Stabilization of Charges and Protonation States in the Active Site of the Protein Tyrosine Phosphatases: A Computational Study†

Valérie Dillet,†,‡ Robert L. Van Etten,§ and Donald Bashford*,†

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

Received: April 25, 2000; In Final Form: August 29, 2000

The initial step of the dephosphorylation of phosphotyrosine by the protein tyrosine phosphatases (PTPases) requires an active-site cysteine residue to be in its deprotonated (thiolate) form so as to make a nucleophilic attack on the phosphate to form a phosphoenzyme intermediate. At the same time, an active-site aspartic acid residue must be in its protonated form to serve as a general acid donating a proton to the leaving tyrosyl moiety. The protein must therefore provide an unusual electrostatic environment in the active site to maintain these protonation states in the free enzyme and in the Michaelis complex with the substrate, whose phosphate group is in the −2 charged form. We present electrostatic calculations for the pK_a of the ionizable side chains of four different PTPases. The study includes Michaelis complexes with the substrate, as well as free enzymes. The calculations consistently predict that in the neutral to mildly acidic pH range, the active-site cysteine residue is deprotonated and the aspartic acid residue is protonated, as required for optimal catalysis. This prediction is made even for the Michaelis complexes. Good quantitative agreement with measured pK_a values of the aspartic acid residues are obtained, while the calculated pK_a values of the cysteine residue are generally much lower than experimental estimates, suggesting that the protonation of the (normally unprotonated) cysteine residue may induce a conformational change in the active site. An analysis is made of the electrostatic factors responsible for maintaining the functionally essential protonation states. The phosphate-binding loop, whose backbone dipolar groups and conserved arginine side chain are oriented so as to stabilize negative charge in the active site, is found to be particularly important. However, structural features that are not conserved across the PTPase family make significant contributions in particular cases. The PTPases use dipoles more than positive charges to stabilize negative charge in the active site, and this helps to prevent the destabilization of the general acid by the longer range interactions with positive charges. The use of a low protein dielectric constant in the calculations is essential to obtaining calculated protonation states that are consistent with the catalytic mechanism. Models using a higher dielectric constant bring the calculated cysteine pK_a values closer to some reported experimental values for the free enzyme, but fail to predict the protonation of the general acid, or the maintenance of the thiolate form of the cysteine residue in the Michaelis complex.

Protein phosphorylation is a common mechanism through which cellular processes are regulated, especially in response to extracellular signals.[1] In particular, phosphorylation of tyrosine residues is implicated in the control of cell growth, proliferation, and death.[2,3] As with serine or threonine phosphorylation, the level of tyrosine phosphorylation is the result of the balance between the action of protein kinases and protein phosphatases. Although lagging behind studies of protein tyrosine kinases, studies of protein tyrosine phosphatases (PTPases) have now emerged as an important topic in the regulation of intracellular processes.[1,4] Recently, the PTPases have been the topic of several theoretical studies concerning reaction pathways, ionization states and molecular dynamics.[5−10] The present study focuses on the ionization states of groups involved in catalysis in the free enzyme and in the Michaelis complex for several diverse members of the PTPase family.

Approximately 100 PTPases are known, and they can be classified into three subfamilies.[4,11,12] (i) High-molecular-weight PTPases: these include both cytosolic and membrane-bound receptor enzymes and are perhaps the best understood in terms of biological function. They all share a conserved catalytic domain of approximately 250 residues, but are otherwise quite diverse. X-ray crystallographic structures have been obtained for two of them, PTP1B[13,14] and Yersinia PTPase,[15−18] both of which are nonreceptor enzymes containing a single catalytic domain. Despite having less than 15% sequence identity, both of these proteins have a similar α/β folding architecture. (ii) Low-molecular-weight ($M_r \simeq 18\,000$) PTPases: the biological function of these enzymes is not as well understood as that of the high-M_r PTPases, but they are ubiquitous in mammalian tissues and are thought to be involved in cell-cycle regulation[4] or cell−cell communication processes.[19,20] Recently, a bacterial PTPase belonging to this family has been characterized, and an endogenous protein kinase has been identified as a substrate.[21] Members of this subfamily are homologous, and exhibit substantial sequence similarity, but apart from a conserved active-site motif, they have very limited sequence similarity to the other PTPase subfamilies. High-resolution X-ray structures are available for bovine (BPTP),[22,23] human[24] and yeast[25] low-M_r PTPases, and the solution structure of the bovine enzyme

† The Scripps Research Institute.

‡ Permanent address: LEDSS, 301, rue de la Chimie, D. U. Saint Martin d'Hères - BP 53, 38000 Grenoble Cedex 9 - France

§ Purdue University.

10.1021/jp001575l CCC: $19.00
Published on Web 11/04/2000

has been studied by NMR.[26] The structures of these enzymes have the classic $\beta\alpha\beta$ motifs of the Rossmann dinucleotide-binding fold. (iii) Dual-specificity PTPases: although these enzymes are active against phosphoserine and phosphothreonine as well as phosphotyrosine, they are structurally and mechanistically related to the PTPases rather than the Ser/Thr phosphatases. They are ubiquitously expressed, and the regulatory function of several of them has been demonstrated.[4] Apart from the active-site motif, they have little sequence similarity to the other PTPase subfamilies. An X-ray structure of one of them, VHR, has been determined, and aspects of structural similarity between VHR, PTP1B and Yersinia PTPase have been discussed.[27]

Although quite diverse in sequence and structure, all PTPases apparently share a common mechanism and a few common structural features.[4,12] Hydrolysis of the phosphotyrosyl residue occurs in two steps. In the first step (enzyme phosphorylation) a covalent phosphoenzyme intermediate is formed by the nucleophilic attack on the substrate phosphorus atom of a strictly conserved cysteine, while an acidic group donates a proton to the oxygen atom of the leaving tyrosyl group. The second step (dephosphorylation) is the hydrolysis of the phosphoenzyme intermediate, releasing an inorganic phosphate anion and restoring the enzyme to its starting state.

All of the PTPases share a common active-site sequence motif: (H/V)**C**X_5**R**(S/T), where **C** indicates the strictly conserved nucleophilic cysteine residue, and **R** is a strictly conserved arginine residue. In the known three-dimensional structures, the residues of this motif form a loop, known as the phosphate binding loop or P-loop, that surrounds the phosphate binding site and is highly superimposable[12] from one PTPase to another (Figure 1). The C-terminus of the P-loop coincides with the N-terminus of an α-helix, which is another structural feature conserved among all PTPases. Crystallographic and NMR solution structure data for the low Mr PTPases indicate that the P-loop is very rigid, while kinetic and site-directed mutagenesis results show that mutations that alter the structure of the P-loop lead to reductions in the catalytic activities of the mutant enzymes.[23,25,26,28,29] The positively charged arginine side chain is absolutely essential for activity.[30] These observations may be understood upon examination of structures such as one of the HEPES substrate analogue complexes (Figure 2). Such structures consistently show that six of the backbone NH groups of the P-loop, and two of the NH groups of the side chain of the conserved arginine are all positioned so as to form strong hydrogen bonds with a phosphate in the active site. Several hydrophobic residues near the entrance of the active site cavity help to bind the substrate and to position it[12,22] with an axial symmetry similar to that which is observed in the trigonal bipyramidal, vanadate ion transition state analogue complex of BPTP.[22,23] Moreover, the oriented hydrogen bonds contributed by the rigid P-loop are positioned such that they are even shorter in the trigonal bipyramidal transition state complex than they are in the reactant complex.[23]

The Ser/Thr at the end of the P-loop is close enough to form a hydrogen bond to the cysteine sulfur atom, and its mutation in BPTP (S19A) or VHR (S131A) has been shown to retard the enzymic phosphorylation step.[29,31] This conserved Ser/Thr reduces the pK_a of the conserved, active-site cysteine residue.[29,31−33] It is known that the mutation of Ser-19 to Ala in BPTP causes only small changes in the positions of the P-loop atoms.[25] Despite this, the kinetic behavior of this mutant is dramatically altered.[29] Values of k_{cat}/K_m are reduced by several orders of magnitude, the rate determining step is changed from dephosphorylation to phosphorylation, and the pH dependence of k_{cat}/K_m clearly shows an ascending limb in the low pH region that cannot be seen with the wild-type enzyme.[29] (In the case of wild-type BPTP, attempts to extend the pH dependence measurements below approximately pH 3.8 result in protein denaturation.) Thus, although the pK_a of the active site cysteine of low M_r PTPases cannot be determined from measurements of k_{cat}/K_m, it is clearly less than 4.

In contrast to the conserved P-loop motif, there is comparatively little sequence similarity across the PTPase family in the region of the general acid. For representative enzymes of each

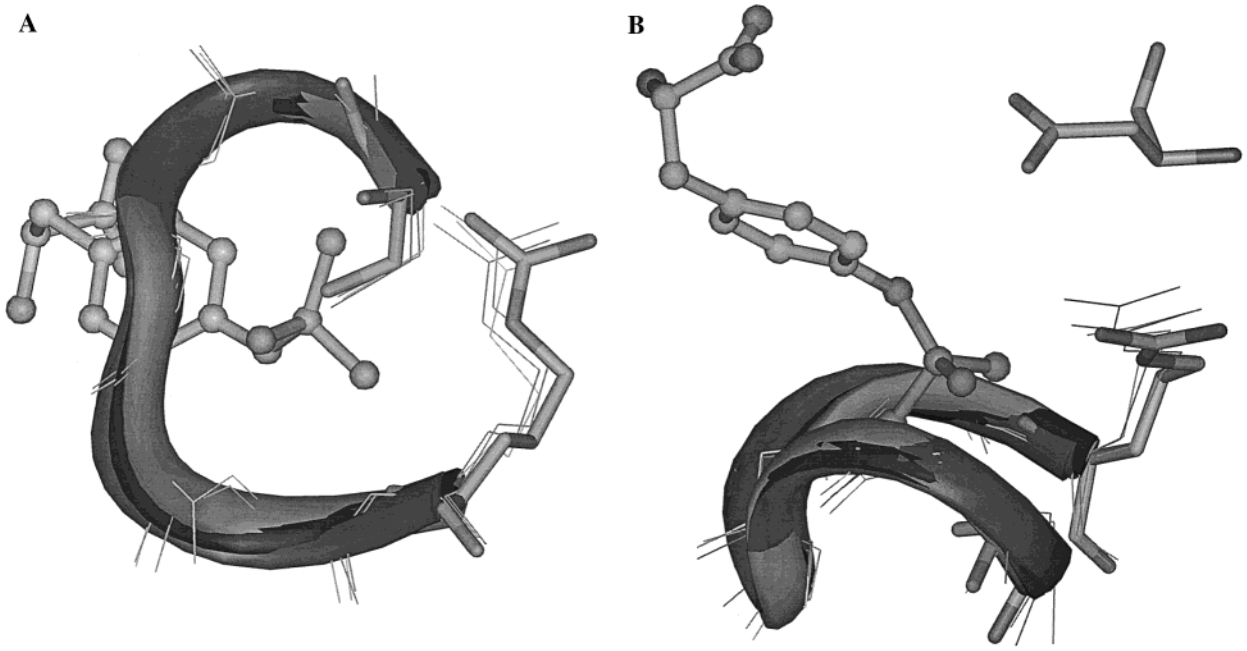

Figure 1. (A) Superposition of the phosphate binding loop (from the nucleophilic cysteine to the invariant arginine) from PTP1B (1ptv.pdb),[14] Yersinia PTPase (1ytw.pdb),[17] BPTP (2pnt.pdb),[23] and VHR (1vhr.pdb/A).[27] The serine (mutated from the nucleophilic cysteine) and the arginine of 1ptv are displayed with heavier sticks, whereas the substrate of this same structure (a phosphotyrosine residue) is displayed with balls and sticks. (B) Side view showing position of the general acid, Asp, (upper right) relative to the phenyl phosphate (upper left) from the 1ptv structure. The superposition and graphics were done using the InsightII program from MSI of San Diego (Version 97.0).

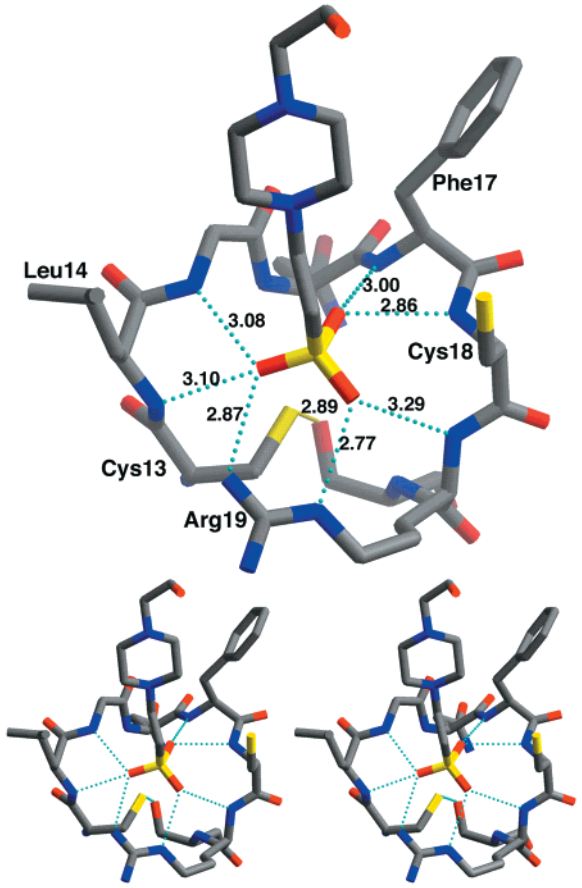

Figure 2. Hydrogen bonding pattern observed for the complex of HEPES buffer with the P-loop residues of the low Mr PTPase from *Saccharomyces cerevisiae.* (A) Hydrogen bond distances and residue identification. Shown are the multiple hydrogen bonds and distances from main-chain backbone NH's and conserved Arg side chain to sulfonate oxygen atoms. Also indicated is the hydrogen bond between the active site nucleophile (Cys-13 in the yeast enzyme) and the hydroxyl group of the conserved Ser residue in the P-loop. (B) A stereoview illustrating the relative orientation of the substrate analogue with respect to the P-loop and the active site nucleophile. Residue numbering is omitted for the sake of clarity. Coordinates are from PDB file 1D1P.[25]

family, the general acid has been identified, and it is always an aspartic acid residue (ref 11 and references therein), but there is no consensus as to its position in the sequence—in the low-M_r PTPases it lies approximately 110 residues to the C-terminal side of the P-loop, whereas in the high-M_r and dual-specificity subfamilies it lies 30 to 50 residues to the N-terminal side of the P-loop. In the high-M_r PTPases, the conserved aspartic acid is part of a flexible loop, termed the WpD loop,[16] which apparently changes from an "open" to a "closed" conformation upon ligand binding.[16,34] For PTP1B and Yersinia PTPase, X-ray crystallographic structures are available for both forms.[13−18] On the other hand, there does not appear to be so much flexibility in the Asp-bearing loop of the low-M_r PTPases, which has no sequence similarity, apart from the Asp, to the corresponding loop of the other PTPases. The only common 3D-structural feature of the conserved Asp across all subfamilies is that it is positioned with its carboxylic acid group within hydrogen-bonding distance from the apical oxygen of the substrate (Figure 1b), and is thus well-placed to protonate the leaving tyrosinate group during the initial step of catalysis.[14,16,17,23,27]

Measurements of k_{cat}/K_m as a function of pH show that for the first step of catalysis to proceed, the phosphate group of the substrate must be deprotonated to the dianionic form, the

TABLE 1: Structural Models

model	enzyme	PDB code of original structure	ligand in original structure[a]
	Unliganded (open) enzymes		
2hnp	PTP1B	2hnp	
1ypt/A	Yersinia PTPase	1ypt	
1ypt/B	Yersinia PTPase	1ypt	
	Enzyme/phenyl phosphate complexes		
1ptv	PTP1B	1ptv	pTyr
1ptt	PTP1B	1ptt	pTyr tetrapep.[b]
1ptu	PTP1B	1ptu	pTyr hexapep.[c]
1pty	PTP1B	1pty	pTyr
1yts	Yersinia PTPase	1yts	sulfate
1ytw	Yersinia PTPase	1ytw	tungstate
1ytn	Yersinia PTPase	1ytn	nitrate
2pnt	BPTP	2pnt	HEPES
1vhr/A	VHR	1vhr	HEPES
1vhr/B	VHR	1vhr	HEPES

[a] Only the ligands lying in the active-site binding pockets are given. In structural models of the complexes, these original ligands are replaced by phenyl phosphate (see text). [b] Phosphotyrosine-containing tetra-peptide (Ac-DEpYL-NH_2). [c] Phosphotyrosine-containing hexa-peptide (DADEpYL-NH_2).

TABLE 2: Numbering of Selected Key Residues[a]

PTP1B	Yersinia PTPase	BPTP	VHR
Arg45	Arg228		-
Glu115	Glu290		Met69
Lys120	Gly297		Asn74
His175	His350		Gly87
Asp181	**Asp356**	**Asp129**	**Asp92**
His214	His402		His123
Cys215	**Cys403**	**Cys12**	**Cys124**
Ser216	Arg404		Arg125
Arg221	**Arg409**	**Arg18**	**Arg130**
Ser222	**Thr410**	**Ser19**	**Ser131**
Arg257	Arg440		Arg158

[a] Based on the sequence alignment of ref 27 for PTP1B, Yersinia PTPase and VHR. Bold: important catalytic residues conserved in the PTPase family (numbering of these residues for BPTP is also included).

active-site cysteine residue must be deprotonated, and the active-site aspartic acid residue must be protonated.[29,31,32,35−37] These studies and others[38] establish that the active-site cysteine pK_a is shifted downward into the acidic range, whereas the aspartic acid pK_a is shifted upward into the neutral to mildly acidic range. Kinetic isotope effect experiments indicate that the tyrosine leaving group departs in its neutral, protonated, form.[39−41] These observations, together with structural studies, lead to the following picture of the initial step of catalysis:[29] In the free enzyme, the cysteine is initially unprotonated and the aspartic acid is protonated. The free phosphotyrosyl substrate moiety is initially unprotonated. These protonation states are maintained as the Michaelis complex is formed. Then, as the aspartic acid donates a proton to the leaving tyrosyl group, the thiolate of the active-site cysteine residue makes a nucleophilic attack on the substrate phosphorus atom. This results in the transfer of the phosphate group from the substrate to the enzyme, to form a covalent phosphocysteinyl enzyme intermediate.[30] (The subsequent dephosphorylation step is outside the scope of the present study.)

The aim of the present study is to give a quantitative account of the means by which the PTPases stabilize the unusual charge states in the free enzyme and in the Michaelis complex that are required for the initial step of protonation. In particular, the requirement that the cysteine thiol group and the substrate phosphate group both be unprotonated means that the enzyme must stabilize a charge of −1 in the free enzyme and −3 in the

TABLE 3: Charges and Radii

	Phenyl Phosphate							
atom	P	O1	O2	O3	O4	C1	C(ar)	H(ar)
charge[a]	0.90	−0.85	−0.85	−0.85	−0.38	0.03	−0.10	0.10
radius[b]	2.15	1.77	1.70	1.70	1.70	1.992	1.992	1.358
radius[c]	1.8	1.4	1.4	1.4	1.4	1.7	1.7	1.0

	Cysteine[d]		
atom	CB	SG	HG1
charge			
protonated form	0.00	−0.29	0.29
deprotonated form	−0.08	−0.92	0.00
radius	2.00	1.85	1.00

	Aspartic Acid[d]		
atom	CG	OD1	OD2	HD2
charge				
protonated form	0.55	−0.495	−0.49	0.435
deprotonated form	0.1	−0.55	−0.55	0.00
radius	1.70	1.40	1.40	1.00

[a] Charges for the deprotonated form (adapted from the CVFF parameters of the protonated form). [b] Radii (in Å) used for the energy minimizations, taken from the CHARMM22 parameters for tyrosine and methyl phosphate. [c] Radii (in Å) used for the electrostatic calculations, taken from the PARSE parameter set[53] for oxygen, carbon and hydrogen atoms and equal to the mean van der Waals radius defined by Bondi[72] for the phosphorus atom. [d] Charges and radii for the MEAD calculations only. These parameters are from the PARSE set, as are parameters for all other atoms not shown.

Michaelis complex. At the same time, a nearby aspartic acid residue must be kept protonated, so as to serve as a general acid. This paper reports the use of a semi-macroscopic electrostatic model to study this problem for set of enzymes representative of the PTPase family. Calculations of pK_a values based on the model are presented, and an analysis of the contributions to the calculated pK_a shifts from various parts of the protein, and across different members of the family is made. We also compare our results to previous theoretical studies of active-site ionization states.[8,9] Because the large electrostatic effects in the PTPases present a greater challenge to electrostatic models than the more common application to the prediction of the pK_a shifts of surface residues, this study also provides a test of the applicability and parametrization of the electrostatic model. In particular, we address the question of the appropriate value for the interior dielectric constant of the protein in such models.

Structural Models

The source and principal characteristics of the PTPase structural models used in the present study are summarized in Table 1. The numbering of homologous functionally important residues in PTP1B, Yersinia PTPase, BPTP, and VHR is listed in Table 2.

Only PTP1B[13] and Yersinia PTPase[15] have been successfully crystallized in the absence of ligand. Therefore, structural models for the unliganded form of the enzyme were derived from these two structures, and no other PTPases were modeled in the unliganded form. Structural models of the Michaelis complex were derived from X-ray structures of PTPase/ligand complexes available for PTP1B,[14,18] Yersinia PTPase,[16,17] BPTP,[23] and VHR.[27]

A common hydrogen building and energy-minimization scheme was used for all models: Crystallographic water molecules were removed from the original PDB files. Coordinates for the hydrogen atoms were generated using the HBUILD[42] module of CHARMM[43] and refined by steepest-descent energy minimization (200 steps) in the absence of electrostatic interactions, followed by conjugate-gradient minimization including electrostatic interactions (13.0 Å cutoff, dielectric constant 4.0). Unless stated otherwise for a particular model, only hydrogen atoms were allowed to move during these minimizations. The CHARMM computer program[43] and the CHARMM22 empirical potential energy function[44] were used for these energy minimizations. During all minimizations, the nucleophilic cysteine was modeled in the deprotonated (thiolate) state, to preserve the hydrogen bond from the hydroxyl of the Ser/Thr conserved residue to the sulfur atom of the cysteine. Coordinates for the thiol hydrogen were subsequently generated using HBUILD. For all of the Michaelis complexes, the phenyl phosphate was modeled in its fully deprotonated state. The charges and radii used for this group during the minimizations are given in Table 3. These charges were adapted from those of the protonated form in the CVFF force field[45] and compare well with the values derived by Hansson et al.[6] for phosphate esters. The radii of the phenyl phosphate atoms were taken from the CHARMM parameters for tyrosine and methyl phosphate.

Models for the Unliganded Enzyme: PTP1B and Yersinia PTPase. Models for the free enzyme were taken from the crystal structure of the unliganded form of PTP1B (2hnp.pdb)[13] and Yersinia PTPase (1ypt.pdb).[15] In the latter case, the two protein molecules present in the crystal unit were treated independently, and the two distinct structural models are referred to here as 1ypt/A and 1ypt/B.

Models for the Michaelis Complex. A number of models for Michaelis complexes of a PTPase and a phenyl phosphate substrate were produced using available crystallographic structures of various PTPases complexed with various ligands as starting points. A first model of the Michaelis complex was derived from the X-ray crystallographic coordinates of the C251S mutant of PTP1B complexed with phosphotyrosine (1ptv.pdb).[14] The mutation inactivates the enzyme by replacing the nucleophilic cysteine side chain with the sterically similar serine side chain, so the structure should closely resemble the Michaelis complex of the native enzyme. In the modeling procedure, Ser215 was changed back to a deprotonated cysteine residue by removing the H^{γ} atom and replacing the O^{γ} atom by a sulfur atom. The C^{β} and backbone atoms of the phosphotyrosine ligand were removed, so that the ligand in the model structure was phenyl phosphate. The structure was further refined as described above, with the difference that (in addition to the hydrogen atoms) the substrate atoms and the sulfur atom of Cys215 were allowed to move during the minimization. This was intended to model an enzyme−substrate geometry more

representative of the catalytically active Michaelis complex, in which the cysteine is deprotonated.

Models for other complexes are based on the crystallographic coordinates of (i) PTP1B (C215S mutant) complexed with a phosphotyrosine tetrapeptide (1ptt.pdb),[14] a phosphotyrosine hexapeptide (1ptu.pdb),[14] or a phosphotyrosine (1pty.pdb);[18] (ii) Yersinia PTPase complexed with a sulfate (1yts.pdb),[17] a tungstate (1ytw.pdb),[16] or a nitrate ligand (1ytn.pdb);[17] (iii) BPTP with a HEPES ligand (2pnt.pdb)[23] and (*iv*) VHR (1vhr.pdb)[27] with a HEPES ligand (Table 1). The two protein molecules present in 1vhr.pdb were treated independently and will be referred to as model structures 1vhr/A and 1vhr/B. For these nine liganded structures, after deletion of the crystallographic water and substrate molecules, the P-loop backbone atoms were superimposed with those of the complex modeled using the 1ptv.pdb structure to obtain the initial phenyl phosphate coordinates. Further processing was as described previously for the 1ptv model, with the sulfur atom of the nucleophilic cysteine and the substrate atoms allowed to move during the minimizations.

Theory and Computations

The calculations are based on the assumptions that any difference between the pH-titration behavior of a group in a protein versus that in an analogous model compound is the result of differences in the electrostatic interactions of the group in the protein versus the model compound, and that these electrostatic interactions can be treated by the MEAD (macroscopic electrostatic with atomic detail) model, in which the protein interior is a medium of dielectric constant ϵ_p, the exterior is a medium with the dielectric constant and ionic strength of the bulk solvent, the boundary between interior and exterior is the Connolly[46] surface, as determined from the atomic coordinates and radii, and the electrostatic interactions within this system can be found by solving the corresponding linearized Poisson−Boltzmann equation.

Methods for the pK_a Calculations. Here we briefly summarize the procedure used to compute the pK_a values of titratable residues from the protein structures. More details are given elsewhere.[47,48] To compute the pK_a of a given residue *i* we first evaluate its intrinsic pK_a (pK_{intr}), defined as the pK_a of that residue when all the other ionizable groups are neutral.[49] This quantity can be written as

$$\mathrm{p}K_{intr} = \mathrm{p}K_{mod} + \Delta\mathrm{p}K_{Born} + \Delta\mathrm{p}K_{back} \qquad (1)$$

where pK_{mod} is the pK_a of the *N*-formyl-*N*-methylamide derivative of the residue (i.e., the model compound), and ΔpK_{Born} and ΔpK_{back} are the p*K* shifts due to the desolvation of the residue and its interaction with the atomic partial charges of the protein (in its neutral form), respectively. To compute the pK_a we also need the charge−charge interactions between residue *i* and the other ionizable groups *j*. In our model this quantity is expressed as a sum of pairwise terms (W_{ij}). The ΔpK_{Born}, ΔpK_{back} and W_{ij} terms are evaluated using a continuum electrostatic model. In this model, the protein (or the model compound) is regarded as a low dielectric medium (with a dielectric constant of 4), in which the protein partial atomic charges are embedded, and the solvent is regarded as a high dielectric medium (with a dielectric constant of 80). The counterion atmosphere in the solvent is modeled using the Debye−Hückel approximation. The ΔpK_{Born}, ΔpK_{back}, and the W_{ij} terms of all the titrating groups can be evaluated by solving the linearized Poisson−Boltzmann equation for various subsets of partial charges (see refs 47, 50 for details). Given these quantities, the fractional protonation of all residues at a given pH can, in principle, be evaluated using a Boltzmann distribution over the possible protonation microstates of the system (for a protein with N ionizable sites, there are 2^N possible protonation microstates). After computing the fractional protonation for all of the ionizable residues over a sufficient range of pH, the ΔpK_a of any titratable group is obtained as the pH value at which this group is half protonated. To permit more detailed analysis, it is useful to split ΔpK_{back} into individual contributions (Δp$K_{back,k}$) from all other residues *k* as

$$\Delta\mathrm{p}K_{back} = \sum_{k=1}^{N_{res}} \Delta\mathrm{p}K_{back,k} \qquad (2)$$

where N_{res} is the total number of residues in the protein. The ΔpK_{back} term can also be subdivided into backbone and side chain contributions.

Parameters and Programs for the Electrostatic Calculations. For the PTP1B and Yersinia PTPase Michaelis complexes, the electrostatic potential was evaluated using the finite-difference method with the grid-focusing technique[51] over three successive grids with spacing of 2, 1.0, and 0.25 Å. In all other cases only two successive grids, with spacing of 1.0 and 0.25 Å, were used. Cubic lattices of 81^3 and 61^3 points were used for the calculations on the enzyme (or Michaelis complex) and the model compounds, respectively. The dielectric boundary between the interior (unliganded enzyme, Michaelis complex or model compound) and the exterior (solvent) media was defined by the contact and reentrant surfaces[52] using a solvent probe radius of 1.4 Å. Dielectric constants of 80 and 4 were used for the exterior and the interior media, respectively. An ionic strength of 0.15 M was used together with an ion exclusion region extending 2.0 Å beyond the atomic radii.

Atomic radii and charges were taken from the PARSE parameter set.[53] The charges and radii used for the phenyl phosphate are those listed in Table 3.

The *N*-formyl *N*-methylamide derivatives of the ionizable amino acids were chosen as model compounds. The pK_{mod} values were taken from Nozaki and Tanford[54] (Asp 4.0, Glu 4.4, His 7.0, Cys 9.5, Tyr 9.6, Lys 10.4, Arg 12.0, C-terminus 3.8, N-terminus 7.5). All protein groups of the above types were included as titrating sites in the calculations. The phenyl phosphate was assumed to be fully deprotonated.

The ΔpK_{Born}, ΔpK_{back} and the charge−charge pairwise terms (W_{ij}) for all of the titrating sites were evaluated with the multiflex module of the MEAD program suite,[47,55] which uses finite-difference methods[56,57] to solve the linearized Poisson−Boltzmann equation. Protonation fractions were calculated for pH values ranging from −4 to 15, by increments of 0.1 pH unit, using the Monte Carlo method developed by Beroza et al.,[58] because a direct summation over all 2^N protonation states is not feasible. In cases where the protein's electrostatic environment strongly stabilizes either the protonated or deprotonated form of a site, the program will predict the site to be mostly deprotonated even at pH −4 or mostly protonated at pH 15. In these cases the results are reported as "<−4.0" or ">15," respectively.

The MEAD program suite and input data files needed to reproduce most of the results presented here are available from the web page, http://www.scripps.edu/bashford. The Beroza Monte Carlo program is available from ftp://ftp.scripps.edu/pub/case/beroza.

TABLE 4: Experimental pK_as[a]

	free enzyme			complexes
	Cys[b]	Cys[c]	Asp[c]	Asp[d]
PTP1[e]	5.57	5.44	4.9	6.7
Yersinia PTPase	4.67[f]		5.1[g]	5.2, 5.8[g,h]
BPTP[i]		<4[i]	5.3	
VHR	5.6[j]	5.4[j]	5.7[k]	7.2[j]

[a] CYS and ASP stand for the nucleophilic cysteine and the general acid, respectively. The numbering of these catalytic residues in the four proteins is given in Table 2. [b] Values obtained from inactivation by iodoacetate. [c] Values obtained from analysis of the pH dependence of the k_{cat}/K_M ratio. [d] pK_as extracted from the pH dependence of k_{cat}. [e] Values from PTP1 32, the rat structural homologue of PTP1B (human). [f] Reference 38. [g] Reference 36. [h] Values reflecting the pK_a of the general acid in the phosphoenzyme intermediate rather than the Michaelis complex (see text). [i] Reference 29. The actual pK_a value for the Cys cannot be measured because of the onset of acid-induced denaturation below pH 3.5. [j] Reference 35. [k] Reference 31.

Review of Experimental Data

Here we briefly review the experimental data pertaining to titration of the active-site residues in the PTPases, and tabulate them (Table 4) for comparison with calculations.

In the Unliganded Enzymes. The pK_a values extracted from the pH dependence of the k_{cat}/K_M ratio refer to the protonation equilibria of catalytically important titratable groups in the free enzyme and the free substrate.[59] For the four PTPases studied here, the log(k_{cat}/K_M) versus pH curves display an ascending slope of +2 in the acidic range (pH < 4−5), and a descending slope of +1 in the basic range (pH > 6).[29,31,32,36] This indicates that enzymatic activity requires two ionizable groups to be in their deprotonated form and one to be in its protonated form. The descending slopes in the basic domain can be unambiguously attributed to Asp181, Asp356, Asp129, and Asp92 in the unliganded forms of PTP1,[32] Yersinia PTPase,[36] BPTP,[29] and VHR,[31] respectively. The pK_as are listed in Table 4. The substrate (generally pNPP) and the nucleophilic cysteine are responsible for the ascending slope in the acidic domain. The cysteine pK_as obtained in this way are listed in Table 4. The cysteine pK_as obtained from the pH dependence of iodoacetate inactivation of the enzymes PTP1 and VHR are also reported[32,35] and are in reasonable agreement with the former estimates.

In the Michaelis Complexes. Because the pH dependence of k_{cat} may reflect the kinetic properties of several different states along the reaction path[59] (in our case, the Michaelis complex or the phosphoenzyme intermediate), the pK_as of the general acid obtained from these curves, and reported in the last column of Table 4, should be interpreted with caution. Burst kinetic experiments suggest that the rate-limiting step of the enzymatic reaction is the hydrolysis of the phosphoenzyme intermediate (i.e., the second step of the catalytic process) in the case of BPTP[60] and Yersinia PTPase,[33] and the formation of the phosphoenzyme (i.e., the first step of the catalytic process) for PTP1[32] and VHR.[31] Consequently, in the case of Yersinia PTPase, the intermediate should accumulate and the pK_a extracted from the pH dependence of k_{cat} should reflect the pK_a of the general acid in the phosphoenzyme intermediate rather than in the Michaelis complex. Conversely, in the cases of PTP1 and VHR, the concentration of the phosphoenzyme intermediate will be much smaller than the concentration of the Michaelis complex, and thus the experimental pK_a values of 6.7 and 7.2 (see Table 4) should be more representative of the pK_a of the general acid in the Michaelis complex. (The log k_{cat} versus pH data[61] do not appear to provide pK_a values for the nucleophilic cysteine in the Michaelis complex.)

Results and Discussion

We first examine the results of the calculations on the unliganded enzymes, PTP1B and Yersinia PTPase. These results are summarized in Table 5. The calculated pK_as for the nucleophilic cysteine and the general acid are analyzed in terms of desolvation (ΔpK_{Born}), dipoles (ΔpK_{back}) and charged side chain contributions. For the Michaelis complexes (PTP1B, Yersinia PTPase, BPTP, and VHR), the same kind of analysis is done and reported in Table 6, with an additional contribution due to the presence of the negatively charged phenyl phosphate. The ΔpK_{back} terms are further broken down into contributions due to the P-loop and the nearby α-helix backbone dipoles, the conserved Ser/Thr side chain of the PTPase signature motif, and the remaining dipoles of the protein (see Table 7). The individual contributions to ΔpK_{back} (based on eq 2) of the cysteine and the general acid in the Michaelis complex of BPTP are displayed in Figure 3. Finally, we compare the results obtained for the four PTPases in the Michaelis complex (PTP1B, Yersinia PTPase, BPTP, and VHR) to bring out the similarities and differences among them with respect to electrostatic interactions within the active site.

Protonation state in the unliganded enzyme: PTP1B and Yersinia PTPase. *pKa of the nucleophilic cysteine.* The calculated pK_as for the nucleophilic cysteine based on the structural models for the unliganded enzyme are shown in Table 5. These calculated values are substantially smaller (always < −4.0) than the experimental values of 5.6/5.4 and 4.7 for PTP1B and Yersinia PTPase, respectively.

Although the calculations are qualitatively correct in predicting that the cysteine residue is deprotonated in the pH range of the enzyme activity, the magnitude of the downward pK_a shift appears to be significantly overestimated. This is probably the result of using a single-conformational model, which does not allow for any acid-induced structural change in the active site coupled to the protonation of the cysteine residue. In general, extreme calculated pK_a values such as these should not be interpreted as literal predictions of pK_a measurements but as predictions that the protein cannot accommodate the change of protonation without undergoing some significant conformational rearrangement.[47,62] In this context, our calculations are consistent with the experimental finding that the cysteine titration in BPTP is obscured by acid denaturation below pH 4.[29]

Both charge−dipole and charge−charge interactions stabilize the deprotonated (negatively charged) form of the reactive cysteine. These influences are more than enough to offset the Born desolvation term, which destabilizes the charged form. The backbone dipoles of the P-loop make the largest contribution to the stabilization of the thiolate form, decreasing the pK_a of the cysteine by −8.3 to −10.1. The backbone dipoles of the first 14 residues of the conserved α-helix located at the C-terminus of the PTPase signature motif provide an additional contribution of about −2.5. Finally, the side-chain dipole of the conserved hydroxyl group (Ser222 in PTP1B and Thr410 in Yersinia PTPase) further stabilizes the thiolate form, lowering its pK_a by −2.9 to −3.6. This last observation is opposite to the result of Peters et al.,[8] namely that the Ser/Thr residue destabilizes the thiolate form. Kinetic studies on the Ser/Thr → Ala mutants, of PTP1B,[32] Yersinia PTPase,[33] BPTP,[29] and VHR,[31] all support the conclusion that the thiolate form is stabilized by the conserved hydroxyl group, in agreement with the present calculations.

Calculations were also done using structural models prepared with an energy minimization protocol that did not constrain any heavy atoms, but was otherwise the same as that described in

TABLE 5: Calculated Δp*K*s and pK_a for the General Acid and the Nucleophilic Cysteine in the Unliganded Enzymes[a]

	general acid			nucleophilic cysteine		
	PTP1B	Yersinia PTPase		PTP1B	Yersinia PTPase	
	2hnp	1ypt/A	1ypt/B	2hnp	1ypt/A	1ypt/B
ΔpK_{Born}	2.0	3.0	2.9	6.1	5.7	5.9
ΔpK_{back}	-0.6	-2.0	-2.3	-15.6	-14.8	-15.0
pK_{intr}	5.4	5.0	4.6	0.0	0.4	0.4
charged side chains	-0.8	0.8	0.6	< -4.0	< -4.4	< -4.4
pK_a	4.6	5.8	5.2	< -4.0	< -4.0	< -4.0

[a] pK_{intr} is obtained by adding the pK_a of the model compound to ΔpK_{Born} and ΔpK_{back}. pK_a − pK_{intr} gives the contribution of the charged side chains. Regarding calculated pK_a values < -4.0, see the description of the methods of pK_a calculation under Models and Theory.

TABLE 6: Calculated Δp*K*s and pK_a for the General Acid and the Nucleophilic Cysteine in the Michaelis Complexes[a]

	PTP1B				Yersinia PTPase			BPTP	VHR	
	1ptv	1ptt	1ptu	1pty	1yts	1ytw	1ytn	2pnt	1vhr/A	1vhr/B
					general acid					
ΔpK_{Born}	4.8	5.2	4.5	3.0	4.3	4.1	4.9	3.3	4.2	3.9
ΔpK_{back}[b]	−4.4	−2.8	−2.9	−4.3	−6.3	−6.5	−6.4	−9.0	−6.2	−5.8
pK_{intr}[b]	4.4	6.4	5.6	2.7	2.0	1.6	2.5	−1.7	2.0	2.1
phenyl phosphate[c]	10.4	8.9	7.9	8.6	9.4	8.5	11.6	11.2	6.7	7.7
charged side chains	>0.2	−9.9	−6.8	−4.4	−2.8	−2.4	−2.1	−0.4	−1.2	−1.3
pK_a	>15.0	5.4	6.7	6.9	8.6	7.7	12.0	9.1	7.5	8.5
					nucleophilic cysteine					
ΔpK_{Born}	10.6	10.4	10.6	10.6	10.7	10.4	10.6	10.0	10.1	10.2
ΔpK_{back}[b]	−31.3	−27.2	−29.4	−29.4	−29.5	−30.2	−29.0	−31.5	−27.2	−26.7
pK_{intr}[b]	−11.2	−7.3	−9.3	−9.3	−9.3	−10.3	−8.9	−12.0	−7.6	−7.0
phenyl phosphate[c]	25.0	24.6	25.1	24.4	25.7	23.8	25.3	21.1	21.1	20.8
charged side chains	<−17.8	−20.4	<−19.8	<−19.1	<−20.4	<−17.5	<−20.4	−11.5	−13.9	−13.9
pK_a	<−4.0	−3.1	<−4.0	<−4.0	<−4.0	<−4.0	<−4.0	−2.4	−0.4	−0.1

[a] pK_{intr} is obtained by adding the pK_a of the model compound to ΔpK_{Born} and ΔpK_{back}. Regarding calculated pK_a values < -4.0 or >15, see the description of the methods of pK_a calculation under Models and Theory. The charged side chains contribution is obtained by subtracting pK_{intr} and the contribution of the phenyl phosphate from the calculated pK_a. [b] Include only contributions due to the dipoles of the protein. Contribution due to the substrate charge is not included. [c] Contribution due to the interaction with the whole charge of the substrate (−2).

TABLE 7: Calculated ΔpK_{Back} for the General Acid and the Nucleophilic Cysteine in the Michaelis Complexes

	PTP1B				Yersinia PTPase			BPTP	VHR	
	1ptv	1ptt	1ptu	1pty	1yts	1ytw	1ytn	2pnt	1vhr/A	1vhr/B
					general acid					
P-loop[a]	−2.8	−2.4	−2.3	−2.3	−2.7	−2.3	−1.5	−2.5	−1.9	−2.2
α-helix[b]	−0.6	−0.5	−0.5	−0.5	−0.6	−0.5	−2.3	−0.6	−0.7	−0.7
remaining	−1.0	0.1	−0.1	−1.5	−3.0	−3.7	−2.6	−5.9	−3.6	−2.9
ΔpK_{back}	−4.4	−2.8	−2.9	−4.3	−6.3	−6.5	−6.4	−9.0	−6.2	−5.8
					nucleophilic cysteine					
P-loop[a]	−22.0	−19.2	−20.6	−21.2	−21.4	−21.6	−21.3	−19.7	−21.3	−20.6
α-heix[b]	−3.8	−3.4	−3.7	−3.8	−3.6	−3.5	−3.5	−3.0	−3.0	−3.0
SER/THR[c]	−3.5	−2.7	−3.6	−3.6	−3.4	−3.0	−3.2	−3.5	−3.1	−3.7
remaining	−2.0	−1.9	−1.5	−0.8	−1.1	−2.1	−1.0	−5.3	0.2	0.6
ΔpK_{back}	−31.3	−27.2	−29.4	−29.4	−29.5	−30.2	−29.0	−31.5	−27.2	−26.7

[a] Contribution due to the interaction with the backbone dipoles of the P-loop (residues 215 to 222 in PTP1B). [b] Contribution due to the backbone dipoles of the α-helix at the C-terminus of the active-site (residues 223 to 236 in PTP1B). [c] Contribution due to the side chain of the last residue of the active-site signature motif (Ser222 in PTP1B).

Structural Models. The all-atom rms displacement due to the minimization was approximately 0.7 Å. The calculated pK_a values for the nucleophilic cysteine of PTP1B and the two Yersina structures, are all < -4.0 (results not tabulated), essentially the same result as obtained with the more conservative minimization protocol.

pK$_a$ of the General Acid. The results for the general acid using the three model structures for the unliganded enzyme are also reported in Table 5. The pK_as computed for the unliganded enzymes (4.6 and 5.8/5.2, for the PTP1B structure and the A/B structures of Yersinia PTPase, respectively) are in good agreement with the experimental values deduced from the pH dependence of the k_{cat}/K_M ratio (4.9[32] and 5.1,[36] for PTP1B and Yersinia PTPase, respectively). For the three structures the Born term (ΔpK_{Born}) tends to stabilize the protonated form of the residue, whereas the contribution due to the interactions with the dipoles of the protein (ΔpK_{back}) tends to destabilize it.

The major contribution to ΔpK_{back} comes from residues Ser287 to Tyr301 in Yersinia PTPase and from the corresponding region in PTP1B (residues Arg112 to Tyr124, according to the structure-based sequence alignment reported by Saper et al.[27]). This contribution is larger for Yersinia PTPase (−1.7 and −2.0 for structure A and B, respectively) than for PTP1B (−0.6). In the case of Yersinia PTPase, Ser287 alone contributes approximately −0.9 (−0.7 for the side chain alone). In PTP1B, the corresponding residue only contributes −0.1, and the largest individual contribution is due to the backbone dipole of Lys291 (−0.2).

The interactions with charged side chains slightly favor the carboxylate form in PTP1B, whereas they destabilize it in

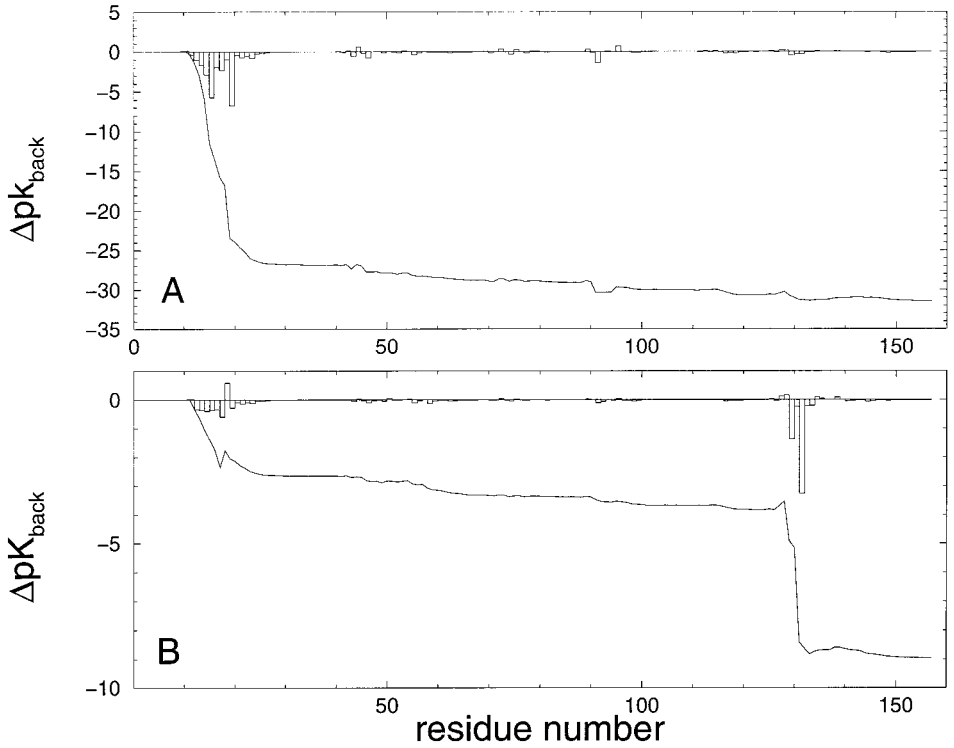

Figure 3. Contributions of individual residues to ΔpK_{back} of Cys12 (A) and Asp129 (B) in BPTP. The solid line refers to the running sum of ΔpK_{back} over the residues.

Yersinia PTPase. This difference between the two enzymes is mainly due to the difference in charge−charge interaction of the general acid with Glu115 in PTP1B (pK_a shift = +0.5), and the corresponding residue in Yersinia PTPase, Glu290 (pK_a shift = +2).

The magnitude of both the desolvation term (ΔpK_{Born}), and the electrostatic interactions (charge−dipole and charge−charge), are smaller in PTP1B than in Yersinia PTPase. This might have been expected since the relative solvent accessibility of the side chain of the general acid is higher in PTP1B (52%) than in Yersinia PTPase (32% and 34%).

The calculations that used the energy minimization protocol that allowed all heavy atoms to move produced similar Born, background and interaction terms for the general acid. The calculated pK_a value was lower by 0.1 for PTP1B, and the values for the A and B structures of Yersinia PTPase were lower by 0.6 and 0.4, respectively. Overall, the choice between the more conservative or more aggressive energy-minimization protocol makes little difference in terms of agreement with experiment, and the variation in calculated pK_a is no greater than the variation associated with choosing molecule A or B from the 1ypt structure file.

Protonation States in the Michaelis Complexes: PTP1B, Yersinia PTPase, BPTP, and VHR. *pK_a of the Nucleophilic Cysteine.* The results for the nucleophilic cysteine using the 10 structural models for the Michaelis complex are reported in Tables 6 and 7. Despite the fact that in all calculations the phenyl phosphate is assumed to be unprotonated and to contribute a charge of −2 (offset, of course, by the +1 charge of the arginine), the calculated pK_a of the nucleophilic cysteine in the complex is always lower than 0.0 (Table 6). Thus, the cysteine is always predicted to be in its thiolate form, as it should be for optimal catalysis. The ΔpK_{back} term is the major contribution (ranging from −26.7 to −31.5) to the stabilization of the thiolate form of the cysteine. These ΔpK_{back} values are significantly more negative than the corresponding terms for the unliganded enzyme. For the two cases where both unliganded enzyme and Michaelis complex structures were available (PTP1B and Yersinia), we carried out calculations on "pseudo-unliganded" structures that were prepared by simply deleting the phenyl phosphate from the active site of the Michaelis complex structures. The resulting ΔpK_{back} values (not shown) are closer to the unliganded enzyme calculations of Table 5 than to the Michaelis complex calculations of Table 6, indicating that the increased magnitude of the ΔpK_{back} terms in the latter case is mostly the result of the displacement by the substrate of high-dielectric solvent that would otherwise screen the background interactions. The interaction due to the charged side chains also favors the deprotonated form, lowering the pK_a of the cysteine by at least 11.5. These two effects more than compensate for the destabilizing effect of the negatively charged phenyl phosphate (between 20.8 and 25.7) and desolvation (10.3 ± 0.3).

As shown in Table 7, the interaction with the P-loop's backbone dipoles makes the largest contribution to ΔpK_{back} (about −19 to −22). The side chain of the conserved Ser/Thr residue of the PTPase signature motif also stabilizes the thiolate form of the cysteine (by about −3) and the backbone dipoles of the nearby α-helix give a contribution of about −3 to −4. The remaining contribution to ΔpK_{back} is substantial in BPTP (−5.3). An analysis using eq 2 for the contributions of individual residues to the ΔpK_{back} of Cys12 in BPTP shows that this remaining contribution of −5.3 is distributed over a large number of residues outside of the active-site sequence (Figure 3A). The largest individual contribution (−0.7) is due to the interaction with Asn95. For the other PTPases, the contribution from residues outside the P-loop and the α-helix is smaller (between 0.6 and −2.1) and distributed over a more restricted set of residues.

As shown in Table 6, the charge−charge interactions with the phenyl phosphate shifts the pK_a of the cysteines up by terms ranging from 20.8 to 25.7, whereas the contribution of the charged side chains is always negative (≤ −11.5), i.e.,

TABLE 8: Calculated p*K* Shifts of the Nucleophilic Cysteine in the Michaelis Complex due to Positively Charged Side Chains[a]

	PTP1B				Yersinia PTPase			BPTP	VHR	
	1ptv	1ptt	1ptu	1pty	1yts	1ytw	1ytn	2pnt	1vhr/A	1vhr/B
Arg45[b]	−1.7	−1.8	−1.6	−1.7	−1.9	−1.3	−1.1			
His270[c]					−1.5	−1.5	−1.5			
Lys120[b]	−0.5	−1.9	−1.2	−0.6						
His173[b]	−1.3	−1.3	−1.2	−1.3						
His175[b]	−1.2	−1.3	−1.2	−1.2	−1.4	−1.6	−1.5			
His214[b]	−3.8	−3.9	−3.7	−3.7	−3.7	−3.6	−3.7		−3.3	−3.4
Arg404[c]					−1.5	−1.5	−1.6		−0.7	−0.8
Arg221[b]	−9.3	−10.6	−8.7	−9.1	−9.4	−10.9	−10.0	−8.0	−8.3	−8.0
Arg257[b]	−4.0	−3.4	−3.9	−3.9	−3.8	−3.5	−3.5		−3.0	−3.0
His72[d]								−1.8		

[a] Only the significant p*K* shifts (lower than −1) are reported. [b] The residue number refers to PTP1B. The number of the corresponding residues for the others PTPases can be found in Table 2. [c] This residue number refers to Yersinia PTPase. [d] This residue number refers to BPTP.

dominated by the contributions due to positively charged groups. The largest contributions caused by neighboring positive residues are gathered in Table 8. The invariant arginine of the PTPase signature motif plays the major role, inducing a p*K* shift between −8.0 and −10.9. Two other positive residues that are conserved in PTP1B, Yersinia PTPase, and VHR (but not in BPTP) lower the pK_a of the attacking cysteine, by about 3.5 each. These residues are His214 and Arg257 in the PTP1B numbering. All of the other charge−charge terms appearing in Table 8 are smaller than 2.0 and are less consistent among the different PTPases.

For the 1ptv structure of PTP1B, calculations were also done using a minimization protocol that allowed all heavy atoms to move, but was otherwise the same as that described in Structural Models. The all-atom rms displacement due to the minimization was 0.76 Å. The result for the nucleophilic cysteine is the same as that obtained with the more conservative protocol: pK_a < −4.0 (results not tabulated).

pK_a of the General Acid. All the calculated pK_as for the general acid in the Michaelis complexes are above 5.4 (Table 6). The interaction with the negatively charged phenyl phosphate and the desolvation term tend to favor the protonated form of the general acid and more than counteract the destabilizing effect due to the charged side chains and the protein dipoles. In most cases, the magnitude of these individual terms is greater than the overall p*K* shift.

In the case of PTP1B, the pK_a values of the general acid, Asp181, calculated using three of the structural models, 1ptt, 1ptu, and 1pty, are 5.4, 6.7, and 6.9, respectively. The experimental pK_a, based on the pH dependence of k_{cat}, is 6.7,[32] which is very close to the 1ptu and 1pty calculations. The variations among the results for these three models is caused primarily by local conformational variations that change the interaction of Asp181 with the nearby charged side chain of Lys120. The p*K* shift due to this interaction is −8.7, −5.5, and −1.8 in 1ptt, 1ptu, and 1pty, respectively. In the calculated results based on a fourth model of PTP1B, 1ptv, the Born term and the interaction with the substrate push the titration of Asp181 into such a high pH range that the conserved arginine residue 221, which would normally help stabilize the negative charge of Asp181, becomes deprotonated. This leads to the erroneous prediction that Asp181 has a pK_a above 15. The calculations based on the 1ptv model in which a more aggressive energy-minimization protocol allowed heavy atoms to move give the same result. This illustrates the potential for error in calculations involving large, pH dependent, counteracting terms, and the need for alternative structural models and care in interpreting the results of calculations.

The calculated pK_as for the general acid in VHR (7.5/8.5 for A/B structures, respectively) are reasonably close to the experimental value of 7.2 extracted from the pH dependence of k_{cat} for this enzyme.[35]

In the case of Yersinia PTPase, the pK_a calculated for the general acid based on the 1ytn structure is much larger (12.0) than the value obtained with the two other structures 1yts (8.6) and 1ytw (7.7). As is evident from Table 6, this difference comes essentially from the charge−charge interaction with the substrate. The PDB structure 1ytn.pdb is considered as a transition-state-like structure.[17] In the corresponding structural model (1ytn) Asp356 is closer to the substrate than in 1yts and 1ytw, explaining the enhanced influence of the phenyl phosphate. More specifically, the C^{γ}−Asp356 P-substrate distance is 4.1 Å in the 1ytn structure and 4.4 Å for both the 1yts and the 1ytw structures. Consequently, the modeled structure 1ytn is certainly a poorer model for the Michaelis complex, and the results calculated using this model should not be regarded as representative of the complex. For the two other structural models, the calculated pK_as (8.6 and 7.7) are much higher than the experimental values reported in Table 4 (5.2 and 5.8[36]). However, as discussed previously, these experimental values are probably more representative of the phosphoenzyme intermediate than of the Michaelis complex.

Upon substrate binding, the solvent accessibility of the general acid is expected to decrease, destabilizing the ionized carboxylate form. Since the negatively charged substrate also destabilizes this ionized form, the pK_a of the general acid can be expected to increase upon substrate binding. Consequently, in the case of Yersinia PTPase and BPTP (for which no experimental values are available for comparison with our results), the experimental value obtained for the general acid in the unliganded enzyme can be regarded as a lower limit for the pK_a of this same residue in the Michaelis complex. The calculated pK_as obtained for the general acid in the Michaelis complexes of Yersinia PTPase and BPTP effectively satisfy this expectation.

Similarities and Differences among PTPases. The detailed breakdown of electrostatic terms contributing to p*K* shifts that is provided by the present calculations allows for a comparative study of the means by which the members of this diverse family of enzymes solve the problem of maintaining the nucleophilic cysteine in a deprotonated state, even in the presence of the −2 charge of the substrate, while maintaining the general acid in the protonated state. Of particular interest are the relative importance of conserved features, such as the P-loop and the proximity of the N-terminal end of an α-helix to the active site, versus nonconserved features of the proteins. In the following, we focus on results obtained for the Michaelis complexes, and compare the electrostatic properties of the active site in PTP1B, Yersinia PTPase, BPTP, and VHR. The results obtained with the modeled structures, 1ptv, for PTP1B, and 1ytn, for Yersinia

PTPase, are not taken into account in this comparison (see previous discussion).

Electrostatic Environment of the Nucleophilic Cysteine. The contributions from the conserved Ser/Thr side chain, the P-loop and the α-helix backbone dipoles to the ΔpK_{back} term for the nucleophilic cysteine do not vary significantly among the PTPases (see Table 7). The dipoles of the residues outside the active-site seem to play a more important role in the stabilization of the reactive thiolate in BPTP (−5.3) than for the other PTPases (between 0.6 and −2.1). Surprisingly, in the case of BPTP, the dominant part (−4.7) of this contribution is distributed over the last 120 residues (see Figure 3A). The interactions with charged side chains vary significantly in magnitude but always favor the deprotonated form of the cysteine (Table 6). The effect is larger for PTP1B and Yersinia PTPase (< −17.5) than for VHR (−13.9) and BPTP (−11.5). The positively charged residues providing the largest contributions are listed in Table 8. The electrostatic interaction of the nucleophilic cysteine with the arginine of the PTPase signature motif is the only one that is conserved among the four PTPases considered, and it is larger for PTP1B and Yersinia PTPase (where it decreases the pK_a of the cysteine by 9 to 11) than for VHR and BPTP (where it decreases the pK_a of the cysteine by approximately 8). Two other positive residues conserved in PTP1B, Yersinia PTPase, and VHR (but not in BPTP) also substantially favor the thiolate form (by 3 to 4 pK units): an arginine located at the N-terminus (Arg257, Arg440, and Arg158 in PTP1B, Yersinia PTPase and VHR, respectively) and the histidine preceding the nucleophilic cysteine in the protein sequence (His214, His402, and His123 in PTP1B, Yersinia PTPase, and VHR, respectively).

It appears that the extra ΔpK_{back} from nonconserved regions in BPTP makes up for the relative lack of charge−charge interactions to stabilize the thiolate form of the nucleophile. In all of the PTPases, most of the thiolate stabilization comes from conserved regions, but nonconserved regions also make significant contributions.

Electrostatic Environment of the General Acid. As reported above, the principal terms tending to maintain the general acid in the catalytically necessary protonated form are the ΔpK_{Born} term (desolvation) and the negative charge of the substrate. On the other hand, many of the same factors that maintain the deprotonation of the active-site cysteine residue also tend to favor deprotonation of the general acid. The contribution to the ΔpK_{back} term for the general acid coming from the P-loop and the α-helix are -2.3 ± 0.4 and −0.6, respectively (Table 7), whereas the conserved Ser/Thr side chain's contribution is insignificant (close to −0.1 in all cases). However, these effects are much weaker on the general acid than on the cysteine residue in all of the PTPases. Residues outside these regions also contribute negative terms to ΔpK_{back}, but the magnitude varies among different PTPases. As was the case with contributions to the cysteine residue pK_a value, BPTP is unique in having a large contribution to the ΔpK_{back} of the general acid from outside the conserved regions (−5.9), and this contribution is spread over a large number of residues (Figure 3B).

In all of the PTPases, the overall effect of interactions with charged side chains is to lower the pK_a of the general acid; this effect is significant for PTP1B, but less so in the other PTPases (Table 6). In all of the PTPases, the conserved arginine residue shifts the general acid pK_a by −3 to −4, and the negative charge of the cysteine nucleophile shifts the general acid pK_a by approximately +2. Contributions from other residues vary considerably among the PTPases, with more similarity observed between PTP1B and Yersinia PTPase than between others.

Concluding Remarks

In this study we computed the pK_a values of the nucleophilic cysteine residue and the general acid for both the unliganded form and the Michaelis complexes of a variety of PTPases in order to better understand the electrostatic properties of the active sites of these enzymes and thus their catalytic activity. The calculated pK_a values for the general acid are mostly in reasonable agreement with available experimental data for both the free enzymes and the Michaelis complexes. The calculated pK_a values for the nucleophilic cysteine residue in the free enzyme are consistent with a thiolate form of this residue in the neutral to mildly acidic range. The extremely low calculated pK_a values amount to a prediction that the protonation of this cysteine residue cannot occur without a significant structural change (see below). For one of the enzymes studied, BPTP, it is known experimentally that this cysteine titration is obscured by acid denaturation, which is consistent with the calculations. For the others, it is not known whether this protonation is accompanied by a conformational change. Experimental estimates of the cysteine pK_a value for the Michaelis complex are not available for comparison, but the calculations predict that the thiolate form, which the catalytic mechanism requires, is preserved. Thus, the predicted protonation states of the groups directly involved in the first step of catalysis are correct for both the free enzymes and the Michaelis complexes. The stabilization of the thiolate form in the free enzyme implies an unusual pK_a shift of the cysteine residue, and the maintenance of this negative charge within an active-site cleft even upon the approach of the −2 charge of the substrate is remarkable.

There is a potential conflict between stabilizing negative charge and keeping the general acid protonated, and this may be the reason that these enzymes primarily use dipole interactions for the former. Specifically, the active site must provide enough positive electrostatic potential near the cysteine residue to allow the −2 charge of the substrate to dock without causing the protonation of the cysteine, but this positive potential might tend to deprotonate the nearby general acid. Dipole interactions have an advantage over monopole interactions in this situation, because their short-range, directional nature allows the resulting potentials to be more narrowly focused.

For the calculations reported above, we have chosen the protein interior dielectric constant, ϵ_p, to be 4.0. Values near 4.0 are supported by powder measurements[63,64] and molecular dynamics studies.[65−67] This choice implies that the protein dielectric response is confined to electronic polarization and relatively small fluctuations of the protein dipoles in a linear response to electrostatic perturbations. Accounting for significant reorganizations of protein dipoles in response to charge changes is beyond the scope of calculations based on a single conformer and a low ϵ_p value. In this context, the correct interpretation of extreme calculated pK_a values (such as negative values) is that if the protein were held near its modeled conformation then titration would not be able to occur in a normal pH range. However, titration could occur if the protein were able to undergo conformational change. In other words, extreme calculated pK_a values should be interpreted as a prediction that titration entails some conformational change, and that a single-conformer, linear-dielectric-response model that lacks any detailed model of this change cannot be expected to provide a quantitative prediction of the pK_a. The conformational change could be a small one, such as a local side chain movement, or

a large one, such as unfolding or a significant change of active-site geometry.

Calculations using multiple sources of structural data, as in the present work, are helpful in distinguishing these cases. For example, the general acid of the Michaelis complex of PTP1B is predicted to titrate in the neutral to mildly acidic range for three of the four structural models, but for one of them, 1ptv, an extreme value is calculated. Structural comparison can then be used to trace the pK_a difference to fairly minor structural differences between 1ptv and the other models. On the other hand, extremely low pK_a values are calculated for the nucleophilic cysteine side chain in all of the structural models of PTP1B, and indeed, over the entire PTPase family, suggesting that the conformational changes linked to the protonation of this group are quite significant. This is consistent with the experimental observation that in the case of PTPT, the cysteine titration is obscured by acid denaturation.

Some workers in the protein electrostatics field have advocated the use of a higher values of ϵ_p, such as 20.0, based on empirical fitting of calculated pK_a values to experiment.[68] In particular, Garcia-Moreno et al.[69] have fit the ϵ_p values of a Born-like model that has some similarity to the MEAD model to measured pK_a shifts of deeply buried residues and obtained ϵ_p values of 10 or more. As these workers acknowledge, the discrepancy between this high value, and the low value of ϵ_p obtained from powder measurements and molecular dynamics simulations probably reflects local conformational changes or the penetration of water into the protein interior upon creation of a buried charge. Thus, the choice of a high value of ϵ_p implies an attempt to account for significant conformational changes through the dielectric constant. In our view, this approach has some difficulties in principle and some associated hazards. It seems unlikely that the local conformational rearrangements that one is attempting to account for by a high ϵ_p value are really in the linear-response regime. The specific details of the rearrangement, which are neglected (typically because they are unknown) might well be determinants of the energetics of titration. A particular hazard of the approach is that it might be quite incorrect in the case where a protein structure has evolved to accommodate charges in the interior, and thus does not suffer conformational changes due to the creation of such charges. Similarly, the high-ϵ_p approach could be expected to fail in cases where evolution has endowed active sites with an unusual degree of structural rigidity. Thus, while the high-ϵ_p model reduces average deviations between calculations and experiment, it could cause one to miss important features of protein active sites. On the other hand, the low-ϵ_p approach will occasionally produce large quantitative errors, but these errors may have an informative interpretation.

To investigate the difference between high- and low-ϵ_p models of the PTPases, we have repeated our calculations using $\epsilon_p = 20$. The results are summarized here but not tabulated in detail. In the unliganded enzymes, the calculated pK_a of the nucleophilic cysteine ranges from 1.7 to 2.7, which is closer to the experimental range than we achieve with a low-ϵ_p model. However, the active-site aspartic acid pK_a values are calculated to be approximately 3.0 in the free enzymes, which is both qualitatively incorrect and implies a protonation state inconsistent with this conserved residue's role as a general acid. When the $\epsilon_p = 20$ model is applied to the Michaelis complexes, the calculated pK_a values for the active-site cysteine and the aspartic acid residues are in the ranges 7−9 and 2−4, respectively. This is completely inconsistent with the former's role as a nucleophile and the latter's role as a general acid in the neutral to mildly acidic pH range. It also shows that the $\epsilon_p = 20$ model fails to account for the extraordinary stabilization of negative charge that the PTPase active site provides. The finding that a low-ϵ_p model gives a better overall account of the electrostatic factors important for the phosphorylation step of catalysis than a high-ϵ_p model implies that the PTPase active site is quite resistant to conformational change, even as the amount of charge in the active site changes by −2. This is consistent with the crystallographic and NMR evidence that the P-loop is quite rigid[22,23,26] and with the strict conservation of its three-dimensional structure across the PTPase family.

Previously, Peters et al.[8] presented MEAD-type calculations for the unliganded forms of the same set of four PTPases studied here, but they did not present calculations of the general acid pK_a values, nor did they study the Michaelis complexes. They used an ϵ_p value of 20 and found good agreement with experimental measurements of the pK_a of the active-site cysteine of the free enzyme. This finding is qualitatively consistent with our finding that raising ϵ_p value tends to move the calculated pK_a values closer to experiment, but since the charge model and the methods of preparing the coordinate sets differ, quantitative agreement between our calculations and theirs is not to be expected. In particular, Peters et al. employed a highly simplified model of titrating sites that treats protonation/deprotonation as the alteration of a single unit point charge. This eliminates certain dipole terms that may be significant when titrating groups are close together or in close proximity to other polar groups, such as the backbone amide groups. Peters et al. predict that mutation of the conserved serine of the P-loop to alanine should stabilize the thiolate form of the active-site cysteine residue; however, the experimental findings are the opposite.[29,31−33] Our findings in the present work are in agreement with experiment on this point (see Results and Discussion). We suspect that this difference between the calculated results is related to the hydrogen-atom-coordinate building procedures used. Our procedure results in the hydroxyl hydrogen atom being oriented toward the cysteine sulfur atom, so that the dipole interaction stabilizes the thiolate form. In a subsequent study by Peters et al.,[9] molecular dynamics trajectories of PTP1B with ligand bound were generated and pK_a values of the catalytic cysteine residue were calculated at a number of points along the trajectory. In contrast to the present results, they found that using $\epsilon_p = 20$ resulted in predictions of a thiolate form at neutral pH. Results for the general acid were not reported, nor did the study include other members of the PTPase family. The differences between these and the present results is probably related to the use of dynamics and to differences in methodologies: Peters et al. again used the single point charge model, and also used a complex scheme that assigns different dielectric constants to the protein at different points in the calculation.[70]

A recent theoretical study[10] of the initial reaction pathway of PTPase was based in part on a somewhat different picture of the protonation states than that used here. It was supposed that the phosphate group is protonated in the Michaelis complex, and that it either binds as a monoanion if the cysteine thiol group is initially deprotonated, or if the thiol is initially protonated, the phosphate group binds as a dianion and the thiol proton is transferred to it.[6] In either of these cases, there is one more proton in the reacting system than in the standard picture. However, this model is not consistent with the structural data and the available experimental data on the pH dependence of the enzyme kinetics, which are reviewed here in the introductory section.

The electrostatic interactions of the nucleophilic cysteine and of the general acid with active-site residues, including charge−dipole interactions with the P-loop backbone, the α-helix backbone, and the conserved Ser/Thr residue, and charge−charge interactions with the substrate and the invariant arginine of the PTPase signature motif, all show quantitative similarity among the PTPases considered. Quantitative differences in interactions are seen mainly with residues outside the active site. Several positively charged residues are involved in the lowering of the cysteine pK_a in PTP1B and Yersinia PTPase, a lesser number in VHR, and fewer still in BPTP. Similarly, the number of charged residues influencing the protonation state of the general acid seems to decrease in the order PTP1B, Yersinia PTPase, VHR, and BPTP. On the other hand, the effect of charge−dipole interactions with residues outside the active site is larger in BPTP than in the three other PTPases. In general, PTP1B and Yersinia PTPase appear to share very similar electrostatic active-site properties, and BPTP seems to differ more significantly from this pair than does VHR. Interestingly, a similar relationship among these PTPases is found in the phylogenetic tree generated by Goldstein[71] using sequence alignment techniques.

Acknowledgment. The authors thank Dr. P. Hünenberger for his critical reading of the manuscript, and Drs. Cynthia Stauffacher and Shuishu Wang for their help in obtaining Figure 2. This work was supported by the National Institutes of Health (GM45607 and GM27003).

References and Notes

(1) Fischer, E. H.; Charbonneau, H.; Tonks, N. K. *Science* **1991**, *253*, 401−406.
(2) Bishop, J. M. *Cell* **1991**, *64*, 235−248.
(3) Hunter, T. *Cell* **1995**, *80*, 225−236.
(4) Barford, D.; Jia, Z.; Tonks, N. K. *Nature Struct. Biol.* **1998**, *2*, 1043−1053.
(5) Tishmack, P. A.; Bashford, D.; Harms, E.; Van Etten, R. L. *Biochemistry* **1997**, *36*, 11984−11994.
(6) Hansson, T.; Nordlund, P.; Aqvist, J. *J. Mol. Biol.* **1997**, *265*, 118−127.
(7) Alhambra, C.; Wu, L.; Zhang, Z.-Y.; Gao, J. *J. Am. Chem. Soc.* **1998**, *120*, 3858−3866.
(8) Peters, G. H.; Frimurer, T. M.; Olsen, O. H. *Biochemistry* **1998**, *37*, 5383−5393.
(9) Peters, G. H.; Frimurer, T. M.; Andersen, J.; Olsen, O. H. *Biophys. J.* **1999**, *77*, 505−515.
(10) Kolmodin, K.; Nordlund, P.; Åqvist, J. *Proteins: Structure, Function and Genetics* **1999**, *36*, 370−379.
(11) Zhang, Z.-Y. *Crit. Rev. Biochemistry Mol. Biol.* **1998**, *33*, 1−52.
(12) Zhang, M.; Stauffacher, C. V.; Van Etten, R. L. *Adv. Prot. Phosphatases* **1995**, *9*, 1−23.
(13) Barford, D.; Flint, A. J.; Tonks, N. K. *Science* **1994**, *263*, 1397−1404.
(14) Jia, Z.; Barford, D.; Flint, A. J.; Tonks, N. K. *Science* **1995**, *268*, 1754−1758.
(15) Stuckey, J. A.; Schubert, H. L.; Fauman, E. B.; Zhang, Z.-Y.; Dixon, J. E.; Saper, M. A. *Nature* **1994**, *370*, 571−575.
(16) Schubert, H. L.; Fauman, E. B.; Stuckey, J. A.; Dixon, J. E.; Saper, M. A. *Protein Sci.* **1995**, *4*, 1904−1913.
(17) Fauman, E. B.; Yuvanimaya, C.; Schubert, H. L.; Stuckey, J. A.; Saper, M. A. *J. Biol. Chem.* **1996**, *271*, 18780−18788.
(18) Puius, Y. A.; Zhao, Y.; Sullivan, M.; Lawrence, D. S.; Almo, S. C.; Zhang, Z.-Y. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 13420−13425.
(19) Stein, E.; Lane, A. A.; Ceretti, D. P.; Schoecklmann, H. O.; Schroff, A. D.; Van Etten, R. L.; Daniel, T. O. *Genes Dev.* **1998**, *12*, 667−678.
(20) Huang, L.; Sankar, S.; Lin, C.; Kontos, C. D.; Schroff, A. D.; Cha, E. H.; Feng, S. M.; Li, S. F.; Yu, Z.; Van Etten, R. L.; Blanar, M. A.; Peters, K. G. *J. Biol. Chem.* **1999**, *274*, 38183−38188.
(21) Grangeasse, C.; Doublet, P.; Vincent, C.; Vaganay, E.; Riberty, M.; Duclos, B.; Cozzone, A. J. *J. Mol. Biol.* **1998**, *278*, 339−347.
(22) Zhang, M.; Van Etten, R. L.; Stauffacher, C. V. *Biochemistry* **1994**, *33*, 11097−11105.
(23) Zhang, M.; Zhou, M.; Van Etten, R. L.; Stauffacher, C. V. *Biochemistry* **1997**, *36*, 15−23.
(24) Zhang, M.; Stauffacher, C. V.; Lin, D.; Van Etten, R. L. *J. Biol. Chem.* **1998**, *273*, 21714−21720.
(25) Wang, S.; Tabernero, L.; Zhang, M.; Harms, E.; Van Etten, R. L.; Stauffacher, C. V. *Biochemistry* **2000**, *39*, 1903−1914.
(26) Logan, T. M.; Zhou, M.-M.; Nettesheim, D. G.; Meadows, R. P.; Van Etten, R. L.; Fesik, S. W. *Biochemistry* **1994**, *33*, 11087−11096.
(27) Yuvaniyama, J.; Denu, J. M.; Dixon, J. E.; Saper, M. A. *Science* **1996**, *272*, 1328−1331.
(28) Wang, S.; Stauffacher, C. V.; Van Etten, R. L. *Biochemistry* **2000**, *39*, 1234−1242.
(29) Evans, B.; Tishmack, P. A.; Pokalsky, C.; Zhang, M.; Van Etten, R. L. *Biochemistry* **1996**, *35*, 13609−13617.
(30) Davis, J. P.; Zhou, M. M.; Van Etten, R. L. *J. Biol. Chem.* **1994**, *269*, 8734−8740.
(31) Denu, J. M.; Dixon, J. E. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 5910−5914.
(32) Lohse, D. L.; Denu, J. M.; Santoro, N.; Dixon, J. E. *Biochemistry* **1997**, *36*, 4568−4575.
(33) Zhang, Z.-Y.; Palfrey, B. A.; Wu, L.; Zhao, Y. *Biochemistry* **1995**, *34*, 16389−16396.
(34) Juszczak, L. J.; Zhang, Z.-Y.; Wu, L.; Gottfried, D. S.; Eads, D. D. *Biochemistry* **1997**, *36*, 2227−2236.
(35) Denu, J. M.; Zhou, G.; Guo, Y.; Dixon, J. E. *Biochemistry* **1995**, *34*, 3396−3403.
(36) Zhang, Z.-Y.; Malachowski, W. P.; Van Etten, R. L.; Dixon, J. E. *J. Biol. Chem.* **1994**, *269*, 8140−8145.
(37) Wu, L.; Zhang, Z.-Y. *Biochemistry* **1996**, *35*, 5426−5434.
(38) Zhang, Z.-Y.; Dixon, J. E. *Biochemistry* **1993**, *32*, 9340−9345.
(39) Hengge, A. C.; Sowa, G. A.; Wu, L.; Zhang, Z.-Y. *Biochemistry* **1995**, *34*, 13982−13987.
(40) Hengge, A. C.; Denu, J. M.; Dixon, J. E. *Biochemistry* **1996**, *35*, 7084−7092.
(41) Hengge, A. C.; Zhao, Y.; Wu, L.; Zhang, Z.-Y. *Biochemistry* **1997**, *36*, 7928−7936.
(42) Brünger, A. T.; Karplus, M. *Proteins Struct. Funct. Genet.* **1988**, *4*, 148−156.
(43) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(44) MacKerrell, 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.
(45) Dauber-Osguthorpe, P.; Roberts, V. A.; Osguthorpe, D. J.; Wolff, J.; Genest, M.; Hagler, A. T. *Protein Sci.* **1988**, *4*, 31−47.
(46) Connolly, M. L. *Science* **1983**, *221*, 709−713.
(47) Bashford, D.; Gerwert, K. *J. Mol. Biol.* **1992**, *224*, 473−486.
(48) Dillet, V.; Dyson, H. J.; Bashford, D. *Biochem* **1998**, *37*, 10298−10306.
(49) Tanford, C.; Kirkwood, J. G. *J. Am. Chem. Soc.* **1957**, *79*, 5333−5339.
(50) You, T. J.; Bashford, D. *Biophys. J.* **1995**, *69*, 1721−1733.
(51) Nicholls, A.; Honig, B. *J. Comput. Chem.* **1991**, *12*, 435−445.
(52) Richards, F. M. *Annu. Rev. Biophys. Bioeng.* **1977**, *6*, 151−176.
(53) Sitkoff, D.; Sharp, K. A.; Honig, B. *J. Phys. Chem.* **1994**, *98*, 1978−1988.
(54) Nozaki, Y.; Tanford, C. *Methods Enzymol.* **1967**, *11*, 715−734.
(55) Bashford, D. In *Scientific Computing in Object-Oriented Parallel Environments*; Ishikawa, Y., Oldehoeft, R. R., Reynders, J. V. W., Tholburn, M., Eds.; *Lecture Notes in Computer Science*. ISCOPE97, Springer, Berlin, 1997; vol. 1343, pp 233−240.
(56) Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. *Numerical recipes in C. The art of scientific computing*. Cambridge University Press: Cambridge, 1988.
(57) Warwicker, J.; Watson, H. C. *J. Mol. Biol.* **1982**, *157*, 671−679.
(58) Beroza, P.; Fredkin, D. R.; Okamura, M. Y.; Feher, G. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 5804−5808.
(59) Fersht, A. *Enzyme structure and mechanism*, W. H. Freeman and Company: New York; 1985; chapter 3.
(60) Zhang, Z.-Y.; Van Etten, R. L. *J. Biol. Chem.* **1991**, *266*, 1516−1525.
(61) Zhang, Z.-Y.; Wang, Y.; Dixon, J. E. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 1624−1627.
(62) Bashford, D.; Karplus, M. *Biochem.* **1990**, *29*, 10219−10225.
(63) Bone, S.; Pethig, R. *J. Mol. Biol.* **1982**, *157*, 571−575.
(64) Bone, S.; Pethig, R. *J. Mol. Biol.* **1985**, *181*, 323−326.
(65) Gilson, M. K.; Sharp, K. A.; Honig, B. H. *J. Computat. Chem.* **1987**, *9*, 327−335.

(66) Simonson, T.; Perahia, D. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 1082−1086.
(67) Simonson, T.; Perahia, D. *J. Am. Chem. Soc.* **1995**, *117*, 7987−8000.
(68) Antosiewicz, J.; McCammon, J. A.; Gilson, M. K. *J. Mol. Biol.* **1994**, *238*, 415−436.
(69) Garcia-Moreno, B.; Dwyer, J. J.; Gittis, A. G.; Lattman, E. E.; Spencer, D. S.; Stites, W. E. *Biophys. Chem.* **1997**, *64*, 211−24.
(70) Demchuk, E.; Wade, R. C. *J. Phys. Chem.* **1996**, *100*, 17373−17387.
(71) Goldstein, B. J. *Protein Profile* **1995**, *2*, 1425−1585.
(72) Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441−451.