 10^4$ fold preference to diester catalysis; (3) the Mg^{2+} site in AP contributes through water-mediated hydrogen-bonding interaction with the transferred phosphoryl group, which bears less negative charge in the case of diesters, to favor the monoester reaction by a $\sim 10^4$ fold.

Despite progress, crucial questions remain to be answered for catalysis in AP and its superfamily members. In particular, a

fundamental hypothesis regarding catalytic promiscuity in AP/NPP, which was motivated by experimental linear free energy relation (LFER¹⁸) and kinetic isotope effect (KIE) data,^{19–23} is that AP and NPP do not alter the nature of phosphoryl transfer TS relative to solution reactions, instead they recognize and stabilize TSs of different nature for cognate and noncognate substrates. This property has been proposed to assist in the evolutionary optimization of promiscuous activities and challenges the traditional notion that an enzyme active site is evolved to stabilize a single type of TS. Recent QM/MM calculations^{24–26} using the AM1(d)-PhoT method²⁷ as the QM level, however, do not seem to support this model. Although the calculations found that phosphate monoester hydrolysis in AP proceeds via a loose TS,²⁴ similar to in aqueous solution, the TS for phosphate diester was found to change from synchronous in solution to very loose in both NPP²⁵ and AP.²⁶ The latter is in contrast to conclusions from LFER analysis for phosphate diester hydrolysis in AP,²² in which the TS is determined to be synchronous, similar to its solution reference. Nevertheless, citing the previous discussion²⁸ of ambiguity in using LFER data to infer the structure of TS, the authors of ref.²⁶ proposed a picture for the evolution (and catalytic promiscuity) of the AP superfamily in which the nature of the TS (loose) is maintained for different substrates (e.g., mono- and diesters);²⁶ this scenario has been established to explain catalytic promiscuity observed for protein phosphatase-1.²⁹ It should be noted, however, that whether the computational method was sufficiently reliable in the recent QM/MM studies is not clear; for example, the Zn^{2+} - Zn^{2+} distance was found to vary greatly during the reaction for both mono- and diester substrates in AP and NPP,^{24–26} reaching 7.0 Å as compared to the value of ~ 4 Å in the crystal structure^{16,30} and other structural characterizations (EXAFS).³¹

To help clarify the situation, we set out to use combined QM/MM potential to systematically investigate the hydrolysis reaction of various cognate/noncognate substrates of AP/NPP in solution and enzymes. In this paper, we focus on the hydrolysis reactions for the same diester substrate studied in

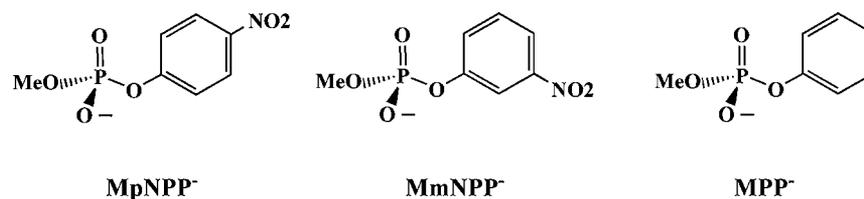


Figure 2. Methyl p-nitrophenyl phosphate (MpNPP⁻) and its two diester analogs studied in this work.

previous QM/MM calculations,^{24,25} MpNPP⁻ (Figure 2), and its phosphorothioate analog (MpNPPS⁻), in solution, two experimentally well-characterized variants of AP (R166S and R166S/E322Y), and the wild type NPP. Since the active sites of AP and NPP are fairly open and readily accessible to solvents (which is what made it possible to carry out LFER studies for these systems), conformational sampling is expected to be crucial. This consideration together with the fairly large size of the bimetallic zinc catalytic center suggest that an appropriate approach is to use the Self-Consistent-Charge Density-Functional-Tight-Binding (SCC-DFTB)³² as QM in a QM/MM framework. With respect to computational efficiency, SCC-DFTB is comparable to widely used semiempirical methods such as AM1 and PM3,³³ i.e., it is 2–3 orders of magnitude faster than popular DFT methods. In terms of accuracy, fairly extensive benchmark calculations have indicated that it is particularly reliable for structural properties, while energetics are generally comparable to AM1 and PM3.^{34–36} For phosphoryl transfer reactions, however, a reaction-specific parametrization based on hydrolysis reactions of model phosphate species, referred to as SCC-DFTBPR,³⁷ appears to be more effective than standard semiempirical methods and has been found successful in several applications to solution and enzyme systems.^{38–40}

Here we further test the reliability of SCC-DFTBPR for MpNPP⁻ in different environments (solution, AP and NPP) by comparing results to higher-level QM (QM/MM) calculations as well as available experimental data. A more systematic comparison with LFER and KIE data requires much more extensive calculations and is left as a separate study. Nevertheless, the encouraging benchmark results obtained so far suggest that the SCC-DFTBPR based QM/MM approach can be used to probe the nature and energetics of phosphoryl transfer TS in AP and NPP at a semiquantitative level. In contrast to recent QM/MM calculations,^{24,25} which found a much looser TS in NPP than in solution, the results here support that the nature of the phosphoryl transfer TS for phosphate diesters is not loosened in neither AP nor NPP relative to solution; in fact, the TS becomes slightly tighter in AP and NPP than in solution, due in part to the geometry of the bimetallic zinc motif. Therefore, our study highlights the importance of using a carefully benchmarked QM/MM model to investigate the nature of phosphoryl transfer TS; moreover, these data provide the first explicit computational support of the hypothesis that the nature of TS for the same substrate is similar in the AP family and in solution.

The paper is organized as follows: in Section II we summarize computational methods and simulation setup. In Section III, we first present results for the reference solution reactions and then analysis of the phosphoryl transfer TS for phosphate diesters in several variants of AP and wild type NPP; we also analyze the effect of thio substitution of the diester substrate, which was used experimentally to probe the orientation of the substrate in the active site. Since the Zn²⁺-

Zn²⁺ distance exhibits rather different behaviors in this and previous QM/MM simulations of AP/NPP,^{24,25} we also explicitly analyze the impact of this fundamental geometrical feature of the bimetallic zinc site on the catalysis. Before concluding in Section IV, we summarize the key differences between our and recent QM/MM studies^{24–26} and also a number of issues that we recommend to examine by future experiments.

II. COMPUTATIONAL METHODS

A. Diester Hydrolysis in Solution with the SCC-DFTBPR Based Implicit Solvent Model. As an important benchmark and reference, we first study the hydrolysis of MpNPP⁻ and two of its analogs (Figure 2) in solution, which have been thoroughly studied experimentally. To focus on the quality of the QM model rather than other technical details such as QM/MM coupling and sampling, we use the implicit solvent model that we have implemented and parametrized for SCC-DFTBPR.⁴¹ In this model, the solute radii are dependent on the charge distribution, which makes it particularly useful for studying solution reactions that involve highly charged species; our previous benchmark calculations suggest that the method has comparable accuracy as the SM6 model,⁴² while being much more efficient (due to the use of SCC-DFTB) and having only a small number of parameters.

The potential energy surface (PES) relevant to the hydrolysis reaction is first explored by adiabatic mapping calculations, in which the reaction coordinates are the P–O^{lg} and P–O^{nu} distances; here hydroxide is the nucleophile, and “O^{lg}” and “O^{nu}” indicate the reactive oxygen in the leaving group and nucleophile, respectively. Each point in the 2D PES is obtained by starting the constrained optimization from several different initial structures and taking the lowest energy value. Following the adiabatic mapping calculations, structures along the approximate reaction path are examined carefully to ensure that the change of geometry is continuous along the path; subsequently, the saddle point is fully optimized by conjugated peak refinement (CPR)⁴³ to obtain more precise TS structure and energy. Finally, frequency calculations are carried out to confirm the nature of the stationary points and to compute the vibrational entropy and zero point energies to obtain approximate activation free energy; although using a harmonic approximation to estimate activation entropy is known to be of limited accuracy, previous studies of phosphate diester hydrolysis found that activation entropy does not differ much between different diesters.²⁸ The vibrational frequencies are also used to estimate ¹⁸O kinetic isotope effects (KIEs) for MpNPP⁻ as an additional benchmark of the methodology (see the Supporting Information).

To correct for intrinsic errors of SCC-DFTBPR energies, we explore corrections based on gas phase single-point energy calculations with both B3LYP/6-311++G** and MP2/6-311++G** at SCC-DFTBPR geometries; the B3LYP level was found to give very similar results to MP2 for simple phosphate hydrolysis reactions³⁷ (however, see below). As discussed in the literature,⁴⁴ such a simple correction may not always improve the energetics for semiempirical methods given the errors in geometry; however, our previous tests^{37,41} indicated that this correction scheme appears useful for SCC-DFTBPR since the method gives fairly reliable structures, even for transition states.

To further facilitate the analysis of sources of errors in both QM and QM/MM calculations, additional analysis for gas-phase/solution

proton affinities (PA) for the leaving groups in MpNPP⁻ and its analogs. QM-only calculations are carried out by Gaussian03,⁴⁵ PCM⁴⁶ and SM6⁴² models are employed to describe solvation effects. To test the accuracy of QM/MM coupling, solution PAs are also calculated using SCC-DFTB/MM based free energy perturbation calculations.⁴⁷

B. Enzyme Model Setup. For the hydrolytic reaction catalyzed by the AP superfamily members, a two-step mechanism is usually followed,⁴⁸ in which an oxygen nucleophile (e.g., Ser or Thr) first attacks the phosphorus/sulfur, then a water (hydroxide) replaces the leaving group in a step that is essentially the reverse of the first; for some family members of the superfamily, however, the mechanism can be more complex.⁴⁹ In this work, to understand the catalytic mechanism of AP with phosphate diesters, we investigate the first step of the hydrolysis reaction of MpNPP⁻ in an *E. coli* AP variant in which Arg166 is mutated to Ser; this is expected to be the rate-limiting step given the experimental leaving group LFER analysis. Experimentally, this mutant was used to avoid inhibition by P_i,²² and it is believed that the mutation does not alter the reaction mechanism of AP since LFERs are similar in the mutant and WT.^{20,50} Moreover, the chemical step is fully rate-limiting in this mutant. We also study a double mutant, R166S/E322Y AP, which was constructed in recent experimental studies to analyze the contribution(s) from the Mg²⁺ site in AP. For NPP, we study the wild type enzyme from *Xanthomonas axonopodis pv citri* (*Xac*).

The enzyme models are constructed based on the X-ray structures for the *E. coli* AP mutant R166S with bound inorganic phosphate at 2.05 Å resolution (PDB code 3CMR³⁰) and *Xac* NPP with bound Adenosine Mono-Phosphate (AMP) at 2.00 Å resolution (PDB code 2GSU¹⁶). The enzyme model for R166S/E322Y AP is constructed based on the crystal structure of E322Y AP (PDB code 3DYC¹⁴) by mutating Arg166 to serine. In each case, starting from the PDB structure, the ligand is first “mutated” to the substrate of interest, MpNPP⁻ or MpNPPS⁻; two possible orientations of the substrates are considered for the AP active site, with the -OMe group oriented toward either the magnesium ion or Ser102 backbone amide (see additional discussions in Section III). Hydrogen atoms are added by the HBUILD module⁵¹ in CHARMM.⁵² All basic and acidic amino acids are kept in their physiological protonation states except for Ser102 and Thr90 in AP and NPP, respectively, which are assumed to be the nucleophiles and deprotonated in the reactive complex. Water molecules are added following the standard protocol of superimposing the system with a water droplet of 27 Å radius centered at Zn1²⁺ (see Figure 1 for atomic labels) and removing water molecules within 2.8 Å from any atoms resolved in the crystal structure.⁵³ Protein atoms in the MM region are described by the all-atom CHARMM force field for proteins,⁵⁴ and water molecules are described with the TIP3P model.⁵⁵ The QM region includes groups most relevant to the reaction: the two zinc ions and their 6 ligands (Asp51, Asp369, His370, Asp327, His412, His331), Ser102 and MpNPP⁻ for R166S AP; for NPP, this includes two zinc ions and their 6 ligands (Asp54, Asp257, His258, Asp210, His363, His214), Thr90 and MpNPP⁻. Only side chains of protein residues are included in the QM region, and link atoms are added between C_α and C_β atoms. A larger QM region also has been tested for R166S AP which further incorporates the entire magnesium site, including Mg²⁺, side chains of Thr155, Glu322 and three ligand water molecules. Comparison of optimized structures using different QM regions indicates fairly similar optimized structures (see the Supporting Information for details), thus the smaller QM region is used for the majority of the calculations. The treatment of the QM/MM frontier follows the DIV scheme in CHARMM; previous benchmark calculations have shown that this scheme generally gives reliable results for structure and energetics in QM/MM calculations provided that the MM charge is small near the QM/MM boundary.⁵⁶ Since the Mg²⁺ near the QM region in AP is treated as a point charge (otherwise the QM region will become substantially larger), to avoid overpolarization of nearby QM groups, a NOE potential is added to the C–O bonds in Asp51, which is coordinated to both Mg²⁺ and Zn²⁺. The NOE potential takes the form

$$E_R = 0.0 \quad R_{\min} < R < R_{\max} \\ = 0.5 \cdot K_{\max} \cdot (R - R_{\max})^2 \quad R_{\max} < R \quad (1)$$

in which R_{\min} and R_{\max} set the interval between which the restraining potential is zero; they are taken to be 0 and 1.28 Å, respectively. K_{\max} is set to be 10⁴ kcal/(mol · Å²).

Due to the fairly large size of the QM region (more than 80 atoms) and extensive sampling required for the open active site of AP and NPP, the SCC-DFTBPR method³⁷ is used for PMF calculations. Extensive benchmark calculations and applications indicate that it is comparable to the best semiempirical method available in the literature for phosphate chemistry.^{27,57}

The generalized solvent boundary potential (GSBP)^{58,59} is used to treat long-range electrostatic interactions in geometry optimizations and MD simulations. The system is partitioned into a 27-Å spherical inner region centered at the Zn1 atom, with the rest in the outer region. Newtonian equations-of-motion are solved for the MD region (within 23 Å), and Langevin equations-of-motion are solved for the buffer regions (23–27 Å) with a temperature bath of 300 K; protein atoms in the buffer region are harmonically constrained with force constants determined from the crystallographic B-factors.⁶⁰ All bonds involving hydrogen are constrained using the SHAKE algorithm,⁶¹ and the time step is set to 1 fs. All water molecules in the inner region are subject to a weak GEO type of restraining potential to keep them inside the inner sphere with the MMFP module of CHARMM. The static field due to outer-region atoms, ϕ_s^i , is evaluated with the linear Poisson–Boltzmann (PB) equation using a focusing scheme with a coarse cubic grid of 1.2 Å spacing and a fine grid of 0.4 Å spacing. The reaction field matrix M is evaluated using 400 spherical harmonics. In the PB calculations, the protein dielectric constant of $\epsilon_p = 1$, the water dielectric constant of $\epsilon_w = 80$, and 0.0 M salt concentration are used; the value of ϵ_p is not expected to make a large difference in this particular case because the active site is already very solvent accessible and the inner/outer boundary is far from the site of interest. The optimized radii of Nina et al.^{62,63} based on experimental solvation free energies of small molecules as well as the calculated interaction energy with explicit water molecules are adopted to define the solvent–solute dielectric boundary. To be consistent with the GSBP protocol, the extended electrostatic model⁶⁴ is used to treat the electrostatic interactions among inner region atoms in which interactions beyond 12 Å are treated with multipolar expansions, including the dipolar and quadrupolar terms.

C. Benchmark Enzyme Calculations Based on Minimizations and Reaction Path Calculations. To further test the applicability of SCC-DFTBPR/MM to AP and NPP, geometry optimization for the reactant (Michaelis) complex is compared to results from B3LYP^{65–67}/MM calculations. The basis set used in the B3LYP/MM calculations is 6-31G*,⁶⁸ and the calculations are carried out with the QChem⁶⁹ program interfaced with CHARMM(c36a2 version).⁷⁰ Due to the rather large size of the QM region and the high cost of *ab initio* QM/MM calculations, atoms beyond 7 Å away from Zn1 are fixed to their crystal positions in these minimizations (Note that these are not fixed in the potential of mean force simulations, see below. Also, test calculations at the SCC-DFTBPR level show that fixing atoms beyond 7 Å from Zn1 in minimizations do not lead to much difference as compared to a fully flexible inner-region calculation within the GSBP framework.). The convergence criteria for geometry optimization are that the root-mean-square (rms) force on mobile atoms is smaller than 0.30 kcal/(mol · Å) and the maximum force smaller than 0.45 kcal/(mol · Å).

The Minimum Energy Path (MEP) calculations are carried out by one-dimensional adiabatic mapping at both SCC-DFTBPR and B3LYP/6-31G* levels; the reaction coordinate is the antisymmetric stretch involving the breaking and forming P–O bonds (PO^g.PO^m), and the step size for the adiabatic mapping is 0.2 Å. At the SCC-DFTBPR level, the transition state is further refined using CPR.

D. 1D and 2D Potential of Mean Force (PMF) Simulations. To study the free energy profile of enzyme reactions, PMF simulations have been carried out for R166S AP, R166S/E322Y AP, and NPP with

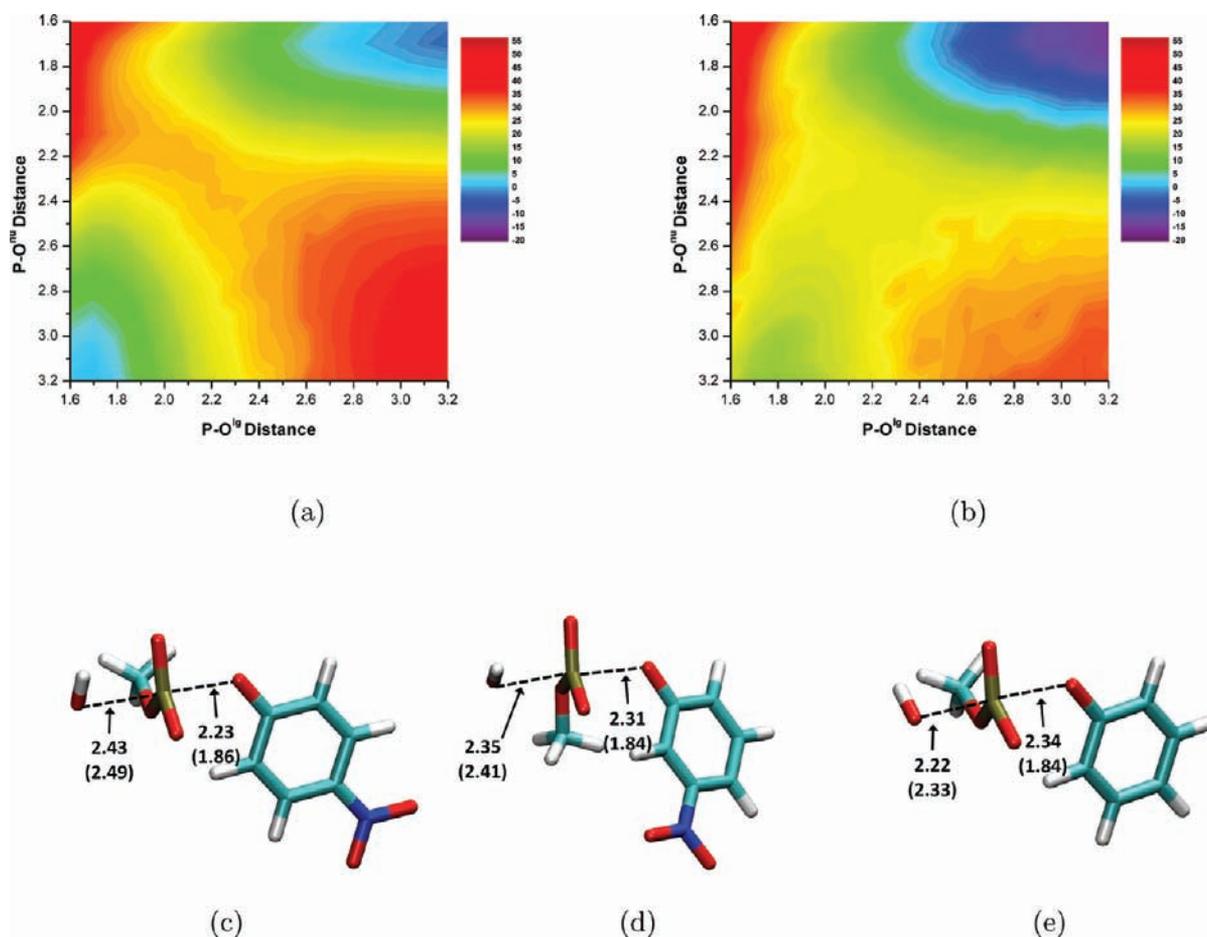


Figure 3. Aqueous hydrolysis of phosphate diesters with hydroxide as the nucleophile. Key distances are labeled in Å and energies are in kcal/mol. (a) Adiabatic mapping results for MpNPP⁻ by SCC-DFTBPR/PB. (b) Adiabatic mapping results for MpNPP⁻ after including single point gas phase correction at the MP2/6-311++G** level. (c-e) Hydrolysis transition state optimized with Conjugate Peak Refinement (CPR) calculations for MpNPP⁻, MmNPP⁻, and MPP⁻. Numbers without parentheses are obtained by SCC-DFTBPR/PB; those with parentheses are taken from ref 28. As shown in the Supporting Information, including the MP2 correction tends to slightly tighten the transition state, especially along P-O^{lg}.

MpNPP⁻ and MpNPPS⁻ as the substrates. After the initial minimizations starting from the relevant crystal structure, the enzyme system is slowly heated to 300 K and equilibrated for 100 ps. The reaction coordinate is defined as P-O^{lg}-P-O^{mu}. The umbrella sampling approach⁷¹ is used to constrain the system along the reaction coordinate by using a force constant of 150 kcal/mol·Å⁻². In total, more than 51 windows are used for each PMF, and 100 ps simulations are performed for each window. The first 50 ps trajectories are discarded, and only the last 50 ps are used for data analysis. Convergence of the PMF is monitored by examining the overlap of reaction coordinate distributions sampled in different windows and by evaluating the effect of leaving out segments of trajectories. The probability distributions are combined together by the weighted histogram analysis method (WHAM)⁷² to obtain the PMF along the reaction coordinate. The averaged key structural properties for each window are calculated and summarized in Table 4.

In a separate set of PMF calculations, the Zn²⁺-Zn²⁺ distance is constrained to be 3.6, 4.1, and 4.6 Å, respectively, by a strong constraint with a force constant of 2,000 kcal/mol·Å⁻², to investigate the impact of this fundamental variable of the bimetallic site on catalysis in AP and NPP. For reference, the Zn²⁺-Zn²⁺ distance found in the various crystal structures for AP and NPP is close to be 4.1 Å.

To verify that the 1D PMFs capture the nature of the phosphoryl transfer transition state, we also carry out 2D PMF calculations for the α orientation of MpNPP⁻. The reaction coordinates are defined as the P-O^{lg} and P-O^{mu} distances, and the range of each distance is similar to that in the 1D PMF calculations. In total, 272 windows are used, and a force constant of 200 kcal/mol·Å⁻² is used for all windows; for

each window, 100 ps simulations are carried out, and only the last 50 ps are used for the subsequent WHAM analysis.

III. RESULTS AND DISCUSSION

A. MpNPP⁻ Hydrolysis in Solution. The hydrolysis of MpNPP⁻ has been studied extensively by experiments and computations. Experimental studies²² determined the activation free energy of 25.7 kcal/mol at 42 °C with hydroxide as the nucleophile, and the mechanism is established as concerted with a synchronous TS based on LFER analysis. Several computational work also studied the same reaction by employing various levels of theory. By using B3LYP/6-31+G* and the PCM model, Rosta and co-workers obtained a fairly loose TS with P-O^{lg} and P-O^{mu} as 1.86 and 2.49 Å, respectively.²⁸ By using B3LYP/6-311+G** with the COSMO continuum model on PCM-minimized geometries and a careful treatment of solute configurational entropy, they obtained an activation free energy barrier of 24.4 kcal/mol. In the more recent QM/MM simulations using explicit solvent (TIP3P) and AM1(d)-PhoT²⁷ as QM, López-Canut et al. obtained a free energy barrier of 20.5 kcal/mol; the transition state was featured with the P-O^{lg} and P-O^{mu} distances of 1.81 and 2.23 Å, respectively, somewhat more compact compared with the PCM result.

With our SCC-DFTB(PR)/PB method and charge dependent atomic radii,⁴¹ the adiabatic map for MpNPP⁻ hydrolysis (Figure 3a) is qualitatively consistent with previous studies and indicates a synchronous TS with an energy barrier of around 30 kcal/mol. After adding higher level (B3LYP or MP2) single point energy corrections, the general landscape of the adiabatic map does not change (Figure 3b); the synchronous TS is still preferred with the barrier lowered to around 25 kcal/mol when MP2 corrections are used. After CPR refinement, the fully optimized TS (Figure 3c) has a P–O^{lg} of 2.23 Å and P–O^{nu} of 2.43 Å; i.e., the P–O^{lg} distance is longer compared with previous theoretical results, while P–O^{nu} is similar. We note, however, the PES is rather flat near the transition state. The free energy barrier by including ZPE and solute configurational entropy is 29.3 kcal/mol at the SCC-DFTBPR level, which is decreased to 24.4 kcal/mol by including gas phase single point energy correction at the MP2/6-311++G** level (Table 1); the latter value compares favorably with experimental value.

Table 1. Energetics for Diester Hydrolysis Reactions in Solution from Experiments and Calculations

diester	pK_a	ΔG_{exp}^a	ΔG_{lit}^b	ΔG_{lit}^c	ΔG_{calc}^d
MpNPP ⁻	7.14	25.7	24.4	20.5	29.3/21.3/24.4
MmNPP ⁻	8.35	26.3	27.3		33.3/28.1/27.2
MPP ⁻	9.95	28.6	29.9		39.8/30.5/30.6

^aExperimental result taken from ref 22. ^bCalculation result taken from ref 28. ^cCalculation result taken from ref 25. ^dFor each entry, the numbers are as follows: SCC-DFTBPR/PB, SCC-DFTBPR/PB result including gas-phase B3LYP/6-311++G** single point energy correction, and SCC-DFTBPR/PB result including gas-phase MP2/6-311++G** single point energy correction.

In addition, we study the hydrolysis of two other related diesters (Figure 2), methyl 3-nitrophenyl phosphate (MmNPP⁻) and methyl phenyl phosphate (MPP⁻), with the approach. Experimentally, the trend is that the hydrolysis barrier increases as the pK_a of the leaving group increases (Table 1).²² This trend has been reproduced by a previous theoretical study²⁸ in which the nature of transition states is found to be synchronous and becomes looser as the pK_a of the leaving group decreases; P–O^{lg} ranged from 1.84 to 1.86 Å and P–O^{nu} ranged from 2.33 to 2.49 Å. With our SCC-DFTBPR/PB method, this trend is also qualitatively reproduced, regardless of whether the gas-phase correction at higher level (B3LYP or MP2) is included. At a quantitative level, however, the SCC-DFTBPR/PB barriers are too high, and the effects of the substitution are overestimated (see Table 1); for example, the barrier difference between MpNPP⁻ and MPP⁻ is only 2.9

kcal/mol according to experiment but 10.5 kcal/mol at the SCC-DFTBPR/PB level. The discrepancy remains fairly large even with B3LYP corrections, while including the MP2 gas-phase corrections significantly improves the agreement with experimental value. The nature of the TS also becomes somewhat tighter (especially P–O^{lg}, by ~0.2 Å) when the MP2 correction is included (see the Supporting Information).

To better understand the quantitative differences between SCC-DFTBPR, B3LYP, and MP2 results, we examine the relative PAs of the leaving groups in the three phosphate diesters in both gas-phase and solution; gas-phase PAs reflect the intrinsic accuracy of the QM method, while solution PA calculations also examine the accuracy of either the implicit solvent model or QM/MM interactions in explicit solvent simulations.⁷³ As shown in Table 2, all DFT methods, which include both SCC-DFTB(PR) and B3LYP, have errors much larger than “chemical accuracy” (1 kcal/mol) for the relative gas-phase PAs, especially concerning the effect of introducing the nitro group; by comparison, MP2 does a much better job. The errors in the relative solution PAs follow the same trend as the relative gas-phase PAs, suggesting that errors in the gas-phase PAs are the major source of error; this is confirmed by the observation that computed relative solvation free energies for the leaving groups considered here are in good agreement with experimental values using both SCC-DFTB(PR) and B3LYP based implicit solvent models, with the exception of B3LYP and UAKS radii (see the Supporting Information).

As additional benchmark for the nature of the TS, ¹⁸O KIE calculations for MpNPP⁻ in solution are carried out with the SCC-DFTBPR transition state and harmonic vibrational frequencies. As shown in Table S2, the trends are in qualitative agreement with experimental results^{74,75} and previous AM1(d)-PhoT and B3LYP results.⁷⁶ On the other hand, we note that SCC-DFTBPR overestimates the magnitude of the KIEs, especially for the effects associated with the leaving group oxygen and the nonbridging oxygen. This is consistent with the trend discussed above that SCC-DFTBPR predicts a solution TS that is looser as compared to previous theoretical calculations, with most notably a weaker and longer P–O^{lg} bond.

In short, the benchmark calculations for the phosphate diesters and their leaving groups suggest that SCC-DFTBPR can provide fairly reliable structural properties of these species and a semiquantitative description of energetics and the nature of hydrolysis transition state, especially for relative trends associated with different substituents on the leaving group.

B. First Step of MpNPP⁻ Hydrolysis in R166S AP.

Based on the crystal structure of the AP R166S mutant complexed with inorganic phosphate, the phosphate ligand is

Table 2. Relative Proton Affinities (in kcal/mol) for Leaving Groups in the Studied Diesters^a

diester	exp ^b	SCC-DFTBPR ^c	SCC-2nd ^c	B3LYP ^d	MP2 ^e
MpNPP ⁻	0	0	0	0	0
MmNPP ⁻	6.6 (1.7)	8.7/9.7 (4.7)	7.6/8.3 (3.3) [4.9]	9.1/11.1 (1.1)	5.5/7.9
MPP ⁻	21.4 (3.8)	28.3/30.0 (10.0)	26.7/27.8 (6.8) [10.1]	25.3/28.2 (2.2)	21.2/24.1

^aSince only relative proton affinities (PAs) are of interest, no zero-point energy or thermal corrections has been included. The numbers without parentheses are gas-phase PAs calculated at gas-phase/solution optimized structures; those with parentheses are solution PAs. Numbers with bracket are obtained by explicit solvent QM/MM free energy perturbation. ^bExperimental values for gas-phase PA are taken from ref 98. The solution PAs are converted based on experimental pK_a differences at 298 K.⁹⁹ ^cThe solution geometries are optimized by SCC-PB⁴¹ at the corresponding level; “SCC-2nd” indicates the standard second-order SCC-DFTB.³² ^dGas-phase geometries are optimized with B3LYP/6-311++G(d,p); solution geometries are optimized with PCM/UAKS at the same level of theory. ^eGas-phase geometries are optimized with B3LYP/6-311++G(d,p); solution geometries are optimized with PCM/UAKS at the same level of theory. Single point energies are calculated with MP2/6-311++G(d,p).

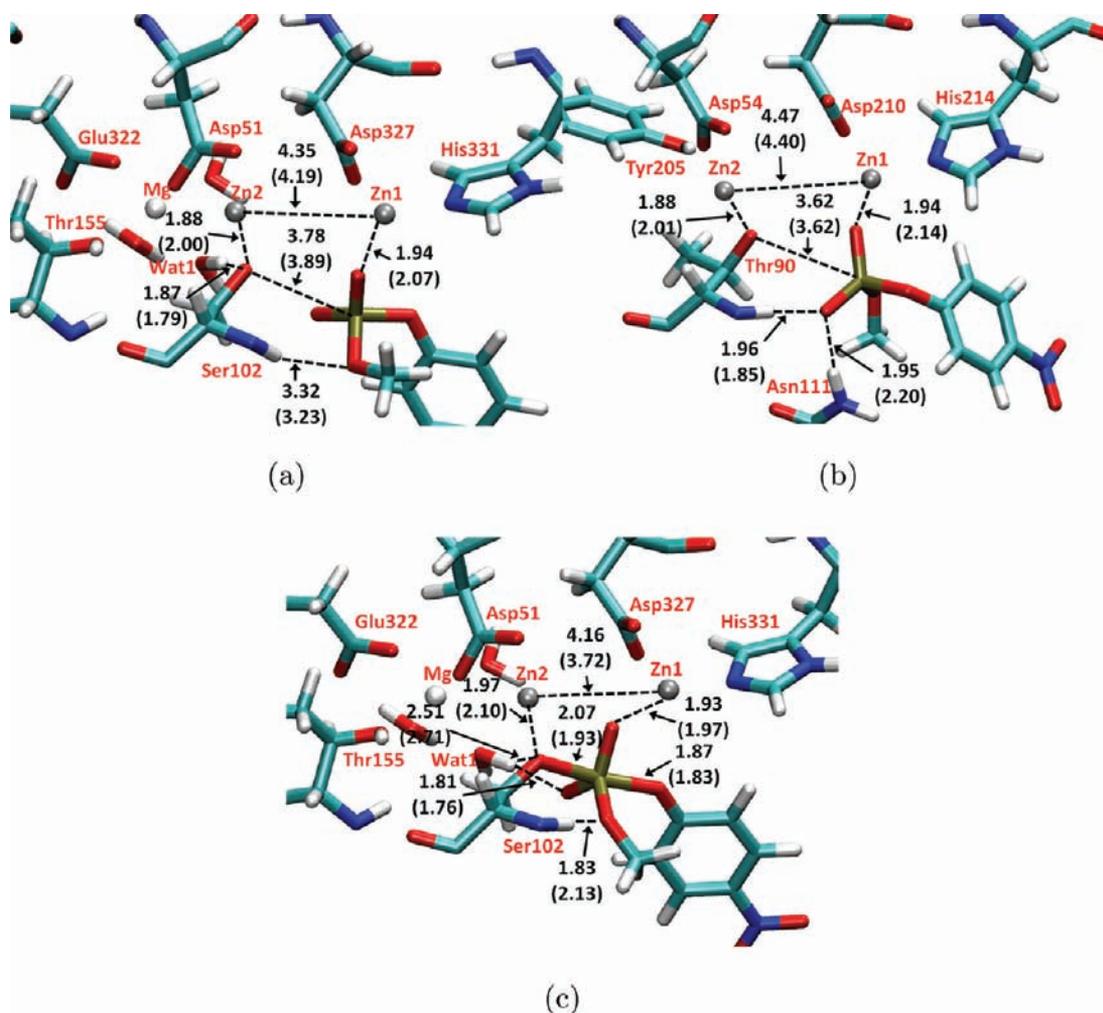


Figure 4. Benchmark calculations for MpNPP⁻ in enzymes. Key distances are labeled in Å. Numbers without parentheses are obtained with B3LYP/6-31G*/MM optimization; those with parentheses are obtained by SCC-DFTBPR/MM optimization. (a) In R166S AP with the substrate methyl group pointing toward Ser102 backbone (the β orientation). (b) In NPP with the substrate methyl group pointing toward the hydrophobic pocket. (c) Comparison of transition state obtained by adiabatic mapping for the β orientation in R166S AP. In (a,c), Asp369, His370, and His412 are omitted for clarity, while in (b), Asp257, His258, and His363 are omitted for clarity.

“mutated” to MpNPP⁻ by adding necessary functional groups to phosphate oxygen. The leaving 4-nitrophenyl group is added to O1 due to the geometrical requirement of the in-line attack from Ser102 (see Figure 1a). The methyl group can be added to O3 or O4, which correspond to two different substrate orientations (denoted as α and β orientations, respectively, following the notation of ref 14). Recent experimental studies¹⁴ using a double mutant AP with the Mg²⁺ site removed and phosphorothioate diesters suggested that the α orientation is preferred over the β orientation. As discussed below (see Section III.C), however, the interpretation of those elegant experiments may not be as clear-cut as presented. Moreover, even if the α orientation is indeed dominant, it is not clear if the discrimination comes from binding or the chemical step. Therefore, it is informative to study both orientations.

The comparison of optimized structures by B3LYP/MM and SCC-DFTBPR/MM shows good agreement between the two levels (Figure 4a). The O_{Ser102}-P distance increases from 3.1 Å in the crystal structure to 3.4 (3.3)/3.8 (3.9) Å in B3LYP/MM (SCC-DFTBPR/MM) optimized structure for the α/β orientation, leading to a stable reactant complex. The O2 of the substrate coordinates to one of the zinc ions and O1 with

the phenyl group is solvated by water molecules. Interestingly, the slight shift of the substrate position also increases the distance between a Mg²⁺-bound water (Wat1) and substrate O3 such that a hydrogen bond is formed between Wat1 and the Zn²⁺-activated Ser102, instead of with an inorganic phosphate oxygen as in the crystal structure (this holds also in MD simulations at the SCC-DFTBPR/MM level; also see discussion below for interactions in the TS). O4 and the nearby Ser102 backbone amide forms the only direct hydrogen bond between the substrate and the enzyme, which is shorter (1.9 Å vs 3.3 Å) with the substrate in the α orientation than in the β orientation. If this is the only major difference between those two orientations, the binding affinity of MpNPP⁻ to the enzyme is likely stronger with the α orientation than with the β orientation, although a more quantitative estimate remains to be carried out with free energy simulations, which we defer to a future study. Many other hydrogen-bonding distances (e.g., between Wat1 and Ser102, MpNPP⁻ and Ser102 backbone amide) and distances involving the zinc ions (e.g., between Ser102/MpNPP⁻ and the zinc ions) are similar at the two levels of theory (see Figure 4). The Zn²⁺-Zn²⁺ distance is generally shorter at the SCC-DFTBPR/MM level (by ~0.1–

Table 3. Barriers and Experimental Rates for the First Step of MpNPP⁻ Hydrolysis in AP Variants and Wild Type NPP

system ^a	substrate	k_{cat}/K_M	exp ^b	α/R_p ^c	β/S_p ^d
R166S AP	MpNPP ⁻	0.48	18.0	23.4 ^f	(13.1 ^g)
	MpNPPS ⁻	1.1×10^{-3}	21.6	25.5 ^f	(15.2 ^g)
R166S/E322Y AP	MpNPP ⁻	0.24	18.4	22.6 ^f	>30 ^f
	MpNPPS ⁻	3.0×10^{-3}	21.0	19.7 ^f	>40 ^f
NPP	MpNPP ⁻	2.3×10^2	14.3		20.2 ^{e,f}
R166S AP cons-3.6	MpNPP ⁻			19.0	
R166S AP cons-4.1	MpNPP ⁻			22.7	
R166S AP cons-4.6	MpNPP ⁻			>29	
NPP cons-3.6	MpNPP ⁻			17.0	
NPP cons-4.1	MpNPP ⁻			19.2	
NPP cons-4.6	MpNPP ⁻			24.9	

^aIn the “cons” simulations, the Zn²⁺-Zn²⁺ distance is constrained to be a specific value. ^bFree energy barrier (kcal/mol) calculated by transition state theory at 300 K based on experimental k_{cat}/K_M value. ^cFor MpNPP⁻, the substrate methyl group points toward the Mg²⁺ site; for MpNPPS⁻, it is the R_p enantiomer. ^dFor MpNPP⁻, the substrate methyl group points toward the Ser102 backbone; for MpNPPS⁻, it is the S_p enantiomer. ^eThe substrate methyl group points toward the hydrophobic pocket of NPP. ^fPMF barrier with SCC-DFTBPR/MM. ^gBarrier from CPR calculations with SCC-DFTBPR/MM. ^hAdiabatic mapping barrier with B3LYP/6-31G*/MM.

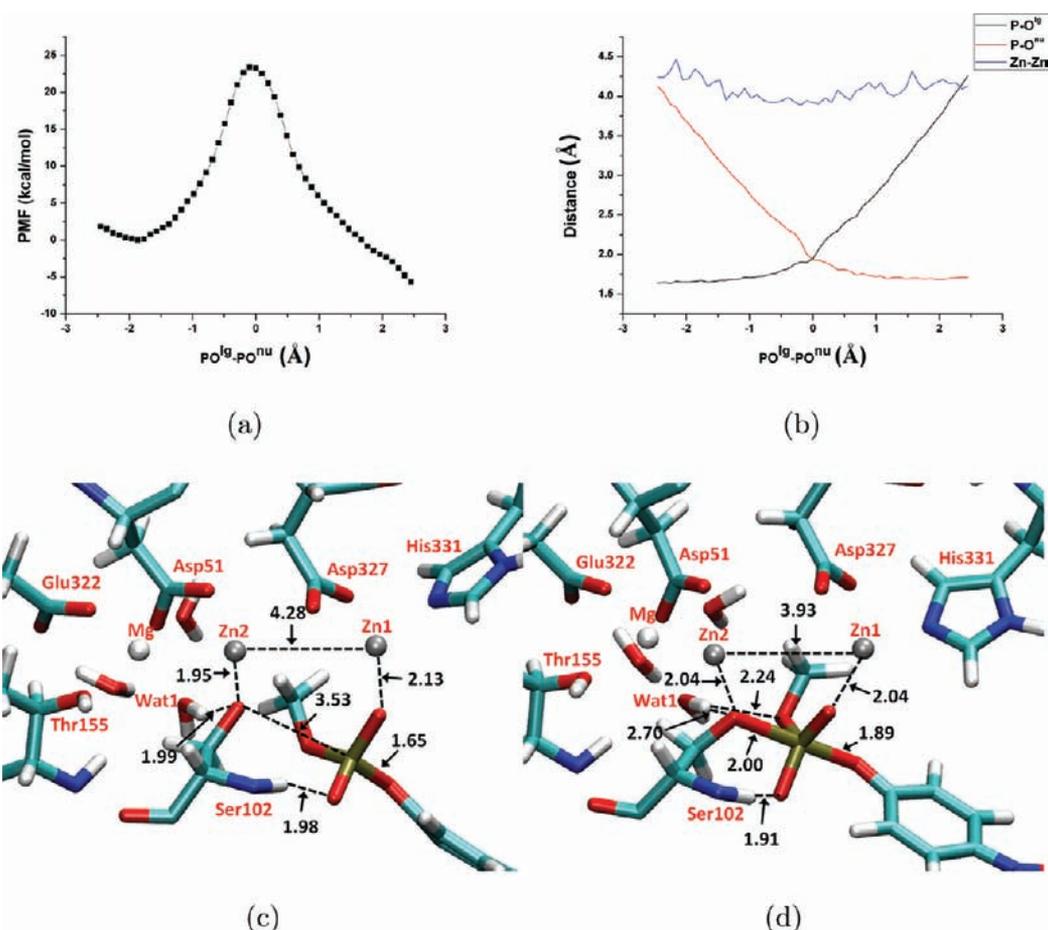


Figure 5. Potential of Mean Force (PMF) calculation results for MpNPP⁻ hydrolysis in R166S AP with the substrate methyl group pointing toward the Mg²⁺ site (the α orientation). Key distances are labeled in Å and energies are in kcal/mol. (a) PMF along the reaction coordinate (the difference between P-O^{lg} and P-O^{nu}); (b) Changes of average key distances along the reaction coordinate; (c) A snapshot for the reactant state, with average key distances labeled. (d) A snapshot for the TS, with average key distances labeled. In (c-d), Asp369, His370, and His412 are omitted for clarity.

0.2 Å), while the distances between Zn and its ligand oxygen are generally longer. Overall, however, the agreement between optimized structures at the two levels of theory is excellent, supporting the use of SCC-DFTBPR/MM.

In addition to the structural similarity in optimized reactants, the MEP results (for β orientation) from adiabatic mapping

also show good agreement (16.9 vs 15.7 kcal/mol) between SCC-DFTBPR/MM and B3LYP/6-31G*/MM calculations, which further supports the use of SCC-DFTBPR/MM; the adiabatic mapping and CPR calculations at the SCC-DFTBPR/MM level give similar transition states (see the Supporting Information), although the barrier height is slightly lower with

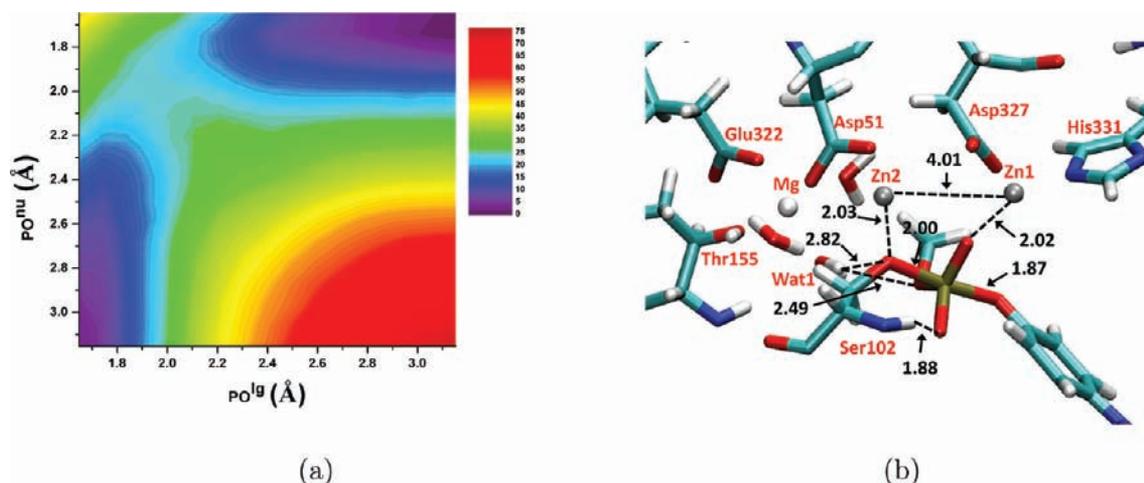


Figure 6. 2D Potential of Mean Force (PMF) calculation results for MnNPP⁻ hydrolysis in R166S AP with the substrate methyl group pointing toward the Mg²⁺ site (the α orientation). Key distances are labeled in Å and energies are in kcal/mol. (a) The 2D PMF along the reaction coordinates; (b) A snapshot for the TS, with average key distances labeled. Asp369, His370, and His412 are omitted for clarity. Note that the 2D PMF results are consistent with the 1D PMF results shown in Figure 5.

Table 4. Calculated Key Structural Properties for the First Step of MnNPP⁻ Hydrolysis in AP Variants and Wild Type NPP

	solution		R166S AP				NPP	
			α^a		β^b		HP ^c	
	reactant	TS	reactant	TS	reactant	TS	reactant	TS
P-O ^{lg}	1.67	2.23	1.65 ± 0.03	1.89 ± 0.07	1.66 ± 0.03	1.86 ± 0.06	1.63 ± 0.03	1.83 ± 0.06
P-O ^{mu}	∞	2.43	3.53 ± 0.07	2.00 ± 0.09	3.83 ± 0.07	2.08 ± 0.08	3.51 ± 0.06	2.03 ± 0.09
RC ^d	-∞	-0.20	-1.88 ± 0.06	-0.11 ± 0.07	-2.17 ± 0.06	-0.22 ± 0.08	-1.88 ± 0.06	-0.20 ± 0.07
TC ^e	∞	4.66	5.18 ± 0.09	3.89 ± 0.14	5.49 ± 0.08	3.94 ± 0.13	5.14 ± 0.08	3.86 ± 0.14
		(4.05)	(5.85)	(5.00)			(6.29)	(5.66)
Zn ²⁺ -Zn ²⁺			4.28 ± 0.22	3.93 ± 0.18	4.21 ± 0.22	3.86 ± 0.16	4.38 ± 0.21	3.92 ± 0.17
Cons-3.6 ^f								
TC ^e			4.65 ± 0.09	3.88 ± 0.13			4.65 ± 0.09	3.85 ± 0.11
Cons-4.1 ^f								
TC ^e			5.31 ± 0.09	3.97 ± 0.18			5.08 ± 0.08	3.97 ± 0.16
Cons-4.6 ^f								
TC ^e			5.83 ± 0.10	4.08 ± 0.25			5.34 ± 0.08	4.09 ± 0.19

^aThe substrate methyl group points toward the Mg²⁺ site (the α orientation). ^bThe substrate methyl group points toward the Ser102 backbone (the β orientation). ^cThe substrate methyl group points toward the hydrophobic pocket. ^dThe Reaction coordinate (RC) is defined as the difference between P-O^{lg} and P-O^{mu}. ^eThe Tightness coordinate (TC) is defined as the sum of P-O^{lg} and P-O^{mu}; in parentheses are values from previous QM/MM simulations. ^f $f_{Zn^{2+}-Zn^{2+}}$ distance constrained at 3.6, 4.1, and 4.6 Å respectively. The rms fluctuations are 0.01 Å.

fully (CPR) optimized saddle point (see Table 3). As shown in Figure 4c, the main differences between SCC-DFTBPR/MM and B3LYP/MM transition states include a shorter Zn²⁺-Zn²⁺ distance at the former level, a shorter P-O_{Ser102} bond length, and hydrogen-bond distances for the interaction between the substrate and nearby groups (e.g., Ser102 backbone amide and Mg²⁺-bound Wat1). We note that, in those MEP calculations with SCC-DFTBPR, the Zn²⁺-Zn²⁺ distance appears somewhat shorter in the TS for the β orientation than the α orientation (see the Supporting Information), although the difference is much smaller in the PMF calculations (see below), again highlighting the importance of sampling protein fluctuations. Nevertheless, there is room to further improve the SCC-DFTBPR method, which may require including complete third-order terms in the SCC-DFTB expansion⁷⁷ and a more systematic refitting of the P-O repulsive potential.⁷⁸

For both substrate orientations, the PMF peaks at the reaction coordinate (PO^{lg}-PO^{mu}) slightly less than 0 Å and then drops in the product region, corresponding to an exothermic

reaction (Figure 5a, 7a). The free energy barriers are 23.4 and 19.6 kcal/mol for the α and β orientations, respectively (see Table 3). It is worth mentioning that these barriers correspond to the free energy difference between the TS and the Michaelis complex, i.e., k_{cat} while experimentally reported values are k_{cat}/K_M , which prevents a direct comparison between calculation and experiment. Nevertheless, the measured k_{cat}/K_M for R166S AP with MnNPP⁻ corresponds to a free energy barrier of 18.0 kcal/mol, which gives the lower bound for the free energy barrier for the chemical step. Therefore, our calculated barriers for the chemical step are qualitatively consistent with the measured k_{cat}/K_M value. The calculations also suggest that the α orientation has a higher barrier for the chemical step. Therefore, for the α orientation to be at least as competitive as the β orientation in terms of k_{cat}/K_M , the corresponding binding free energy should be at least 3.8 kcal/mol stronger, which is qualitatively consistent with the above observation of a stronger hydrogen bonding interaction between the enzyme and MnNPP⁻ in the α orientation. Although further binding

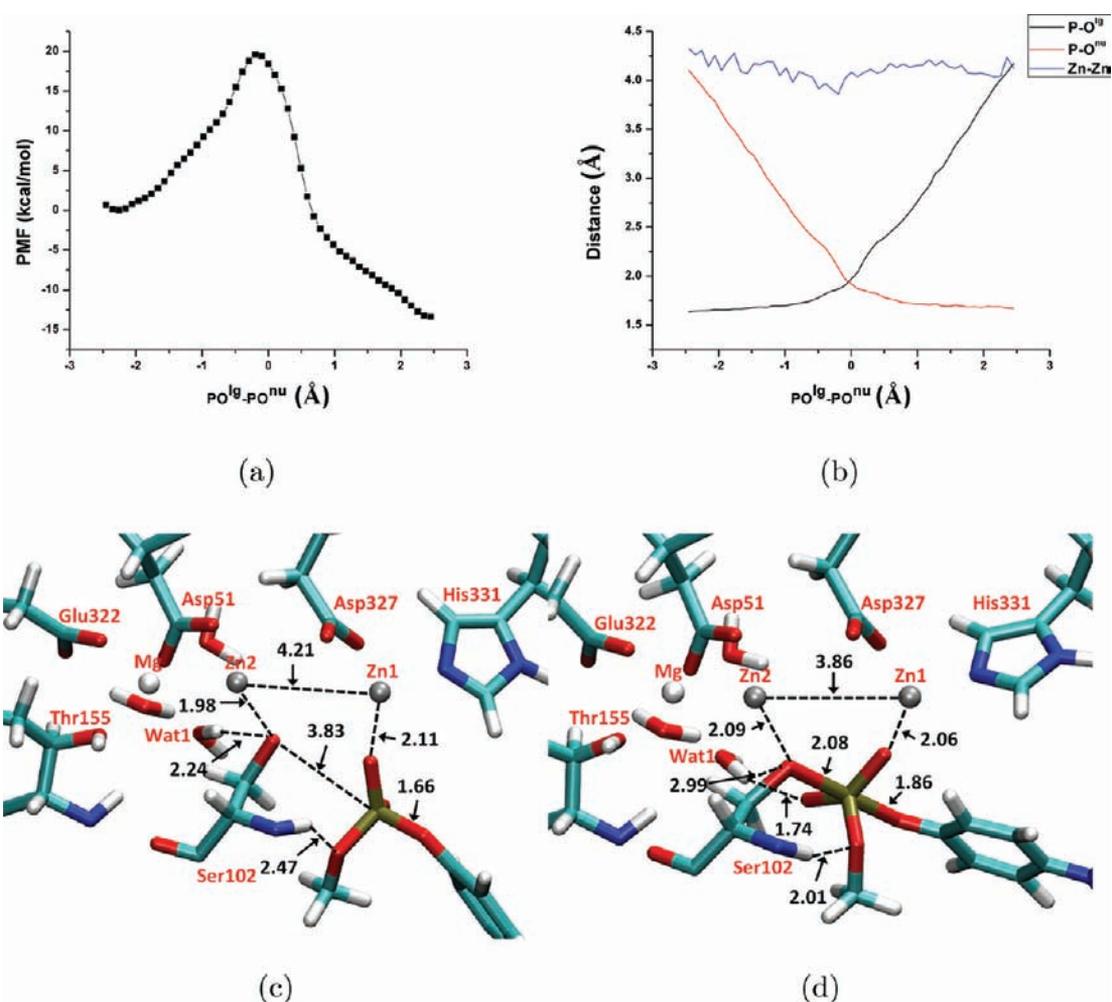


Figure 7. Potential of Mean Force (PMF) calculation results for MpNPP⁻ hydrolysis in R166S AP with the substrate methyl group pointing toward Ser102 backbone (the β orientation). All other format details follow Figure 5.

free energy calculations need to be carried out, the results suggest that the model¹⁴ in which the α orientation is the only productive binding mode seems oversimplified (see Section III.C below for additional discussions).

The key structural properties of the active site are averaged over the trajectory of each window and plotted as functions of the reaction coordinate (Figure 5b, 7b). The changes of P–O^{lg} and P–O^{nu} clearly show that the concerted pathway with a synchronous TS is operative for both substrate orientations and similar to the reaction in aqueous solution, supporting that AP does not significantly alter the nature of hydrolysis TS for phosphate diesters; this is in qualitative agreement with LFER data for R166S AP and a series of substituted methyl phenyl phosphate diesters.²² The nature of the calculated TS is not sensitive to the definition of the 1D reaction coordinate, since the 1D (Figure 5) and 2D (Figure 6) PMF calculations show very consistent results in terms of both structures and energetics. In terms of the tightness coordinate (TC=PO^{lg}+PO^{nu}), it decreases from 4.66 Å in solution to 3.89 Å (α) and 3.94 Å (β) in R166S AP (Table 4); this decrease is likely due to the bimetallic zinc motif in AP, since both MpNPP⁻ and Ser102 are coordinated with Zn²⁺ ions, which are separated by \sim 4 Å throughout the reaction. We note that the degree of tightening is likely not as large as the values for TC imply since our calculations appear to overestimate the value of

TC for solution reactions as compared to previous calculations.^{25,28}

We note that recent QM/MM simulation of WT AP with both mono- and diesters^{24,26} showed rather large structural changes compared to the crystal structure; for example, the Zn²⁺-Zn²⁺ distance increases up to 7 Å, while no such major distortion is observed here (also see discussions below in Section III.E). In the TS for the β orientation (Figure 7d), Wat1 of the Mg²⁺ breaks the hydrogen bond with Ser102 and forms a new one with O4 of MpNPP⁻, which is presumably due to the larger reduction of charge on the Ser102-O than the substrate O4; the average Mulliken charges for Ser102 O and O4 are -0.79 and -0.92 , respectively, in the reactant, -0.56 and -0.89 , respectively, in the TS. This change of hydrogen-bonding interactions likely helps lower the reaction barrier compared with the α orientation in which Wat1 interacts loosely with both Ser102-O and O3 in the TS (see Figure 5d); the average Mulliken charges for Ser102 O and O3 are -0.79 and -0.32 , respectively, in the reactant and -0.50 and -0.40 , respectively, in the TS.

Another interesting observation is that the leaving group does not form a direct interaction with any zinc ion in either the reactant state or TS; rather, it is “solvated” by water molecules accessible to the fairly open active site. This binding mode (especially in the TS) is in contrast to that observed for

vanadate, a widely used transition state analog for phosphoryl transfers, in the crystal structures for AP-vanadate and NPP-vanadate complexes,^{16,79} these structures suggest a binding mode in which one nonbridging oxygen and the leaving oxygen interact directly with one of the zinc ions. Benchmark calculations suggest that our QM/MM protocol is able to reproduce the binding mode of vanadate in the active site of both AP and NPP and that SCC-DFTBPR describes the interaction between the dimetallic zinc motif and phosphate diesters in good agreement with B3LYP (see the Supporting Information). Moreover, MD simulations starting from the vanadate binding mode with the reaction coordinate constrained to be zero converge to the same binding mode shown in Figures 5-7. Collectively, these results suggest that the binding mode observed in the current work is unlikely an artifact of the computational methodology and indeed energetically favorable for systems studied here. We note that the TS for diesters in AP (and NPP, see below) is rather tight in nature, thus the leaving group oxygen does not bear any significant formal charge. Therefore, the binding mode captured in the vanadate structures better reflects the situation for monoesters, which feature a much looser TS in which the leaving group is substantially more charged. Our preliminary calculations for monoesters in AP indeed find tighter interactions between the zinc ion and the leaving group (Hou and Cui, work in progress).

C. Additional Analysis of Substrate Orientation: Activity in the Double Mutant (R166S/E322Y) and Thio Effects in R166S AP.

To further clarify the issue of substrate orientation in AP, we carry out simulation studies to analyze two sets of experiments that were designed to answer the same question. In the first set of experiments, Zalatan and co-workers constructed mutants with the Mg^{2+} site removed (E322Y, E322A, R166S/E322Y) and measured the catalytic activities for phosphate monoester, phosphate diester, and sulfate monoester in these mutants.¹⁴ Based on the observed large detrimental effects of Mg^{2+} removal on phosphate monoester and sulfate monoester hydrolysis but negligible effect on phosphate diester hydrolysis, they concluded that the α orientation is preferred over the β orientation (also see the discussion on thio effects below). However, these observations alone do not rule out the possibility that the two orientations are in fact similar in activity in the WT (and R166S) enzyme, which is what we have observed in this study (see Table 3). In a mutant where one pathway is significantly perturbed, the other can still provide an alternative route, which may explain the only 2 fold decrease in k_{cat}/K_M for R166S/E322Y AP as compared to R166S AP (both with MpNPP⁻ as the substrate). To support this, we explicitly carry out calculations for the double mutants.

As mentioned in Computational Methods, the R166S/E322Y double mutant simulations are prepared based on the crystal structure of E322Y AP,¹⁴ in which the hydroxyl group of Tyr322 occupies the region corresponding to the Mg^{2+} site in WT AP and forms a hydrogen bond with Asp51; the bimetallic zinc site is largely unaffected. Based on the comparison of results for the α and β orientations in R166S AP (Figures 5-7), we expect that mutating away the Mg^{2+} site, which turns off interactions between the Mg^{2+} -bound water and substrate oxygen, will result in a large detrimental effect on the β orientation but a much smaller effect on the α orientation. This is exactly what we observe from the double mutant calculations; as shown in Table 3, the reaction barrier is slightly decreased to 22.6 kcal/mol for the α orientation but is increased to be over

30 kcal/mol for the β orientation. These results directly support the important role of the Mg^{2+} site in reducing the barrier for the β orientation, and by inference, for the hydrolysis of phosphate monoesters, as observed experimentally.¹⁴ Analysis of structures from PMF calculations also indicates that a water molecule penetrates into the double mutant active site for the β orientation to further stabilize the nucleophile in the reactant state, thus also contributing to the significant increase of the barrier; no such water penetration is observed for the α orientation since the hydrophobic -OMe group in the substrate helps block additional water from the active site. The double mutant calculations also explicitly support that the α orientation is not affected much by the Mg^{2+} site, in agreement with the experimental observation¹⁴ that the activity of phosphate diesters remains largely unperturbed in the Mg^{2+} -site mutants. In short, our model that both α and β orientations are productive binding modes in R166S AP while only α is reactive in the Mg^{2+} -site mutants is consistent with available experimental data.

In the second set of experimental studies, Zalatan and co-workers analyzed the reactivities of phosphorothioate diesters in several variants of AP.¹⁴ Take MpNPPS⁻ as an example (Figure 8), the key observation was that R166S AP reacts with

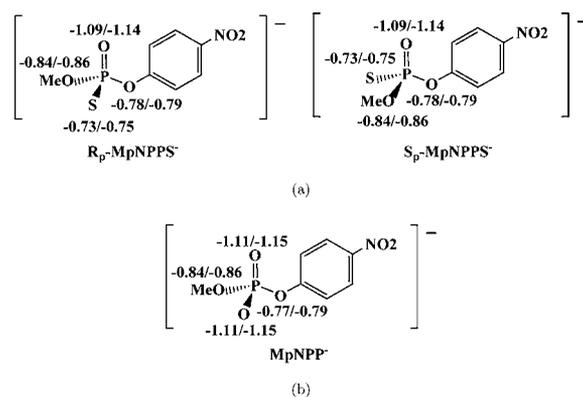


Figure 8. NBO charge analysis for MpNPPS⁻ and MpNPP⁻ in gas phase and solution. Geometries are optimized in gas phase by B3LYP/6-311++G(d,p). Solvation effects are added by PCM with UAKS radii. Numbers before/after slash are gas-phase/solution NBO charges. (a) Enantiomers of MpNPPS⁻; (b) MpNPP⁻.

the R_p enantiomer at least 10^2 times faster than with the S_p enantiomer, suggesting that the binding mode with the R' (Me) group toward the Mg^{2+} site and the sulfur toward the Ser102 backbone amide is dominant; a relevant piece of information here is that previous experiments in the same group established that the sulfur is not placed between the Zn^{2+} ions.¹⁵ We note, however, these discussions rely on the assumption that phosphorothioate diesters behave, in terms of hydrogen bonding with active site groups, similarly to phosphate diesters, which may not be as clear-cut as commonly believed. For example, NBO charge analysis in both gas phase and solution (see Figure 8) shows that in MpNPPS⁻ the oxygen bonded with the methyl group is in fact more negatively charged than the sulfur; i.e., it is not obvious that the R_p enantiomer of MpNPPS⁻ reflects the α -orientation of MpNPP⁻. Therefore, although the experimental observation that removal of the Mg^{2+} site (R166S/E322Y AP vs R166S) has a similar impact on the hydrolysis of MpNPPS⁻ and MpNPP⁻ suggests that interactions from the Mg^{2+} site are similar in the phosphor-

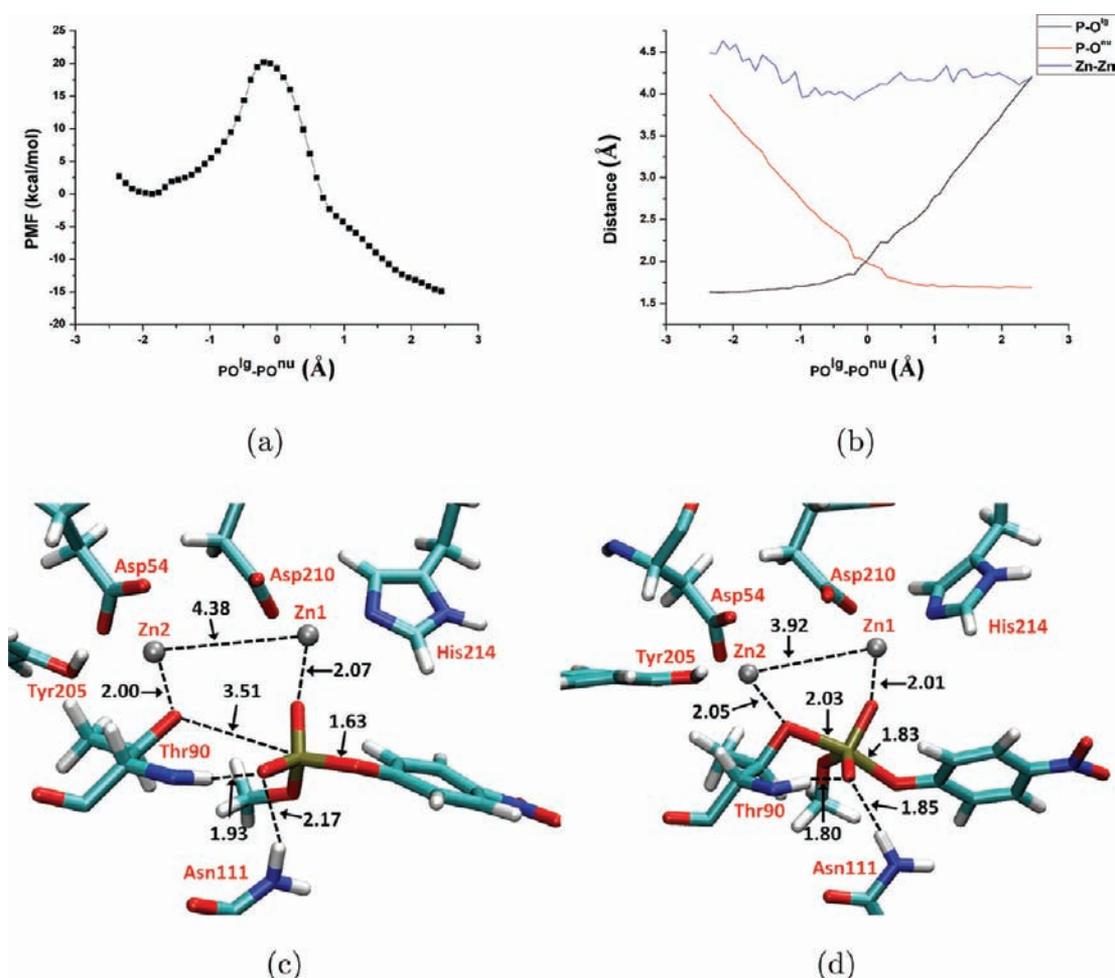


Figure 9. Potential of Mean Force (PMF) calculation results for MnNPP⁻ hydrolysis in NPP with the substrate methyl group pointing toward the hydrophobic core. Other format details follow Figure 5. In (c-d), Asp257, His258, and His363 are omitted for clarity.

othioate and phosphate esters, it is not clear if the thio effects can unambiguously infer the binding mode of phosphate diester.

To help better understand the thio effects, we calculate the PMF for MnNPPS⁻ hydrolysis in both R166S AP and R166S/E322Y AP. The results support that in R166S AP, the R_p enantiomer is indeed favored over S_p enantiomer by 7.9 kcal/mol for the chemical step (see Table 3), in qualitative agreement with experimental findings. The experimental thio effect (the ratio of rate constant for phosphate ester substrate over that for the phosphorothioate analog) corresponds to a free energy barrier difference of 3.6 kcal/mol, while our calculated value is 2.1–5.9 kcal/mol, depending on whether the α or β orientation for MnNPP⁻ is used as reference; note again, however, the experimental value is based on k_{cat}/K_M while our values are based on the chemical step only. One point worth mentioning is that due to the weaker substrate binding following thio substitution, the Solvent Accessible Surface Area (SASA) of the sulfur in MnNPPS⁻ is much larger than that for the corresponding oxygen in MnNPP⁻, especially for the S_p enantiomer (see Table 2 in the Supporting Information); this does not occur at the transition state. The larger SASA provides extra solvent stabilization for the reactant state and probably accounts partially for the much larger thio effects calculated for the S_p enantiomer (13.8 kcal/mol higher in barrier relative to the β orientation of MnNPP⁻). The PMF

profiles for MnNPPS⁻ also peak at where the value of the reaction coordinate is slightly less than 0, while the position of the reactant state is decreased from ~ 2 Å in MnNPP⁻ to ~ -2.5 to -3 Å, reflecting the larger substrate size and weaker binding interactions with the active site than MnNPP⁻ (see the Supporting Information); the transition state is still synchronous in nature and slightly tighter than MnNPP⁻.

The effect of removing the Mg²⁺ site (in R166S/E322Y AP) on the hydrolysis of MnNPPS⁻ is expected to be small for both enantiomers since the interactions between Mg²⁺-water and either the -OMe or -S⁻ are fairly weak. Indeed, for the R_p enantiomer, the barrier actually decreases in the R166S/E322Y mutant relative to the R166S AP to 19.7 kcal/mol; this value is slightly lower than the experimental k_{cat}/K_M value of 21.0 kcal/mol. For the S_p enantiomer, surprisingly, the barrier increases to be over 40 kcal/mol. A closer examination of the simulation snapshots shows that for the reactant state with the S_p enantiomer a solvent water penetrates into the active site and forms a hydrogen bond with the nucleophile (Ser102 oxygen), as illustrated by the comparison of integrated water distribution near the nucleophilic oxygen in Ser102 in the reactant state (see the Supporting Information). As mentioned above for the β orientation of MnNPP⁻ in R166S/E322Y AP, which also features a significantly increased barrier (Table 3), a similar water penetration is observed as well. The water penetration to Ser102 only happens for these two cases and is reproducible in

simulations in which the penetrated water is first deleted and then the system further equilibrated. We note that in R166S AP the active site is already rather open to solvent molecules, thus additional water penetration into the active site is not unexpected when the Mg^{2+} site is removed; the β orientation and the S_p enantiomer are particularly susceptible to water penetration since they lack the bulky methyl group near Ser102. Nevertheless, water penetration in AP mutants remains an interesting issue that deserves in-depth analysis from future experimental and computational studies.

Considering the results for MpNPP^- and those for the R_p enantiomer for MpNPPS^- in the two AP variants, our calculations qualitatively reproduced key experimental observations concerning the effects of Mg^{2+} site removal and thio substitution, further supporting the argument in the last subsection that experimental data so far can not be used to unambiguously determine the orientation of diester substrates in the AP active site. In the broader context, as we mentioned above, the charge distributions for MpNPP^- and MpNPPS^- bear some nontrivial differences (Figure 8); in addition, the possibilities of water penetrating into the active site for certain orientation of the (thio-substituted) substrate and that different substrate orientations are dominant in different variants of AP (R116S vs R116/E332Y) further complicate interpretation of the observed thio effects.

D. First Step of MpNPP^- Hydrolysis Reaction in NPP.

The hydrophobic groove in NPP has been suggested to contribute at least 10^4 -fold to the catalysis of phosphate diester reactions¹⁶ by favorable interactions with the extra R' group in diesters (Figure 1b). Therefore, only one orientation is studied here for MpNPP^- in which the methyl group points toward the hydrophobic pocket. Similar to the comparisons made above for AP, SCC-DFTBPR/MM minimizations for MpNPP^- -NPP also give similar reactant complex structure to B3LYP/MM calculations (Figure 4b). The $\text{O}_{\text{Thr90}}\text{-P}$ distance increases from 3.2 Å in the crystal, which contains AMP as the inhibitor, to 3.6 (3.6) Å at the SCC-DFTBPR/MM (B3LYP/MM) level. The substrate O2 coordinates with Zn1, while O4 forms hydrogen bonds with Asn111 and the backbone amide of Thr90. The optimized $\text{Zn}^{2+}\text{-Zn}^{2+}$ distance is 4.47 (4.40) at the B3LYP/MM (SCC-DFTBPR/MM) level. The two hydrogen bonds formed between O4-Asn111 and O4-Thr90-backbone-amide are also in decent agreement at different levels of theory (Figure 4b).

The PMF calculation shows that the free energy profile corresponds to an exothermic process with the barrier located at $\text{PO}^{\text{g}}\text{-PO}^{\text{nu}} \sim -0.20$ Å and a barrier height of 20.2 kcal/mol. The measured k_{cat}/K_M corresponds to 14.3 kcal/mol at 300 K¹⁶ and sets the lower limit for the chemical step barrier. Compared with AP, the calculated barrier for NPP is close to the β orientation but lower than that in the α orientation. Since MpNPP^- is the cognate substrate of NPP and therefore expected to bind tighter to NPP than to AP, the calculated barrier implies a higher k_{cat}/K_M value for NPP than for AP, which is consistent with experimental observations. In other words, the calculations explicitly support that although diesters are cognate substrates for NPP and promiscuous substrate for AP, the chemical step in NPP is not much accelerated over (R166S) AP.

The changes of P-O^{g} and P-O^{nu} (Figure 9b) show that the concerted pathway is also operative for NPP, with a TS similar to that in aqueous solution and AP. For example, the tightness coordinate is 3.86 Å (Table 4), as compared to the values of

3.89(α)/3.94(β) and 4.66 Å, respectively, in R166S AP and solution, respectively. Considering that the tightness coordinate for solution TS seems overestimated by our method compared to previous work,^{25,28} our calculations support the idea motivated by LFER data²² that, instead of altering transition state structure, NPP and AP catalyze phosphoryl transfer reactions by recognizing and stabilizing transition states similar to those in aqueous solution.

E. Comparison to Recent QM/MM Simulations.^{25,26} As just stated, our calculations find that the transition states for diester hydrolysis in AP and NPP are similar and slightly tighter than that in solution. This is in direct contrast to the recent QM/MM studies^{25,26} which found that the TS in AP/NPP is much looser in nature with the tightness coordinate of 5.66 (NPP)/5.00 (AP) vs a value of 4.05 in solution. Several pieces of evidence suggest that those calculations are less reliable than our SCC-DFTBPR/MM calculations. First, as noted above for AP, their calculations led to large structural distortions in the bimetallic zinc motif relative to the crystal structure, while no such distortions occur in our calculations; the same trends hold for NPP calculations, and the loose TS found in previous work^{25,26} might be a result of the substantially elongated $\text{Zn}^{2+}\text{-Zn}^{2+}$ distance. Another example concerns the hydrogen-bond between Asn111 and MpNPP^- in NPP, which is observed in the crystal structure (with AMP and vanadate) and throughout our simulations; by contrast, this interaction was broken in the recent QM/MM simulations.²⁵ Second, our calculated barrier heights are consistently higher than the barriers that correspond to experimentally measured k_{cat}/K_M values, while this is not the case for recent QM/MM calculations for both NPP²⁵ and AP.²⁶ For example, for NPP with the same substrate, their best estimate of the barrier²⁵ was ~ 11 kcal/mol, which is even lower than the barrier estimated based on experimental k_{cat}/K_M , casting further doubt on the quantitative nature of their result.

What could be the origin for the differences between our and the recent QM/MM calculations?^{25,26} Both SCC-DFTBPR and AM1(d)-PhoT are approximate in nature and parametrized rather specifically for phosphate hydrolysis reactions. For solution reaction, the AM1(d)-PhoT approach gives encouraging results in terms of both energetics and KIEs,⁷⁶ while SCC-DFTBPR, once corrected for gas-phase energies based on MP2, gives good barriers but too loose transition states as reflected by the computed KIE values. In the context of zinc containing enzymes, however, the SCC-DFTBPR approach has been tested with more benchmark calculations, including those discussed here. Indeed, although AM1/PM3 has been used successfully to describe a number of metalloenzymes that catalyze phosphoryl transfers,^{80,81} combining AM1 for zinc and AM1(d)-PhoT for phosphoryl transfers in zinc enzymes has not been carefully tested. Therefore, we suspect that the main cause for the large structural variations in the recent QM/MM simulations^{25,26} is the use of AM1 for zinc. In this regard, although it is possible that the crystal structure with an inhibitor (e.g., inorganic phosphate) does not capture all structural features (e.g., variation in the zinc-zinc distance) of the transition state, the fact that the active site undergoes little change with a transition state analogue (vanadate, see discussions in Section III.B and also the Supporting Information) in both high-resolution crystallography and EXAFS studies³¹ suggests that large variations in the zinc-zinc distance seen in the recent QM/MM simulations^{25,26} (including in the Michaelis complexes!) are unlikely realistic. Since the large increase in the zinc-zinc distance occurs in their

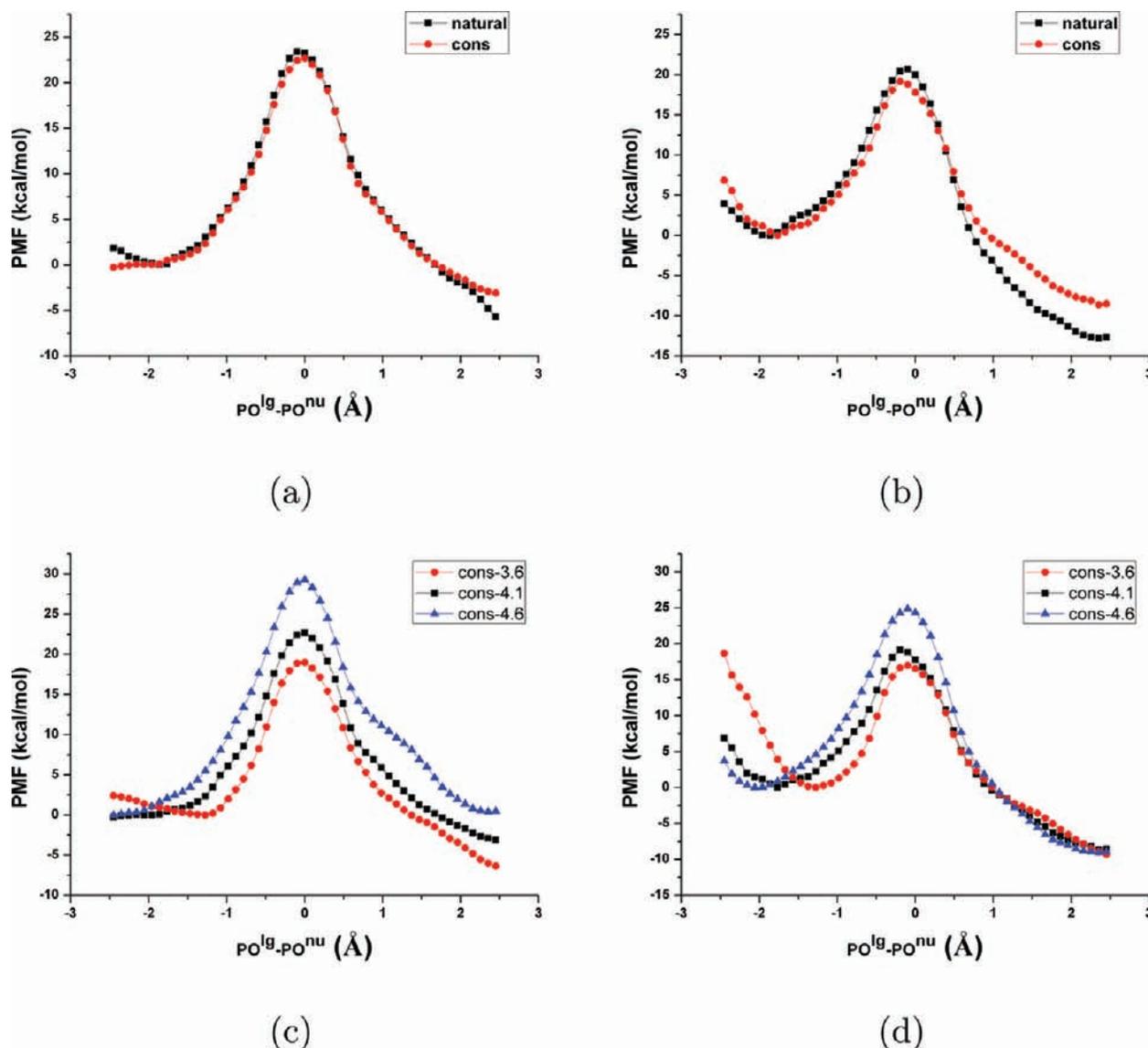


Figure 10. Potential of Mean Force (PMF, in kcal/mol, along the reaction coordinate defined as the difference between $P-O^{lg}$ and $P-O^{nu}$) comparisons for MpNPP⁻ hydrolysis in R166S AP and NPP with the Zn^{2+} - Zn^{2+} distance constrained at different values. (a) Between unconstrained and constrained (4.1 Å) simulations for R166S AP. (b) Between unconstrained and constrained (4.1 Å) simulations for NPP. (c) Between constrained simulations at 3.6, 4.1, and 4.6 Å for R166S AP. (d) Between constrained simulations at 3.6, 4.1, and 4.6 Å for NPP. For structural information, see Table 4 and the Supporting Information.

calculations of several AP and NPP variants,^{25,26} the effect likely cancels out for relative trends; this explains why mutation effects were adequately captured in ref 26. Finally, we noted that the QM/MM boundary in refs 25 and 26 cuts across fairly polar covalent bonds, yet the simple single link-atom scheme was used. As emphasized by several researchers,^{56,82–84} extra care needs to be exercised when the QM/MM boundary involves polar bonds. This technical detail likely also contributes to the uncertainty of the results in refs 25 and 26.

F. Why Is the Nature of TS for Phosphate Diesters in AP and NPP Similar to That in Solution? As another way to evaluate the conflicting findings in the current and previous QM/MM studies, we further dissect whether our observation that the nature of TS for phosphate diesters in AP and NPP is similar to that in solution is consistent with other known experimental facts. We do so with the scheme outlined in Figure 11, which is qualitatively similar to that used by Herschlag and co-workers.^{18,22,23}

For aryl phosphate diester hydrolysis in solution, the measured reaction barrier, $\Delta G^{\ddagger(aq)}$ is ~ 26 kcal/mol (e.g., see Table 1 for MpNPP⁻); the corresponding barrier in the enzyme, $\Delta G^{\ddagger(E)}$, is, for R166S AP, 18.0 kcal/mol (see k_{cat}/K_M in Table 3). Therefore, the enzyme binds to the TS, which is shown here to be of rather similar synchronous nature in solution and AP/NPP (Table 4), by about 8 kcal/mol ($\Delta\Delta G_{syn\ddagger}^b$).

For the enzyme to shift the nature of TS from synchronous to loose, the driving force needs to be large enough to overcome the binding energy for the synchronous TS ($\Delta\Delta G_{syn\ddagger}^b$) plus the energy gap between these two kinds of structures in solution ($\Delta\Delta G_{syn/loose}^{\ddagger(aq)}$). The latter, although not measurable directly with experiments, can be estimated based on calculations; our calculations shown in Figure 3 give a value ~ 8 kcal/mol (the “loose” structure is taken to have a tightness coordinate of ~ 5.7 Å as found for the TS in NPP in previous QM/MM calculations²⁵). In other words, the enzyme needs to bind to

the loose TS by more than 16 kcal/mol ($\Delta\Delta G_{\text{syn}^\ddagger}^b + \Delta\Delta G_{\text{syn}^\ddagger/\text{loose}}^{\ddagger(\text{aq})} = 8 + 8$) to make the loose TS more favorable in the enzyme than a synchronous one.

Considering what we know about the activity of AP toward phosphate monoesters, however, we argue that such a strong binding is unlikely for phosphate diesters. For a phosphate monoester related to the diesters studied here, such as pNPP²⁻ (*p*-nitrophenyl phosphate), the solution barrier is about 32 kcal/mol¹⁸ and the barrier in R166S AP is 10.6 kcal/mol.⁸⁵ Since LFER data indicate that the nature of TS is loose in both AP and solution,^{20,85} these results suggest that R166S AP binds a loose TS for monoesters by ~ 21 kcal/mol. Since diesters feature less charge and are promiscuous substrates of (R166S) AP, we expect that the binding energy of a loose TS for diesters to R166S AP is substantially lower than 21 kcal/mol. Therefore, we do not expect that R166S AP is able to shift the TS for diesters to be much looser in nature, in agreement with our QM/MM calculations.

For NPP, the diesters are cognate substrates, thus it is conceivable that the active site has been evolved to optimize the catalysis of diester hydrolysis, which might involve modifying the nature of TS relative to solution. However, we note that at least for the diester substrate studied here, the calculated chemical step barrier in NPP (20.2 kcal/mol) is not significantly lower than that in solution (25.7 kcal/mol). Therefore, it is not unreasonable that the nature of TS in NPP is not significantly changed relative to solution and a significant component of rate enhancement over solution (k_{cat}/K_M relative to k_n) is due to substrate binding.

G. The Effects of Zn²⁺-Zn²⁺ Distance on Reaction Energetics. Since the bimetallic zinc site is a prevalent catalytic motif,⁸⁶⁻⁹⁰ it is of interest to establish what features (e.g., structural vs electrostatic) are important to the catalytic proficiency. In this work, motivated in part by the fact that the Zn²⁺-Zn²⁺ distances behave rather differently in the current and previous QM/MM calculations^{24,25} of AP and NPP, we examine the effect of this fundamental structural feature on the catalysis in AP and NPP.

Specifically, we design two sets of simulations for R166S AP and NPP with MpNPP⁻ (α orientation only) to study the effects of Zn²⁺-Zn²⁺ distance fluctuation and variation. In the first set, we constrain the Zn²⁺-Zn²⁺ distance close to the value in crystal structures of AP and NPP (4.1 Å). Due to this constraint, the Zn²⁺-Zn²⁺ distance fluctuations in PMF simulations are significantly damped (root-mean-square-fluctuation, rmsf, of ~ 0.01 Å) as compared to unconstrained simulations (rmsf ~ 0.2 Å). As shown in Figure 10a-b, the PMFs for AP and NPP do not exhibit much change, especially in the barrier height, due to the constraint; the nature of the TS (see the Supporting Information) also does not change. These observations suggest that fluctuation of the Zn²⁺-Zn²⁺ distance *during* the reaction is not critical to the reaction barrier or the nature of the TS. The significant overlaps of the PMFs also indirectly support the reproducibility and convergence of our PMF simulations.

In the second set of simulations, the Zn²⁺-Zn²⁺ distances are constrained at 3.6 and 4.6 Å, respectively, which represent the two extreme values observed in unconstrained simulations. The changes of PMF are similar in AP and NPP (see Figure 10c-d): the reactant state position is shifted to a less negative value and the barrier height is reduced as the Zn²⁺-Zn²⁺ distance is decreased from 4.6 to 3.6 Å. In terms of the nature of TS, there is no qualitative change as reflected by the essentially invariant

peak position of the PMFs. At the quantitative level, there are variations as reflected by the average tightness coordinate (see Table 4) and other structural parameters (see the Supporting Information); for example, the tightness coordinate (TC) increases, as expected, as the Zn²⁺-Zn²⁺ distance increases.

In short, these calculations explicitly demonstrate that the Zn²⁺-Zn²⁺ distance of the bimetallic zinc site plays an important role in tuning the catalysis. This is not unexpected since the distance between the zinc ions influences the electrostatic properties in the active site, which are crucial to the phosphoryl transfers.⁹¹ Nevertheless, the results clearly underlines the importance of reproducing geometrical properties of a bimetallic site for a meaningful analysis of its catalytic properties.

H. Issues Worthwhile Investigating with Future Experiments. The key objective of this work is to characterize the nature of the hydrolysis TS of phosphate diesters in AP and NPP so as to evaluate different hypotheses^{22,26} regarding the catalytic promiscuity in these enzymes. To evaluate the main findings of this work, we propose that the following experimental studies are worthwhile.

First, unlike the recent QM/MM simulations,^{25,26} our calculations suggest that the nature of diester hydrolysis TS is largely similar in AP, NPP, and solution. Along this line, LFER and KIE studies for diester hydrolysis in NPP will be highly informative.

Second, our calculations find that the leaving group does not interact strongly with the zinc ion in either the reactant or TS for diester hydrolysis in AP/NPP. This is rather unexpected considering the crystal structures of vanadate-AP/NPP complexes^{16,79} and therefore worth further investigations. Note that this result does not suggest that the second zinc ion does not play an essential role in catalysis since it coordinates with the bridging oxygen. The leaving group is stabilized by active site solvent molecules, suggesting that the dependence of the catalytic rate with respect to the substitution of the leaving group, including both chemical and isotope substitutions, should be close to that in solution.

We have extensively discussed in Section III.C the issue of different binding orientations of the substrate and relation of current calculations with available experimental studies. As highlighted by the NBO charges shown in Figure 8, one complicating factor for the interpretation of thiol experiments is that phosphorothioate and phosphate esters have rather different charge distributions; in MpNPPS⁻ the oxygen bonded with the methyl group is in fact more negatively charged than the sulfur. Therefore, it will be valuable to explore the reactivity of thiol substrates with both well-defined stereochemistry and relative charge distributions closer to the phosphate esters, such as by replacing both O3 and O4 in MpNPP by sulfur.

Finally, as alluded to in Section III G, the metal-metal distance appears to be essential to both the barrier height and the nature of the transition state; the barrier is higher and the nature of TS looser with a longer metal-metal distance. This can be tested by substituting the zinc ions with other metals of different size (e.g., replacing Zn²⁺ by Co²⁺) and performing kinetic and LFER analyses. Along this line, analysis of TS nature with synthetic analogs of dimetallic zinc motifs^{92,93} will be highly informative, since the structure of such artificial catalysts can be better controlled and monitored.

IV. CONCLUDING REMARKS

In this work, we have studied the hydrolysis of MpnPP^- and its several analogs in solution, two experimentally well characterized variants of AP (R166S and R166S/E332Y) and wild type NPP using SCC-DFTBPR/MM simulations. The main goal is to investigate whether the nature of phosphoryl transfer transition state for the same substrate is significantly different in these enzymes and in solution, a question that has a direct implication to the remarkable catalytic promiscuity exhibited by members of the AP superfamily.

Overall, our results are consistent with available experimental observations. In solution, we show that our calculations are able to capture trends in the hydrolysis barrier for MpnPP^- and its two analogs, and the reaction proceeds through a synchronous TS, consistent with expectation based on experimental LFER data. For several variants of AP, our simulations support that the nature of TS is not perturbed significantly relative to solution, with a small degree of tightening due presumably to the interaction with the bimetallic zinc motif; these are also in qualitative agreement with LFER data for AP variants. Therefore, our calculations support the picture that although the native function of AP is to catalyze the hydrolysis of phosphate monoesters through a loose transition state, its active site accommodates the tighter transition state for diesters. Such active site “plasticity” has been proposed to be related to the catalytic promiscuity of AP. Our analysis (Figure

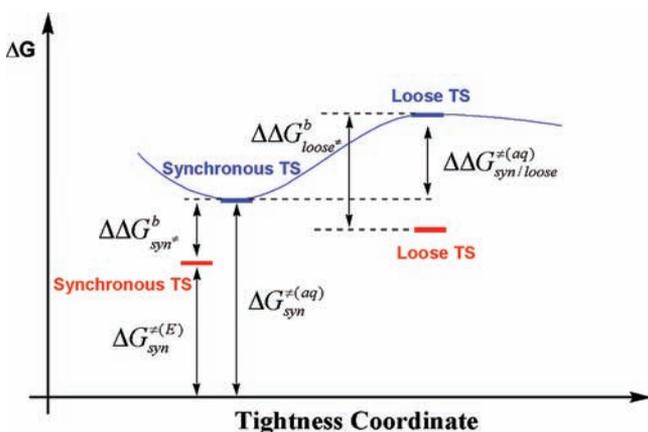


Figure 11. A scheme that illustrates how relative energetics of synchronous and loose transition states in the enzyme (in red) compare to those in solution (in blue). $\Delta G_{syn}^{\ddagger(aq/E)}$ gives the free energy barrier (relative to infinitely separated substrate and nucleophile) in solution/enzyme; $\Delta\Delta G_{syn/loose}^b$ gives the binding free energy of a syn/loose TS structure to the enzyme; $\Delta\Delta G_{syn/loose}^{\ddagger(aq)}$ is the free energy difference between the synchronous and loose transition state structures in solution. For the enzyme to shift the nature of TS from synchronous to loose, $\Delta\Delta G_{loose}^b$ needs to be larger than $\Delta\Delta G_{syn}^b + \Delta\Delta G_{syn/loose}^{\ddagger(aq)}$ which we argue is unlikely for AP and diesters (see text for discussions).

11) of the free energy surfaces for phosphate diester hydrolysis in solution and AP/NPP supports a simple interpretation of such functional plasticity: the binding of AP to diester substrates is simply not strong enough, presumably due to their lower charge compared to monoesters, to significantly shift the nature of the transition state from synchronous to lose.

For NPP, for which diesters are cognate substrates, our calculations also do not support any major change in the nature of TS relative to solution, in contrast to the recent QM/MM

calculations.²⁵ Since both structural features of the active site and energetics from our calculations are in better agreement with available experimental data, we expect that the nature of TS from the current work is more realistic. The lack of any significant change in the nature of TS relative to solution is not unexpected considering that the calculated chemical step barrier in NPP is not significantly lower than that in solution. Nevertheless, we hope the current work will stimulate additional LFER studies for NPP to further confirm the nature of TS.

The calculations also reveal several features and underlying complexity of AP catalysis not thoroughly recognized by previous work. For example, concerning the orientation of a diester substrate in the AP active site, our calculations for two variants of AP and two diester substrates collectively indicate that the experimental data alone cannot be used to unambiguously show that a single orientation (α) is the only reactive binding mode. In fact, we find that the β orientation with the substrate methyl group pointing toward the Ser102 backbone amide has a reaction barrier lower than that for the α orientation, which has the substrate methyl group pointing toward the Mg^{2+} site; however, it is possible that the binding free energy for the α orientation is larger to make the overall free energy profile at least comparable to the β orientation. We also argue that the thio-substitution experiments are not always straightforward to interpret, because there is nontrivial differences in the charge distributions for phosphorothioate and phosphate esters; the possibilities of water penetrating into the active site for certain orientation of the (thio-substituted) substrate and that different substrate orientations dominate in different variants of AP (R116S vs R116S/E332Y) further compromise the clarity of interpretation of thio effects. Finally, we discuss results supporting that the Zn^{2+} - Zn^{2+} distance plays a significant role in modulating the energetics, especially barrier, of phosphoryl transfer in AP and NPP. This result can be probed experimentally by metal substitution and underlines the importance of reproducing geometrical properties of a bimetallic site for a meaningful computational analysis of its catalytic properties.

For a more thorough understanding of catalytic promiscuity and functional evolution of the AP superfamily, it is crucial to carry out similar systematic benchmark and analyses for monoester hydrolysis in solution and AP family enzymes. Since phosphate monoesters bear higher charges than diesters, and that SCC-DFTBPR has been developed based mainly on diesters, it is likely that further improvements of the computational methodology are needed.^{77,78} Once validated for AP and NPP, the computational approach coupled with other methodological advances^{94–97} is potentially applicable in the prediction and rational design of catalytic promiscuity in other enzyme families.

■ ASSOCIATED CONTENT

📄 Supporting Information

Additional calculation results and discussions on the following subjects are included: additional benchmarks in both gas-phase (active site cluster model), solution (implicit solvent, kinetic isotope effects) and enzyme active sites (AP bound to inorganic phosphate, vanadate, or different binding modes of MpnPP^-), additional results for the double mutants, with thio-substituted MpnPP^- and constrained zinc–zinc distance. The complete reference for refs 45,54, 69, and 70 are also included. This

material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

cui@chem.wisc.edu

ACKNOWLEDGMENTS

We acknowledge stimulating discussions with Professor D. Herschlag and Drs. J. Lassila, L. D. Andrews, and J. Zalatan and their critical reading of the manuscript. This work is supported by NIH grant R01-GM084028. Computational resources from the National Center for Supercomputing Applications at the University of Illinois and the Centre for High Throughput Computing (CHTC) at UW-Madison are greatly appreciated; computations are also supported in part by National Science Foundation through a major instrumentation grant (CHE-0840494).

REFERENCES

- O'Brien, P. J.; Herschlag, D. *Chem. Biol.* **1999**, *6*, R91–R105.
- Copley, S. D. *Curr. Opin. Chem. Biol.* **2003**, *7*, 265–272.
- Schmidt, D. M. Z.; Mundorff, E. C.; Dojka, M.; Bermudez, E.; Ness, J. E.; Govindarajan, S.; Babbitt, P. C.; Minshull, J.; Gerlt, J. A. *Biochemistry* **2003**, *42*, 8387–8393.
- Zalatan, J. G.; Herschlag, D. *Nat. Chem. Biol.* **2009**, *5*, 516–520.
- Jonas, S.; Hollfelder, F. *Pure Appl. Chem.* **2009**, *81*, 731–742.
- Jensen, R. A. *Annu. Rev. Microbiol.* **1976**, *30*, 409–425.
- Aharoni, A.; Gaidukov, L.; Khersonsky, O.; Gould, S. M.; Roodveldt, C.; Tawfik, D. S. *Nat. Genet.* **2005**, *37*, 73–76.
- Khersonsky, O.; Roodveldt, C.; Tawfik, D. S. *Curr. Opin. Chem. Biol.* **2006**, *10*, 498–508.
- Khersonsky, O.; Tawfik, D. S. *Annu. Rev. Biochem.* **2010**, *79*, 471–505.
- O'Brien, P.; Herschlag, D. *J. Am. Chem. Soc.* **1998**, *120*, 12369–12370.
- O'Brien, P.; Herschlag, D. *Biochemistry* **2001**, *40*, 5691–5699.
- van Loo, B.; Jonas, S.; Babbitt, A. C.; Benjdia, A.; Berteau, O.; Hyvönen, M.; Hollfelder, F. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 2740–2745.
- Lad, C.; Williams, N. H.; Wolfenden, R. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 5607–5610.
- Zalatan, J.; Fenn, A.; Herschlag, D. *J. Mol. Biol.* **2008**, *384*, 1174–1189.
- Lassila, J.; Herschlag, D. *Biochemistry* **2008**, *47*, 12853–12859.
- Zalatan, J.; Fenn, T.; Brunger, A.; Herschlag, D. *Biochemistry* **2006**, *45*, 9788–9803.
- Stec, B.; Holtz, K.; Kantrowitz, E. *J. Mol. Biol.* **2000**, *299*, 1303–1311.
- Lassila, J. K.; Zalatan, J. G.; Herschlag, D. *Annu. Rev. Biochem.* **2011**, *80*, 669–702.
- Hollfelder, F.; Herschlag, D. *Biochemistry* **1995**, *38*, 12255–12264.
- O'Brien, P.; Herschlag, D. *J. Am. Chem. Soc.* **1999**, *121*, 11022–11023.
- Nikolic-Hughes, I.; Rees, D.; Herschlag, D. *J. Am. Chem. Soc.* **2004**, *126*, 11814–11819.
- Zalatan, J.; Herschlag, D. *J. Am. Chem. Soc.* **2006**, *128*, 1293–1303.
- Zalatan, J.; Catrina, I.; Mitchell, R.; Grzyska, P.; O'Brien, P.; Herschlag, D. *J. Am. Chem. Soc.* **2007**, *129*, 9789–9798.
- López-Canut, V.; Martí, S.; Bertrán, J.; Moliner, V.; Tuñón, I. *J. Phys. Chem. B* **2009**, *113*, 7816–7824.
- López-Canut, V.; Roca, M.; Bertrán, J.; Moliner, V.; Tuñón, I. *J. Am. Chem. Soc.* **2010**, *132*, 6955–6963.
- López-Canut, V.; Roca, M.; Bertrán, J.; Moliner, V.; Tuñón, I. *J. Am. Chem. Soc.* **2011**, *133*, 12050–12062.
- Nam, K.; Cui, Q.; Gao, J.; York, D. *J. Chem. Theory Comput.* **2007**, *3*, 486–504.
- Rosta, E.; Kamerlin, S. C. L.; Warshel, A. *Biochemistry* **2008**, *47*, 3725–3735.
- McWhirter, C.; Lund, E. A.; Tanifum, E. A.; Feng, G.; Sheikh, Q. I.; Hengge, A. C.; Williams, N. H. *J. Am. Chem. Soc.* **2008**, *130*, 13673–13682.
- O'Brien, P.; Lassila, J.; Fenn, T.; Zalatan, J.; Herschlag, D. *Biochemistry* **2008**, *47*, 7663–7672.
- Boby, 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. *J. Mol. Biol.* **2011**, In press.
- Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. *Phys. Rev. B* **1998**, *58*, 7260–7268.
- Thiel, W. *Adv. Chem. Phys.* **1996**, *93*, 703–757.
- Kruger, T.; Elstner, M.; Schiffels, P.; Frauenheim, T. *J. Chem. Phys.* **2005**, *122*, 114110.
- Sattelmeyer, K. W.; Tirado-Rives, J.; Jorgensen, W. *J. Phys. Chem. A* **2006**, *110*, 13551–13559.
- Otte, N.; Scholten, M.; Thiel, W. *J. Phys. Chem. A* **2007**, *111*, 5751–5755.
- Yang, Y.; Yu, H.; York, D.; Elstner, M.; Cui, Q. *J. Chem. Theory Comput.* **2008**, *4*, 2067–2084.
- Yang, Y.; Yu, H.; Cui, Q. *J. Mol. Biol.* **2008**, *381*, 1407–1420.
- Yang, Y.; Cui, Q. *J. Phys. Chem. A* **2009**, *113*, 12439–12446.
- Yang, Y.; Cui, Q. *J. Phys. Chem. B* **2009**, *113*, 4930–4933, NIHMS:103392.
- Hou, G. H.; Zhu, X.; Cui, Q. *J. Chem. Theory Comput.* **2010**, *6*, 2303–2314.
- Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Theory Comput.* **2005**, *1*, 1133–1152.
- Fischer, S.; Karplus, M. *Chem. Phys. Lett.* **1992**, *194*, 511–527.
- Cui, Q.; Karplus, M. *J. Phys. Chem. B* **2002**, *106*, 1768–1798.
- Frisch, M. J. et al. "Gaussian 03", 2003.
- Miertus, S.; Tomasi, J. *Chem. Phys.* **1982**, *65*, 239–245.
- Li, G. H.; Cui, Q. *J. Phys. Chem. B* **2003**, *107*, 14521–14528.
- Coleman, J. *Annu. Rev. Biophys. Biomol. Struct.* **1992**, *21*, 441–483.
- Jonas, S.; van Loo, B.; Hyvönen, M.; Hollfelder, F. *J. Mol. Biol.* **2008**, *384*, 120–136.
- Holtz, K. M.; Catrina, I. E.; Hengge, A. C.; Kantrowitz, E. R. *Biochemistry* **2000**, *39*, 9451–9458.
- Brunger, A.; Karplus, M. *Proteins: Struct., Funct., Genet.* **1988**, *4*, 148–156.
- Boorks, B.; Bruccoleri, R.; Olafson, B.; States, D.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.
- Boorks, C.; Karplus, M. *J. Chem. Phys.* **1983**, *79*, 6312–6325.
- MacKerell, A.; et al. *J. Chem. Phys.* **1998**, *102*, 3586–3616.
- Jorgensen, W.; Chandrasekhar, J.; Madura, J.; Impey, R.; Klein, M. *J. Chem. Phys.* **1983**, *79*, 926–935.
- Konig, P.; Hoffmann, M.; Frauenheim, T.; Cui, Q. *J. Phys. Chem. B* **2005**, *109*, 9082–9095.
- Arantes, G.; Loos, M. *Phys. Chem. Chem. Phys.* **2006**, *8*, 347–353.
- Im, W.; Bernffche, S.; Roux, B. *J. Chem. Phys.* **2001**, *114*, 2924–2937.
- Schaefer, P.; Riccardi, D.; Cui, Q. *J. Chem. Phys.* **2005**, *123*, 014905.
- Brooks, C.; Karplus, M. *J. Mol. Biol.* **1989**, *208*, 159–181.
- Ryckaert, J.; Ciccotti, G.; Berendsen, H. *J. Comput. Phys.* **1977**, *23*, 327–341.
- Nina, M.; Beglov, D.; Roux, D. *J. Phys. Chem. B* **1997**, *101*, 5239–5248.
- Nina, M.; Im, W.; Roux, D. *Biophys. Chem.* **1999**, *78*, 89–96.
- Steinbach, P.; Brooks, B. *J. Comput. Chem.* **1994**, *15*, 667–683.
- Becke, A. *Phys. Rev. A* **1988**, *38*, 3098–3100.
- Becke, A. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- Lee, C.; Yang, W.; Parr, R. *Phys. Rev. B* **1988**, *37*, 785–789.

- (68) Petersson, G.; Bennett, A.; Tensfeldt, T.; Allaham, M.; Shirley, W.; Mantzaris, J. J. *Chem. Phys.* **1988**, *89*, 2193–2218.
- (69) Shao, Y.; et al. *Phys. Chem. Chem. Phys.* **2006**, *27*, 3172–3191.
- (70) Brooks, B. R.; et al. *J. Comput. Chem.* **2009**, *30*, 1545–1614.
- (71) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187–199.
- (72) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011–1021.
- (73) Riccardi, D.; Schaefer, P.; Cui, Q. *J. Phys. Chem. B* **2005**, *109*, 17715–17733.
- (74) Hengge, A. C.; Tobin, A. E.; Cleland, W. W. *J. Am. Chem. Soc.* **1995**, *117*, 5919–5926.
- (75) Harris, M. E.; Cassano, A. G.; Anderson, V. E. *J. Am. Chem. Soc.* **2002**, *124*, 10964–10965.
- (76) Tunon, I.; Lopez-Canut, V.; Ruiz-Pernia, J.; Ferrer, S.; Moliner, V. *J. Chem. Theory Comput.* **2009**, *5*, 439–442.
- (77) Gaus, M.; Cui, Q.; Elstner, M. *J. Chem. Theory Comput.* **2011**, *7*, 931–948.
- (78) Gaus, M.; Chou, C. P.; Witek, H.; Elstner, M. *J. Phys. Chem. A* **2009**, *113*, 11866–11881.
- (79) Holtz, K. M.; Stec, B.; Kantrowitz, E. R. *J. Biol. Chem.* **1999**, *274*, 8351–8354.
- (80) Wong, K. Y.; Gao, J. L. *Biochemistry* **2007**, *46*, 13352–13369.
- (81) Wong, K. Y.; Gao, J. L. *FEBS J.* **2011**, *278*, 2579–2595.
- (82) Das, D.; Eurenus, K. P.; Billings, E. M.; Sherwood, P.; Chafffield, D. C.; Hodoscek, M.; Brooks, B. R. *J. Chem. Phys.* **2002**, *117*, 10534–10547.
- (83) Gao, J. L.; Ma, S. H.; Major, D. T.; Nam, K.; Pu, J. Z.; Truhlar, D. G. *Chem. Rev.* **2006**, *106*, 3188–3209.
- (84) Senn, H. M.; Thiel, W. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198–1229.
- (85) O'Brien, P.; Herschlag, D. *Biochemistry* **2002**, *41*, 3207–3225.
- (86) Kim, E. E.; Wyckoff, H. W. *J. Mol. Biol.* **1991**, *218*, 449–464.
- (87) Strater, N.; Lipscomb, W. N.; Klabunde, T.; Krebs, B. *Angew. Chem., Int. Ed.* **1996**, *35*, 2024–2055.
- (88) Steitz, T. A.; Steitz, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 6498–6502.
- (89) Tesmer, J. J. G.; Sunahara, R. K.; Johnson, R. A.; Gosselin, G.; Gilman, A. G.; Sprang, S. R. *Science* **1999**, *285*, 756–760.
- (90) Jedrzejewski, M. J.; Setlow, P. *Chem. Rev.* **2001**, *101*, 607–618.
- (91) Nikolic-Hughes, I.; O'Brien, P.; Herschlag, D. *J. Am. Chem. Soc.* **2005**, *127*, 9314–9315.
- (92) Gao, H.; Ke, Z.; DeYonker, N. J.; Wang, J.; Xu, H.; Mao, Z.; Phillips, D. L.; Zhao, C. *J. Am. Chem. Soc.* **2011**, *133*, 2904–2915.
- (93) Fan, Y. B.; Gao, Y. Q. *Acta Phys.-Chim. Sin.* **2010**, *26*, 1034–1042.
- (94) Hermann, J. C.; Ghanem, E.; Li, Y.; Raushel, F. M.; Irwin, J. J.; Shoichet, B. K. *J. Am. Chem. Soc.* **2006**, *128*, 15882–15891.
- (95) Hermann, J. C.; Marti-Arbona, R.; Fedorov, A. A.; Fedorov, E.; Almo, S. C.; Shoichet, B. K.; Raushel, F. M. *Nature* **2007**, *448*, 775–779.
- (96) Toscano, M. D.; Woycechowsky, K. J.; Hilvert, D. *Angew. Chem., Int. Ed.* **2007**, *46*, 3212–3236.
- (97) Jiang, L.; Althoff, E. A.; Clemente, F. R.; Doyle, L.; Rothlisberger, D.; Zanghellini, A.; Gallaher, J. L.; Betker, J. L.; Tanaka, F.; Barbas, C. F.; Hilvert, D.; Houk, K. N.; Stoddard, B. L.; Baker, D. *Science* **2008**, *319*, 1387–1391.
- (98) Fujio, M.; Mciver, R. T.; Taft, R. W. *J. Am. Chem. Soc.* **1981**, *103*, 4017–4029.
- (99) *CRC Handbook Chemistry and Physics*; Lide, D. R., Ed.; CRC Press: 85 ed.; 2005.