J|A|C|S
ARTICLES

Published on Web 04/29/2010

Theoretical Study of Phosphodiester Hydrolysis in Nucleotide Pyrophosphatase/Phosphodiesterase. Environmental Effects on the Reaction Mechanism

Violeta López-Canut,[‡] Maite Roca,[‡] Juan Bertrán,[†] Vicent Moliner,*[,§,∫] and Iñaki Tuñón*[,‡]

Departament de Química Física, Universitat de València, 46100 Burjassot, Spain, Departament de Química Física i Analítica, Universitat Jaume I, 12071 Castellón, Spain, Institute of Applied Radiation Chemistry, Technical University of Lodz, 90-924 Lodz, Poland, and Departament de Química; Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain

Received October 2, 2009; E-mail: ignacio.tunon@uv.es; moliner@uji.es

Abstract: We here present a theoretical study of the alkaline hydrolysis of methyl *p*-nitrophenyl phosphate ($MpNPP^-$) in aqueous solution and in the active site of nucleotide pyrophosphatase/phosphodiesterase (NPP). The analysis of our simulations, carried out by means of hybrid quantum mechanics/molecular mechanics (QM/MM) methods, shows that the reaction takes place through different reaction mechanisms depending on the environment. Thus, while in aqueous solution the reaction occurs by means of an A_ND_N mechanism, the enzymatic process takes place through a D_NA_N mechanism. In the first case, we found associative transition-state (TS) structures, while in the enzyme TS structures have dissociative character. The reason for this change is rationalized in terms of the very different nature of the electrostatic interactions established in each of the environments: while the aqueous solution reduces the repulsion between the negatively charged reacting fragments, assisting their approach, the NPP active site stabilizes the charge distribution of dissociative TS structures, allowing the reaction to proceed with a significantly reduced free energy cost. Interestingly, the NPP active site is able to accommodate different substrates, and it seems that the nature of the TSs depends on their electronic characteristics. So, in the case of the $MpNPP^-$ substrate, the nitro group establishes hydrogen-bond interactions with water molecules and residues found in the outer part of the catalytic site, while the leaving group oxygen atom does not coordinate directly with any of the zinc atoms of the active site. If methyl phenyl phosphate is used as substrate, then the charge on the leaving group is supported to larger extent by the oxygen atom and the phenolate anion can be then coordinated to one of the two zinc atoms present in the active site.

1. Introduction

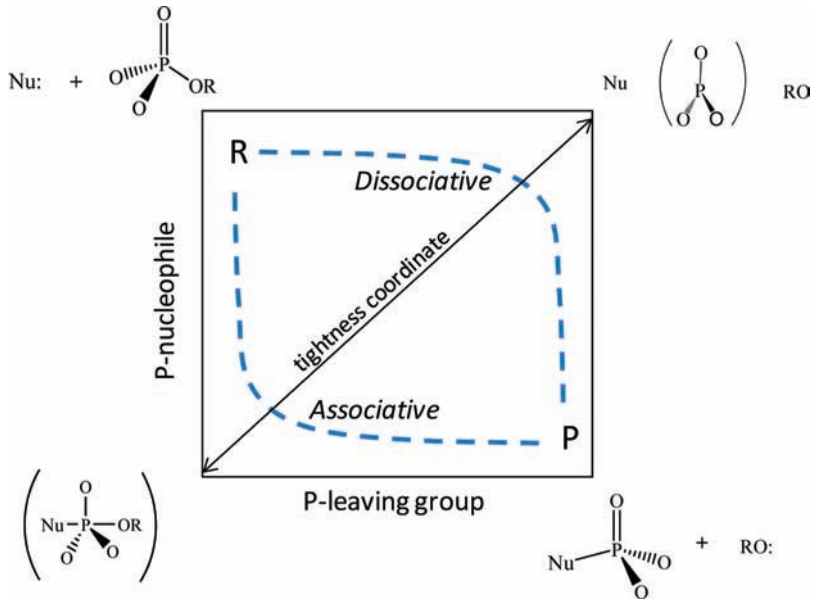
Phosphoester hydrolysis is a fundamental process in biological systems. Many biological molecules contain phosphate and reactions in which phosphorus−oxygen bonds are broken are important for energy flow, signal transduction, and genetic inheritance.[1−8] The kinetic stability of the phosphorus−oxygen bond in aqueous solution imposes the use of enzymes to reach chemical rates compatible with life.[9] In fact, enzymes involved in the catalysis of these reactions, such as kinases, ATPases, and phosphatases, have some of the largest catalytic activities known.[10]

In spite of their fundamental importance, the reaction mechanism of phosphoesters hydrolysis is still a controversial question. In principle, phosphoesters hydrolysis can proceed through different reaction mechanisms.[11] Two different limits can be traced on a More−O'Ferrall−Jencks diagram describing two stepwise mechanisms (see Scheme 1): (i) a dissociative reaction mechanism in which cleavage of the phosphorus leaving group bond precedes the formation of the new phosphorus−nucleophile bond ($D_N + A_N$ mechanism) or (ii) an associative mechanism in which there is no significant bond cleavage to the leaving group while the new bond to the nucleophile is established ($A_N + D_N$ mechanism). Concerted processes can also be traced on the diagram as a diagonal from reactants to products, and depending on the synchronicity of the breaking and forming bonds (i.e., deviations from the diagonal), the mechanism would be classified as dissociative or associative.

[‡] Universitat de València.
[†] Universitat Autònoma de Barcelona.
[§] Universitat Jaume I.
[∫] Technical University of Lodz.

(1) Vetter, I. R.; Wittinghofer, A. *Q. Rev. Biophys.* **1999**, *32*, 1–56.
(2) Cleland, W. W.; Hengge, A. C. *Chem. Rev.* **2006**, *106*, 3252–3278.
(3) Westheimer, F. H. *Science* **1987**, *235*, 1173–1178.
(4) Benkovic, S. J., Schray, K. J. In *The Enzymes*, Vol. VIII; Boyer, P. D. Ed; Academic Press: New York, 1973, pp. 201−238.
(5) Kirby, J. A.; Warren, S. G. *The Organic Chemistry of Phosphorous*; Elsevier: Amsterdam, 1967.
(6) Westheimer, F. H. *Chem. Rev.* **1981**, *81*, 313–326.
(7) Mildvan, A. S. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1979**, *49*, 103–126.
(8) Cox, J. R., Jr.; Ramsay, O. B. *Chem. Rev.* **1964**, *64*, 317–352.
(9) Lad, C.; Williams, N. H.; Wolfenden, R. *Proc. Nat. Acad. Sci. U.S.A.* **2003**, *100*, 5607–5610.
(10) Zalatan, J. G.; Herschlag, D. *J. Am. Chem. Soc.* **2006**, *128*, 1293–1303.
(11) Zalatan, J. G.; Fenn, T. D.; Brunger, A. T.; Herschlag, D. *Biochemistry* **2006**, *45*, 9788–9803.

Scheme 1



The transition state (TS) of these two types of concerted paths essentially differs in the sum of the distances from the phosphorus atom to the leaving group and to the nucleophile oxygen atoms (also called tightness coordinate):[10] the value of this coordinate is smaller in the associative TS than in the dissociative one. Gradual changes of the reaction mechanism for different phosphoesters can then be described using this coordinate, which will depend on several factors, including the nature of the leaving group and the effect of the environment.

Alkaline phosphatase (AP) superfamily is an evolutionarily related group of enzymes that catalyzes phosphoryl and sulfuryl transfer reactions. Members of this group include phosphomonoesterases, phosphodiesterases, phosphoglycerate mutases, phosphopentomutases, and sulfatases.[12−16] AP, the best characterized member of the AP superfamily, catalyzes phosphate monoesters hydrolysis, but it also shows secondary activity catalyzing phosphodiester, phosphite, and sulfate monoester hydrolysis.[10,16,17] The *Escherichia coli* AP active site contains two zinc ions and an arginine residue that interact with the substrate and a deprotonated serine acting as a nucleophile (see Figure 1A). A third metallic ion, a magnesium ion, could play a fundamental role in the discrimination between phosphate monoesters and diesters.[18] A highly related member of this superfamily is the nucleotide pyrophosphatase/phosphodiesterase (NPP). This enzyme hydrolyzes extracellular phosphodiesters affecting several biological processes,[19,20] but it also has secondary activity as phosphomonoesterase[11] and sulfatase.[21] The structural relationship between NPP and AP is evident when describing the active site of the former: it contains two zinc ions with the same coordination shell as AP and the nucleophile is now a deprotonated threonine (see Figure 1B). The main differences are the substitution of an arginine (Arg166 in AP) by an asparagine (Asn111 in NPP) and the absence of a lysine (Lys328 in AP) and of the magnesium ion. NPP also has a pocket able to accommodate the nonleaving hydrophobic group present in phosphodiesters (R′ in Figure 1B).[11] Structural similarities and promiscuous activities can be vestiges of evolutionary events with a common ancestor displaying broader catalytic specificity.[10]

The similarity of the active sites has raised the question of the relationship between structure and activity in this family of enzymes.[10] It seems to be well-established, both from experimental[12,14] and theoretical studies,[22] that AP catalyzes phosphomonoester hydrolysis through a dissociative mechanism. This suggests that NPP should also catalyze phosphodiester hydrolysis through a dissociative TS that could be complementary to the electrostatic features of the highly conserved active site. However, Zalatan and Herschlag,[10] from analysis of kinetic isotope effects (KIEs) and linear free energy relationships (LFERs), determined that the nature of the TS of AP-catalyzed phosphodiester hydrolysis is more probably associative. These authors then suggest that the AP enzyme is able to recognize and stabilize different TSs. The active site would then be flexible enough to be adapted to the characteristics of the chemical system, avoiding the energy penalty associated to a change in the nature of the TS.[10] The opposite way has been proposed to explain the promiscuous activity of protein phosphatase-1 (PP-1) with phosphonates. While the hydrolysis of these compounds proceeds through an associative mechanism in solution, the actual mechanism in the

(12) Zalatan, J. G.; Catrina, I.; Mitchell, R.; Grzyska, P. K.; O'Brien, P. J.; Herschlag, D.; Hengge, A. C. *J. Am. Chem. Soc.* **2007**, *129*, 9789–9798.
(13) Catrina, I.; O'Brien, P. J.; Purcell, J.; Nikolic-Hughes, I.; Zalatan, J. G.; Hengge, A. C.; Herschlag, D. *J. Am. Chem. Soc.* **2007**, *129*, 5760–5765.
(14) Nikolic-Hughes, I.; Rees, D. C.; Herschlag, D. *J. Am. Chem. Soc.* **2004**, *126*, 11814–11819.
(15) Yang, K. C.; Metcalf, W. W. *Proc. Nat. Acad. Sci. U.S.A.* **2004**, *101*, 7919–7924.
(16) O'Brien, P. J.; Herschlag, D. *Biochemistry* **2001**, *40*, 5691–5699.
(17) Coleman, J. E. *Annu. Rev. Biophys. Biomolec. Struct.* **1992**, *21*, 441–483.
(18) Zalatan, J. G.; Fenn, T. D.; Herschlag, D. *J. Mol. Biol.* **2008**, *384*, 1174–1189.
(19) Stefan, C.; Jansen, S.; Bollen, M. *Trends Biochem. Sci.* **2005**, *30*, 542–550.
(20) Goding, J. W.; Grobben, B.; Slegers, H. *Biochim. Biophys. Acta-Mol. Basis Dis.* **2003**, *1638*, 1–19.
(21) Lassila, J. K.; Herschlag, D. *Biochemistry* **2008**, *47*, 12853–12859.
(22) López-Canut, V.; Martí, S.; Bertrán, J.; Moliner, V.; Tuñón, I. *J. Phys. Chem. B* **2009**, *113*, 7816–7824.

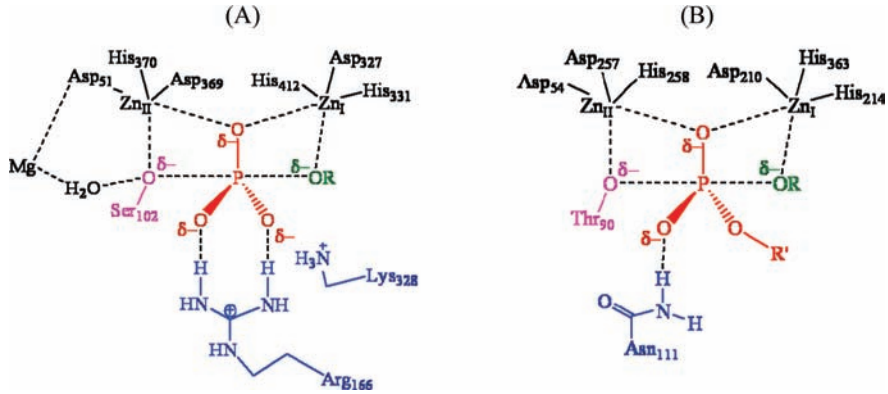


Figure 1. Schematic representation of the putative TS of the phosphomonoester and phosphodiester hydrolysis in the AP (A) and in NPP (B) active sites, respectively.

active site of PP-1 seems to be highly dissociative.[23] Thus, in this proposal, the chemical system is adapted to the protein environment. Both alternative scenarios imply different evolutionary strategies to develop new functions in existing enzymes: the enzymes of the superfamily evolved either by adapting the catalytic site to different reaction mechanisms or by preserving the same mechanism. Finally, it should be taken into consideration that interpretation of experimental KIE and LFER data can be ambiguous due to the inherent difficulties associated with the flat nature of the free energy surface of phosphoesters hydrolysis.[24] In fact, it has been shown that similar predictions, compatible with experimental data, can be derived from different mechanistic paths, indicating that theoretical simulations of the reaction process are needed to determine the reaction mechanism in a conclusive way.[24]

Regarding the reaction mechanisms of phosphodiester hydrolysis in alkaline solution, which should be an adequate counterpart process for phosphodiester hydrolysis in the AP superfamily, experimental studies based on KIEs suggest that this process takes place through a concerted associative A_ND_N mechanism.[25,26] Theoretical studies demonstrated that the reaction proceeds through a concerted A_ND_N mechanism which becomes progressively more dissociative as the pK_a value of the leaving group is reduced, both for mono- and diesters.[24,27]

We present herein the first comparative theoretical study of the reaction mechanism of phosphodiester hydrolysis in aqueous solution and in the NPP active site. Hybrid quantum mechanics/molecular mechanics (QM/MM) molecular dynamics (MD) simulations are carried out to obtain the reaction free energy profiles and the free energy changes associated to variations along the tightness coordinate in the TSs in both environments. Our results suggest a mechanistic change, from associative to dissociative, when going from the solution to the enzyme. Theoretical elucidation of the reaction mechanism employed by NPP can be crucial to understand the evolutionary paths followed to develop and improve new activities from a previously existing function in the AP superfamily. Our results are compatible with an evolutionary scenario where the reaction mechanism remains unaltered in the AP family.

2. Methodology

2.1. Building the Systems. The initial coordinates were taken from the X-ray crystal structure of *Xanthomonas axonopodis pv citri* nucleotide pyrophosphatase/phosphodiesterase in complex with vanadate ion (VO_4^{3-}), with PDB entry 2GSO.[11] NPP is a homodimer with two active sites; each one was found to contain two Zn^{2+} atoms. The inorganic vanadate was used as reference to replace it by the methyl *p*-nitrophenyl phosphate diester ($MpNPP^-$), which is the substrate used in this study. Hydrogen atoms on titrable residues were added using the cluster method as implemented by Field and co-workers[28,29] assuming a pH = 8, which is the experimental condition used to obtain k_{cat}/K_M.[11] Thr90 residue was considered to be unprotonated (as Ser102 residue in AP).[11,22,30] We solvated the system with a water molecules sphere of 40 Å of radius centered on the phosphorus atom of the substrate. Water molecules placed at less than 2.8 Å than any other non-hydrogen atom of the system were deleted.

As found in a previous theoretical study in AP, a large QM subsystem is needed to reproduce the charge-transfer effects between the reactive system and the neighboring residues.[22] The QM region contains the Thr90 residue, the Asn111 residue, the substrate ($MpNPP^-$), and the two zinc ions and their coordination spheres (Asp54, Asp257, His258, Asp210, His214, His363) (see Figure 1B). Because of the small size of the threonine side chain, we decided to include part of the previous and next residues of the protein sequence (Leu89 and Phe91) in the QM region. To saturate the valence of the QM/MM frontier we used the link atoms procedure,[31,32] placing these link atoms between the C_α and C_β atoms of each residue. The number of QM atoms then became 112, while the final system contains a total of 31885 atoms.

Because we define a large QM subsystem and the large number of structures that must be sampled during simulations, we employed a semiempirical Hamiltonian to reduce the computational cost. The effect of using standard semiempirical quantum methods, such as AM1,[33] PM3,[34] or MNDO,[35] could be dramatic for the system

(23) 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.
(24) Klahn, M.; Rosta, E.; Warshel, A. *J. Am. Chem. Soc.* **2006**, *128*, 15310–15323.
(25) Hengge, A. C.; Tobin, A. E.; Cleland, W. W. *J. Am. Chem. Soc.* **1995**, *117*, 5919–5926.
(26) Cassano, A. G.; Anderson, V. E.; Harris, M. E. *J. Am. Chem. Soc.* **2002**, *124*, 10964–10965.
(27) Rosta, E.; Kamerlin, S. C. L.; Warshel, A. *Biochemistry* **2008**, *47*, 3725–3735.
(28) Antosiewicz, J.; McCammon, J. A.; Gilson, M. K. *J. Mol. Biol.* **1994**, *238*, 415–436.
(29) Field, M. J.; Amara, P.; David, L.; Rinaldo, D. Personal communication.
(30) Gijsbers, R.; Ceulemans, H.; Stalmans, W.; Bollen, M. *J. Biol. Chem.* **2001**, *276*, 1361–1368.
(31) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1986**, *7*, 718–730.
(32) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700–733.
(33) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909.
(34) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209–220.
(35) Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4899–4907.

under study because they are not sufficiently accurate to model reactions involving phosphorus atoms. This is essentially due to the fact that these methods do not incorporate "d" orbitals. For this reason, we use a new semiempirical Hamiltonian, AM1/d-PhoT (hereafter simply named as AM1d).[36] This Hamiltonian incorporates d-extension for the phosphorus atom and modified AM1 parameters for oxygen and hydrogen atoms, while the remaining atoms are described at the AM1 level.[36] This semiempirical treatment has been shown to provide results in very good agreement with high-level DFT calculations in the study of the hydrolysis of phosphorus compounds.[22,36−38] In particular, we employed this Hamiltonian in a previous QM/MM study of the alkaline hydrolysis of $MpNPP^-$ in aqueous solution, and the reported kinetic isotopic effects (KIEs) were in excellent agreement with experimental results.[37] Thus, the AM1d semiempirical Hamiltonian was selected to describe the QM region, and the MM region was described by means of the OPLS-AA potential[39,40] implemented in DYNAMO library.[41,42]

The quantum level used to describe the Zn ions deserves some comments. These ions were described with the standard AM1 Hamiltonian because specific parameters for the new AM1d Hamiltonian have still not been developed. In recent work by Gao and co-workers, the AM1 method was employed to treat the two zinc ions located in the active site of a phosphotriesterase (PTE), providing a reasonable picture of the reaction process.[43] The same combination of theoretical treatments (AM1d, AM1, and OPLS) was successfully tested in the study of the catalytic activity of *E. coli* AP.[22] In that work, we showed that the employed quantum treatment (AM1d−AM1) provided geometries in good agreement with DFT results for a complex presenting a Zn coordination similar to that present in AP or NPP.

A similar protocol was used to study the same reaction in solution, placing the QM subsystem, the substrate ($MpNPP^-$), and a hydroxyl ion (OH^-) in a 55.8 Å side length box of TIP3P[44] water molecules. These 24 QM atoms were described by means of an AM1d method. Details of the simulations in solution are given elsewhere.[37] The van der Waals parameters used for the QM atoms, in solution and in the enzyme, were taken from the OPLS potential.[39,40]

2.2. Simulations. The enzymatic system was optimized during 6000 steps using a combination of steepest descent and conjugate gradients methods. After exploration of the potential energy surface, a TS structure was located. We then constrained the distances between the phosphorus atom (P) and the two key oxygen atoms, the oxygen atom of the leaving group (O_{lg}), and the one of the nucleophile (O_{nuc}), applying a harmonic potential with a force constant of 3000 $kJ\cdot mol^{-1}\cdot Å^{-2}$. Because of the size of the system, all residues further than 22 Å from the phosphorus atom were kept frozen (23563 atoms from a total of 31885 atoms) The system was heated by means of 80 ps of molecular dynamics (MD) using a time step of 1 fs. Finally, the system was equilibrated at 300 K using the NVT ensemble and the Langevin−Verlet integrator with a time step of 1 fs during a total simulation time of 200 ps. A switched cutoff, from 14 to 16 Å, was employed for all MM interactions (including the frozen atoms), while the QM region was allowed to interact with every flexible MM atom. All the QM/MM calculations were carried out using the DYNAMO program.[41]

The starting structure in aqueous solution was taken from a TS structure optimized by means of a polarizable continuum model (PCM) approach at B3LYP level and the 6-31+G* basis set.[37] The entire system was equilibrated by means of MD simulations applying a constraint on the O···P···O angle (attacking angle), to be close to linear, and on the involved distances of the phosphoester hydrolysis to remain in the TS region. We defined as flexible region all those atoms belonging to residues or molecules found at a distance less than or equal to 22 Å of the phosphorus atom, and the rest of the water molecules (10386 atoms) were kept frozen. The QM subsystem was allowed to interact with all the MM atoms in the flexible region, while a cutoff switched from 11 to 13 Å was employed for all MM interactions. A time step of 1 fs was also used.

To analyze the energetics of the chemical reaction in the enzyme and in aqueous solution, we obtained the corresponding potentials of mean force (PMFs)[45] in which the antisymmetric combination of the distances describing the breaking and forming bonds ($RC = d(P-O_{lg}) - d(P-O_{nuc})$) was employed as distinguished reaction coordinate. The umbrella sampling approach[46] was used to constrain the system along the reaction profile, employing a force constant of 2500 $kJ\cdot mol^{-1}\cdot Å^{-2}$, and the probability distributions were put together by means of the weighted histogram analysis method (WHAM)[47,48] to obtain the full probability distribution along the reaction coordinate. The values of the force constant used for the harmonic umbrella sampling were determined to allow a full overlapping of the different windows traced in the PMFs evaluation, but without losing control over the selected coordinate. Each window consisted in 10 ps of equilibration followed by 15 ps of production. The total number of windows employed to cover the whole range of the reaction coordinate in the enzyme and aqueous solution were of 208 and 153, respectively.

As discussed in the next section, the PMFs obtained in solution and in the enzyme correspond to different reaction mechanisms. In order to verify that the mechanism obtained in solution was not also possible for the enzymatic reaction we decide to explore a free energy surface (or two-dimensional PMF, 2D-PMF) obtained as a function of the two distances corresponding to the breaking and forming bonds ($d(P-O_{lg})$ and $d(P-O_{nuc})$) treated independently. To obtain this 2D-PMF we ran 723 simulation windows, each one consisting of 10 ps of equilibration and 15 ps of production. This gave a total simulation time larger than 18 ns. The force constant applied to each of the two distances was of 2000 $kJ\cdot mol^{-1}\cdot Å^{-2}$. In all the simulations carried out to obtain the PMFs the reference temperature was 300 K.

3. Results and Discussion

3.1. Free Energy Profiles. Figure 2 shows the PMFs obtained for the alkaline hydrolysis of $MpNPP^-$ in aqueous solution and in the enzyme at the AM1d/MM level as a function of the antisymmetric combination of the distances describing the breaking and forming bonds. The nucleophile in water is a hydroxide anion, while the enzyme contains a deprotonated threonine (Thr90). While both PMFs display a maximum at a similar value of the reaction coordinate (−0.60 Å in NPP and −0.43 Å in aqueous solution) the features of the reaction profiles are quite different in both environments. The aqueous solution PMF displays a sharp maximum with a well-defined bell-shaped curve. Instead, the enzymatic PMF shows a rather flat region

(36) Nam, K.; Cui, Q.; Gao, J.; York, D. M. *J. Chem. Theory Comput.* **2007**, *3*, 486–504.
(37) López-Canut, V.; Ruiz-Pernía, J.; Tuñón, I.; Ferrer, S.; Moliner, V. *J. Chem. Theory Comput.* **2009**, *5*, 439–442.
(38) Nam, K. H.; Gao, J.; York, D. M. *J. Am. Chem. Soc.* **2008**, *130*, 4680–4691.
(39) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657–1666.
(40) Pranata, J.; Wierschke, S. G.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1991**, *113*, 2810–2819.
(41) Field, M. J. *A Practical Introduction to the Simulation of Molecular Systems*, 1st ed.; Cambridge University Press: Cambridge, U.K., 1999.
(42) Field, M. J.; Albe, M.; Bret, C.; Proust-de Martin, F.; Thomas, A. *J. Comput. Chem.* **2000**, *21*, 1088–1100.
(43) Wong, K. Y.; Gao, J. *Biochemistry* **2007**, *46*, 13352–13369.
(44) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
(45) Roux, B. *Comput. Phys. Commun.* **1995**, *91*, 275–282.
(46) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187–199.
(47) Ferrenberg, A. M.; Swedsen, R. H. *Phys. Rev. Lett.* **1989**, *63*, 1195–1198.
(48) Kumar, J.; Bouzida, D.; Swedsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011–1021.

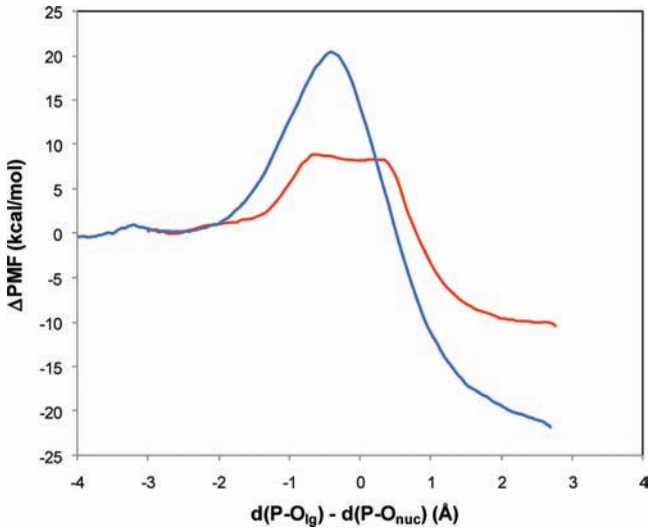


Figure 2. PMFs for the alkaline hydrolysis reaction of MpNPP$^-$ in aqueous solution (blue) and in NPP (red).

just after the maximum. The free energy barrier obtained as the difference between the maximum of the PMF and the reactants minimum is 20.5 kcal·mol^{-1} for the bimolecular reaction in water (corrected for 1 M standard state)[37] and 8.9 kcal·mol^{-1} for the enzymatic catalyzed reaction.

Figure 3 shows representative snapshots of the reactants and TS in NPP, where averaged values of significant distances are provided. Some relevant averaged distances are also given in Table 1, together with relative free energies. The nature of the TSs in solution and in the enzyme, located at the maxima of the corresponding PMFs, is quite different, reflecting the different nature of the reaction in both environments. As shown in Table 1, the averaged distances of the phosphorus atom to the leaving and nucleophilic oxygen atoms in the aqueous solution TS are 1.81 and 2.23 Å, respectively. In the enzyme, the averaged values of these distances are much larger, 2.53 and 3.13 Å, respectively. Thus, and in spite of the similar value of the reaction coordinate at the maxima of the PMFs shown in Figure 2, the two TSs differ significantly in the value of the tightness coordinate (d(P−O$_{lg}$) + d(P−O$_{nuc}$)), being 4.05 Å in solution and 5.66 Å in the enzyme (see Table 1). As seen in Figure 3, the substrate is accommodated into the active site through a lengthening of the Zn−Zn distance, up to 6.23 Å. Zn(I) is coordinated to one of the nonbridging oxygen atoms of the phosphate group, while Zn(II) stabilizes the deprotonated hydroxyl group of Thr90. The distance of this oxygen atom to the phosphorus atom (d(P−O$_{nuc}$)) is 4.59 Å in the reactant's state, while the distance to the leaving group (d(P−O$_{lg}$)) is ∼1.71 Å. In the TS, the oxygen atom of the leaving group is not coordinated to Zn(I), as Figure 1 suggests.[11,18,21] Instead, Zn(I) is coordinated to the same oxygen atom as in the reactant state, reducing the Zn−Zn distance up to 5.50 Å. As we discuss below, the coordination of the leaving group oxygen atom to Zn(I) may actually depend on the nature of the leaving group. The *p*-nitrophenolate anion is a very good leaving group because of the ability of the nitro group to delocalize the charge. In fact, we have found that, in the TS, this group makes hydrogen-bond interactions with polar residues found in the outer side of the catalytic cavity, such as Gln287. Water-mediated hydrogen bond interactions are also established with Arg354 and Arg360. Obviously, these interactions depend on the nature of the leaving (see below). In addition, a water molecule was found to interact with the O$_{lg}$ in the TS (averaged H$_w$−O$_{lg}$ distance of 1.97 Å). As observed in Figure 3, the TS is also stabilized by means of the hydrogen-bond contact established between a nonbridging oxygen atom of the phosphate group and the backbone NH group of Thr90. Finally, it is worth mentioning that we did not observe any specific interaction between Asn111 and the substrate in the reactant state or in the TS.

The predicted activation free energy in aqueous solution compares quite reasonably with the value derived from the experimental rate constant measured in solution at 42 °C, 25.9 kcal/mol.[10] The agreement can be quantitatively improved when systematic errors of the semiempirical treatment are corrected by means of higher level calculations. In particular, we obtained corrected free energy barriers for the enzymatic and in solution processes by applying the following equation:

$$\Delta G^{\ddagger}_{corr} = \Delta G^{\ddagger}_{AM1d/MM} + (\langle E_{TS} - E_R \rangle_{HL} - \langle E_{TS} - E_R \rangle_{AM1d}) \quad (1)$$

The averaged energy differences appearing in the above expression have been obtained by means of single-point calculations of a set of 10 different reactants and TS structures localized in aqueous solution and the enzyme using a micro-macroiteration optimization scheme[49−51] at the AM1d/MM level. In our previous analysis in solution, the use of 10 different structures was enough to obtain converged values of KIEs.[37] These structures are optimized using a Hessian-based search for the QM atoms while the MM atoms are fully optimized at each step of this search using only gradients.[49−51] In eq 1, HL stands for the high level chosen to obtain the corrected estimation of the activation free energy. In our case B3LYP/6-311+G**[52,53] has been used, a functional that has been shown to provide good results in previous studies of phosphorus hydrolysis.[54,55] HL calculations are carried out under the effect of the field created by the MM environment, where the electrostatic coupling between the QM and MM subsystems is calculated using point charges on the QM atoms.[49−51] As expected, the corrected free energy obtained in aqueous solution (28.2 kcal·mol^{-1}) compares better with the experimental value, although it should be kept in mind that the simulation and experimental temperatures are different. This agreement makes us confident in the theoretical method chosen to represent the reaction. Unfortunately, there is no experimental estimation of k_{cat} for the reaction catalyzed by NPP but only k_{cat}/K_M (2.3×10^2 M^{-1}·s^{-1} at 298 K)[11] has been measured. k_{cat}/K_M provides a measure of the free energy difference between the TS and the separated reactant species (solvated substrate and protein), while k_{cat} refers to the free energy difference between the TS and the Michaelis complex, which would be the quantity to be compared with our PMF barrier. The free energy barrier estimated from the enzymatic second-order rate constant (k_{cat}/K_M) would be 14.2 kcal·mol^{-1} (1 M standard state). Assuming a negative binding free energy, this value should be a lower limit for the free energy

(49) Moliner, V.; Turner, A. J.; Williams, I. H. *Chem. Commun.* **1997**, 1271–1272.
(50) Turner, A. J.; Moliner, V.; Williams, I. H. *PCCP Phys. Chem. Chem. Phys.* **1999**, *1*, 1323–1331.
(51) Martí, S.; Moliner, V.; Tuñón, I. *J. Chem. Theory Comput.* **2005**, *1*, 1008–1016.
(52) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
(53) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.
(54) Kamerlin, S. C. L.; Florian, J.; Warshel, A. *ChemPhysChem* **2008**, *9*, 1767–1773.
(55) Kamerlin, S. C. L.; Williams, N. H.; Warshel, A. *J. Org. Chem.* **2008**, *73*, 6960–6969.

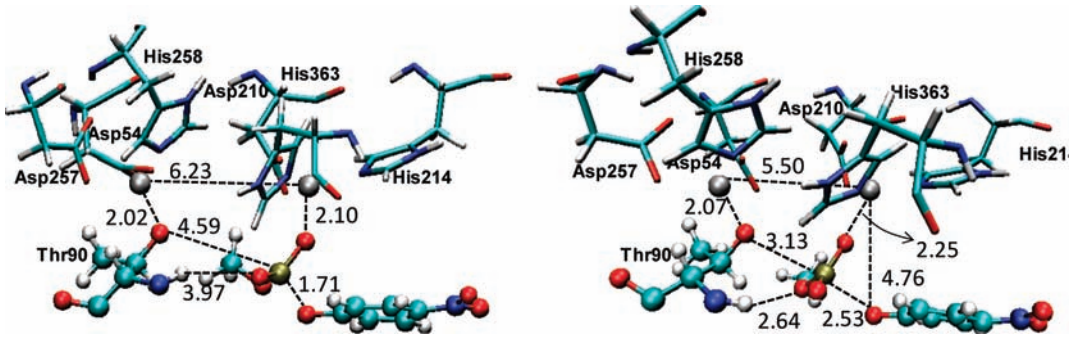


Figure 3. Snapshots for the reactants state (left) and transition state (right) of the phosphodiester hydrolysis catalyzed by NPP. Key averaged distances are given in angstroms.

Table 1. Estimated Free Energy Barriers Computed at AM1d/MM level, $\Delta G^{\ddagger}_{AM1d/MM}$, and Corrected at B3LYP/MM Level, $\Delta G^{\ddagger}_{corr}$, for the Phosphodiester Hydrolysis in Solution and in the Npp Active Site[a]

	aqueous solution		NPP	
	reactants	TS	reactants	TS
$\Delta G^{\ddagger}_{AM1d/MM}$		20.5		8.9
$\Delta G^{\ddagger}_{corr}$		28.2		11.1
$d(P-O_{lg})$	1.67 ± 0.03	1.81 ± 0.05	1.71 ± 0.03	2.53 ± 0.10
$d(P-O_{nuc})$	∞	2.23 ± 0.06	4.59 ± 0.05	3.13 ± 0.10
$d(O_{nuc}-O_{lg})$	∞	4.03 ± 0.10	6.03 ± 0.11	5.57 ± 0.19
$d(P-O_{lg}) - d(P-O_{nuc})$	$-\infty$	-0.43	-2.88	-0.60
$d(P-O_{lg}) + d(P-O_{nuc})$	∞	4.05	6.29	5.66

[a] Relevant geometrical parameters are reported in Å while energies are in kcal·mol^{-1}.

barrier measured from the Michaelis complex. Our corrected estimation of the free energy barrier, 11.1 kcal·mol^{-1}, is not far away from this limit. In any case, this value is consistent with the large catalytic ability displayed by this enzyme.

3.2. Comparative Analysis of Reaction Mechanisms. The Tightness Coordinate. According to the results derived from the PMFs shown before, it seems that the reaction takes place through much more dissociative TS structures in the active site of NPP than in solution. This difference in the reaction mechanisms can be analyzed by means of the averaged $P-O_{lg}$ and $P-O_{nuc}$ distances obtained along the free energy profile, as shown in Figure 4. These averaged distances have been drawn on the 2D-PMF obtained for the enzymatic reaction, as described in the previous section. Figure 4 can be now compared with Scheme 1, confirming the fact that the reaction mechanism in solution is associative, while it is dissociative in the enzymatic active site. In both environments, the reaction, at its first stages, can be described as an approach of the nucleophile to the phosphorus atom while the leaving group remains strongly bonded. The differences appear when the nucleophile group approaches the phosphorus atom at distances shorter than 2.7 Å. In aqueous solution the reaction continues without appreciable changes in the $P-O_{lg}$ distance up to the moment in which both distances, $P-O_{lg}$ and $P-O_{nuc}$, are about 1.9 Å. After the formation of this transient compact pentacoordinated structure, the bond between the phosphorus atom and the leaving group oxygen atoms is broken. In the enzyme, the reaction proceeds in a completely different way; the cleavage of the $P-O_{lg}$ bond begins in a much earlier stage of the reaction, in such a way that the system reaches regions where both distances ($P-O_{nuc}$ and $P-O_{lg}$) correspond to broken bonds, forming a transient PO_3CH_3 moiety. Afterward, this group is transferred to the nucleophile. During this transfer, the leaving group

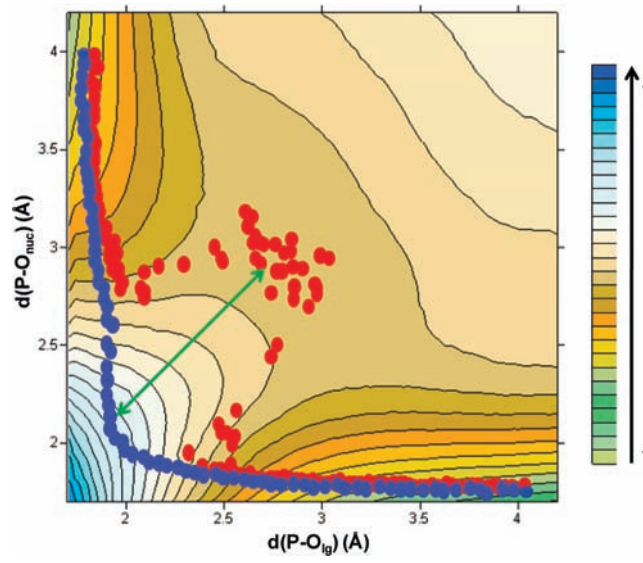


Figure 4. Averaged values of the $P-O_{nuc}$ and $P-O_{lg}$ distances found during the reaction PMFs in solution (blue) and in the NPP active site (red) projected on the 2D-PMF obtained for the enzymatic reaction. The green arrow qualitatively shows the differences in the tightness coordinate between the TS structures obtained in the enzyme and in solution.

follows, at some extent, the motion of the PO_3CH_3, as reflected by the slight reduction of the $P-O_{lg}$ distance in this stage. Once the transfer has been completed, the distance between the phosphorus atom and the leaving group increases.

The isoenergetic lines obtained from the 2D-PMF in the enzyme clearly show that the only feasible reaction mechanism in the enzyme is the dissociative one. The averaged distances corresponding to the bond-breaking and bond-forming processes in the enzyme, obtained from the one-dimensional PMF presented in Figure 2, clearly follow the minimum free energy regions of the 2D-PMF. Instead, the averaged distances obtained from the associative reaction mechanism obtained in solution cross much higher free energy regions on the enzymatic 2D-PMF, and then this mechanism is clearly disfavored.

As noted previously, both the enzymatic and solution TSs appear at similar values of the reaction coordinate but at very different values of the tightness coordinate, $d(P-O_{lg}) + d(P-O_{nuc})$, as qualitatively shown in Figure 4 with a green arrow. Thus, in order to get a deeper insight into the differences between the aqueous solution and enzymatic TSs, we performed the corresponding one-dimensional PMFs in both media, sampling this tightness coordinate from 3.7 Å (corresponding to an associative or tighten TS) to 6.0 Å (corresponding to a dissociative or loose TS). The reaction coordinate was kept

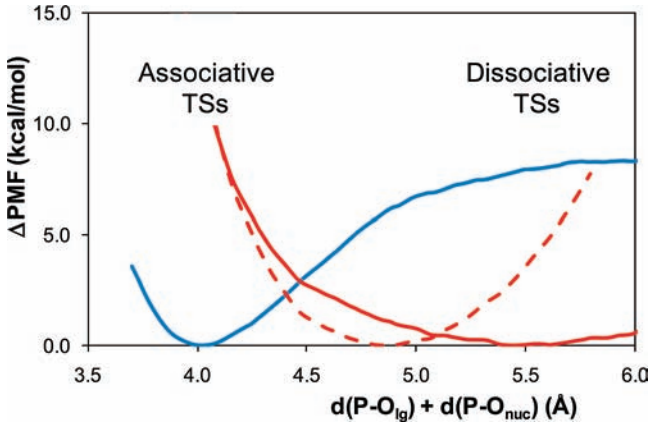


Figure 5. AM1d/MM PMFs obtained for the aqueous (blue) and enzymatic (red) TSs along the tightness coordinate. The dashed line corresponds to the enzymatic curve corrected at the B3LYP/6-311+G** level.

constrained at the value corresponding to the respective TSs (−0.60 Å in NPP and −0.43 Å in solution). The resulting PMFs are presented in Figure 5. According to these results, the associative TS is preferred in solution, while in NPP the dissociative TS structures are more stable than the associative ones, in agreement with the 2D-PMF depicted in Figure 4. The minimum of the aqueous solution PMF appears at about 4.0 Å, which corresponds to the TS characterized in Table 1. In the NPP, the minimum appears at about 5.5 Å, again in agreement with the features of the TS described above.

In general, the final balance between associative and dissociative TSs depend on two factors: the environment and the substrate (the leaving group). Moreover, in the case of theoretical studies, the final answer may also depend on the level of theory chosen to represent the system. Figure 5 also shows the enzymatic free energy curve corrected at the B3LYP/6-311+G** level. To obtain this curve, we used the correction procedure explained above, employing several structures optimized from the AM1d/MM curve obtained along the tightness coordinate. The corrected PMF shows a minimum at a value of the tightness coordinate of about 4.8 Å. This value corresponds to a dissociative TS, although slightly more tighten than the one predicted at the AM1d level. In any case, even if the final determination of the associative/dissociative character of the TS of the enzymatic hydrolysis of a phosphodiester may require further experimental and theoretical work, our analysis clearly shows that the relative effect of the enzymatic environment with respect to the aqueous solvent is to favor more dissociative TS structures.

The results presented in Figure 5 are a direct computational estimation of the energetic cost of changing the mechanism in each environment. In aqueous solution, going from reactants to products crossing through the dissociative part of the energy surface would imply an increase of the free energy barrier between 6 and 8 kcal·mol^{-1}. We got additional evidence of this extra free energy cost of the dissociative pathway in solution by using the free energy perturbation (FEP) method.[56] Following the dissociative mechanism in solution (see the Supporting Information), we obtained a free energy profile with a free energy barrier 6.8 kcal·mol^{-1} larger than for the associative one, confirming the results derived from the PMFs presented in Figure 5.

(56) Kollman, P. *Chem. Rev.* **1993**, *93*, 2395–2417.

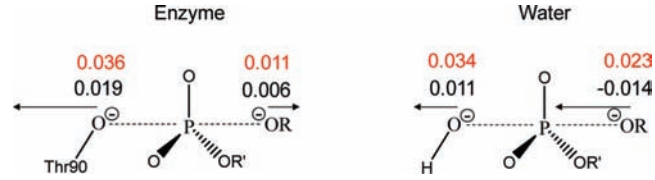


Figure 6. Electrostatic field (in au) created by the environment on the O_{nuc} and O_{lg} atoms at the corresponding TSs in enzyme and in solution. The modulus is presented in red and the projection along their respective P−O bond vectors in black. The arrows represent the electrostatic forces acting on the two oxygen atoms.

In the enzyme the opposite picture appears. A displacement toward smaller values of the tightness coordinate, i.e., more associative TS, implies a free energy cost of at least 5 kcal·mol^{-1}. Analysis of the changes found in the enzymatic structure along the tightness coordinate shows that the Zn−Zn distance should be shortened in order to accommodate more compact TS structures for this substrate. The averaged value of the Zn−Zn distance is reduced from 5.6 Å in the dissociative limit to about 4.8 Å in the associative one. However, as we will discuss below, not necessarily a more compact TS implies a smaller Zn−Zn distance. This tendency, observed for Mp-NPP$^-$, depends on the particular mode of coordination of each possible leaving group.

Thus, according to these findings, associative TSs for the alkaline hydrolysis of MNpPP$^-$ are expected in aqueous solution and not in the enzyme. The reason for this different behavior of the chemical system in water environment is due to the reaction field created by the solvent molecules on the substrate. Effectively, water molecules are able to displace and rotate according to the charge distribution of the solute. In this case, the water solution effectively diminishes the electrostatic repulsion between the two negatively charged reacting fragments (the hydroxyl anion and the substrate). In a continuum treatment of the solvent, this repulsion is reduced by a factor equal to the dielectric constant. If the electrostatic repulsion is diminished, the bond forces dominate the chemical process, favoring associative TS structures. Instead, the electric field created by the enzyme is not a reaction field: even if the active site shows some degree of flexibility and plasticity, being able to accommodate even different substrates, the charges and dipoles appearing in the active site cannot be rotated or displaced freely because their motions are restrained by the protein structure. This different behavior is reflected in the values of the electric field computed in the TS in both environments. Figure 6 shows the averaged values of the electric field calculated on the nucleophilic oxygen and the leaving group oxygen atoms, both its modulus and its projection along the respective P−O bonds. These electric fields have been computed from the contribution of all the atoms of the system, except those of the nucleophile and the substrate.[57] The environmental electrostatic forces acting on the nucleophilic and the leaving group oxygen atoms are qualitatively displayed as arrows in Figure 6. It is interesting to notice that the values of the electric field modulus, are quite similar, clearly showing that the effect of the enzyme is not simply a desolvation effect. Moreover, the effect of the electrostatic forces due to the environment on the leaving and

(57) In the enzyme, some atoms of the environment are part of the QM subsystem (zinc ions and their coordination shells and Asn111). The contribution of these QM atoms to the environmental electric field was computed through the use of point charges derived from the polarized wave function.

nucleophilic group in the TS is completely different in solution and in the enzyme. In aqueous solution, the resulting force favors an associative transition structure, with a smaller value of the tightness coordinate. According to the forces depicted in Figure 6 the solvent produces a diminution of the $d(P-O_{lg}) + d(P-O_{nuc})$ coordinate and a displacement of the TS structure toward the reactants (where the charge is localized on the smaller hydroxide anion). While the reaction field created by the aqueous solvent assists the approach of the two negatively charged reaction species, resulting in a more associative TS, in the enzyme these forces clearly act favoring dissociative TS structures, with larger values of the $d(P-O_{lg}) + d(P-O_{nuc})$ coordinate.

Thus, the environment seems to play a fundamental role determining the nature of the TS in the alkaline hydrolysis of phosphodiesters. The reaction field created by the aqueous solvent favors more associative TS structures and the electric field created by the enzyme stabilizes the charge separation found in dissociative TS structures, instead. Our results then point toward a scenario where the AP superfamily evolved stabilizing dissociative transition structures, both for phosphomonoesters and phosphodiesters.

3.3. Coordination of the Leaving Group. Role of Zn Ions. The role of Zn(II), as deduced from our results, agrees with the traditional view of the two-center active site in the AP superfamily.[2,13,18,58−60] This metal center stabilizes the deprotonated form of the nucleophilic residue (Thr90). However, while Zn(I) has been proposed to stabilize the leaving group,[2,13,18,59,60] our simulations indicate that the coordination of the leaving group in the active site depend on the nature of the leaving group. Effectively, in the case of the $MpNPP^-$ substrate, Zn(I) only coordinates one of the nonbridging oxygen atoms bonded to the phosphorus atoms, both in the reactants and in the TS. As pointed out above, the nitro group in *para* position is able to delocalize the negative charge very efficiently. In fact, we found that the nitro group establishes several hydrogen bonds with water molecules and enzymatic residues in the TS.

To investigate this point we performed simulations of methyl phenyl phosphate substrate (MPP^-) in the active site of NPP by replacing the nitro group of the leaving group by a hydrogen atom. We characterized the TS for the hydrolysis reaction of this substrate in NPP by analyzing structures selected from the maximum of the corresponding PMF. A representative snapshot of the TS is presented in Figure 7, where some averaged distances are displayed. As observed, the TS is still dissociative, but the value of the tightness coordinate (4.88 Å) is smaller than in the case of $MpNPP^-$ (5.66 Å). However, the most relevant change is the coordination of the leaving group. In the case of the MPP^- the charge of the leaving group cannot be delocalized by the nitro group and thus a close interaction is now established between the O_{lg} atom and the Zn(I) atom. In this case, this atom increases its coordination number during the reaction process because one of the nonbridging oxygen atoms bonded to the phosphorus is still coordinated to the metal center in the TS. In the case of the $MpNPP^-$ the leaving group was not coordinated to Zn(I) atom during the whole reaction process. Thus, to accommodate the TS of the parent compound

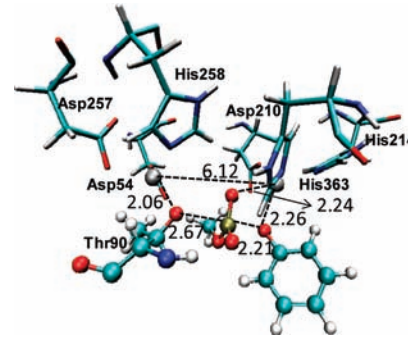


Figure 7. Representative snapshot and averaged distances (in Å) of the TS of the hydrolysis of MPP^- in NPP.

the Zn−Zn distance must be slightly increased, from 5.50 Å in $MpNPP^-$ to 6.12 Å in MPP^-. This result is counterintuitive, as one could expect, in principle, to have larger Zn−Zn distances to accommodate more dissociative TS structures. The change in this tendency is due to the change in the coordination of the Zn ion when the leaving group develops a high negative charge on the O_{lg} atom.

This ability of the active site to adapt its geometry to accommodate different TSs could explain the plasticity of this enzyme, which is able to recognize and catalyze the hydrolysis of different substrates, not only phosphate diesters, but also monoesters and sulfates.[11,19−21] Effectively, while most probably all these reactions take place through a D_NA_N mechanism, with dissociative TS structures, the specific characteristics of each residue may require slight changes in the active site probably tuning the Zn−Zn distance.

4. Conclusions

We have here presented a comparative QM/MM study of the alkaline hydrolysis of methyl *p*-nitrophenyl phosphate ($MpNPP^-$) in aqueous solution and in the active site of nucleotide pyrophosphatase/phosphodiesterase (NPP). The free energy barriers obtained from our calculations seem to be in reasonable agreement with estimations from experimental measurements, when available, especially if the theoretical results are corrected for the deficiencies of the semiempirical Hamiltonian, using higher level calculations.

The analysis of our MD simulations clearly shows that the reaction takes place through different reaction mechanisms depending on the environment. Thus, in aqueous solution the reaction occurs by means of an A_ND_N mechanism, with associative or tighten TS structures. These structures are closer to the reactants than to the products, appearing at negative values of the reaction coordinate. In the enzyme, the reaction takes place through a D_NA_N mechanism, with dissociative or loose TS structures, also appearing at negative values of the reaction coordinate. The reason for this change is due to the very different nature of the electrostatic interactions established in each of the environments. The aqueous solution reduces the repulsion between the negatively charged reacting species (nucleophile and substrate), allowing a more efficient orbital overlap between them. Instead, the electrostatic features of the NPP active site stabilize the charge distribution of dissociative TS structures, allowing the reaction to proceed with a significantly reduced free energy cost. The AP superfamily could then be specialized in the catalysis of phosphoesters through dissociative mechanisms.

Interestingly, the NPP is able to accommodate different substrates in its active site, as demonstrated by the fact that

(58) Stec, B.; Holtz, K. M.; Kantrowitz, E. R. *J. Mol. Biol.* **2000**, *299*, 1303–1311.
(59) O'Brien, P. J.; Herschlag, D. *Biochemistry* **2002**, *41*, 3207–3225.
(60) O'Brien, P. J.; Lassila, J. K.; Fenn, T. D.; Zalatan, J. G.; Herschlag, D. *Biochemistry* **2008**, *47*, 7663–7672.

this enzyme possesses secondary activities. In particular, analysis of the TS structures of other substrates in the active site of NPP, as for example methyl phenyl phosphate (MPP^-), shows that different leaving groups are stabilized in different ways. Thus, the nitro group of the *p*-nitrophenolate anion of $MpNPP^-$ establishes hydrogen-bond interactions with residues found in the outer part of the catalytic site and water molecules, while the oxygen atom originally bonded to the phosphorus atom does not coordinate directly with the Zn atoms. In absence of the nitro group, the charge of the leaving group of the MPP^- is supported to larger extent by this oxygen atom. In this case the leaving group is coordinated to the Zn(I) atom. Although it is apparently surprising, this change results into less dissociative TS structures but with larger Zn−Zn distances.

Acknowledgment. This work was supported by the Ministerio de Ciencia e Innovación, project CTQ2009-14541-C02, by Generalitat Valenciana, Prometeo/2009/053, and by Universitat Jaume I-Bancaixa foundation, project P1·1B2008-38. V.L.-C. and M.R. thank Ministerio Ciencia e Innovación for a doctoral grant and a "Juan de la Cierva" contract, respectively. V.M. thanks the Spanish Ministry Ministerio de Educación for traveling financial support, project PR2009-0539. We acknowledge computational facilities of the Servei d'Informàtica de la Universitat de València in the "Tirant" supercomputer, which is part of the Spanish Supercomputing Network.

Supporting Information Available: Computational details and a figure presenting the results of the FEP calculation for the dissociative mechanism in solution and coordinates of the QM subsystem for the transition structures localized in aqueous solution and in the enzyme in PDB format. This material is available free of charge via the Internet at http://pubs.acs.org.

JA908391V