JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

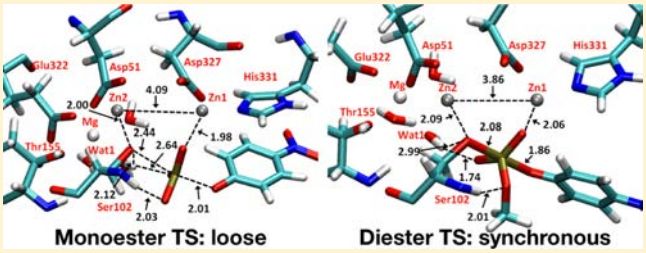
Stabilization of Different Types of Transition States in a Single Enzyme Active Site: QM/MM Analysis of Enzymes in the Alkaline Phosphatase Superfamily

Guanhua Hou and Qiang Cui*

Department of Chemistry and Theoretical Chemistry Institute, University of Wisconsin-Madison, 1101 University Avenue, Madison, Wisconsin 53706, United States

Supporting Information

ABSTRACT: The first step for the hydrolysis of a phosphate monoester ($pNPP^{2-}$) in enzymes of the alkaline phosphatase (AP) superfamily, R166S AP and wild-type NPP, is studied using QM/MM simulations based on an approximate density functional theory (SCC-DFTBPR) and a recently introduced QM/MM interaction Hamiltonian. The calculations suggest that similar loose transition states are involved in both enzymes, despite the fact that phosphate monoesters are the cognate substrates for AP but promiscuous substrates for NPP. The computed loose transition states are clearly different from the more synchronous ones previously calculated for diester reactions in the same AP enzymes. Therefore, our results explicitly support the proposal that AP enzymes are able to recognize and stabilize different types of transition states in a single active site. Analysis of the structural features of computed transition states indicates that the plastic nature of the bimetallic site plays a minor role in accommodating multiple types of transition states and that the high degree of solvent accessibility of the AP active site also contributes to its ability to stabilize diverse transition-state structures without the need of causing large structural distortions of the bimetallic motif. The binding mode of the leaving group in the transition state highlights that vanadate may not always be an ideal transition state analog for loose phosphoryl transfer transition states.



INTRODUCTION

Alkaline phosphatase (AP) superfamily contains a set of evolutionarily related enzymes whose natural function is to catalyze the hydrolytic reactions of phosphates or sulfates;[1,2] an oxygen nucleophile (e.g., Ser or Thr) first attacks the phosphorus/sulfur (Figure 1a), then a water (hydroxide) replaces the leaving group in a step that is essentially the reverse of the first. In recent years, the AP enzymes have attracted much attention because they exhibit extraordinary catalytic promiscuity toward a broad class of phosphates and sulfates:[3,4] their charges range between 0 and −2 and the solution reactions involve transition states of a different nature (e.g., loose vs tight/synchronous), attack at two different reaction centers (P and S), and have diverse intrinsic reactivities (with half-lives between 20 and 85 000 years under near neutral solution conditions).[5] The level of promiscuous activity can be rather high. For example, AP's catalytic proficiency (i.e., k_{cat}/K_m relative to the rate constant of uncatalyzed solution reaction) toward cognate substrates of phosphate monoester is up to 10^{27}, and the proficiency toward noncognate substrates, such as phosphate diesters, are still as high as 10^{11}. Similarly, another family member, the nucleotide pyrophosphatase/phosphodiesterase (NPP, see Figure 2a,b for a comparison of active sites) catalyzes the hydrolysis of phosphate diesters with a speed up to 10^{16} relative to solution, and the promiscuous activity toward phosphate monoesters is around 10^{10} higher than uncatalyzed reactions in solution.

Since the hydrolysis of mono- and diesters in solution is known to adopt transition states of different nature,[7] it is interesting to ask whether the AP enzymes significantly modify the nature of the transition state for these substrates to achieve a significant degree of catalysis for both cognate and noncognate substrates. Along this line, Herschlag and co-workers have carried out extensive[8−16] linear free energy relationship (LFER) and kinetic isotope effects (KIE) analyses for the hydrolysis of phosphate monoesters and diesters in AP and NPP. The results led to the interesting interpretation that AP and NPP do not significantly alter the nature of the phosphoryl transfer transition state for both phosphate monoesters and diesters relative to their solution counterparts, i.e., the transition state for monoesters remains loose, while the transition state for diesters remains tight/synchronous. In other words, the active site of AP enzymes appears to have evolved to be able to recognize and stabilize transition states of different nature. This challenges the traditional paradigm[17,18] that enzyme active sites have evolved to stabilize a single, specific type of transition state.

The experimental investigations also stimulated computational studies,[6,19−23] motivated in part by the concern that it may

Received: April 4, 2013
Published: June 20, 2013

   dx.doi.org/10.1021/ja403293d | J. Am. Chem. Soc. 2013, 135, 10457−10469

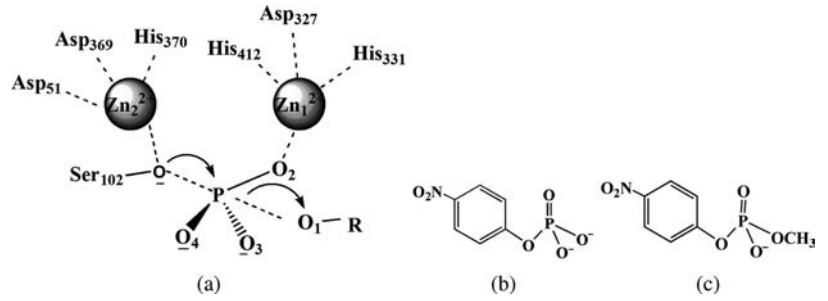


Figure 1. Basic reaction mechanism of the AP enzymes and the substrates studied here and in previous works. (a) The first step of phosphate monoester hydrolysis catalyzed by AP. (b,c) Phosphate monoester ($pNPP^{2-}$) and diester ($MpNPP^{-}$) studied in this and previous[6] works, respectively.

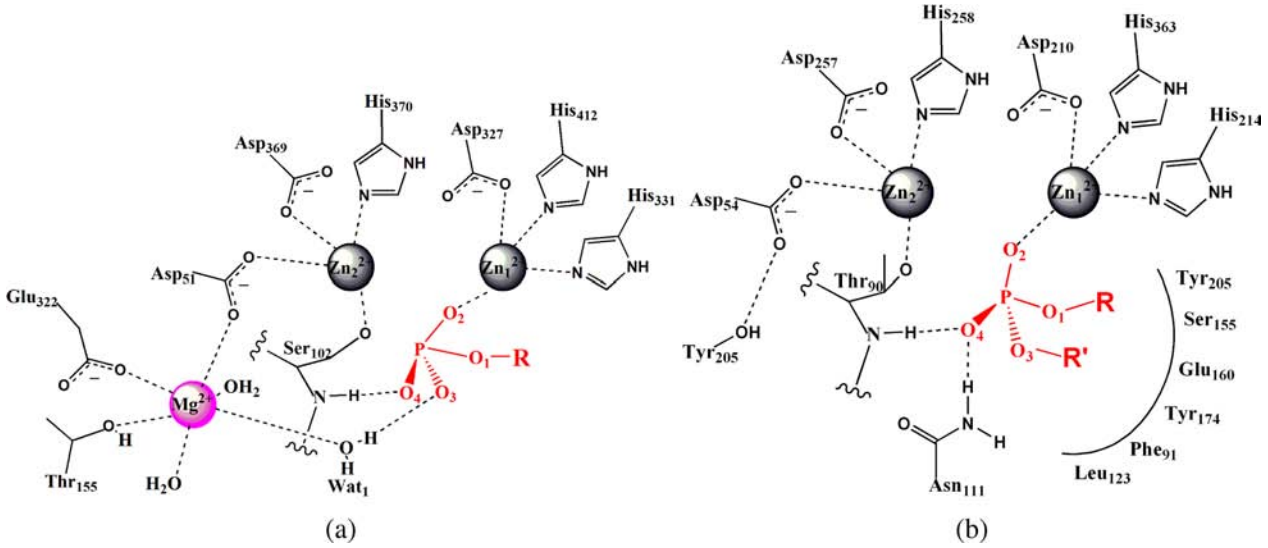


Figure 2. Active sites of AP and NPP are generally similar, with a few distinct differences. (a) *E. coli* R166S AP and (b) *Xac* NPP active sites. Cognate substrates for AP and NPP are phosphate monoesters and diesters, respectively.

not be straightforward to infer the precise nature of transition state based on LFER and KIE data alone.[24] For example, Tuñón and co-workers carried out QM/MM simulations, using the AM1(d)-PhoT method[25] as the QM level, to study phosphate monoester hydrolysis in AP[19] and diester hydrolysis in AP and NPP.[20,21] Those calculations provided rather different descriptions for the reaction mechanism from the LFER/KIE studies:[12,14] AP was found to catalyze monoester hydrolysis via a two-step mechanism with a metastable penta-valent intermediate rather than a one-step mechanism; for phosphate diesters, a very loose transition state was found in both AP and NPP, in contrast to a synchronous transition state found in solution.[7] Therefore, calculations in refs 19−21 suggested that AP and NPP alter the nature of transition state for both phosphate mono- and diesters relative to the solution reactions. It is worth noting that major structural changes were observed in the enzyme active site in those simulations; e.g., the zinc−zinc distance for the catalytic bimetallic motif was found to increase from ∼4 Å in crystal structures to up to 7 Å. Since notable structural changes in bimetallic motifs have been observed in previous computational[26−28] and experimental[29,30] studies of other enzymes, it is difficult to dispute the results of refs 19−21 based solely on the computed structural changes of the active site.

In our recent work,[6] we carried out systematic QM/MM studies of phosphate diester hydrolysis in AP and NPP using an approximate DFT method, SCC-DFTBPR,[31] as the QM level. Compared to the AM1(d)-PhoT model,[25] SCC-DFTB(PR) has been better tested for zinc,[32−34] including bimetallic zinc motifs.[6,32,35] Our studies of a series of phosphate diester reactions[6,36] suggested that neither AP nor NPP significantly alters the nature of transition state for the diester hydrolysis; instead, the hydrolysis goes through a synchronous transition state that is slightly tighter compared to the solution case. As discussed in ref 6, this observation is qualitatively consistent with known catalytic/binding properties of AP and NPP toward different substrates and the free energy surface of diester hydrolysis in solution, i.e., AP/NPP does not have sufficiently strong binding to a diester-like species to overcome the intrinsic energetic difference between synchronous and loose transition-state-like configurations in solution. At the structural level, our SCC-DFTBPR/MM simulations did not observe any significant distortion in the bimetallic zinc motif.

In our analysis[6] of diester hydrolysis in the AP enzymes, by calculating the potential of mean force (PMF) for the phosphoryl transfer with the zinc−zinc distance constrained to specific values, we found that the nature of the transition state (as measured by the "tightness coordinate", the sum of the distances of the breaking and forming P−O bonds) correlates with the zinc−zinc distance; this partially explained why a very loose transition state was observed for diester hydrolysis in the AM1(d)-PhoT/MM simulations,[20,21] which sampled large increases in the zinc−zinc distance. If such correlation also applies to the hydrolysis of phosphate monoesters, since only

dx.doi.org/10.1021/ja403293d | J. Am. Chem. Soc. 2013, 135, 10457−10469

limited structural fluctuations in the bimetallic motif were observed in our (unconstrained) QM/MM simulations, one might predict that the transition state of monoester hydrolysis becomes substantially tighter in the AP enzymes than in solution. This prediction appears to be contradictory to both LFER/KIE analysis[12,14] and AM1(d)-PhoT/MM simulations.[19] Therefore, it is imperative to explicitly carry out computational analysis of phosphate monoester hydrolysis in the AP enzymes using our QM/MM approach. Due to the highly charged nature of phosphate monoesters, such calculations have not become possible until very recently due to the development of an improved QM-MM electrostatic model[37,38] for SCC-DFTB based QM/MM simulations.

In this work, taking advantage of the newly developed SCC-DFTBPR/MM Hamiltonian,[38] we explore the hydrolysis of a particular phosphate monoester, *p*-nitrophenyl phosphate ($pNPP^{2-}$, Figure 1b) in R166S AP and wild-type (WT) NPP; $pNPP^{2-}$ is a phosphate monoester whose hydrolysis has been widely studied in solution[7] and AP/NPP,[12,14] and it is a natural counterpart to the phosphate diester (methyl *p*-NPP, $MpNPP^{-}$, Figure 1c) that we have studied in ref 6. With a simple correction scheme for the systematic errors of SCC-DFTBPR for phosphate monoesters, our QM/MM simulations lead to observations in good agreement with available experimental data for both solution and enzyme systems. Our results suggest that $pNPP^{2-}$ hydrolysis proceeds through a loose transition state in both AP and NPP, despite the relatively rigid bimetallic zinc motif, and that the nature of the transition state in these enzymes is fairly close to that in solution. Therefore, together with our previous analysis[6] of diester hydrolysis in the AP enzymes, our QM/MM calculations provide further support to the hypothesis[12,14] that AP enzymes are able to recognize and stabilize multiple types of transition states in a single active site.

The paper is organized as follows: First we summarize computational methods and simulation setup. Next, we briefly review the reference solution reactions and then present analysis of $pNPP^{2-}$ hydrolysis (the first step) in R166S AP and WT NPP based on SCC-DFTBPR/MM simulations; we also present benchmark calculations using ab initio QM/MM calculations. Finally, we draw a few conclusions and discuss a number of remaining challenges.

COMPUTATIONAL METHODS

Enzyme Model Setup. We investigate the first step of the hydrolysis reaction of $pNPP^{2-}$ in an *E. coli* AP variant in which Arg166 is mutated to Ser and in the WT NPP (Figure 2). These systems are chosen because the chemical steps are fully rate-limiting in these enzymes and the mutation does not change the nature of the transition state.[9,11,16]

The enzyme models used in this work are similar to those in our previous study.[6] Therefore, we only briefly summarize the key features. The starting structure for the R166S AP mutant is based on the X-ray structure (PDB code 3CMR,[16] 2.05 Å resolution) for the enzyme with inorganic phosphate bound in the active site; for the *Xac* NPP, the crystal structure (PDB code 2GSU,[13] 2.00 Å resolution) is bound with adenosine monophosphate (AMP). In each case, starting from the PDB structure, the ligand is first "mutated" to $pNPP^{2-}$; the orientation of the nitrophenyl group is chosen randomly because, as shown below, it fluctuates greatly during the MD simulation. Hydrogen atoms are added by the HBUILD module[39] in CHARMM.[40] 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 neucleophiles and deprotonated in the Michaelis complex. Water molecules are added following the standard protocol of superimposing the system with a water droplet of 27 Å radius centered at $Zn1^{2+}$ (see Figure 2 for atomic labels) and removing water molecules within 2.8 Å from any atoms resolved in the crystal structure.[41] Protein atoms in the MM region are described by the all-atom CHARMM force field for proteins,[42] and water molecules are described with the TIP3P model.[43] The QM region includes groups most relevant to the reaction: the two zinc ions and their six ligands (Asp51, Asp369, His370, Asp327, His412, His331), Ser102 and $pNPP^{2-}$ for the R166S AP; for NPP, this includes two zinc ions and their six ligands (Asp54, Asp257, His258, Asp210, His363, His214), Thr90 and the substrate. Only the side chains of protein residues are included in the QM region and link atoms are added between C_α and C_β atoms. The treatment of the QM/MM frontier follows the DIV scheme in CHARMM.[44] A NOE potential is added to the C−O bonds in Asp51, which is coordinated to both Mg^{2+} and Zn^{2+}, to avoid over polarization.[6]

Due to the fairly large size of the QM region (>80 atoms) and extensive sampling required for the solvent-accessible active site of AP and NPP, the SCC-DFTBPR method[31] is used for PMF calculations. Benchmark calculations[6,31,45] and applications[46−49] indicate that it is comparable to the best semiempirical method available in the literature for phosphate chemistry,[25,50] including the hydrolysis of $MpNPP^{-}$ and $pNPP^{2-}$ in solution.[6,38] Nevertheless, as described below, we carry out explicit benchmark calculations for $pNPP^{2-}$ in AP and NPP.

The generalized solvent boundary potential (GSBP)[51−53] 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 restrained with force constants determined from the crystallographic B-factors.[54] All bonds involving hydrogen are constrained using the SHAKE algorithm,[55] 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, $\phi_s^{(o)}$, is evaluated with the linear Poisson−Boltzmann (PB) equation[56] 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 $\varepsilon_p = 1$, the water dielectric constant of $\varepsilon_w = 80$, 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.[57] The optimized radii of Nina et al.[58,59] 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 model[60] 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.

QM/MM Interaction Scheme and PMF Simulations. Due to the high charge of phosphate monoesters, the original QM/MM electrostatic Hamiltonian[61] in which the QM atoms interact with the MM atoms as point charges can lead to significant errors especially as the QM and MM atoms approach each other. In our recent work, we have developed an alternative scheme based on the Klopman−Ohno (KO) approximation[62−65] to the two-electron integrals:

$$H_{\mathrm{elec,KO}}^{\mathrm{QM/MM}} = \sum_{\alpha I} \frac{\Delta q_\alpha Q_I}{\sqrt{R_{\alpha I}^2 + a_\alpha \left(\frac{1}{U_\alpha(\Delta q_\alpha)} + \frac{1}{U_I}\right)^2 e^{-b_\alpha R_{\alpha I}}}} \quad (1)$$

where $U_\alpha(\Delta q_\alpha)$ depends explicitly[31,66] on the Mulliken charge of the QM atom (Δq_α); for the MM atom *I*, the Hubbard parameter takes the value based on atomic calculations and does not depend on charge. Effectively, this KO based QM/MM electrostatic scheme treats both QM and MM atoms as spherical charges of finite width (determined by the corresponding Hubbard parameters) and therefore takes charge penetration effect[67−69] into consideration. As demonstrated in our

recent work,[38] the KO scheme introduces only two element-dependent parameters (a_α,b_α) and is very effective at describing the interaction between highly charged solute and solvent molecules, as compared to full SCC-DFTB(PR) calculations. For example, with the KO scheme, the hydrolysis of two phosphate monoesters in solution can be well described with SCC-DFTBPR/MM simulations, while the error in the barrier height is >10 kcal/mol with the original QM/MM electrostatic Hamiltonian. The key results for $pNNP^{2-}$ hydrolysis in solution (free energy barrier and nature of transition state) are also summarized in Tables 1 and 2.

Table 1. $pNPP^{2-}$ Hydrolysis Barrier in Solution, R166S AP, and NPP from Experiments and Calculations

	exp[a]	SCC-DFTBPR/MM[b]	+ M06/MM correction[c]
solution	31.8	32.0	
R166S AP	18.0/*12.1*	13.5 (13.7/8.2)	18.7 [18.5]
NPP	−/*17.5*	14.0 (13.6/8.5)	19.0

[a]Free energy barriers (kcal/mol) calculated by transition state theory at 300 K based on experimental rate constants; values after the slash in italics are based on k_{cat}/K_M values. [b]Values without parentheses are from SCC-DFTBPR/MM PMF simulations; values with parentheses are adiabatic mapping barriers with M06/6-31+G**/MM (before slash) and SCC-DFTBPR/MM (after slash). Single point M06/MM calculations with the 6-311++G(d,p) basis set lead to very minor changes in the adiabatic barrier heights on the order of 0.1−0.3 kcal/mol. All SCC-DFTBPR/MM results for $pNPP^{2-}$ are obtained with the KO scheme[38] for QM/MM interactions. [c]Numbers without brackets are PMF results plus a barrier correction taken to be the difference in barriers from adiabatic mapping calculations at the M06/6-311+G(d,p)/MM and SCC-DFTBPR/MM levels; the value with brackets is the PMF barrier that includes a M06/6-31+G(d,p)/MM perturbative correction (eq 3) based on ∼400 snapshots (see Supporting Information).

To characterize the transition state and energetics of the phosphoryl transfer in the enzymes, PMF simulations are carried out with umbrella sampling.[70] 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 $PO^{lg} - PO^{nu}$, i.e., the bond length difference between the breaking and forming P−O bonds; our previous comparison[6] of such 1D PMF results to a 2D PMF using both PO^{lg} and PO^{nu} as the reaction coordinates confirmed that the 1D PMF is sufficient for our purpose. More than 51 windows are sampled to cover the relevant ranges of the reaction coordinate; each window is sampled for 100 ps with only the last 50 ps used for data analysis, and the force constant in the umbrella sampling is 150 kcal/mol·Å^{-2} for all windows. The probability distributions are combined together with the weighted histogram analysis method[71] to obtain the PMF along the reaction coordinate. 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 averaged key structural properties for each window are calculated and summarized in Table 2.

Benchmark Calculations Based on Minimizations and Reaction Path Calculations in the Enzyme. To explicitly test the applicability of SCC-DFTBPR/MM to phosphate monoester hydrolysis in AP and NPP, optimized structures for the Michaelis complex are compared to results from both B3LYP[72−74]/MM (see Supporting Information) and M06[75]/MM calculations; the calculations are carried out with the QChem[76] program interfaced with CHARMM (c36a2 version).[77] The basis set used in the B3LYP/MM calculations is 6-31G*[78] and 6-31+G** in the M06/MM calculations; these basis sets are the typical ones employed in ab initio QM/MM simulations of zinc-containing enzymes for structural properties.[79−82] 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. 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).

In addition, adiabatic mapping calculations using $PO^{lg}-PO^{nu}$ as the reaction coordinate are also carried out at both SCC-DFTBPR/MM, B3LYP/MM, and M06/MM levels to compare the approximate transition states on the potential energy surfaces. The protocols are similar to those used in our recent study.[6] To examine the effect of basis set on the energetics, single point energy calculations are carried out with M06/MM using the 6-311++G(d,p) basis.

M06/MM Correction of the SCC-DFTBPR/MM PMF Barrier. As indicated in our benchmark calculations (see below and Table 1), SCC-DFTBPR/MM tends to underestimate the reaction barriers of $pNPP^{2-}$ hydrolysis in the AP enzymes, therefore it is necessary to include high-level QM corrections to better compare with experimental data for the phosphoryl transfer barrier. A simple but fairly effective approach is to include a barrier correction based on adiabatic mapping calculations at different QM/MM levels,[83,84] i.e., the correction is the difference in barriers from adiabatic mapping calculations using QM/MM with lower- (SCC-DFTBPR) and higher-level (M06) QM methods. The basic assumption is that the "entropic contribution" is adequately estimated by comparing PMF and adiabatic mapping calculations with low-level QM/MM simulations, while the major energetic contribution is captured with reaction path/adiabatic mapping calculations using a high-level QM/MM potential function. Alternatively, we also explore using M06/MM in a one-step free energy perturbation scheme for the

Table 2. Key Structural Properties (all in Å) for the Transition State of the First Step of Phosphate Monoester and Diester Hydrolysis in Solution, AP and NPP, from SCC-DFTBPR/MM Simulations[a] and Other Computational Work (in italics)

	substrate	RC[b]	TC[c]	$P-O^{lg}$	$P-O^{nu}$	$Zn^{2+}-Zn^{2+}$
solution	$pNPP^{2-}$	−0.31	4.21	1.95	2.26	
	$MpNPP^-$	−0.20	4.66	2.23	2.43	
		−0.43(−0.63)[d]	*4.05 (4.35)*	*1.81 ± 0.05 (1.86)*	*2.23 ± 0.06 (2.49)*	
R166S AP	$pNPP^{2-}$	−0.41 ± 0.07	4.50 ± 0.19	2.04 ± 0.11	2.46 ± 0.10	4.10 ± 0.21
	$MpNPP^-$[e]	−0.11 ± 0.07	3.89 ± 0.14	1.89 ± 0.07	2.00 ± 0.09	3.93 ± 0.18
		0.36	*5.00*	*2.68 ± 0.11*	*2.32 ± 0.11*	*5.68*
NPP	$pNPP^{2-}$	−0.41 ± 0.07	4.63 ± 0.23	2.11 ± 0.13	2.52 ± 0.11	4.11 ± 0.21
	$MpNPP^-$	−0.20 ± 0.07	3.86 ± 0.14	1.83 ± 0.06	2.03 ± 0.09	3.92 ± 0.17
		−0.60	*5.66*	*2.53 ± 0.10*	*3.13 ± 0.10*	*5.50*

[a]The solution SCC-DFTBPR results[6,38] are based on an implicit solvent model with charge-dependent atomic radii.[45] For $MpNPP^-$ in the AP enzymes, SCC-DFTBPR/MM results are taken from ref 6; those for $pNPP^-$ are from the current work. [b]RC is defined as the difference between $P-O^{lg}$ and $P-O^{nu}$. [c]TC is defined as the sum of $P-O^{lg}$ and $P-O^{nu}$. [d]For values from previous work, numbers without parentheses are from refs 20 and 21 and those with parentheses are from ref 94. [e]The two substrate orientations result in very similar structural properties,[6] therefore only results for one orientation are included.

dx.doi.org/10.1021/ja403293d | J. Am. Chem. Soc. 2013, 135, 10457−10469

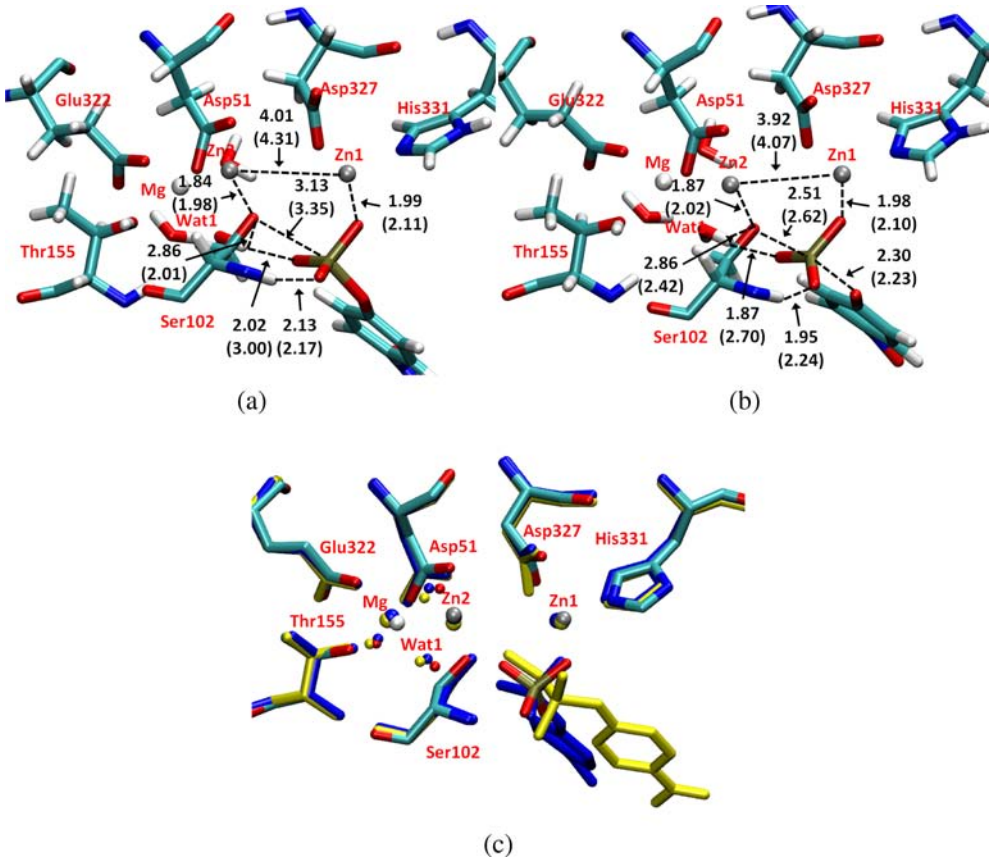


Figure 3. Benchmark calculations for $pNPP^{2-}$ in R166S AP. Key distances are labeled in Å. Numbers without parentheses are obtained with M06/6-31+G**/MM optimization; those with parentheses are obtained by SCC-DFTBPR/MM optimization with KO scheme. Asp369, His370, and His412 are omitted for clarity. (a) Reactant state in R166S AP; (b) transition state in R166S AP by adiabatic mapping; and (c) overlay of crystal structure with PO_4^{3-} (colorful), M06/6-31+G**/MM optimized structures with $pNPP^{2-}$ (blue) and $MpNPP^-$ (yellow). Hydrogen atoms are omitted. For additional comparisons of DFT(M06 or B3LYP)/MM and SCC-DFTBPR/MM structures, see Supporting Information.

SCC-DFTBPR/MM results for selected PMF windows (the Michaelis complex and the transition-state region):

$$\Delta G_{\text{M06-SCC}} = -k_B T \ln < e^{-\beta(U_{\text{M06/MM}} - U_{\text{SCC/MM}})} >_{\text{SCC/MM}} \quad (2)$$

Since only a relatively small (~400) number of snapshots from SCC-DFTB/MM trajectories are used, a second-order cumulant expansion is used to improve the numerical stability of the perturbation calculation:

$$\Delta G_{\text{M06-SCC}} = <U_{\text{M06/MM}} - U_{\text{SCC/MM}} >_{\text{SCC/MM}} - \frac{\beta}{2}[<(U_{\text{M06/MM}} - U_{\text{SCC/MM}})^2 >_{\text{SCC/MM}} - <U_{\text{M06/MM}} - U_{\text{SCC/MM}} >^2_{\text{SCC/MM}}] \quad (3)$$

As discussed extensively in the literature,[85,86] such one-step perturbation is effective only if the configuration space distributions at the two levels overlap significantly; this is assumed to be the case considering the previous observation[31] that higher level (e.g., MP2) single point energy calculations at SCC-DFTBPR geometries give similar results as calculations at geometries optimized at the DFT level. To control the convergence of such perturbative correction, the structure of the QM region is fixed to that in a snapshot which is close to the average structure from the relevant umbrella sampling window; the MM environment is resampled over 100 ps of SCC-DFTBPR/MM MD simulations, and up to 450 snapshots are taken from the last 50 ps of these trajectories to compute the perturbative correction according to eq 3. The similar perturbative scheme has been found successful in combining MM and QM/MM simulations for metalloenzymes[87,88] as well as correcting semiempirical simulations with ab initio QM methods.[89] Nevertheless, the correction in the current work should be considered qualitative and additional studies are in progress to improve the statistical behaviors of multilevel QM/MM calculations. For NPP, since the experimental k_{cat} data are not available and difficult to determine in general, we have only pursued the barrier correction based on adiabatic mapping calculations at different QM/MM levels.

RESULTS AND DISCUSSION

Before discussing the results from our study in the context of available experimental data, it is worth emphasizing that enzyme catalyzed reactions involve multiple steps,[90] and it is often not straightforward to map experimentally measured k_{cat} and k_{cat}/K_m values to the free energy barrier(s) calculated in a computational study.[91] As mentioned above, we study $pNPP^{2-}$ and $MpNPP^-$ hydrolysis in R166S AP because the chemical step is rate-limiting. Even so, k_{cat} was measured only for a few cases,[16] and most kinetic data reported for the AP enzymes are k_{cat}/K_m values; the K_m values do not correspond precisely to substrate binding/dissociation equilibrium constants since there is a change in rate-limiting step in the progression from subsaturating to saturating conditions (Lassila, private communication).

First Step of $pNPP^{2-}$ Hydrolysis in R166S AP. $pNPP^{2-}$ is a cognate substrate of AP. The experimentally measured phosphoryl transfer barrier, which includes the substrate binding process (k_{cat}/K_m), is equal to 12.1 kcal/mol for R166S AP (Table 1) at 300K. Compared to the similar diester, $MpNPP^-$, for which k_{cat}/K_m corresponds to a free energy barrier of 18.0 kcal/mol, R166S AP favors the hydrolysis of $pNPP^{2-}$ by 5.9 kcal/mol. Since the AP active site features several positively charged motifs, e.g.,

dx.doi.org/10.1021/ja403293d | J. Am. Chem. Soc. 2013, 135, 10457−10469

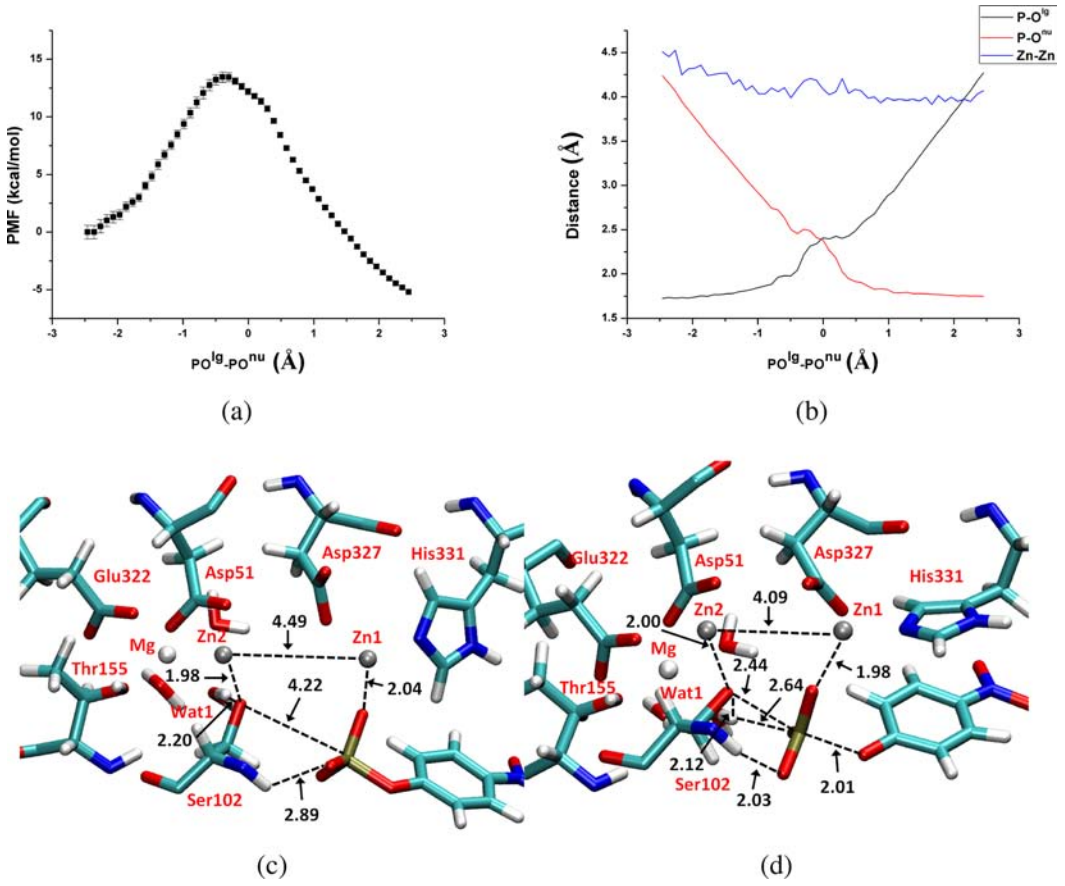


Figure 4. PMF calculation results for $pNPP^{2-}$ hydrolysis in R166S AP with SCC-DFTBPR/MM. Key distances are labeled in Å and energies are in kcal/mol. (a) PMF along the reaction coordinate with error bar included. (b) Changes of average key distances along the reaction coordinate. (c) Snapshot for the reactant state, with average key distances labeled. (d) Snapshot for the transition state, with average key distances labeled. Asp369, His370, and His412 are omitted for clarity.

the magnesium ion, it is likely that the binding free energy of a phosphate monoester is greater than that for a diester. Therefore, the barrier difference for the actual chemical step between $pNPP^{2-}$ and $MpNPP^{-}$ is likely <5.9 kcal/mol. Along this line, the free energy barrier for the chemical step (k_{cat}) has been measured only for $pNPP^{2-}$ as 18.0 kcal/mol in R166S AP,[16] while no value has been reported for diesters.

The comparison of optimized structures for the Michaelis complex by B3LYP/MM (see Supporting Information), M06/MM and SCC-DFTBPR/MM shows generally good agreement between the three levels (Figure 3, also see Supporting Information), including the Zn−Zn distance; the similar observation was made in an independent study with the solvent boundary potential included for both QM/MM levels.[92] The O_{Ser102}−P distances in the optimized structures are 3.1 (3.4) Å in M06/MM (SCC-DFTBPR/MM), close to the value of 3.1 Å in the crystal structure. The O2 of the substrate (see Figure 2 for labels) coordinates to one of the zinc ions, and O1 with the phenyl group is solvated by water molecules. O4 and the nearby Ser102 backbone amide form a hydrogen bond. At the M06/MM level, a magnesium bound water (Wat1) forms a shorter hydrogen bond with O3 of $pNPP^{2-}$ than with Ser102; the hydrogen bonding distances are 2.0 and 2.9 Å, respectively. The opposite situation was observed in our previous work[6] involving a phosphate diester, $MpNPP^{-}$, in the R166S AP active site; there Wat1 only formed a hydrogen bond with Ser102 but not with the substrate oxygen in the Michaelis complex. The stronger interaction between Wat1 and $pNPP^{2-}$ is consistent with the different charge states of $pNPP^{2-}$ and $MpNPP^{-}$. Indeed, the structural features observed for $pNPP^{2-}$ binding are also similar to those seen in the crystal structure with an inorganic phosphate (most likely PO_4^{3-} after transferring a proton to the deprotonated Ser102)[93] bound as an inhibitor (see Figure 3c for the comparison).

Curiously, Wat1 has a stronger interaction with Ser102 than with O3 of $pNPP^{2-}$ at the SCC-DFTBPR/MM level, suggesting that the KO scheme still has room for further improvement (however, see discussions below). For most structural features, SCC-DFTBPR/MM and M06/MM lead to the same general trends in the approximate transition state from adiabatic mapping calculations. As shown in Figure 3b, compared to the M06/MM result, the main differences in the approximate transition state at the SCC-DFTBPR/MM level include a slightly tighter P−O^{lg} distance, weaker interactions between oxygen and zinc and a weaker hydrogen bond between Wat1 and the metaphosphate. The activation barrier at the SCC-DFTBPR/MM level is lower than that calculated with M06/MM (8.2 vs 13.7 kcal/mol). Nevertheless, the important point is that the nature of the approximate transition state as reflected by the tightness coordinate (TC) is similar at the two QM/MM levels; both point to a rather loose structure, with a TC value of 4.81 and 4.85 Å, respectively. Therefore, the SCC-DFTBPR/MM approach provides a semiquantitative description for the phosphoryl transfer of monoesters in the active site of R166S

dx.doi.org/10.1021/ja403293d | J. Am. Chem. Soc. 2013, 135, 10457−10469

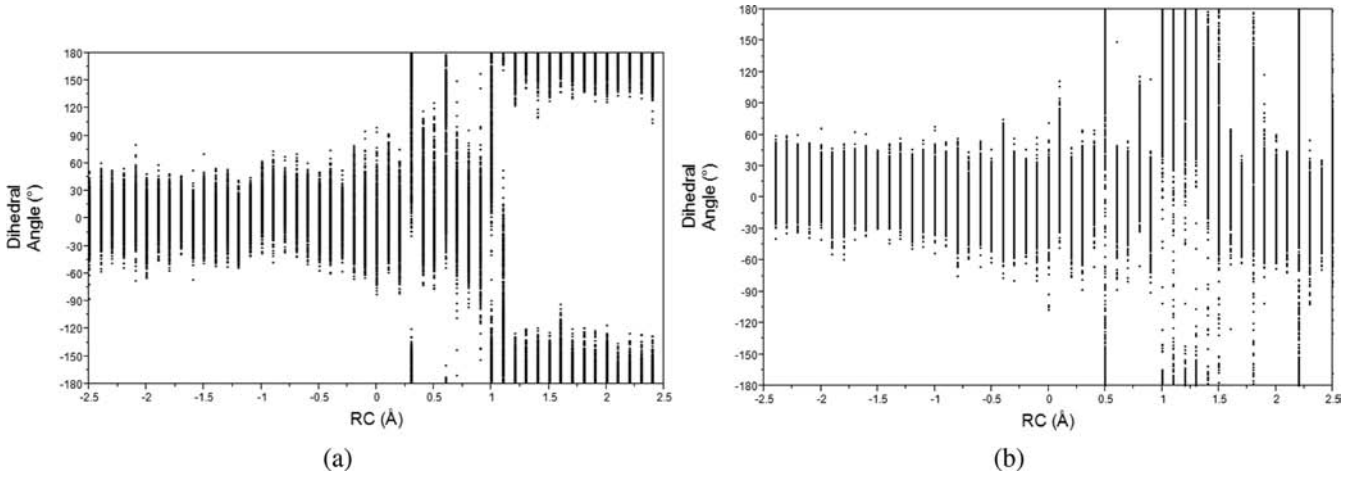


Figure 5. Leaving group (*p*-nitrophenyl) adopts diverse orientations during PMF simulations for the hydrolysis of $pNPP^{2-}$ in both (a) R166S AP and (b) NPP. Distributions of the dihedral angle (P−O1−C−C, see Figure 2 for labels) in all umbrella sampling windows are plotted.

AP, although further improvements in the methodology are still needed for more quantitative results.

The PMF profile at the SCC-DFTBPR/MM level (Figure 4a) indicates a single step exothermic reaction with the barrier at the reaction coordinate (RC) being near −0.4 Å; this is qualitatively different from the two-step mechanism observed in previous AM1(d)-PhoT/MM studies.[19−21] The free energy barrier is 13.5 kcal/mol, which is lower than the experimental estimate[16] based on k_{cat} by 4.5 kcal/mol; this is expected since the adiabatic mapping benchmark discussed above finds that SCC-DFTBPR/MM underestimates the barrier compared to M06/MM by ∼5 kcal/mol (Table 1). With the single-step M06/MM free energy perturbation using ∼400 snapshots, the estimated correction for the barrier is also around 5 kcal/mol, although the convergence is clearly not sufficient for a quantitative correction (see Supporting Information). Considering the ∼5 kcal/mol barrier correction, relative to the calculated barrier of 24.4 kcal/mol for the chemical step of $MpNPP^-$ in our previous work,[6] the monoester reaction is favored in the R166S AP by ∼5−6 kcal/mol, which is qualitatively consistent with available experimental data on these substrates; as discussed above, the experimental k_{cat}/K_m data suggest 5.9 kcal/mol as the upper limit for the barrier difference in the chemical step between $pNPP^{2-}$ and $MpNPP^-$.

Several important structural properties during the PMF calculations are plotted as functions of the RC (Figure 4b). The bond lengths of $P-O^{lg}$ and $P-O^{nu}$ change smoothly and intersect at RC around 0 Å. The Zn−Zn distance fluctuates around 4 Å and remains close to the value observed in the various crystal structures;[16] the largest value (∼4.5 Å) is observed for the reactant state, and the zinc−zinc distance compacts slightly in the transition state. The transition state is located at RC ∼ −0.4 Å (Table 2), more negative than for $MpNPP^-$ (∼ 0 Å), and the averaged $P-O^{lg}$ and $P-O^{nu}$ bond lengths are 2.04 and 2.46 Å, respectively. Compared to the transition state computed[6] for $MpNPP^-$, both bonds are elongated for $pNPP^{2-}$, and the TC increases from 3.89 to 4.50 Å. Therefore, $pNPP^{2-}$ hydrolysis goes through a substantially looser transition state compared to the comparable diester.

In the thermally equilibrated reactant state (Figure 4c), similar to the situation found in the minimized structure (Figure 3a), $pNPP^{2-}$ binds with Zn1 via a nonbridging oxygen and forms a hydrogen bond with a backbone amide. Different from the expectation based on phosphate-bound crystal structure,[15] Wat1 forms a hydrogen bond with the deprotonated Ser102, instead of with the substrate. This is similar to the observation for $MpNPP^-$ binding and probably due to the increased PO^{nu} distance in the reactant state compared to when an inorganic phosphate is bound to the enzyme; however, as discussed above, SCC-DFTBPR/MM and M06/MM appear to have different hydrogen-bonding patterns for Wat1 in the reactant state. Nevertheless, in the transition state (Figure 4d), as Ser102 attacks the substrate, Wat1 partially breaks the hydrogen bond with Ser102 and forms a weak hydrogen bond with a $pNPP^{2-}$ nonbridging oxygen; the average distance to the nonbridging oxygen is still long, ∼2.7 Å, which is similar in the approximate transition state from adiabatic mapping calculations. This was also observed[6] for the $MpNPP^-$ reactions and is consistent with the proposed role of Wat1 in lowering the reaction barrier.[15] Therefore, we expect that SCC-DFTBPR/MM simulations are able to capture the main role of Wat1 during the reaction despite the uncertainty associated with its hydrogen-bonding pattern in the Michaelis complex.

Another interesting and somewhat unexpected observation in the calculated transition state (Figure 4d) is that the leaving group oxygen does not directly coordinate with Zn1 but gets solvated by water molecules accessible to the active site. Again, this is similar to our previous observation for $MpNPP^-$ and at odds with the coordination pattern of vanadate,[95] a popular transition-state analog for phosphoryl transfers, in the crystal structure; the vanadate structure was used to suggest that the leaving group oxygen directly interacts with a zinc ion in the transition state. To clarify this point, we carry out one independent calculation with the initial structure prepared so that the leaving group oxygen is constrained to bind with Zn1 and the RC is constrained at 0.0 Å (see Supporting Information for an illustration); a constraint is applied to RC so that the charge distribution of the reactive fragment reflects that expected for the transition state. Following 50 ps of equilibration, the constraint in Zn1−O1 is removed, and the system is further equilibrated with 50 ps. The final structure is very similar to that shown in Figure 4d, i.e., the leaving group quickly dissociates from Zn1 and becomes solvated by water once the constraint is removed. Although even more extensive benchmark for zinc-leaving group interaction at the SCC-DFTBPR/MM level (or improved parametrization with DFTB3)[96] remains an active effort in our group, we note that the loose transition state for

dx.doi.org/10.1021/ja403293d | J. Am. Chem. Soc. 2013, 135, 10457−10469

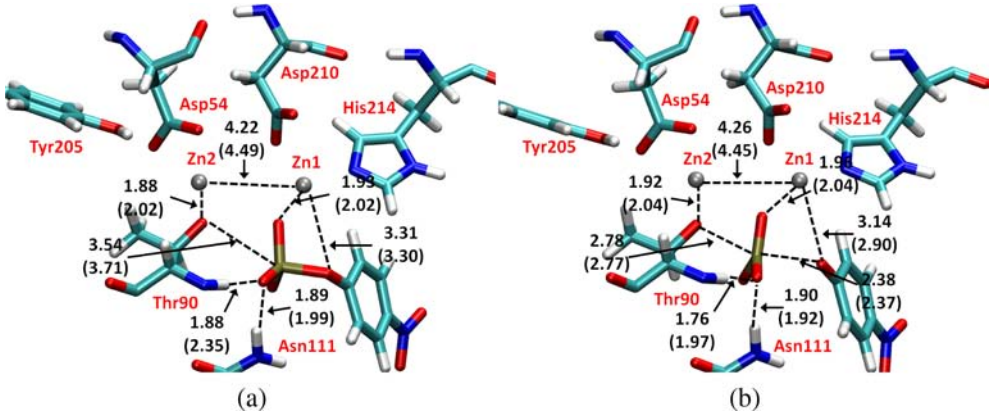


Figure 6. Benchmark calculations for $pNPP^{2-}$ in NPP. Key distances are labeled in Å. Numbers without parentheses are obtained with M06/6-31+G**/MM optimization; those with parentheses are obtained by SCC-DFTBPR/MM optimization with KO scheme. (a) Reactant state in NPP and (b) transition state in NPP by adiabatic mapping. Asp257, His258, and His363 are omitted for clarity. For additional comparisons of DFT(M06 or B3LYP)/MM and SCC-DFTBPR/MM structures, see Supporting Information.

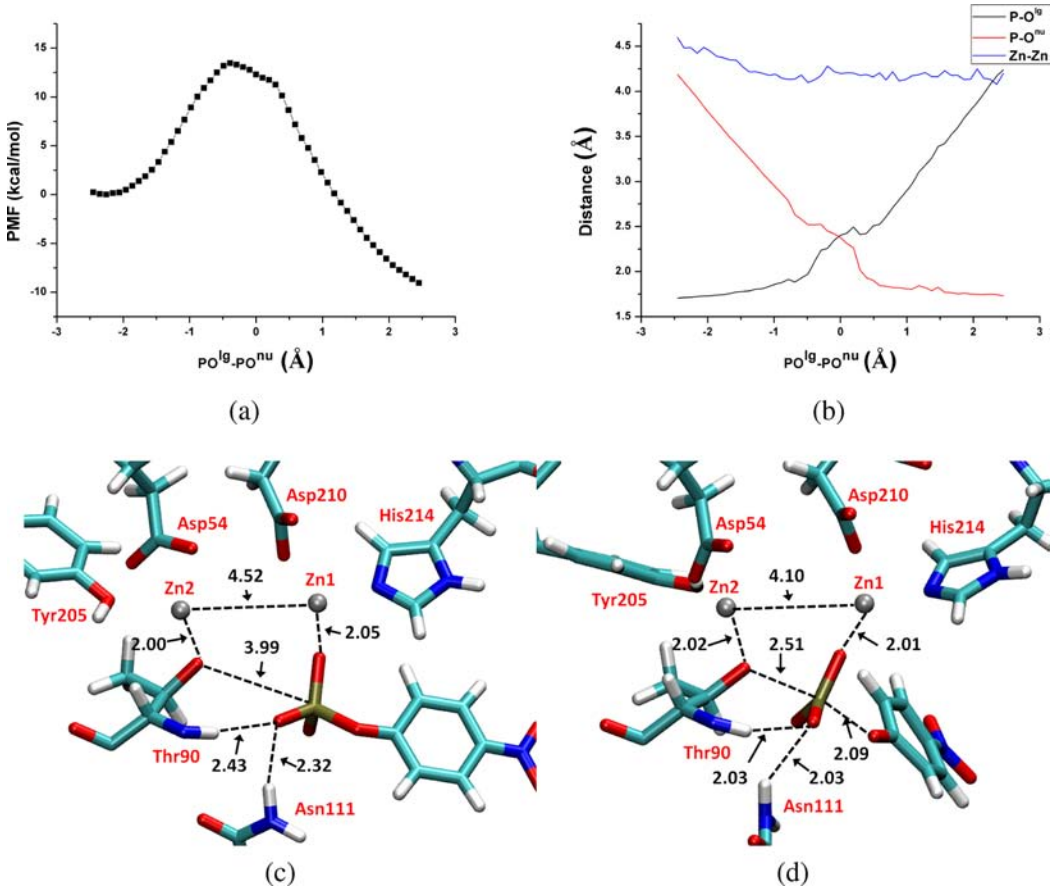


Figure 7. PMF calculation results for $pNPP^{2-}$ hydrolysis in NPP with SCC-DFTBPR/MM. Key distances are labeled in Å, and energies are in kcal/mol. (a) PMF along the reaction coordinate. (b) Changes of average key distances along the reaction coordinate. (c) Snapshot for the reactant state, with average key distances labeled. (d) Snapshot for the transition state, with average key distances labeled. Asp257, His258, and His363 are omitted for clarity.

$pNPP^{2-}$ is not fully compatible with the structure of the bimetallic motif. Indeed, the TC for the transition state is ~4.5 Å, which is substantially longer than the average zinc−zinc distance (4.10 Å) in the transition state; by comparison, the corresponding O−V bonds sum to 3.64 Å. Therefore, it is not unreasonable that the binding mode of vanadate is different from that of the actual transition state.

The dihedral angle distributions that characterize the orientation of the leaving group (*p*-Nitrophenyl) during the PMF simulations are shown in Figure 5. It is evident that the leaving group adopts multiple orientations throughout the reaction, with the width of the dihedral distribution close to or larger than 120°. This degree of orientational flexibility for the leaving group is not surprising because the substrate binding site is large and solvent accessible.

dx.doi.org/10.1021/ja403293d | *J. Am. Chem. Soc.* 2013, 135, 10457−10469

First step of $pNPP^{2-}$ Hydrolysis in NPP. For NPP, $pNPP^{2-}$ is a noncognate substrate. The experimental barrier based on k_{cat}/K_m is 17.5 kcal/mol, slightly higher than the value of 14.3 kcal/mol for $MpNPP^-$. There is no available experimental data for the chemical step (k_{cat}).

Similar to the comparisons made above for AP, SCC-DFTBPR/MM minimizations for $pNPP^{2-}$ in NPP give overall similar results to M06/MM calculations (Figure 6a,b). The O_{Thr90}−P distance increases from 3.2 Å in the crystal structure, which contains AMP as the inhibitor, to 3.5 (3.7) Å at the M06/MM (SCC-DFTBPR/MM) level. The substrate O2 coordinates with Zn1, while O4 forms hydrogen bonds with Asn111 and the backbone amide of Thr90; the hydrogen bonding distances are comparable at the two QM/MM levels (Figure 6a). The optimized Zn−Zn distance is 4.2 (4.5) Å at the M06/MM (SCC-DFTBPR/MM) level. For the results of adiabatic mapping calculations, similar to the situation in AP, SCC-DFTBPR/MM underestimates the barrier compared to M06/MM by a few kcal/mol (8.5 vs 13.6 kcal/mol). Nevertheless, the transition-state geometries are very consistent at the two levels of theory (Figure 6b). Compared to the approximate transition states obtained for R166S AP (Figure 3c), the ones calculated for $pNPP^{2-}$ hydrolysis in NPP appear to adopt a slightly different configuration such that the leaving group forms a very weak interaction with Zn1; the O−Zn1 distance is 3.14 (2.90) Å at the M06/MM (SCC-DFTBPR/MM) level, in comparison to the values of ∼3.30 Å in the minimized reactant. Nevertheless, the nature of the transition state remains loose in NPP; the TC values are 5.16 and 5.14 Å at the M06/MM and SCC-DFTBPR/MM levels, respectively. These values, especially $P-O^{nu}$ distances, are somewhat longer than those in the R166S AP (Figure 3b).

The calculated PMF (Figure 7) at the SCC-DFTBPR/MM level indicates an exothermic reaction with the transition state at RC ∼ −0.4 Å. The calculated free energy barrier is 14.0 kcal/mol, similar to the value observed for R166S AP (13.5 kcal/mol); as discussed above, this is likely underestimated by 4−6 kcal/mol, due to the systematic error of SCC-DFTBPR; with the barrier correction based on adiabatic mapping calculations at different QM/MM levels, an improved estimate for the free energy barrier is ∼19.0 kcal/mol (Table 1). Since there is no experimental k_{cat} value for $pNPP^{2-}$ hydrolysis in NPP and we focus mainly on the nature of the transition state in this study, we have not pursued the M06/MM perturbative correction (eq 3) for the energetics in NPP. Nevertheless, we note that experimental k_{cat}/K_m values indicate that an upper limit for the barrier difference between R166S AP and NPP for $pNPP^{2-}$ hydrolysis is 5.4 kcal/mol; since $pNPP^{2-}$ is expected to have a substantially higher binding affinity to R166S AP than to NPP, the barrier difference for the chemical step is expected to be quite smaller than this upper limit, qualitatively consistent with the similar barriers calculated for $pNPP^{2-}$ hydrolysis in the two enzymes (Table 1).

In terms of the key features of the transition state (Figure 7d), they are very similar to those in the R166S AP; the RC is ∼ −0.4 Å at the transition state, and the corresponding TC is 4.63 Å (Table 2), indicating a much looser transition state than $MpNPP^-$ in NPP (with a TC of 3.86 Å),[6] although the TC value from the PMF simulations is notably lower than that in the approximation transition state from adiabatic mapping calculations (∼5.1 Å). Another difference from the adiabatic mapping result is that the leaving group oxygen does not directly interact with Zn1 (∼4.25 vs ∼3 Å in Figure 6b) in the transition state but is solvated by water molecules accessible to the active site, similar to the situation observed for the R166S AP. These differences appear to be correlated with the orientation of the Thr90 side chain, which is shifted outward in the PMF simulations than in the minimizations, which do not include thermal fluctuations of the enzyme (see Supporting Information for a comparison of active site structures from minimization and MD simulations). Similar to the case of R166S AP, the leaving group is observed to adopt a diverse set of orientations during the PMF simulations, with the width of the relevant dihedral angle distribution close to or >120° (Figure 5b).

Comparison of Phosphate Mono- And Diester Hydrolysis in the AP Enzymes. Together with our previous study of phosphate diester hydrolysis in solution and the AP enzymes[6] and the more recent study of phosphate monoester hydrolysis in solution,[38] our current work on phosphate monoesters in the AP enzymes provides a comprehensive view for the key features of phosphoryl transfers in different environments. Since AP and NPP have different catalytic specificities[12,15] and phosphate mono/diesters feature transition states of different nature in solution,[7] there are three levels of comparison regarding the nature of the phosphoryl transfer transition state in different environments: (1) AP vs NPP for the same substrate; (2) different substrates in the same enzyme active site; and (3) enzyme vs solution for the same substrate. Before making these comparisons using SCC-DFTBPR/MM results, we note that benchmarks presented for diester hydrolysis in ref 6 and for monoester hydrolysis in ref 38 and here reveal that the SCC-DFTBPR/MM model has systematic errors for both structural and energetic properties of phosphoryl transfer transition states. However, as discussed below, since we focus on the relative trends motivated by the three levels of comparison, we expect that the findings based on SCC-DFTBPR/MM results are qualitatively meaningful, while more quantitative insights require further developments in the QM/MM methodology.

First, despite their rather distinct catalytic specificities, R166S AP and NPP feature phosphoryl transfer transition states of very similar nature for the same substrate in our calculations, i.e., the transition state is synchronous for diesters and much looser for monoesters. Both TC values (Table 2) and other key features for the transition states (Figures 4 and 7), such as the orientation and level of solvation of the leaving group, are similar in the two enzyme active sites. This is not entirely surprising since the bimetallic zinc motif adopts similar structures in the two enzymes, and a previous study[15] showed that the specificities of the AP enzymes have a significant contribution from their distinct binding preferences to different substrates; as a matter of fact, as the calculations from this work indicate, the chemical step barrier is rather similar in R166S AP and NPP for a phosphate monoester. The R166S AP active site has an additional positively charged magnesium site relative to NPP, although, apparently, this site has a minimal impact on the nature of the transition state. Therefore, our calculations support the argument that stabilizing electrostatic interactions do not necessarily induce a significant change in the nature of the transition state.[11]

Second, for a given enzyme active site, our calculations show that different substrates go through transition states of a rather different nature. The difference in the TC values between a diester and a comparable monoester (e.g., $MpNPP^-$ vs $pNPP^{2-}$) is consistently ∼0.6−0.7 Å for both R166S AP and NPP. Therefore, we explicitly support the model that a single active site is able to recognize and stabilize transition states of different nature. Comparison of the transition-state structures suggests fairly similar configurations for the bimetallic zinc motif for the

dx.doi.org/10.1021/ja403293d | J. Am. Chem. Soc. 2013, 135, 10457−10469

hydrolysis of different substrates; the Zn−Zn distance is ∼3.9 Å for diester transition state and slightly expanded to ∼4.1 Å for monoester transition state, while the magnitude of the thermal fluctuation is ∼0.2 Å in both cases. Therefore, the "plasticity" of the bimetallic zinc motif is only one of the reasons that the AP enzymes can recognize multiple types of transition state. Our calculations suggest that the significant degree of solvent accessibility of the active site is another reason that the AP enzymes are able to accommodate transition states of different nature. With the leaving group being surrounded by solvent molecules but still close to the zinc ions (see Figures 4d and 7d), a loose transition state is stabilized without causing large structural distortions of the bimetallic site.

Finally, regarding the comparison of transition states in the AP enzymes and in solution, our calculations suggest that the trends depend on the type of substrate. For phosphate diesters, the computed transition state becomes slightly tighter in the AP enzymes compared to solution; the opposite trend is observed here, however, for the phosphate monoesters. For the diester $MpNPP^-$, SCC-DFTBPR calculations with an implicit solvent model predicts the TC as 4.66 Å, which is ∼0.3−0.4 Å longer than results from previous calculations[94] (Table 2; note though the potential energy surface is relatively flat along these degrees of freedom near the transition state);[6] in the AP enzymes, the TC values decrease to ∼3.9 Å with SCC-DFTBPR/MM.[6] For the monoester $pNPP^{2-}$, the prediction from SCC-DFTBPR/PB is 4.2 Å, while the value is slightly shorter by ∼0.2 Å with explicit solvent SCC-DFTBPR/MM simulations;[38] these values are shorter compared to the SCC-DFTBPR/MM values of 4.5−4.6 Å in the AP enzymes. Therefore, the magnitude of the change from solution to the enzyme is modest (∼0.4 Å) when uncertainty in our QM/MM methodology is considered, although the trends appear robust.

Since experimental studies[12,14] focused on the leaving group properties, it is worth noting that the changes in the $P-O^{lg}$ distance in the computed transition state are rather small between enzyme and solution cases. For $pNPP^{2-}$, the change is ∼0.1−0.2 Å. For $MpNPP^-$, the changes appear more significant, but benchmark calculations indicated that SCC-DFTBPR tends to overestimate the $P-O^{lg}$ distance for diester hydrolysis transition state in solution by 0.3−0.4 Å relative to other theoretical calculations. Therefore, considering these limitations of our current QM/MM methodology, it is most sensible to conclude that solution and enzyme transition states have, in fact, largely comparable $P-O^{lg}$ distances for both phosphate monoesters and diesters. This is qualitatively consistent with the interpretation of experimental LFER and KIE data,[12,14] although as discussed above, our calculations suggest that there are actually subtle but consistent differences between enzyme and solution transition states.

Another interesting observation from our recent studies[6,38] is that the transition state for phosphate monoesters in solution is not necessarily looser than diesters. In fact, the computed transition state for $pNPP^{2-}$ is tighter than that for $MpNPP^-$ (TC of 3.94 vs 4.66 Å), even when considering the overestimated $P-O^{lg}$ distance for diester transition states[6] by SCC-DFTBPR/PB. This may not be entirely surprising considering the following: The two hydrolysis reactions feature different nucleophiles, OH^- for diesters but water for monoesters. For the hydrolysis of monoesters, however, the water first transfers one proton to the phosphate monoester, which effectively becomes a diester-like substrate for the subsequent nucleophilic attack by OH^-. Since the extra methyl group in $MpNPP^-$ is electron pushing, the transition state is expected to be looser for $MpNPP^-$ with longer $P-O^{lg/nu}$ distances; this is what we observe computationally. Along this line, we note that a phosphate monoester is not protonated in the AP enzymes (since the nucleophile is a deprotonated Ser or Thr), thus its effective charge is indeed more different from that of a diester as compared to the situation in solution. As a result, it is reasonable that the difference between monoester/diester transition states is more significant in the enzyme (Table 2) than in solution.

■ CONCLUDING REMARKS

In this work, we study the first step of $pNPP^{2-}$ hydrolysis in R166S AP and WT NPP using SCC-DFTBPR/MM simulations and a recently introduced QM/MM interaction scheme.[38] Together with our previous studies of phosphate monoester reactions in solution[38] and diester reactions in solution and the AP enzymes,[6] the current work helps provide a fairly complete model of the catalytic mechanisms in the AP enzymes at a semiquantitative level.

Our calculated free energy barriers for the chemical step in the AP enzymes are qualitatively consistent with available experimental data. The direct comparison of phosphoryl transfer transition states for AP and NPP indicates that similar loose transition states are involved in both enzymes, despite the fact that phosphate monoesters are the cognate substrates for AP but promiscuous substrates for NPP. The computed loose transition states are clearly different from the more synchronous ones for diester reactions calculated in solution and the AP enzymes.[6] Therefore, our results explicitly support the proposal that AP enzymes are able to recognize and stabilize different types of transition states. Analysis of the structural features of the active site in different transition states indicates that the plastic nature of the bimetallic site plays a small role in accommodating multiple transition states and that the solvent accessibility of the active site in the AP enzymes also contributes to their ability to stabilize diverse transition state structures without causing large structural distortions of the bimetallic motif. Regarding the comparison to solution reactions, our calculations suggest that the transition states in the AP enzymes are generally similar to the solution counterparts, especially for the $P-O^{lg}$ distance; this is qualitatively consistent with available experimental LFER and KIE data. On the other hand, our calculations also indicate that there are subtle but consistent differences between enzyme and solution transition states; for diesters, the transition state becomes slightly tighter in the enzyme, while the opposite is observed for monoesters. The opposite trends are due in part to the fact that phosphate monoester gets protonated in solution prior to the nucleophilic attack by OH^-, while this does not occur in the enzyme active site.

Regarding the comparison to previous QM/MM simulations using the AM1(d)-PhoT/MM approach, our monoester results are fundamentally different from the two-step mechanism observed thereof for an alkyl phosphate monoester in AP,[19] and we do not observe large structural deviation of the bimetallic zinc motif from available crystal structures with either phosphate monoesters or diesters. Collectively, our recent[6,38] and current works suggest that the observation of significantly modified transition states in refs 19−21 is likely an artifact due to the sampling of a significantly distorted bimetallic zinc motif.

Another interesting observation concerns the use of vanadate as phosphryl transition-state analog. For reactions in the AP enzymes, crystal structures with a bound vanadate were used to suggest that the leaving group oxygen directly interacts with one

zinc ion in the transition state. In our previous diester studies,[6] this direct interaction was not observed, and we speculated that the reason is mainly due to the different charges: a diester only bears −1 charge, while vanadate has a high charge of −3. In this study, the phosphate monoester $pNPP^{2-}$ bears a −2 charge and therefore is more similar to vanadate. However, the TC in the transition state is >4.5 Å, which is substantially longer than the corresponding value for vanadate (∼3.6 Å) and the zinc−zinc distance (∼4.0 Å) in AP/NPP. Therefore, from a structural perspective, vanadate is also not an ideal analog, which might explain that the direct interaction between the leaving group and zinc ion is not observed in the computed $pNPP^{2-}$ hydrolysis transition state. Although more extensive analysis of zinc-ligand interactions at the SCC-DFTB(PR)/MM level are being carried out as part of our effort to parametrize DFTB3[96] for metal ions, our results do highlight the potential caveats of using vanadate to infer the binding mode of phosphoryl transfer transition state in enzymes, especially when the active site is flexible and/or solvent accessible.

In the future, it is important to go beyond the semiquantitative descriptions in the recent[6] and current analyses. More systematic and quantitative comparison with available experimental KIE and LFER data is valuable to firmly establish the nature of the phosphoryl transition state for different substrates in different environments; in addition to properties, such as the TC value, it is important to probe the binding mode of the leaving group with more systematic leaving group KIE analysis. These are not simple tasks and tackling them requires pushing forward the SCC-DFTB approach for phosphate chemistry and metal ions, adaptive QM/MM partitioning[97,98] as well as a tighter integration with high-level QM/MM methods for both structural and energetic properties.

Finally, it is interesting to ponder to what degree the observations from our study of the AP enzymes are applicable to other enzymes that exhibit significant promiscuous activities. Due to the potential connection to the mechanism of enzyme evolution[99,100] and rational enzyme (re)engineering,[101−103] the analysis of catalytic promiscuity has become more systematic in recent years.[5,104−106] Structural flexibility of the active site appears to be a common feature for enzymes with a high degree of catalytic promiscuity; although this is hardly surprising if an enzyme is to bind diverse substrates, it is worth noting that subtle structural variations, such as displacement of a metal binding site, may underlie the promiscuous catalytic activity.[107] A high degree of solvent accessibility, however, is not always observed. Our analysis of AP and NPP points to different transition states for different substrates, although the same set of protein residues participates explicitly in the catalysis. In other systems, the enzyme may employ different catalytic groups for different substrates[23] or impose transition states of similar nature for different substrates.[108] Therefore, there appears to be diverse mechanisms that underlie the phenomena of catalytic promiscuity in enzymes,[5,104−106] highlighting the need of employing an integrated experimental/computational approach for understanding and predicting promiscuous enzyme activities. Another intriguing research direction is to couple such mechanistic analysis with advances in ancestral gene reconstruction,[109,110] which allows one to explore how the interplay between catalytic promiscuity and specificity during evolution shapes the functional landscape of "modern" enzymes.[111]

■ ASSOCIATED CONTENT

Supporting Information

Results from B3LYP/MM minimization and adiabatic mapping calculations and convergence of M06/MM correction for the SCC-DFTBPR/MM PMF are included. Full references for refs 42, 76 and 77 are included. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

cui@chem.wisc.edu

Notes

The authors declare no competing financial interest.

■ 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 Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant no. OCI-1053575, are greatly appreciated; computations are also supported in part by National Science Foundation through a major instrumentation grant (CHE-0840494) to the Chemistry department and by the Center for High Throughput Computing (CHTC) at UW-Madison.

■ REFERENCES

(1) Galperin, M.; Bairoch, A.; Koonin, E. *Protein Sci.* **1998**, *7*, 1829−1835.
(2) Galperin, M.; Hedrzejas, M. *Proteins: Struct., Funct., Bioinf.* **2001**, *45*, 318−324.
(3) Lassila, J.; Herschlag, D. *Biochemistry* **2008**, *47*, 12853−12859.
(4) van Loo, B.; Jonas, S.; Babtie, A. C.; Benjdia, A.; Berteau, O.; Hyvönen, M.; Hollfelder, F. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 2740−2745.
(5) Jonas, S.; Hollfelder, F. *Pure Appl. Chem.* **2009**, *81*, 731−742.
(6) Hou, G. H.; Cui, Q. *J. Am. Chem. Soc.* **2012**, *134*, 229−246.
(7) Lassila, J. K.; Zalatan, J. G.; Herschlag, D. *Annu. Rev. Biochem.* **2011**, *80*, 669−702.
(8) O'Brien, P.; Herschlag, D. *J. Am. Chem. Soc.* **1998**, *120*, 12369−12370.
(9) O'Brien, P.; Herschlag, D. *J. Am. Chem. Soc.* **1999**, *121*, 11022−11023.
(10) O'Brien, P.; Herschlag, D. *Biochemistry* **2001**, *40*, 5691−5699.
(11) Nikolic-Hughes, I.; Rees, D.; Herschlag, D. *J. Am. Chem. Soc.* **2004**, *126*, 11814−11819.
(12) Zalatan, J.; Herschlag, D. *J. Am. Chem. Soc.* **2006**, *128*, 1293−1303.
(13) Zalatan, J.; Fenn, T.; Brunger, A.; Herschlag, D. *Biochemistry* **2006**, *45*, 9788−9803.
(14) Zalatan, J.; Catrina, I.; Mitchell, R.; Grzyska, P.; O'Brien, P.; Herschlag, D. *J. Am. Chem. Soc.* **2007**, *129*, 9789−9798.
(15) Zalatan, J.; Fenn, A.; Herschlag, D. *J. Mol. Biol.* **2008**, *384*, 1174−1189.
(16) O'Brien, P.; Lassila, J.; Fenn, T.; Zalatan, J.; Herschlag, D. *Biochemistry* **2008**, *47*, 7663−7672.
(17) Fersht, A. *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*; W.H. Freeman and Company: New York, 1999.
(18) Jencks, W. *Catalysis in chemistry and enzymology*; Dover Publications: New York, 1987.
(19) López-Canut, V.; Marti, S.; Bertrán, J.; Moliner, V.; Tuñón, I. *J. Phys. Chem. B* **2009**, *113*, 7816−7824.
(20) López-Canut, V.; Roca, M.; Bertrán, J.; Moliner, V.; Tuñón, I. *J. Am. Chem. Soc.* **2010**, *132*, 6955−6963.

(21) López-Canut, V.; Roca, M.; Bertrán, J.; Moliner, V.; Tuñón, I. *J. Am. Chem. Soc.* **2011**, *133*, 12050−12062.
(22) Luo, J. H.; van Loo, B.; Kamerlin, S. C. L. *Proteins: Struct., Funct., Bioinf.* **2012**, *80*, 1211−1226.
(23) Luo, J. H.; van Loo, B.; Kamerlin, S. C. L. *FEBS Lett.* **2012**, *586*, 1622−1630.
(24) Åqvist, J.; Kolmodin, K.; Florián, J.; Warshel, A. *Chem. Biol.* **1999**, *6*, R71−R80.
(25) Nam, K.; Cui, Q.; Gao, J.; York, D. *J. Chem. Theory Comput.* **2007**, *3*, 486−504.
(26) Wong, K. Y.; Gao, J. L. *Biochemistry* **2007**, *46*, 13352−13369.
(27) Garcia-Viloca, M.; Alhambra, C.; Truhlar, D. G.; Gao, J. L. *J. Am. Chem. Soc.* **2002**, *124*, 7268−7269.
(28) Rosta, E.; Nowotny, M.; Yang, W.; Hummer, G. *J. Am. Chem. Soc.* **2011**, *133*, 8934−8941.
(29) Norotny, M.; Gaidamakov, S. A.; Crouch, R. J.; Yang, W. *Cell* **2005**, *12*, 1005−1016.
(30) Nowotny, M.; Yang, W. *EMBO J.* **2006**, *25*, 1924−1933.
(31) Yang, Y.; Yu, H.; York, D.; Elstner, M.; Cui, Q. *J. Chem. Theory Comput.* **2008**, *4*, 2067−2084.
(32) Elstner, M.; Cui, Q.; Munih, P.; Kaxiras, E.; Frauenheim, T.; Karplus, M. *J. Comput. Chem.* **2003**, *24*, 565−581.
(33) Chakravorty, D. K.; Wang, B.; Lee, C. W.; Giedroc, D. P.; Merz, K. M., Jr. *J. Am. Chem. Soc.* **2012**, *134*, 3367−3376.
(34) Riccardi, D.; Cui, Q. *J. Phys. Chem. A* **2007**, *111*, 5703−5711.
(35) Xu, D.; Guo, H.; Cui, Q. *J. Am. Chem. Soc.* **2007**, *129*, 10814−10822.
(36) Hou, G. QM/MM studies of phosphoryl transfer reactions in the alkaline phosphatase superfamily. Ph. D. Thesis, University of Wisconsin-Madison: Madison, WI, 2012.
(37) Riccardi, D.; Zhu, X.; Goyal, P.; Yang, S.; Hou, G.; Cui, Q. *Sci. China: Chem.* **2012**, *55*, 3−18.
(38) Hou, G.; Zhu, X.; Elstner, M.; Cui. *J. Chem. Theory Comput.* **2012**, *8*, 4293−4304.
(39) Brunger, A.; Karplus, M. *Protein: Struct., Funct., Genet.* **1988**, *4*, 148−156.
(40) Brooks, B.; Bruccoleri, R.; Olafson, B.; States, D.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(41) Brooks, C. L., III.; Karplus, M. *J. Chem. Phys.* **1983**, *79*, 6312−6325.
(42) MacKerell, A.; et al. *J. Chem. Phys.* **1998**, *102*, 3586−3616.
(43) Jorgensen, W.; Chandrasekhar, J.; Madura, J.; Impey, R.; Klein, M. *J. Chem. Phys.* **1983**, *79*, 926−935.
(44) Konig, P.; Hoffmann, M.; Frauenheim, T.; Cui, Q. *J. Phys. Chem. B* **2005**, *109*, 9082−9095.
(45) Hou, G. H.; Zhu, X.; Cui, Q. *J. Chem. Theory Comput.* **2010**, *6*, 2303−2314.
(46) Yang, Y.; Yu, H.; Cui, Q. *J. Mol. Biol.* **2008**, *381*, 1407−1420.
(47) Yang, Y.; Cui, Q. *J. Phys. Chem. B* **2009**, *113*, 4930−4933.
(48) Yang, Y.; Cui, Q. *J. Phys. Chem. A* **2009**, *113*, 12439−12446.
(49) Yang, Y.; Miao, Y. P.; Wang, B.; Cui, G. L.; Merz, K. M., Jr. *Biochemistry* **2012**, *51*, 2606−2618.
(50) Arantes, G.; Loos, M. *Phys. Chem. Chem. Phys.* **2006**, *8*, 347−353.
(51) Im, W.; Bernèche, S.; Roux, B. *J. Chem. Phys.* **2001**, *114*, 2924−2937.
(52) Schaefer, P.; Riccardi, D.; Cui, Q. *J. Chem. Phys.* **2005**, *123*, 014905.
(53) Riccardi, D.; Schaefer, P.; Yang, Y.; Yu, H.; Ghosh, H.; Prat-Resina, X.; König, P.; Li, G.; Xu, D.; Guo, H.; Elstner, M.; Cui, Q. *J. Phys. Chem. B* **2006**, *110*, 6458−6469.
(54) Brooks, C.; Karplus, M. *J. Mol. Biol.* **1989**, *208*, 159−181.
(55) Rychaert, J.; Ciccotti, G.; Berendsen, H. *J. Comput. Phys.* **1977**, *23*, 327−341.
(56) Im, W.; Beglov, D.; Roux, B. *Comput. Phys. Commun.* **1998**, *111*, 59−75.
(57) Lu, X.; Cui, Q. *J. Phys. Chem. B* **2013**, *117*, 2005−2018.
(58) Nina, M.; Beglov, D.; Roux, D. *J. Phys. Chem. B* **1997**, *101*, 5239−5248.
(59) Nina, M.; Im, W.; Roux, D. *Biophys. Chem.* **1999**, *78*, 89−96.
(60) Steinbach, P.; Brooks, B. *J. Comput. Chem.* **1994**, *15*, 667−683.
(61) Cui, Q.; Elstner, M.; Kaxiras, E.; Frauenheim, T.; Karplus, M. *J. Phys. Chem. B* **2001**, *105*, 569−585.
(62) Klopman, G. *J. Am. Chem. Soc.* **1964**, *86*, 4550−4557.
(63) Ohno, K. *Theor. Chim. Acta* **1964**, *2*, 219−227.
(64) Thiel, W. *Adv. Chem. Phys.* **1996**, *93*, 703−757.
(65) Elstner, M. *J. Phys. Chem. A* **2007**, *111*, 5614−5621.
(66) Yang, Y.; Yu, H.; York, D.; Cui, Q.; Elstner, M. *J. Phys. Chem. A* **2007**, *111*, 10861−10873.
(67) Das, D.; Eurenius, K. P.; Billings, E. M.; Sherwood, P.; Chatfield, D. C.; Hodoscek, M.; Brooks, B. R. *J. Chem. Phys.* **2002**, *117*, 10534−10547.
(68) Wang, B.; Truhlar, D. G. *J. Chem. Theory Comput.* **2010**, *6*, 3330−3342.
(69) Stone, A. J. *The theory of intermolecular forces*; Oxford University Press: Oxford, U.K., 1997.
(70) Torrie, G. M.; Valleau, J. P. *J. Chem. Phys.* **1977**, *23*, 187−199.
(71) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011−1021.
(72) Becke, A. *Phys. Rev. A* **1988**, *38*, 3098−3100.
(73) Becke, A. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(74) Lee, C.; Yang, W.; Parr, R. *Phys. Rev. B* **1988**, *37*, 785−789.
(75) Truhlar, D. G.; Zhao, Y. *Theor. Chem. Acc.* **2008**, *120*, 215−241.
(76) Shao, Y.; et al. *Phys. Chem. Chem. Phys.* **2006**, *27*, 3172−3191.
(77) Brooks, B. R.; et al. *J. Comput. Chem.* **2009**, *30*, 1545−1614.
(78) Petersson, G.; Bennett, A.; Tensfeldt, T.; Allaham, M.; Shirley, W.; Mantzaris, J. *J. Chem. Phys.* **1988**, *89*, 2193−2218.
(79) Wu, R.; Lu, Z.; Cao, Z.; Zhang, Y. *J. Am. Chem. Soc.* **2011**, *133*, 6110−6113.
(80) Wu, R.; Wang, S.; Zhou, N.; Cao, Z.; Zhang, Y. *J. Am. Chem. Soc.* **2010**, *132*, 9471−9479.
(81) Xu, D. G.; Guo, H. *J. Am. Chem. Soc.* **2009**, *131*, 9780−9788.
(82) Rao, L.; Cui, Q.; Xu, X. *J. Am. Chem. Soc.* **2010**, *132*, 18092−18102.
(83) Marti, S.; Moliner, V.; Tuñón, I. *J. Chem. Theory Comput.* **2005**, *1*, 1008−1016.
(84) Claeyssens, F.; Harvey, J. N.; Manby, F. R.; Mata, R. A.; Mulholland, A. J.; Ranaghan, K. E.; Schutz, M.; Thiel, S.; Thiel, W.; Werner, H. J. *Angew. Chem., Int. Ed.* **2006**, *45*, 6856−6859.
(85) Trajbl, M.; Hong, G. Y.; Warshel, A. *J. Phys. Chem. B* **2002**, *106*, 13333−13343.
(86) Plotnikov, N. V.; Kamerlin, S. C. L.; Warshel, A. *J. Phys. Chem. B* **2011**, *115*, 7950−7962.
(87) Rod, T. H.; Ryde, U. *J. Chem. Theory Comput.* **2005**, *1*, 1240−1251.
(88) Heimdal, J.; Kaukonen, M.; Srnec, M.; Rulisek, L.; Ryde, U. *ChemPhysChem.* **2011**, *12*, 3337−3347.
(89) Retegan, M.; Martins-Costa, M.; Ruiz-Lopez, M. F. *J. Chem. Phys.* **2010**, *133*, 064103.
(90) Cook, P. F.; Cleland, W. W. *Enzyme kinetics and mechanism*; Garland Science: New York, 2007.
(91) Cui, Q.; Karplus, M. *J. Am. Chem. Soc.* **2001**, *122*, 2284−2290.
(92) Zienau, J.; Cui, Q. *J. Phys. Chem. B* **2012**, *116*, 12522−12534.
(93) Andrews, L. D.; Deng, H.; Herschlag, D. *J. Am. Chem. Soc.* **2011**, *133*, 11621−11631.
(94) Rosta, E.; Kamerlin, S. C. L.; Warshel, A. *Biochemistry* **2008**, *47*, 3725−3735.
(95) Bobyr, E.; Lassila, J. K.; Wiersma-Koch, H. I.; Fenn, T. D.; Lee, J. J.; Nikolic-Hughes, I.; Hodgson, K. O.; Rees, D. C.; Hedman, B.; Herschlag, D. *J. Mol. Biol.* **2011**, *415*, 102−117.
(96) Gaus, M.; Cui, Q.; Elstner, M. *J. Chem. Theory Comput.* **2011**, *7*, 931−948.
(97) Nielsen, S. O.; Bulo, R. E.; Moore, P. B.; Ensing, B. *Phys. Chem. Chem. Phys.* **2010**, *12*, 12401−12414.
(98) Park, K.; Götz, A. W.; Walker, R. C.; Paesani, F. *J. Chem. Theory Comput.*. **2012**, *8*, 2868−2877.
(99) Jensen, R. A. *Annu. Rev. Microbiol.* **1976**, *30*, 409−425.
(100) O'Brien, P. J.; Herschlag, D. *Chem. Biol.* **1999**, *6*, R91−R105.
(101) Kazlauskas, R. J. *Curr. Opin. Chem. Biol.* **2005**, *9*, 195−201.

(102) Hult, K.; Berglund, P. *Trends Biotechnol.* **2007**, *25*, 231−238.
(103) Nobeli, I.; Favia, A. D.; Thornton, J. M. *Nat. Biotechnol.* **2009**, *27*, 157−167.
(104) Zalatan, J. G.; Herschlag, D. *Nat. Chem. Biol.* **2009**, *5*, 516−520.
(105) Khersonsky, O.; Tawfik, D. S. *Annu. Rev. Biochem.* **2010**, *79*, 471−505.
(106) Mohamed, M. F.; Hollfelder, F. *Biochim. Biophys. Acta, Proteins Proteomics* **2013**, *1834*, 417−424.
(107) Ben-David, M.; Wieczorek, G.; Elias, M.; Silman, I.; Sussman, J. L.; Tawfik, D. S. *J. Mol. Biol.* **2013**, *425*, 1028−1038.
(108) 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.
(109) Thornton, J. W. *Nat. Rev. Genet.* **2004**, *5*, 366−375.
(110) Harms, M. J.; Thornton, J. W. *Curr. Opin. Struct. Biol.* **2010**, *20*, 360−366.
(111) Huang, R. Q.; Hippauf, F.; Rohrbeck, D.; Haustein, M.; Wenke, K.; Feike, J.; Sorrelle, N.; Piechulla, B.; Barkman, T. J. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 2966−2971.