J|A|C|S
ARTICLES
Published on Web 09/11/2004

On the Role of the Conserved Aspartate in the Hydrolysis of the Phosphocysteine Intermediate of the Low Molecular Weight Tyrosine Phosphatase

D. Asthagiri,[†,‡] Tiqing Liu,[†,§] Louis Noodleman,[†] Robert L. Van Etten,[||] and Donald Bashford*[,†,¶]

Contribution from the Department of Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, and Department of Chemistry, Purdue University, West Lafayette, Indiana 47907-1393

Received March 9, 2004; E-mail: Don.Bashford@stjude.org

Abstract: The usual rate-determining step in the catalytic mechanism of the low molecular weight tyrosine phosphatases involves the hydrolysis of a phosphocysteine intermediate. To explain this hydrolysis, general base-catalyzed attack of water by the anion of a conserved aspartic acid has sometimes been invoked. However, experimental measurements of solvent deuterium kinetic isotope effects for this enzyme do not reveal a rate-limiting proton transfer accompanying dephosphorylation. Moreover, base activation of water is difficult to reconcile with the known gas-phase proton affinities and solution phase pK_a's of aspartic acid and water. Alternatively, hydrolysis could proceed by a direct nucleophilic attack by a water molecule. To understand the hydrolysis mechanism, we have used high-level density functional methods of quantum chemistry combined with continuum electrostatics models of the protein and the solvent. Our calculations do not support a catalytic activation of water by the aspartate. Instead, they indicate that the water oxygen directly attacks the phosphorus, with the aspartate residue acting as a H-bond acceptor. In the transition state, the water protons are still bound to the oxygen. Beyond the transition state, the barrier to proton transfer to the base is greatly diminished; the aspartate can abstract a proton only after the transition state, a result consistent with experimental solvent isotope effects for this enzyme and with established precedents for phosphomonoester hydrolysis.

1. Introduction

Protein tyrosine phosphatases (PTPases) catalyze the dephosphorylation of tyrosine side chains in proteins, complementing the tyrosine kinases, which catalyze phosphorylation. The level of tyrosine phosphorylation, which is regulated by the balance between the activities of these enzymes, is in turn a regulator of many cellular processes, including cell growth, survival, and division.[1,2] The mammalian low molecular weight PTPase, which is the focus of the present study, has recently been implicated in cell transformation through its regulation of the kinase, EphA2,[3] and is therefore of interest for cancer research. Although the PTPase family is very diverse in sequence and global fold, they share a strictly conserved active site comprising the "P-loop" residues (H/V)**C**(X)$_5$**R**(S/T) and a conserved acidic residue.[4] In all structurally characterized PTPases to date, the three-dimensional structure of these active-site components is also highly conserved suggesting a common catalytic mechanism.[4,5]

In vitro studies based on model substrates, such as phenyl phosphate or *p*-nitrophenyl phosphate, have provided much of the information on the mechanistic aspects of the catalysis. In particular, it is well established that the enzyme completes its action in two major steps. In the first step, the phosphoryl group from the substrate is transferred to the nucleophilic cysteine, forming a phosphoenzyme intermediate. In the second step, this intermediate is hydrolyzed, leading to the regeneration of the enzyme and release of an inorganic phosphate. Although this two-step mechanism is well-established (see, for example, ref 6 and the topical review by Jackson and Denu[4]), various mechanistic aspects still need to be clarified. Regarding the first step, there have been questions about the protonation states of key residues in the free enzyme and the Michaelis complex, as well as about the associative or dissociative nature of the reaction pathway. In previous computational studies of low molecular weight PTP, we have addressed these questions using a combination of continuum electrostatic models and high-level

[†] The Scripps Research Institute.
[‡] Current address: Los Alamos National Laboratory, Group T-12, MS−BB268, Los Alamos, NM 87545.
[§] Current address: Department of Chemistry and Biochemistry, Montana State University, Bozeman, MT 59717.
[||] Purdue University.
[¶] Current address: Department of Molecular Biotechnology, Saint Jude Children's Research Hospital, Memphis, TN 38105.

(1) Fischer, E. H.; Charbonneau, H.; Tonks, N. K. *Science* **1991**, *253*, 401−406.
(2) Hunter, T. *Cell* **2000**, *100*, 113−127.
(3) Kikawa, K. D.; Vidale, D. R.; Van Etten, R. L.; Kinch, M. S. *J. Biol. Chem.* **2002**, *277*, 39274−39279.
(4) Jackson, M. D.; Denu, J. M. *Chem. Rev.* **2001**, *101*, 2313−2340.
(5) Zhang, Z.-Y. *Crit. Rev. Biochem. Mol. Biol.* **1998**, *33*, 1−52.
(6) Zhang, Z.-Y.; Van Etten, R. L. *J. Biol. Chem.* **1991**, *266*, 1516−1525.

density functional theory (DFT), and the results have supported a mechanism with both the enzyme cysteine and the substrate phosphate group deprotonated and a dissociative reaction pathway in which a conserved aspartic acid assists dissociation of a metaphosphate-like ion from the substrate, and this ion is then captured by the nucleophilic cysteinate.[7,8]

In the present study, we consider the second step, namely, the phosphoenzyme hydrolysis reaction. In many of the PTPases, including the low-M_r PTPase, this step is rate-determining, and a conserved aspartic acid residue that acts as a general acid in the first step may also play a role in the hydrolysis step.[5,9−12] It is sometimes proposed that the role of such a residue in the phosphoenzyme hydrolysis step is to serve as a general base facilitating the attack of a water molecule on the phosphate group. In the present case, there are reasons to suspect this proposal in that it implies a partial proton transfer from water to the aspartate residue in the transition state, which is thought to be dissociative in the enzymatic reaction.[13] Consideration of gas-phase proton affinities suggest that proton transfer from water to aspartate would be endothermic by about 42 kcal/mol,[14] while in bulk water the pK_a difference disfavors this proton transfer by nearly 15 kcal/mol. Thus, a proton transfer, to any extent, would not seem to be particularly favorable in a dissociative transition state where there is little bond formation between the phosphorus atom and the attacking water oxygen atom. Furthermore, a proton transfer during the transition state would be expected to lead to a significant solvent deuterium kinetic isotope effects. The measured solvent deuterium effect is 1.5 in both PTP1B[15] and the PTPase from *Yersinia*,[16] while in the dual-specificity enzyme, VHR, it is 1.15.[17] In both the low-M_r bovine enzyme studied here[9,18] and the yeast low-M_r PTPase LTP1,[19] no solvent deuterium isotope effect is seen with aryl substrates. For example, between pH 5.5 and 7.5, the rate-limiting dephosphorylation rate constants in H_2O and D_2O are identical, and below pH 5.5, the two curves show only an expected 0.4 unit shift characteristic of a weak acid ionization in D_2O versus H_2O (see Figure 5 in ref 19). Thus, it appears that there may be little or no general base catalysis in the hydrolysis step,[12,20] but if this is so, what is the role of the aspartate residue in the hydrolysis step?

Here, we present a study of the phosphoenzyme hydrolysis step for the bovine low molecular weight PTPase (BPTP), thus complementing our previous study of the enzyme-phosphorylation step.[8] To study the hydrolysis reaction in the protein environment, we combine high-level DFT calculations for a cluster of atoms selected for their likely importance in the mechanism, with a dielectric continuum model of the surrounding solvent and the protein (with atom-centered charges in the protein). We focus particularly on the attack by water/hydroxide upon the phosphate group and the energetics of proton transfer from the attacking water to aspartate. Linear transit (LT) techniques are used to map out the reaction pathway in terms of the attack by water on the phosphate and the proton transfer from water to the active-site aspartate. The free-energy barrier is calculated and the transition-state structure is analyzed as to its associative or dissociative character. The extent and timing of proton transfer during the reaction is analyzed and discussed in the context of kinetic isotope studies.

2. Methodology

Structural Model. A model of the covalent phosphocysteine intermediate of wild-type bovine low-M_r PTPase was derived from an early version of the 1.9 Å-resolution X-ray crystallographic structure of Zhang et al.[21] (PDB code 1dg9) kindly provided to us by Dr. Marie Zhang. Crystallographic waters and the HEPES buffer in the active site were removed, and Cys-12 was changed to phosphocysteine. The bond lengths and bond angles for this residue were taken from an ab initio geometry optimization of methyl thiophosphate. The force field parameters for this residue were adapted from the methyl phosphate group already in the CHARMM22 parameter set.[22] All molecular mechanics methodologies were performed using version c26b2 of the CHARMM package.[23]

To estimate the disposition of water molecules in the active site, we performed a stochastic boundary molecular dynamics (SBMD) simulation, following the procedure laid out by Alhambra et al.[24] Briefly, a droplet of water of radius 20 Å was overlaid on the active site and water molecules that overlap the protein were deleted. This system was then subjected to SBMD, wherein only the water molecules and the protein atoms within a set radius around the active site are free to move (see Alhambra et al.[24] for details). This allowed computational effort to be focused on pertinent regions of the protein molecule rather than requiring simulation of a much larger system. Unlike Alhambra et al.,[24] our parameters for the phosphocysteine residue were the above-mentioned CHARMM-derived ones rather than parameters based more directly on Hartree−Fock calculations. The SBMD simulations were carried for 200 ps after an initial thermalization period.

Starting from the MD structure, only water molecules within an 8 Å shell around the phosphorus atom were retained. This structure was energy-minimized following the same procedure as in our earlier study.[8] At the end of this energy minimization, by visual inspection we identified a water molecule that had a good hydrogen-bonding interaction with the aspartate and the phosphocysteine residue. In the next stage, we restrained the distance of this water and the nearest aspartate oxygen to be around 2.8 Å. This structure was further energy-minimized as before. Finally, all the water molecules, except the aforementioned water molecule, were deleted and the system was energy-minimized without any restraints, that is, all the atoms were

(7) Dillet, V.; Van Etten, R. L.; Bashford, D. *J. Phys. Chem. B* **2000**, *104*, 11321−11333.
(8) Asthagiri, D.; Dillet, V.; Liu, T.; Noodleman, L.; Van Etten, R. L.; Bashford, D. *J. Am. Chem. Soc.* **2002**, *124*, 10225−10235.
(9) Zhang, Z.-Y.; Van Etten, R. L. *Biochemistry* **1991**, *30*, 8954−8959.
(10) Denu, J. M.; Lohse, D. L.; Vijayalakshmi, J.; Saper, M. A.; Dixon, M. E. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 2493−2498.
(11) Lohse, D. L.; Denu, J. M.; Santoro, N.; Dixon, J. E. *Biochemistry* **1997**, *36*, 4568−4575.
(12) Zhang, Z.; Harms, E.; Van Etten, R. L. *J. Biol. Chem.* **1994**, *269*, 25947−25950.
(13) Zhao, Y.; Zhang, Z.-Y. *Biochem.* **1996**, *35*, 11797−11804.
(14) Bartmess, J. E. Negative ion energetics data. In *NIST chemistry webbook*; Linstrom, P. J., Mallard, W. G., Eds.; NIST standard reference database number 69; NIST: Gaithersberg, MD, 2001; http://webbook.nist.gov.
(15) Zhang, Z.-Y. *J. Biol. Chem.* **1995**, *270*, 11190−11204.
(16) Zhang, Z.-Y.; Malachowski, W. P.; Van Etten, R. L.; Dixon, J. E. *J. Biol. Chem.* **1994**, *269*, 8140−8145.
(17) Zhang, Z.-Y.; Wu, L.; Chen, L. *Biochem.* **1995**, *34*, 16088−16096.
(18) Zhang, Z. Mechanistic, kinetic, and chemical modification studies of low molecular weight protein tyrosine phosphatases. Thesis, Purdue University, 1994.
(19) Wang, S.; Stauffacher, C. V.; Van Etten, R. L. *Biochemistry* **2000**, *39*, 1234−1242.
(20) Wu, L.; Zhang, Z.-Y. *Biochem.* **1996**, *35*, 5426−5434.
(21) Zhang, M.; Zhou, M.; Van Etten, R. L.; Stauffacher, C. V. *Biochemistry* **1997**, *36*, 15−23.
(22) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiókiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(23) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(24) Alhambra, C.; Gao, J. *J. Comput. Chem.* **2000**, *21*, 1192−1203.

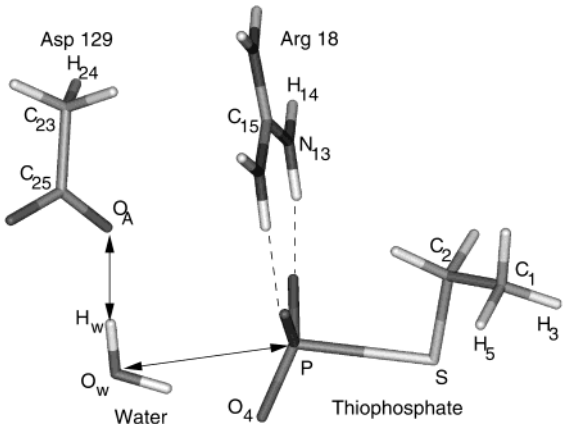


Figure 1. Schematic of the quantum cluster. The solid lines with arrows depict the two reaction coordinates used in the study.

free to move. The structure thus obtained was the starting point for the quantum cluster and protein environment effect calculations.

Quantum Cluster. The set of atoms selected for full quantum-mechanical treatment are shown in Figure 1. The phosphocysteine residue was modeled as an ethane thiophosphate by replacing the bonds between C_α (C_1 in Figure 1) and the backbone N and C atoms by $C_\alpha-H$ bonds. Arg 18 was modeled as a guanidinium ion; the bond between N_ϵ (N_{13} in Figure 1) and C_δ was replaced by a $N_\epsilon-H$ bond. Asp 129 was modeled as acetate, with the bond between the C_β (C_{23} in Figure 1) and C_α being replaced by a $C_\beta-H$ bond. In total, the cluster comprised 32 atoms. The initial geometry for all atoms but the linking hydrogen atoms was taken from the above-described energy-minimized structure. The replacements with hydrogen-linking atoms were performed within the InsightII modeling program, and the heavy-atom link hydrogen bond length was set to 1 Å.

The quantum cluster obtained after the molecular mechanics optimization was further energy-minimized using the PM3[25] semi-empirical Hamiltonian within Gaussian,[26] and this energy-minimized structure served as the starting structure for further higher level calculations. Some constraints were imposed during the PM3 energy minimization to account for the attachment of these moieties to the rest of the protein. Atoms of the ethane−thiolate part of the ethane thiophosphate group and atoms of the acetate group were not allowed to move. The internal degrees of freedom of the guanidinium ion were frozen, but rotation about the $N_{13}-H_{14}$ bond was allowed and the angle formed by $C_{15}-N_{13}-H_{14}$ was unconstrained. The position of the phosphorus atom and the oxygens bonded to it were fully unconstrained, except that the dihedral formed by $P-S-C_2-C_1$ was constrained to its initial value. All atoms of the water nucleophile could move, but the $P-O_W$ distance was constrained to the value obtained with molecular mechanics minimization.

DFT Calculations. The Amsterdam Density Functional (ADF) program (version 2.3)[27−29] was used for all the reaction path and energy calculations (see below). In the present study, we chose the VWN parametrization[30] for the local density approximation and Becke's[31] gradient correction for the nonlocal exchange and Perdew's[32] gradient correction for the nonlocal correlation. The basis sets in our calculations were double-ζ Slater type orbitals (STO) for H, C, and N, and triple-ζ STOs for the O, S, and P atoms and the water protons. A polarization function (single-ζ, 3d) was added for the S, P, and O atoms and the water protons. The inner core orbitals (1s for C, N, O and 1s, 2s, 2p for S and P) were described by the frozen core approximation. This level of basis set was used for all our linear transit/reaction path calculations and single-point energy evaluations. In some preliminary tests on model reactions, the geometry optimizations and zero-point energy calculations were performed without the polarization functions, but heats of reaction were evaluated with the full basis.

Reaction Path Calculations. The linear transit (LT) method was used for mapping the energy surface for the reaction. In this method, a particular internal coordinate, such as a bond length, is designated as the LT parameter, which is stepped from an initial to a final value during the calculation. At each transit point, the LT parameter is constrained and the remaining coordinates optimized, subject to the above-mentioned protein-attachment constraints. Thus, an LT calculation is a sequence of geometry optimizations along some selected reaction coordinates. As described below under Results, a strategy involving the O_W-P distance as a primary reaction coordinate and the O_A-H_W distance as a secondary reaction coordinate (Figure 1) was used to map out a 2-D energy surface.

In the LT optimization, a quasi-Newton approach with the Broyden−Fletcher−Goldfarb−Shanno scheme for Hessian updating[33] was used. As in the earlier work,[8] convergence tolerance required that the change in energy, energy gradients, bond lengths, and angles be simultaneously less than 10^{-3} Hartree, $5\cdot10^{-3}$ Hartree/Å, 0.01 Å, and 0.5 °, respectively. The SCF convergence was accepted when the maximum element of the commutator of the Fock matrix and the density matrix fell below 10^{-4}.

Environment Effects. The self-consistent reaction field (SCRF) approach was used to account for the environmental (protein and solvent) effect on the energetics of the quantum cluster.[34] Computational aspects and parameter sets that were used were as in the earlier work.[8] Briefly, the quantum cluster is merged back into the protein in the appropriate orientation, retaining the link hydrogen atoms in the cluster and removing the corresponding heavy atom from the protein. Alternative methods of treating the link atoms, namely, deleting them before merging the quantum cluster into the protein, gives very similar results.[8] However, the present approach of retaining link hydrogens is simpler to implement.

The protein, with its partial charge distribution, and the solvent are described as continuous media with dielectric constants of 4 and 80, respectively. The quantum cluster has a dielectric constant of 1. The reaction field due to the charges in the quantum cluster and the protein field make additional contributions to the cluster Hamiltonian. An updated electron density is computed with this new Hamiltonian and the cycle is repeated to self-consistency. The quantity of interest is the free energy at some linear transit point i relative to that at the first transit point:

$$W_i = E_i + \Delta G_{env}(i) - [E_1 + \Delta G_{env}(1)] \qquad (1)$$

where E_i is the gas-phase electronic energy, and ΔG_{env} is the free energy of interaction with the environment.

(25) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209−220.
(26) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, Ö.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mannucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Salvador, P.; Dannenberg, J. J.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komáromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*; Gaussian, Inc.: Pittsburgh, PA, 1998.
(27) Guerra, C. F.; Snijders, J. G.; te Velde, G.; Baerends, E. J. *Theor. Chem. Acc.* **1998**, *99*, 391−403.
(28) Te Velde, G.; Bickelhaupt, F. M.; Baerends, E. J.; Fonseca Guerra, C.; Van Gisbergen, S. J. A.; Snijders, J. G.; Ziegler, T. *J. Comput. Chem.* **2001**, *22*, 931−967.
(29) ADF 2.3, SCM, Theoretical Chemistry, Vrije Universiteit: Amsterdam, The Netherlands; http://www.scm.com.
(30) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200−1211.
(31) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.
(32) Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822−8824.
(33) Head, J. D.; Zerner, M. C. *Adv. Quant. Chem.* **1988**, *20*, 1−100.
(34) Li, J.; Nelson, M. R.; Peng, C. Y.; Bashford, D.; Noodleman, L. *J. Phys. Chem. A* **1998**, *102*, 6311−6324.

3. Results and Discussions

We first present calculations on simple model reactions to validate our choice of the density functional and basis set. Next, we present the results of the initial structural modeling and compare them with an experimentally determined structure. Then, the linear transit calculations for the reaction are presented and discussed in terms of the energetics and mechanism of the reaction. Last, we address proton transfer and the role of the conserved aspartate in the catalytic process.

3.1 Validation of Chosen Density Functional. Heats of reaction were computed for the following model reactions whose chemistries are similar to those in the enzymatic system under study:

$$H_2O \rightleftharpoons HO^- + H^+ \qquad \text{(rxn 1)}$$

$$CH_3COOH \rightleftharpoons CH_3COO^- + H^+ \qquad \text{(rxn 2)}$$

$$H_2O + CH_3COO^- \rightleftharpoons HO^- + CH_3COOH \qquad \text{(rxn 3)}$$

$$PO_3^- + H_2O \rightleftharpoons H_2PO_4^- \qquad \text{(rxn 4)}$$

The calculated gas-phase heats of reaction are shown in Table 1. For reaction (rxn) 1, the gas-phase proton affinity is somewhat overestimated without adjustment for basis set superposition errors (BSSE), but BSSE-corrected values are in very good agreement with experimental results. For rxn 2, again the agreement with experiments is excellent. For this reaction, BSSE is quite modest and hence disregarded. For rxn 4, there is some question about the experimental estimate. In the table, we used the most favorable experimental value for the heat of formation of PO_3^-, although others which differ by as much as 40 kcal/mol have also been reported.[14] As Keese and Castleman have noted,[35] a heat of formation around −230 to −240 kcal/mol for PO_3^- is most reasonable, and we find that this choice also leads to heats of reaction for rxn 4 that are quite close to theoretical predictions. (Again BSSE is not considered as it is likely small.) Earlier theoretical predictions for the heat of reaction for rxn 4 are quite close to the value we calculate. Wu and Houk[36] compute a value of −23.2 using MP4/6-31+G**//6-31+G*, whereas Ma et al.[37] compute −18.2 using DZP+diff SCF. Reaction 3 is just rxn 1 − rxn 2. The calculated gas-phase heat of reaction for this reaction is 48 kcal/mol, which is in excellent agreement with the experimental value based on rxn 1 and rxn 2 (Table 1).

For rxn 3, we also calculated the free-energy change in aqueous solution using the self-consistent reaction field approach[43] to calculate the electrostatic component of the solvation free energies. This method is the solvent-only analogue of the technique used here to model the protein/solvent environment.

Table 1. Heats of Model Reactions (kcal/mol)

	expt	this work
rxn 1	390.8[a]	391.3[d]
		394.7[e]
rxn 2	343.2[b]	343.3[f]
rxn 3	47.6	48
rxn 4	−19.1[c]	−19.5

[a] References 14 and 38. [b] References 14 and 39. [c] Based on the following heats of formation: −312.7 ± 6.4 for $H_2PO_4^-$, −235.8 ± 6 for PO_3^-, and −57.8 for H_2O (refs 14, 40, 41). [d] Corrected for basis set superposition errors (BSSE). [e] Non-BSSE value. [f] Non-BSSE value, although in this case BSSE is small.[42]

Table 2. Selected Geometrical Parameters of the Starting Structure (Å and Degrees)

	LTP1[a]	MM[b]	PM3[c]
$P-O_W$	3.3	3.4	3.4
O_A-H_W	2.8	3.0	2.8
O_4-O_W	2.4	3.2	2.8
$\angle O_A H_W O_W$	173.3	166.7	172.1

[a] Values from the LPT1 crystal structure[19] with a crystallographic water molecule. [b] Molecular mechanics minimization of whole protein. [c] Minimization of the cluster using the semiempirical PM3 Hamiltonian (see text).

For CH_3COOH, CH_3COO^-, H_2O, and HO^-, we calculate solvation free energies of −9.4, −81.0, −8.6, and −112.5 kcal/mol, respectively. The calculated free-energy change for rxn 3 in solution is then 15.7 kcal/mol. Using an experimental pK_a of 15.7 for H_2O and 4.8 for CH_3COOH, the corresponding experimental free-energy change is 14.9 kcal/mol. Thus, it is seen that our solution-phase values are also in good agreement with experiment for rxn 3.

Hu and Brinck[44] have compared a similar level of DFT (B3LYP) with a sophisticated level of post-Hartree−Fock theory (CCSD(T)) in calculations of barrier heights for methyl phosphate hydrolysis and found that DFT underestimates barrier heights by 2.4−4.2 kcal/mol.

3.2 Initial Structural Model. At the end of the SBMD molecular dynamics run, but before the subsequent minimization of water positions, there were at least three water molecules in the active site, one of which was hydrogen-bonded to both the Asp-129 side chain and the phosphate group, a result similar to that obtained by Alhambra et al.[24] This water molecule was the one that was finally retained after minimization to obtain an active-site model with water positioned between Asp129 and the phosphate group. The pertinent structural information from this active-site model is listed in Table 2.

Although there is no crystallographic structure of wild-type BPTP trapped in the covalent phosphoenzyme state, there is, for the yeast low molecular weight phosphatase LTP1, a crystal structure of an inorganic phosphate complexed with a Cys-to-Ala mutant enzyme in the presence of an adenine molecule[19] (PDB code 1d2a). In this crystal structure, it was found that there is a water molecule that is hydrogen-bonded to the conserved aspartate residue and to one of the oxygen atoms of the phosphate group. This oxygen is distal to the arginine residue that anchors the phosphate via two hydrogen bonds. The final oxygen is pointed in the direction of cysteine residue, although in the actual crystal structure there is an alanine in place of the cysteine. The adenine tends to H-bond and anchor the water in

(35) Keese, R. G.; Castleman, A. W., Jr. *J. Am. Chem. Soc.* **1989**, *111*, 9015−9018.
(36) Wu, Y. D.; Houk, K. N. *J. Am. Chem. Soc.* **1993**, *115*, 11997−12002.
(37) Ma, B.; Xie, Y.; Shen, M.; Schleyer, P. v. R.; Schaefer, H. F., III. *J. Am. Chem. Soc.* **1993**, *115*, 11169−11179.
(38) Schulz, P. A.; Mead, R. D.; Jones, P. L.; Lineberger, W. C. *J. Chem. Phys.* **1982**, *77*, 1153.
(39) Muftakhov, M. V.; Vasil'ev, Y. V.; Mazunov, V. A. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 1104−1108.
(40) Morris, R. A.; Knighton, W. B.; Viggiano, A. A.; Hoffman, B. C.; Schaefer, H. F., III. *J. Chem. Phys.* **1997**, *106*, 3545−3547.
(41) Rudnyi, E. B.; Vovk, O. M.; Sidirov, L. N.; Sorokin, I. D.; Alikhanyan, A. S. *High Temp.* **1986**, *24*, 56.
(42) Richardson, W. H.; Peng, C.-Y.; Bashford, D.; Noodleman, L.; Case, D. A. *Int. J. Quantum Chem.* **1997**, *61*, 207−217.
(43) Chen, J. L.; Noodleman, L.; Case, D. A.; Bashford, D. *J. Phys. Chem.* **1994**, *98*, 11059−11068.
(44) Hu, C.-H.; Brinck, T. *J. Phys. Chem. A* **1999**, *103*, 5379−5386.

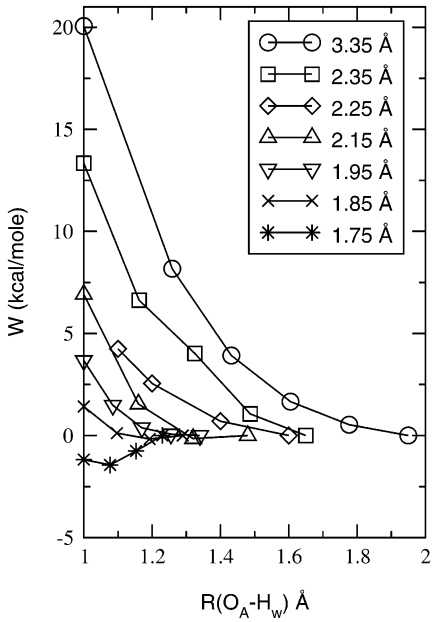


Figure 2. Energy profiles along the proton-transfer pathway for various values of $R = R(P-O_W)$. The proton transfers from right to left. As one progresses from the top curve to the bottom curve, the reaction of water with the phosphorus center progresses. The profile for $R = 2.25$ Å could not be completed down to a $R(O_A-H_W) = 1.0$, as it proved difficult to get a credible optimum geometry.

place. Key distances and angles from this structure are also presented in Table 2, where it can be seen that they are quite similar to those of the energy-minimized structures.

3.3 Reaction Path and Energy Surface. The $P-O_W$ distance was chosen as the primary reaction coordinate, consistent with an attack by either water or hydroxide (see Figure 1 for atom labels). Since the extent of proton transfer from water to the base discriminates between the extremes of attack by water and by hydroxide, the O_A-H_W distance was chosen as a secondary reaction coordinate. Thus, $P-O_W$ and O_A-H_W map out a two-dimensional energy surface that provides information on the reaction surface. For the sake of tractability, this approach limits the dimensionality of the space to be searched, and it is possible that alternative paths are missed. However, it is clear that the attack of the a water oxygen atom on the phosphorus and the cleavage of a water O−H bond must be central to the hydrolysis reaction, so it is likely that our approach captures the main features of the reaction. Starting from the optimized geometries resulting from selected steps along the primary LT path, secondary LTs were done using the O_A-H_W distance as the LT coordinate, while keeping the primary LT coordinate fixed. The energies of the resulting geometries were then calculated in the protein active-site environment using the SCRF method. The resulting energy profiles are shown in Figure 2. The energy gradient opposing proton transfer is initially quite steep but decreases as the water oxygen nears the electrophilic phosphorus center. Beyond $R(P-O_W) = 2.15$ Å, the free-energy changes for the proton transfer are within about 5 kcal/mol. The profiles corresponding to $R(P-O_W) = 1.85$ and 1.75 Å are substantially flattened with respect to proton transfer, and for $R(P-O_W) = 1.75$, the final LT point in the present calculations, proton transfer becomes favorable.

An overall energy profile in the space of the primary and secondary reaction coordinates was compiled by locating, for each value of the primary reaction coordinate, the value of the secondary coordinate for which the calculated free energy was lowest. This profile is shown as a function of the primary coordinate in Figure 3. This procedure predicts that the transition state occurs at an $R(P-O_W) = 2.25$ Å. The discontinuity of the slope at that point appears to involve the difficulty of relaxing a large number of degrees of freedom under multiple constraints and the difference between the vacuum model used in geometry optimization and the protein/solvent environment model used for the single-point calculations making up the profiles. The calculated barrier height of 17.6 kcal/mol is close to the experimentally observed value of 16 kcal/mol for the hydrolysis of the phosphoenzyme intermediate.[6] Given the tendency of DFT to slightly underestimate barriers, and the approximate treatment of the environment, this level of agreement is better than expected.

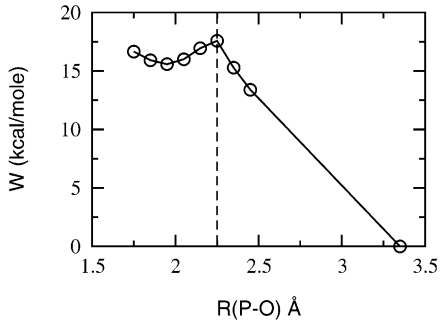


Figure 3. Energy profile as the water nucleophile reacts with the phosphorus center. Note the reactant state is in the lower right corner of the plot.

3.4 Structural Evolution. Figure 4a shows the evolution of the P−S distance as the transfer of metaphosphate to the water nucleophile progresses (reading from right to left in the graphs). From a $P-O_W$ distance of 2.45 Å to the final point at 1.75 Å, the P−S distance increases as the $P-O_W$ distance decreases. At the transition state we predict $R(P-S) = 2.66$ Å and $R(P-O_W) = 2.25$ Å. For comparison, in our previous study of the first step of the catalytic pathway we found $R(P-S) = 3.0$ Å and $R(P-O_p) = 2.7$ Å at the transition state, where O_p is the bridging oxygen in the phenyl phosphate group. The analogous plot in that case showed a nearly linear, highly dissociative pathway. In the present plot of the hydrolysis reaction, we predict a somewhat abrupt transition at around $R(P-O_W) = 2.25$ Å, beyond which the profile is once again linear.

Figure 4b shows the evolution of the O_W-H_W distance along the optimal reaction path, that is, the set of primary and secondary reaction coordinates corresponding to a local free-energy minimum in the secondary coordinate. Since the calculated energy surface is flat in the proton-transfer direction near some values of the primary reaction coordinate $R(P-O_W)$, we attempt to quantify its effect on the proton (H_W) position as follows. Proton positions (or O_W-H_W distances) for which the energy is 1 RT (0.59 kcal/mol) above the minimum are considered accessible and shown with the bars in Figure 4b. The principal conclusion of this approximate analysis is that the proton−O_W bond is substantially elongated after the transition state, but not before.

For proton-transfer pathways, vibrational effects are expected to lower the activation barrier but present methods do not allow a rigorous calculation of these effects or of proton-tunneling

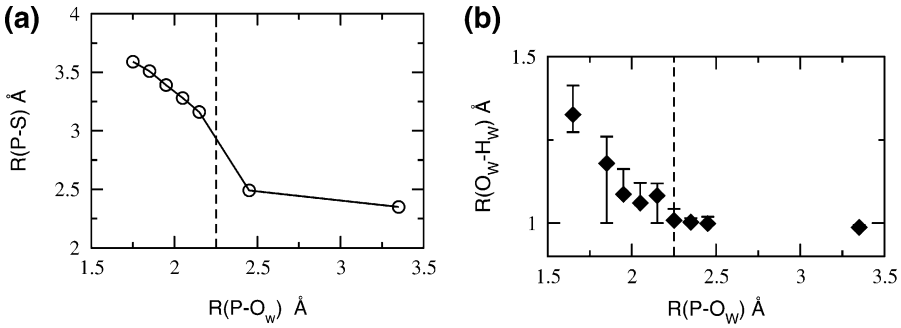


Figure 4. Variation of the P−S bond length (a) and the O_W−H_W bond length (b) along the primary reaction coordinate, the P−O_W bond length. The hydrolysis reaction proceeds from right to left in the plots. The transition-state point is indicated by the vertical broken line in both plots. The error bars in (b) specify the uncertainty in the proton's position based on 1 RT energy variation from the minimum-energy position (see text).

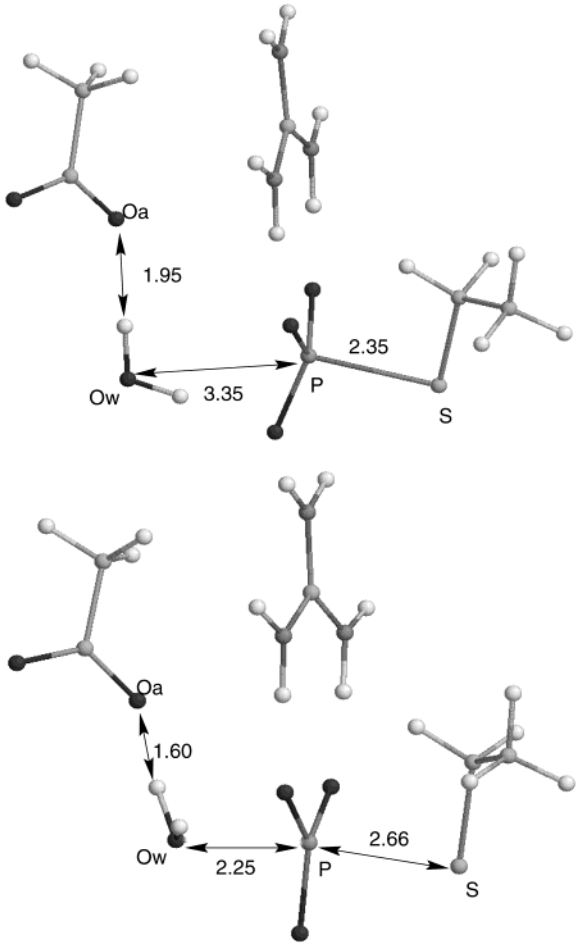


Figure 5. Structure of the quantum cluster at the initial state (top figure) and the transition state (bottom figure).

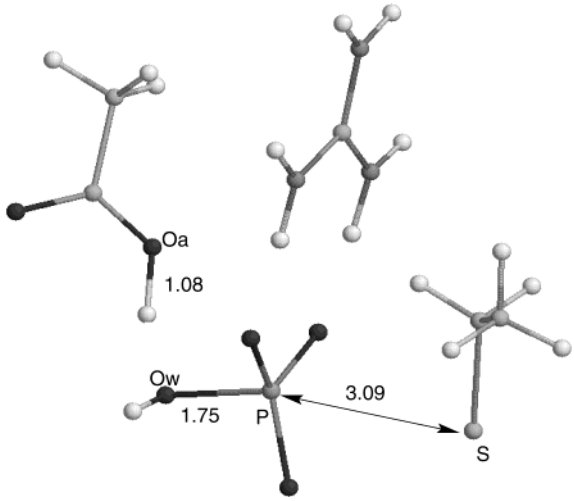


Figure 6. Structure of the quantum cluster at the final state.

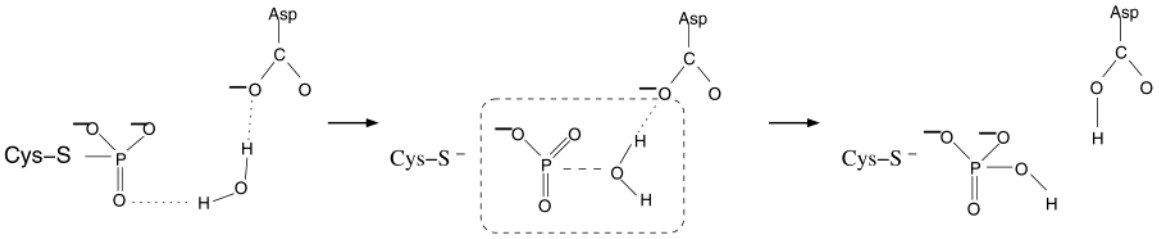
effects. For the proton transfer to the acid along a nearly linear H-bond, vibrational effects can lower the barrier by as much as 1−3 kcal/mol. Regardless of the exact number, which is likely less than 3−4 kcal/mol, it is clear that beyond the transition state, proton transfer is favored, but before the transition state, the proton transfer to the acid is strongly disfavored (Figure 2).

Figures 5 and 6 depict the cluster configuration for the initial, transition, and final states. Despite the somewhat shorter P−O and P−S bond lengths compared to those observed in the highly dissociative first step,[8] the figures are consistent with a dissociative metaphosphate-like intermediate that transfers from the phosphocysteine residue to the water nucleophile. Moreover, the charge on the PO_3 fragment in the phosphocysteine intermediate changes by about +0.2 as the system moves from the initial state to the transition state. This reduction in (negative) charge is again suggestive of a dissociative process leading to the formation of the metaphosphate (see, for example, ref 45).

That the pathway found in the present calculation is only modestly dissociative is likely attributable, in part, to omission from the quantum cluster of the Ser-19 residue that H-bonds the thiolate nucleophile. (Vibrational effects will also play a part.) Although Ser-19 is included in the classical region, the effects of hydrogen bonding are likely to be underestimated since delocalization of charge across the H-bond cannot occur, and electrostatic effects are screened by the protein dielectric. Such a hydrogen-bonding interaction would tend to lower the orbital energies of the thiolate, which, in the present context, would make the thiolate a better leaving group, resulting in a more dissociative mechanism. Such a facilitating role of Ser-19 in BPTP is seen in kinetic experiments using S19A mutants.[46] Residues that occupy a position analogous to S19 in other phosphatases also play a similar role in facilitating the easy removal of the metaphosphate from the phosphoenzyme intermediate. For example, in the *Yersinia* PTPase, removal of the conserved H-402 (analogous to S19 in BPTP) resulted in an enhanced nucleophilicity of the thiolate.[47] In the dual specificity protein tyrosine phosphatase VHR, the S131A mutation actually

(45) Maegley, K. A.; Admiraal, S. J.; Herschlag, D. *Proc. Natl. Acad. Sci.* **1996**, *93*, 8160−8166.
(46) Evans, B.; Tishmack, P. A.; Pokalsky, C.; Zhang, M.; Van Etten, R. L. *Biochemistry* **1996**, *35*, 13609−13617.
(47) Zhang, Z.-Y.; Dixon, J. E. *Biochemistry* **1993**, *32*, 9340−9345.

Scheme 1



inverted the rate-determining step, making the breakdown of the intermediate rate-limiting.[48] Cdc25, a PTPase that appears to regulate the progression of cells through the cell cycle, lacks this conserved serine/threonine residue (see, for example, refs 4 and 5). This protein is also several order of magnitude slower than other PTPases, when measured with a common substrate such as pNPP.[5]

3.5 Isotope Effects and Role of the Aspartate. As noted earlier, solvent deuterium kinetic isotope effect experiments on the low molecular weight enzymes find little or no solvent isotope effect in the hydrolysis of aryl substrates.[9,18] The present result that the proton is not transferred to Asp 129 until after the transition state is consistent with these experimental findings, since this would clearly lead to a lack of kinetic isotope signature. There could be other explanations for the lack of KIE, such as coupling of proton motions to heavy-atom motions, but the lack of positive experimental evidence for a general base role of Asp 129 and the lack of proton transfer in the transition state found here by computation leads us to conclude that general base catalysis by the aspartate is not a likely mechanism.

During most of the molecular dynamics simulation, a water molecule is H-bonded to both Asp 129 and the phosphate. Moreover, during the linear transit calculations along the O_W-P reaction coordinate, an O−H bond of the water molecule remains oriented toward one of the O_δ atoms of the aspartate. Thus, we propose that the importance of this aspartate is first to position the water nucleophile for reaction with the transferring metaphosphate, then to keep the water molecule properly oriented during the reaction, and finally, to serve as a sink for the proton once the metaphosphate transfers to nucleophilic water oxygen. Such a result is entirely consistent with the proposed explanation of the increased k_{cat} of low molecular weight PTPase in the presence of selected purines, including adenine.[19] The crystal structure of an adenine and phosphate complex of the yeast PTPase (Figure 1 in ref 19) showed adenine in an extensively hydrogen-bonded complex that positioned a water molecule precisely in position to react with the phosphate (metaphosphate) leaving group.

The positioning of Asp 129 proved important for the reaction. During the quantum cluster calculations described above, the aspartate is constrained to remain in its initial orientation obtained from the molecular dynamics simulation and energy minimization. This constraint is consistent with geometric features seen throughout the dynamics and energy-minimization procedure. Thus, it appears that the protein provides a steric context in which Asp 129 remains oriented so that favorable H-bonding to the attacking water is enhanced. If, however, in the quantum cluster the carboxylate is allowed to rotate about its C−C bond, thereby oversimplifying the steric context in which the residue exists within the protein, the linear transit calculations lead to a chemically unproductive state in which the water oxygen is forced close to the phosphorus by the LT procedure, but neither water hydrogen dissociates from the oxygen (results not shown). Thus, excessive movement of the Asp 129 carboxylate is predicted to reduce catalytic efficiency. The positioning of the base is also thought be be important in the PTPase from *Yersinia* where a hydrogen bond between the base and a nearby aspartic acid residue appears to play a functional role in keeping the base properly oriented.[16]

4. Concluding Remarks

The results of the present study can be summarized by Scheme 1, which features a relatively dissociative, metaphosphate-like transition state, followed by proton transfer from the water molecule to the carboxylate moiety.

Mildvan[49] has suggested the use of Pauling's formula for fractional bond number as a function of bond length as a quantitative measure of the associative or dissociative character of transition states. By this measure, the transition state found here is 80 dissociative. The role of the aspartate is not to facilitate any energetically meaningful general base catalysis, but to position the attacking water and accept a proton from water after the transition state is passed. Not shown in Scheme 1 is the release of phosphate from the active-site pocket to form free enzyme, completing the catalytic cycle. The energy barrier for this final step is significantly smaller than barrier for hydrolysis studied here.[6] The deprotonated state of Asp 129 predominates in the free enzyme at neutral pH ($pK_a = 5.3$), while the protonation state of the aspartate in the product-bound form is not known. The negatively charged phosphate could perturb the Asp 129 pK_a causing it to remain protonated (an analogous effect in the Michaelis complex has been seen both experimentally and computationally in PTPases[7]), or Asp 129 could deprotonate prior to the release of phosphate. This question would effect the energetics of the final product release but not the rate-limiting hydrolysis step studied here. Scheme 1 is consistent with experimental studies on the low-M_r PTPases, particularly kinetic isotope effect studies. The importance of positioning the water molecule is also suggested by the increase of k_{cat} in the presence of adenine and structural studies showing adenine involved in positioning the water molecule.[19] The spatiotemporal hypothesis, advanced earlier by Menger,[50,51] emphasizes the role of such positioning in catalysis.

The mechanism our calculations support is similar to the Scheme 2 mechanism proposed over a decade ago by Herschlag and Jencks[52] for the phosphorylation of glucose. A dissociative, metaphosphate-like character for the transition state was also

(48) Denu, J. M.; Dixon, J. E. *Proc. Natl. Acad. Sci.* **1995**, *92*, 5910−5914.
(49) Mildvan, A. S. *Proteins* **1997**, *29*, 401−416.
(50) Menger, F. M. *Acc. Chem. Res.* **1985**, *18*, 128−134.
(51) Menger, F. M. *Biochemistry* **1992**, *31*, 5368−5373.
(52) Herschlag, D.; Jencks, W. P. *J. Am. Chem. Soc.* **1989**, *111*, 7587−7596.

found in our previous study of the preceding step in which a phosphate group is transferred from the phenyl phosphate substrate moiety to the nucleophilic cysteinate, Cys 12. These results are consistent with the general expectation that the enzymatic transition states and pathways do not differ greatly from analogous phosphoryl transfer reactions in aqueous solution, which generally proceed through such a transition state.[52−54]

A dissociative character for the dephosphorylation step of PTPase has also been inferred from analysis using linear free-energy relationships of phosphate transfer between the phosphoenzyme and alcohols of varying pK_a.[13] Linear free-energy relationships have also been used to explore the dissociative or associative character of the first step of alkaline phosphatase (AP) which hydrolyzes phosphate monoesters by a two-step mechanism. It had been proposed that the thio effect in AP was indicative of an associative mechanism similar to that of phosphotriesters rather than phosphomonoesters in solution, but measurements of Brønsted β_{lg} in the wild type and a mutant indicated that the thio effect did not characterize the transition state of AP.[55]

Several computational studies relevant to the first step of PTPase have been discussed in our previous publication on that step.[8] There have been some other computational studies of other enzymatic phosphoryl transfers, somewhat analogous to the one studied here. A high-level DFT study of phosphoryl transfer from ATP to a substrate serine analogue in the protein kinase, cAPK has found a reaction path analogous to that of the present study.[56] A dissociative transition state with metaphosphate poised between the β-phosphate and the oxygen atom of the still-protonated substrate is found, and a conserved active-site aspartate acts as a "proton trap" well after the transition state is passed. More recent calculations on cAPK using DFT found a similar path involving Asp 166 as a base but also located a path with a much higher barrier involving proton transfer directly from the substrate serine OH to the transferred phosphoryl group.[57] The latter (disfavored) pathway is more associative and is similar to one proposed earlier as the principal pathway on the basis of semiempirical calculations.[58,59] Comparison of model-compound calculations at various levels of theory indicated that the DFT methods are more reliable for phosphoryl transfer reactions than the semiempirical methods used in the earlier calculations.[57] A different DFT study found an associative pathway with proton transfer directly to the phosphate in agreement with the semiempirical results,[60] but the calculations omitted Asp 166 altogether, so a lower-barrier path involving it as a proton acceptor could not have been found. Overall, the higher-level calculations on cAPK have supported a dissociative mechanism with proton transfer to an aspartate after the transition state. The cAPK active site complex involves one or two Mg(II) ions, and the reaction being catalyzed involves different attacking and leaving groups, so its applicability to the present PTPase mechanism may be limited.

Acknowledgment. D.A. thanks Michael Thompson (Scripps) for helpful discussions. The authors gratefully acknowledge funding from the NIH as follows: D.A. and D.B., grant no. GM-45607; R.L.V.E., grant no. GM-27003; L.N., grant no. GM-43278 and GM-39914.

JA048638O

(53) Thatcher, G. R. J.; Kluger, R. *Adv. Phys. Org. Chem.* **1989**, *25*, 99−265.
(54) Admiraal, S. J.; Herschlag, D. *J. Am. Chem. Soc.* **2000**, *122*, 2145−2148.
(55) Holtz, K. M.; Catrina, I. A.; Hengge, A. C.; Kantrowitz, E. R. *Biochemistry* **2000**, *39*, 9451−9458.
(56) Valiev, M.; Kawai, R.; Adams, J. A.; Weare, J. H. *J. Am. Chem. Soc.* **2003**, *125*, 9926−9927.
(57) Díaz, N.; Field, M. J. *J. Am. Chem. Soc.* **2004**, *126*, 529−542.
(58) Hart, J. C.; Sheppard, D. W.; Hillier, I. H.; Burton, N. A. *Chem. Commun.* **1999**, 79−80.
(59) Hutter, M. C.; Helms, V. *Protein Sci.* **1999**, *8*, 2728−2733.
(60) Hirano, Y.; Hata, M.; Hoshino, T.; Tsuda, M. *J. Phys. Chem. B* **2002**, *106*, 5788−5792.