Articles

High-Resolution Structure of the *Yersinia pestis* Protein Tyrosine Phosphatase YopH in Complex with a Phosphotyrosyl Mimetic-Containing Hexapeptide[‡]

Jason Phan,[#] Kyeong Lee,[§] Scott Cherry,[#] Joseph E. Tropea,[#] Terrence R. Burke, Jr.,[§] and David S. Waugh*,[#]

Macromolecular Crystallography Laboratory and Laboratory of Medicinal Chemistry, Center for Cancer Research, National Cancer Institute at Frederick, Frederick, Maryland 21702

Received June 19, 2003

ABSTRACT: *Yersinia pestis*, the causative agent of bubonic plague, secretes a eukaryotic-like protein tyrosine phosphatase (PTPase) termed *Yersinia* outer protein H (YopH) that is essential for virulence. We have determined, for the first time, the crystal structure of the YopH PTPase domain in complex with a nonhydrolyzable substrate analogue, the hexapeptide mimetic Ac-DADE-F_2Pmp-L-NH_2. As anticipated, the mode of ligand binding in the active site is similar to the way in which the corresponding phosphohexapeptide binds to the structurally homologous human PTP1B. Unexpectedly, however, the crystal structure also revealed a second substrate-binding site in YopH that is not present in PTP1B. The mode of binding and structural conformation of the hexapeptide analogue is quite different in the two sites. Although the biological function of the second substrate-binding site remains to be investigated, the structure of a substrate analogue in the active site of *Y. pestis* YopH opens the door for the structure-based design and optimization of therapeutic countermeasures to combat this potential agent of bioterrorism.

Yersinia pestis utilizes a contact-dependent (type III) secretion apparatus to inject six cytotoxic effector proteins (YopE, YopH, YopM, YopJ/P, YopT, and YpkA) directly into the cytosol of mammalian cells, where they conspire to defeat the innate immune response of the infected organism by interfering with signaling pathways that regulate cytoskeletal dynamics and inflammation (*1*−*3*). One of the effectors, YopH, is a powerful eukaryotic-like protein tyrosine phosphatase (PTPase). Upon translocation into mammalian cells, YopH disrupts signal transduction pathways required for the attachment of cells to the extracellular matrix and phagocytosis by dephosphorylating a variety of proteins associated with the focal adhesion, including focal adhesion kinase (p125[FAK]), paxillin, p130[cas], p105[casL], Fyn-binding protein, and the Src-associated adaptor protein SKAP−HOM (*4*−*9*). Because its PTPase activity is essential for virulence (*10*), YopH is a valid molecular target for antiplague therapeutics that could be used to combat weaponized strains that have been engineered to be resistant to conventional antibiotics.

YopH is a modular protein that is composed of two independently folded domains. The N-terminal domain (residues 1−130) binds tyrosine-phosphorylated proteins in a phosphoryl-dependent manner and evidently plays a role in the recognition of at least some substrates in vivo (*5*, *11*). The

structure of the N-terminal domain of YopH (*12*, *13*) does not resemble those of the eukaryotic Src homology 2 (SH2) or phosphotyrosine-binding (PTB) domains, nor does it bind phosphotyrosine in a similar manner (*14*). The C-terminal domain of YopH (residues 164−468) is the seat of its PTPase activity (*15*). The structure of the *Yersinia* YopH PTPase domain is very similar to the structures of eukaryotic PTPases (*16*), suggesting that it may have been acquired by lateral gene transfer.

The X-ray crystal structure of the catalytic domain of YopH was among the first PTPase structures to be solved (*17*). Phosphate-, tungstate-, nitrate-, and vanadate-bound structures were subsequently determined (*17*−*20*). Although these structures provided valuable information about enzyme−anion interactions and the reaction mechanism, they did not reveal how YopH recognizes its phosphotyrosyl substrates. Such information is required for the structure-based design of more potent and specific inhibitors of the *Yersinia* PTPase. In this report, we describe the structure of the catalytic domain of YopH in complex with a hexapeptide mimetic (Ac-DADE-F_2Pmp-L-NH_2) derived from an autophosphorylation site of the epidermal growth factor receptor cytoplasmic domain.

MATERIALS AND METHODS

Synthesis of the Phosphotyrosyl Mimetic-Containing Hexapeptide. The nonhydrolyzable substrate analogue Ac-DADE-F_2Pmp-L-NH_2 was first reported as a high affinity inhibitor of PTP1B (*21*). The peptide employed in the present

---

study was synthesized by using the F$_2$Pmp[1] residue as its free phosphonic acid, according to a modification of the procedure described by Gordeev (*22*). This involved introduction of the F$_2$Pmp residue using BOP, DIEA, HOBt, and DMAP. Subsequent residues were added in the same fashion. The finished product was cleaved from the resin and purified by reverse-phase HPLC.

Protein Expression and Purification. The PTPase domain of YopH (residues 164−468) was overproduced in *Escherichia coli* as described (*15*). The cell pellet was resuspended in 100 mM sodium acetate (pH 5.7), 50 mM NaCl, and 1 mM ethylenediaminetetraacetic acid (EDTA) (buffer A), and then disrupted with an APV Gaulin model G1000 homogenizer at 10 000 psi for three passes. After centrifugation at 15 000 rpm for 30 min, the supernatant was applied to a carboxymethyl (CM) Sepharose column equilibrated in buffer A. The column was washed extensively with buffer A, after which the bound protein was eluted with a gradient of 0.05− 0.5 M NaCl in buffer A. The peak fractions containing YopH were subsequently run on a Sephacryl S-100 gel filtration column equilibrated with 100 mM sodium acetate (pH 5.7), 100 mM NaCl, and 1 mM EDTA. The S-100 fractions containing pure YopH PTPase were pooled and concentrated by diafiltration to 10.3 mg/mL in 100 mM sodium acetate (pH 5.7), 100 mM NaCl, and 1 mM EDTA.

Crystallization and Data Collection. The protein−ligand complex was prepared by adding a 10-fold molar excess of the substrate analogue to a concentrated solution of YopH PTPase (10.3 mg/mL). The first crystals were grown by the hanging drop, vapor-phase diffusion method at 18 °C by mixing 2 μL of the protein−ligand complex with 2 μL of a precipitant solution consisting of 100 mM Hepes (pH 7.5), 20% (w/v) PEG 4000, and 10% (v/v) 2-propanol. The crystals initially grew in clusters that were not suitable for X-ray diffraction. However, a nice looking rod was dissected from the bunch and crushed for microseeding. The concentrated seed solution (10 μL) was diluted 200-fold, and then 0.5 μL was added to drops that had been preequilibrated for 3−8 h. The composition of these drops was either the same as the original conditions from which the seed source was obtained; 100 mM Hepes (pH 7.5), 20% PEG 3350, and 200 mM NaCl; or 100 mM Bicine (pH 9.0), 20% PEG 4000, and 5% 2-propanol. Single, rod-shaped crystals measuring approximately 0.05 × 0.06 × 0.20 mm were observed after a few days. During the optimization process, it was found that this crystal form could also be obtained from a solution containing 100 mM Tris-HCl (pH 8.0), 18% PEG 8000, 200 mM NaF, and 5−15% (v/v) dioxane.

In any scenario, a single prismatic rod was immersed in either 20% PEG 4000 and 15% (w/v) glycerol or 20% PEG 8000 and 15% glycerol, and then flash-frozen at 100 K for data collection at the SER−CAT insertion device beamline (22-ID, Advanced Photon Source, Argonne National Laboratory). The data were reduced using the HKL2000 suite of programs (*23*). The protein−ligand complex crystallized in the *P*1 space group as a dimer in the unit cell with specific

---

[1] Abbreviations: BOP, benzotriazole-1-l-oxy-tris-(dimethylamino)-phosphonium hexafluorophosphate; DIEA, *N,N*-diisopropylethylamine; DMAP, 4-(dimethylamino) pyridine; F$_2$Pmp, difluorophosphonomethylphenylalanine; HOBt, 1-hydrozybenzotriazole; TPI, [1,1-difluoro-1-((6-carboxamidoglutamic)-naphth-2-yl)] methylphosphonic acid.
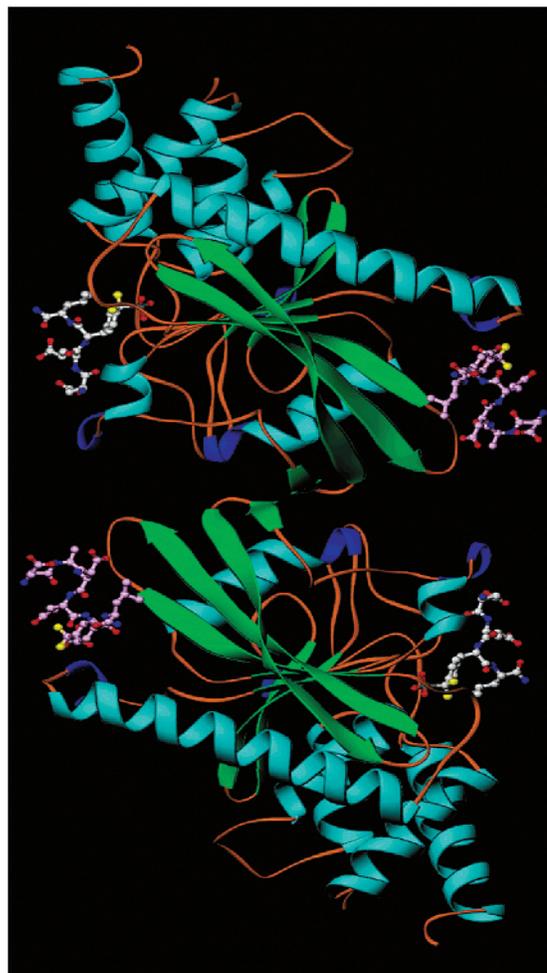
FIGURE 1: Ribbon representation of a crystallographic dimer of YopH complexed with the hexapeptide analogue Ac-DADE-F$_2$Pmp-L-NH$_2$ (shown as ball-and-stick models). The carbon atoms are colored white and purple in the ligands bound in the active site and the second substrate-binding site, respectively. Fluorine and phosphorus atoms are colored yellow and white, respectively. Drawn with RIBBONS and rendered with POV-ray (*45*).

volume Vm = 2.20 Å^3/Da (*24*) and ∼45% solvent content. The cell dimensions (in angstroms) were a = 47.29, b = 53.45, c = 69.06, α = 109.61, β = 104.76, and γ = 89.98.

Structure Determination and Refinement. The structure was solved by molecular replacement with the AMoRe program (*25*) using the coordinates of unliganded YopH PTPase (PDB code: 1YPT) as the search model. Two solutions were found after rigid body fitting with a correlation factor of 72.6 and R-factor of 31.8, corresponding to two molecules in the asymmetric unit (Figure 1). The dimer model was refined with CNS (*26*), using maximum likelihood targets and noncrystallographic symmetry (NCS) restraints. Analysis and manual correction of the structure were carried out with the graphics program O (*27*). A summary of the crystallographic and refinement statistics is presented in Table 1.

Strong and continuous electron density was observed for amino acids 168−468 of YopH. An F_o−F_c simulated annealing OMIT map revealed stretches of density corresponding to the structure of the substrate analogue in the active site and also in a second binding site on the opposite side of the protein. The two N-terminal residues of the hexapeptide mimetic in the active site protrude into solvent

Table 1: Data Collection, Molecular Replacement, and Refinement Statistics

Data Collection	
X-ray source	ID-22 (APS)
wavelength (Å)	0.97148
space group	$P1$
resolution (Å)	25−1.5
unique reflections:	87,736
completeness (%)	89.3
$R_{sym}{}^{a}$ (%)	0.038
B-factor from Wilson Plot (Å^2)	16.24
Molecular Replacement	
R-factor (%)	31.8
correlation factor (%)	72.6
Refinement	
no. of reflections	
working set/test set	59,855/6,702
$R_{cryst}{}^{b}$ (%)	18.8
$R_{free}{}^{c}$ (%)	20.9
r.m.s.^{d} deviation from ideal geometry	
lengths (Å)	0.006
angles (°)	1.2
no. of molecules	
peptide analogues	2
water	500

$^{a}R_{sym} = (\Sigma_h|I_h - \langle I \rangle|)/(\Sigma_h I_h).$ $^{b}R_{cryst} = (\Sigma_h|F_{obs} - F_{cal}|)/(\Sigma_h F_{obs}).$ $^{c}R_{free}$ = crystallographic R-factor for test set as implemented in CNS_1.1 (*26*). ^{d} Root mean-square.

and are disordered. However, all six residues of the hexapeptide in the second binding site have strong electron density (Figure 2). No electron density corresponding to the N-terminal acetyl group of the ligand was observed in either binding site.

RESULTS

Cocrystallization of the YopH PTPase Domain with a Substrate Analogue. Using site-directed mutagenesis and chemical modification to analyze the contributions of each amino acid surrounding an autophosphorylation site of an epidermal growth factor receptor-derived sequence (DADEp-YLIPQQG) to substrate binding and catalysis, Zhang and co-workers concluded that the hexapeptide DADEpYL is the smallest derivative of this sequence that is still an efficient substrate for YopH (*28, 29*). It was also shown that the non-hydrolyzable phosphotyrosine mimetic phosphonomethyl-phenylalanine (Pmp) could be incorporated into the hexapeptide without distorting the geometry of the substrate, but the Pmp-containing peptide binds with reduced affinity because the methylene carbon atom cannot make the hydrogen bonds that are normally formed by the tyrosyl oxygen (*30*). Accordingly, the addition of two fluorine atoms to the methylene carbon of the Pmp moiety enhanced the potency of the hexapeptide inhibitor (*21*). Although the epidermal growth factor receptor is not known to be one of the natural targets of YopH in vivo, its autophosphorylation site is unquestionably the best-characterized substrate in vitro. Additionally, the DADE-pY-L-NH$_2$ hexapeptide had previously been cocrystallized with a catalytically inactive mutant of human PTP1B (*31*). For these reasons, a nonhydrolyzable analogue of this hexapeptide (Ac-DADE-F$_2$Pmp-L-NH$_2$) was prepared and cocrystallized with the PTPase domain of YopH.

Mode of Ligand Binding in the Active Site. The substrate analogue Ac-DADE-F$_2$Pmp-L-NH$_2$ binds to the active-site
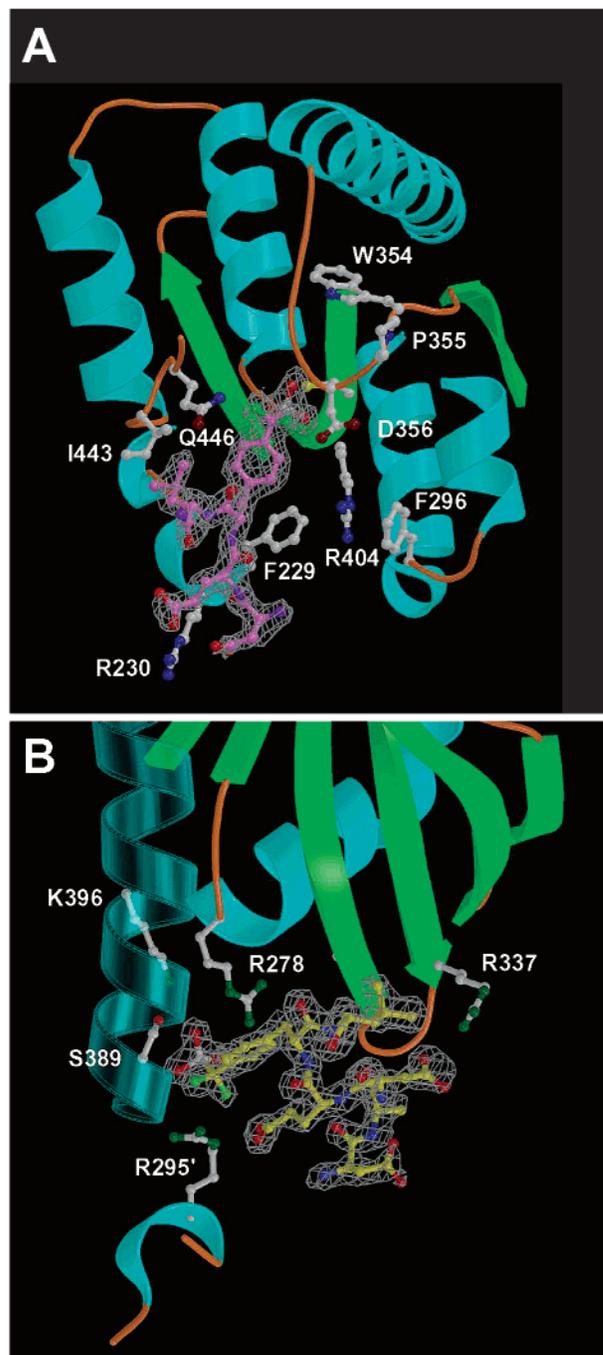
FIGURE 2: Electron density of the ligand in the active site (A) and the second substrate-binding site (B) superimposed on the refined models in ball-and-stick format. The $F_o−F_c$ OMIT map density is contiguous for all six residues of the hexapeptide mimetic in the second binding site and only the last four residues in the active site (*46, 47*).

pocket of YopH in a manner that is very similar to the way in which the corresponding phosphohexapeptide binds to the catalytically inactive human PTP1B C215S mutant (*31*). The phosphotyrosyl mimetic is inserted into a ~10 Å long crevice that is capped by a loop at the end of strand $\beta 8$ (nomenclature as defined in ref *17*) and the first turn of helix $\alpha 5$, which has been termed the phosphate-binding loop or P-loop (Figures 3 and 4A). This highly conserved structural motif is shared by nearly all PTPases and characterized by the sequence (H/V)CX$_5$R(S/T) (*32−34*). The phosphate moiety is surrounded by the main chain amide groups of the P-loop
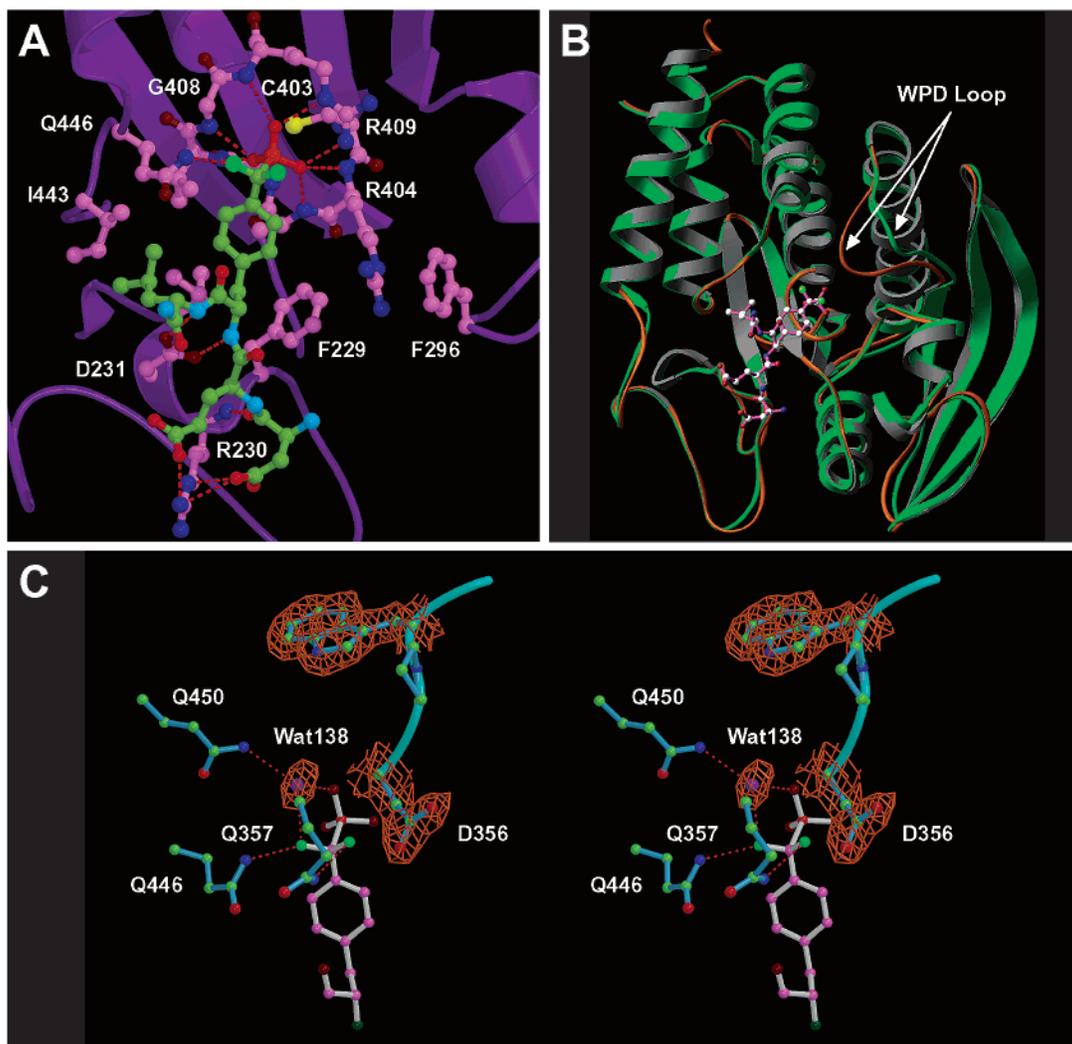
FIGURE 3: Interactions between the substrate analogue and the active site of YopH. (A) Hydrogen bonding interactions between the enzyme and the ligand. (B) Structural alignment of the unliganded YopH structure (green) with that of the enzyme/hexapeptide complex (gray ribbons and copper loops), showing the two different conformations of the WPD loop. (C) Electron density of the WPD loop residues and the trapped water molecule. The fluorine atoms of the F$_2$Pmp moiety interact with invariant glutamines and Wat138 (*45−47*).

(residues 404−409), which donate six hydrogen bonds, the side chain of Arg409, which forms two salt bridges with the phosphate oxygen atoms, and a structural water molecule, Wat138, that is hydrogen-bonded to oxygen O2P (Table 2). At the center of the P-loop lies the catalytic cysteine, Cys403, whose thiolate group is located 4.2 Å from the phosphorus atom. A highly flexible loop (residues 350−360) containing the invariant Trp-Pro-Asp (WPD) triplet (*35*) is in the closed conformation, effectively sealing the ligand in the active site. The aromatic part of the pTyr mimetic is nestled in an aliphatic corridor created by invariant hydrophobic amino acids that line the crevice just outside the P-loop. The leucine in the +1 position of the hexapeptide rests in a shallow cleft formed by residues Ile232, Ile443, Gln446, Asp231, and Arg205 of YopH, and the edge of the phosphotyrosine analogue. The carboxyl side chain of Asp231 forms two hydrogen bonds with the backbone amide groups at the F$_2$Pmp and +1 positions of the hexapeptide. The acidic residues in the −1 and −2 positions of the substrate analogue are stabilized by ionic interactions with the guanidino group of Arg230, whereas the residues in the −3 and −4 positions protrude into the solvent and are disordered in the structure.

This mode of substrate binding in the active site of YopH is consistent with the observation that deleting residues −3 and −4 from the DADEpYL hexapeptide had only a modest effect (5-fold reduction) on catalytic efficiency (*28*).

Conformational Changes Induced by Ligand Binding. The structure of the apoenzyme (PDB code: 1YPT) aligns very closely with that of the enzyme−hexapeptide complex, with an overall Cα RMSD of only 0.6 Å for all residues except amino acids 350−360 (Figure 3B). The latter residues correspond to the flexible WPD loop with the general acid/base Asp356 that swings into the active site to participate in substrate sequestering and catalysis (*18, 36*). The carboxyl side chain of Asp356, which is equivalent to Asp92 in *Vaccinia* H1-related phosphatase (VHR) and Asp181 in PTP1B, acts as a general acid to donate a proton to the leaving tyrosyl hydroxyl group, generating the phosphoenzyme intermediate. In the second step of the reaction, where the cysteinyl-phosphate intermediate is hydrolyzed, kinetic data for the VHR D92N mutant suggested that the conserved aspartic acid behaves as a general base to activate a nearby water molecule for nucleophilic attack on the phosphorus (*18*). Upon binding of the nonhydrolyzable substrate ana-
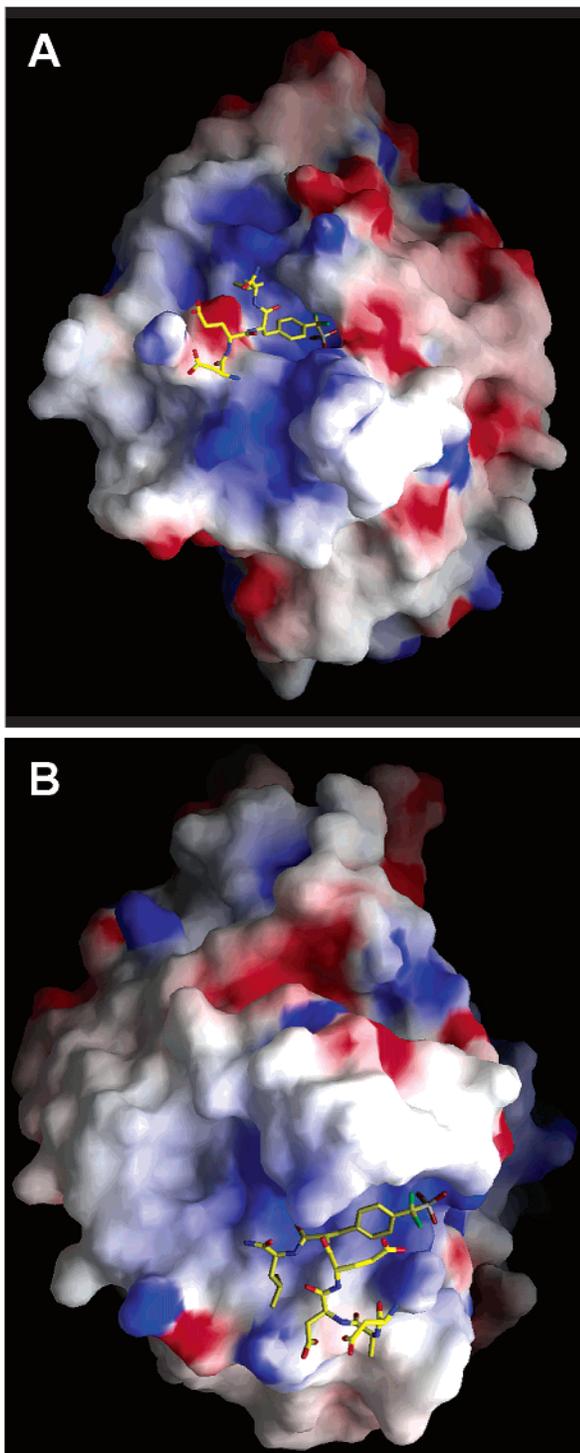
FIGURE 4: Electrostatic potentials (*48*) mapped onto the molecular surface of YopH with the ligand represented as a ball-and-stick model in (A) the active site and (B) the second substrate-binding site.

logue to YopH, Asp356, and Gln357 move 6.3 and 7 Å closer to the active site, respectively. However, the side chain of Asp356 rotates away from the ligand due to the presence of the fluorine atoms on carbon CH of the F_2Pmp moiety (Figure 3C). In addition, the guanidino group of Arg205 moves 4 Å to form a hydrogen bond with the carbonyl of Gln446, which also shifts its position to create part of the binding site for the phenyl moiety of the pTyr mimetic.

Comparison with Anion-Bound Structures. The structure of the YopH PTPase/substrate analogue complex can be aligned with the structures of the enzyme−nitrate, −tungstate (*20*), and −sulfate (*19*) complexes with an average RMSD of only ∼0.4 Å for all Cα atoms. Their flexible WPD loops are all in the "closed" conformation, poised for catalysis. The side chains that need to rotate to accommodate the new loop position also adopt similar conformations in all of these structures. Movement of the α3 helix (residues 286−297) relative to the unliganded structure is also observed in the anion-bound enzymes, but to a lesser extent. The carboxylate group of Asp231 has a similar conformation in the tungstate- and hexapeptide-bound structures, whereas in the two other anion-bound structures its conformation corresponds to that observed in the apoenzyme. This implies that both rotamers probably exist in solution because neither interacts with the ligands in any of these structures. The ligand-binding side chain of Arg230 adopts quite different conformations in all four structures, with the greatest difference occurring in the structure of the enzyme−hexapeptide complex. The side chains of Gln446 and Asp356 have different orientations in the hexapeptide complex because of the presence of the fluorine atoms in the phosphotyrosine mimetic. In all four structures, Arg205 adopts a conformation that permits hydrogen bonding with the backbone carbonyl of Gln446, which does not occur in ligand-free enzyme. Finally, the tungstate ion is almost superimposable with the phosphoryl moiety of the phosphotyrosine mimetic. Its apical oxygen (O4) coincides with the methylene carbon, while the sulfate ion extends 0.6 Å further into the P-loop of the C403S mutant. The dissimilarity can be explained by the chemistry of the C403S substitution (*19*).

Comparison with the Human PTP1B/Ac-DADEpYL-NH₂ Complex. Although YopH and human PTP1B share only ∼20% sequence identity in the catalytic domain (*35*), their three-dimensional structures are remarkably similar (*17, 37*). When the P-loop and WPD-loop from the PTP1B/Ac-DADE-pY-L-NH₂ complex (PDB code: 1PTU) are aligned with their counterparts in the YopH/Ac-DADE-F_2Pmp-L-NH₂ structure, the overall Cα RMSD is only 0.350 Å. The phosphoryl moiety in both complexes is similarly positioned in its respective phosphate-binding loop. Inevitably, however, there are a number of differences between the two structures because their amino acid sequences are so dissimilar. In the YopH complex, Gln357 interacts with the *pro-R* fluorine (Figure 3C), while the corresponding Phe182 in PTP1B forms part of the binding site for the phenyl ring of phosphotyrosine. Phe182 pushes Gln262 away from the phosphate group such that it now lines the pocket for the phosphotyrosine phenyl group. The corresponding residue in YopH is Gln446, which also moves to create the binding site for the F_2Pmp moiety, but instead of rotating away from the phosphate headgroup, it turns upward to form hydrogen bonds with the *pro-S* fluorine (Figure 3C). In the PTP1B/TPI complex (PDB code: 1BZC), which also contains the F_2Pmp moiety, Gln262 adopts the same rotamer as in the PTP1B/hexapeptide structure and does not interact with the fluorine atom (*38*). Nonetheless, unlike its counterpart in the PTP1B/hexapeptide complex, Asp181 in the PTP1B/TPI complex rotates away from the F_2Pmp moiety in much the same way as Asp356 does in the structure of the YopH/hexapeptide complex.

Table 2: Interactions between YopH and the Substrate Analog Ac-DADE-F$_2$Pmp-L-NH$_2$

residues				residues	
active site (503−506)				second site (601−606)	
ligand	enzyme	hydrogen bonds distance (Å)		enzyme	ligand
			2.85	Thr 343 N	Ala 602 O
			2.82	Wat 21	
Asp 503 O	Arg 230 N	2.76	2.89	Arg 337 NH1	Asp 603 OD1
Asp 503 OD2	Arg 230 NH2	2.99	2.65	Thr 343 OG1	
	Arg 230 NE	2.78	2.65	Wat 307	Asp 603 O
	Wat 325	2.82			
Asp 503 OD1	Arg 230 NE	3.19			
Glu 504 OE1	Arg 230 NE	3.10	3.08	Lys 386 NZ	Glu 604 O
	Arg 230 NH2	2.84	2.85	Lys 342 NZ	Glu 604 OE1
Glu 504 N	Wat 291	3.22	2.85	Wat 87	Glu 604 OE2
Glu 504 O	Wat 526	2.82	2.84	Wat 219	
O1P	Arg 409 NH2	2.96	2.82	Arg 278 NH2	O1P
	Arg 404 N	3.05	2.76	Wat 134	
	Ala 405 N	2.96	2.89	Wat 517	
O2P	Arg 409 N	3.06	2.89	Ser 389 OG	O2P
	Arg 409 NE	2.92	2.92	Ser 389 N	
	Wat 138	2.80	2.59	Wat 210	
O3P	Gly 408 N	2.76	2.74	Ser 388 OG	O3P
	Val 407 N	3.04	3.10	Arg 278 NE	
	Gly 406 N	3.35	2.73	Wat 123	
Fpy 505 FH[a]	Gln 357 N	3.05	3.36	Lys 342 NZ	[a]Fpy 605 FH
Fpy 505 FH2	Gln 450 NE2	3.00			
Fpy 505 N	Asp 231 OD1	2.97	2.88	Wat 21	Fpy 605 N
Leu 506 N	Asp 231 OD2	3.19	2.72	Wat 307	Leu 606 O
			2.94	Wat 91	
			2.86	Wat 154	Leu 606 NH2

[a] Fpy = F$_2$Pmp = pTyr mimetic.
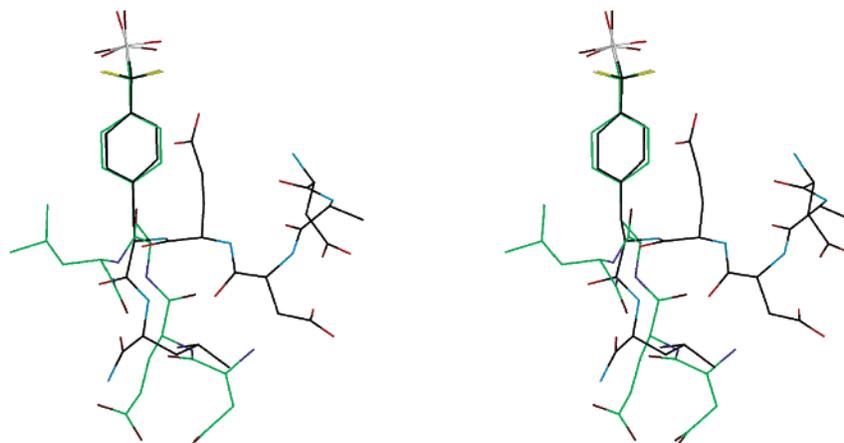
FIGURE 5: Stereoview of the superposition of the hexapeptide analogue in the second substrate-binding site (black carbons and cyan nitrogens) onto that bound in the active site (green carbons and blue nitrogens). Phosphorus and fluorine atoms are colored white and yellow, respectively (*47*, *49*).

Description of the Second Substrate-Binding Site. The second, noncatalytic substrate binding site is located in a long, shallow groove alongside helix α4 and the β6/β7 hairpin, on the opposite side of the molecule from the active site (Figure 4B). Structural alignment of the F$_2$Pmp moieties demonstrates that the ligands adopt very different conformations in the two binding sites (Figure 5). In the noncatalytic binding site, the phosphoryl moiety forms hydrogen bonds with the side chains of Ser388, Ser389, and the backbone amide group of Ser389 (Figure 6). A bidentate salt bridge is formed between the guanidinium ion of Arg278 and phosphate oxygens O1P and O3P. The fluorine atom FH is 3.36 Å from atom NZ of Lys342 (Table 2). The major interaction with the difluoro moiety is contributed by Arg295 from a crystallographic symmetry mate of the same subunit type,

such as B−B′ or A−A′ (Arg295′). The plane of the Arg295′ guanidino group is perpendicular to the plane formed by the methylene carbon and the fluorine atoms. The *pro-R* fluorine FH is hydrogen-bonded to atom NE of Arg295′ (3.13 Å), and the *pro-S* fluorine FH2 makes hydrogen bonds with atoms NH1 and NH2 at 3.33 and 3.37 Å, respectively. Arg295′ also forms a bidentate salt bridge with the carboxyl side chain of the glutamic acid at position −1 in the hexapeptide (Glu604), and hydrogen bonds with the backbone carbonyl oxygens of Lys386 and Gly387. The phenyl ring of the F$_2$Pmp is packed against the side chain of Tyr383 and the aliphatic portions of Lys386 and Glu604, where it also makes hydrophobic contacts with Lys342, Ile344, and Met382. The leucine at position +1 in the substrate analogue (Leu606) engages in hydrophobic interactions with the CG2
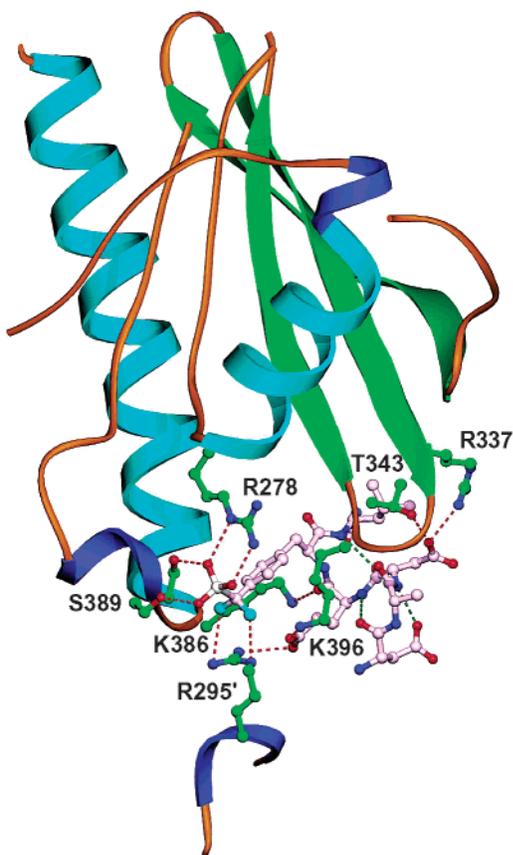
FIGURE 6: Amino acid residues of YopH that compose the second, noncatalytic substrate-binding site. The side chains of the enzyme that interact with the ligand are colored green and the ligand carbons are colored white. Hydrogen bonds are represented by dashed lines, with intermolecular and intrapeptide bonds depicted in red and green, respectively (45).

atoms of Thr335 and Thr343. The side chain of Glu604 is hydrogen-bonded to that of Lys342, and its main chain carbonyl oxygen accepts a proton from Lys386 NZ. The carboxyl side chain of the aspartic acid in position −2 of the hexapeptide (Asp603) forms hydrogen bonds with the OH and NH1 groups of Thr343 and Arg337, respectively. The residues at positions −4 (Asp601) and −3 (Ala602) of the hexapeptide do not interact directly with the protein. Instead, they make intrapeptide interactions with the side chain of Asp601 hydrogen bonding to backbone amides of Asp603 and Ala602. In another intrapeptide interaction, hydrogen bonds are accepted by the side chain of Asp603 from the main chain amide groups of F_2Pmp (FPY605) and Leu606. Interestingly, Asp603 effectively mimics the function of the enzyme's Asp231, which forms similar hydrogen bonds with the substrate analogue in the active site (Figure 3A). In the structure of human PTP1B/Ac-DADE-pY-L-NH$_2$ complex, the corresponding Asp48 also stabilizes the backbone conformation through the same hydrogen-bonding pattern (32).

DISCUSSION

Structural Insights into Substrate Specificity. Kinetic studies with many different types of substrates have shown that the PTPase domain of YopH is not a highly selective enzyme (28, 29, 39). Indeed, given sufficient time, YopH will dephosphorylate a wide range of proteins in eukaryotic cells (40, 41). Nevertheless, YopH clearly exercises some degree of substrate selectivity because it preferentially targets certain proteins associated with the focal adhesion (e.g., p130^{Cas}, paxillin, and FAK). All of these proteins contain multiple phosphorylation sites in a variety of sequence contexts and in no case is it known precisely which phosphates are removed by YopH. Circumstantial evidence suggests that YopH may favor sequences of the general form pY-X$_{aa}$-X$_{aa}$-P (6, 29), but thus far there has been no experimental confirmation of this conjecture. The lack of specific information about the sites that are dephosphorylated by YopH in its "natural" targets discouraged us from attempting to cocrystallize the enzyme with substrate analogues derived from these proteins.

The phosphotyrosine-containing peptide DADE-pY-LIPQQG, which corresponds to an autophosphorylation site (Tyr992) in the epidermal growth factor receptor, is by far the best characterized and most efficient substrate reported for YopH, with a rate constant (2.23×10^7 M^{-1} s^{-1}) that is near the diffusion-controlled limit (28). A systematic analysis of structure−activity relationships within this phosphopeptide revealed that the hexapeptide derivative DADE-pY-L is dephosphorylated with nearly equal efficiency (28). It therefore seemed plausible that the structure of YopH in complex with a nonhydrolyzable form of this hexapeptide would afford a suitable starting point for the structure-based design and optimization of more potent and specific YopH inhibitors. Accordingly, the difluorophosphonomethylphenylalanine derivative Ac-DADE-F$_2$Pmp-L-NH$_2$ was prepared for cocrystallization with the enzyme.

The mode of ligand binding in the active site of YopH is generally consistent with the results of structure−activity studies conducted by Zhang and co-workers (28, 29). The residues in the −3 and −4 positions of the substrate analogue protrude into the solvent in the cocrystal structure and do not make contact with the enzyme. Removing them from the hexapeptide reduces catalytic efficiency only 5-fold (28). The acidic residues in the −1 and −2 positions of the substrate analogue are stabilized by ionic interactions with the guanidino group of Arg230. Accordingly, removing these residues or replacing them with alanine results in more dramatic effects on catalytic efficiency (28, 29). The Glu at position −1 appears to be especially critical; an otherwise identical hexapeptide with alanine in this position is a 126-fold less efficient substrate for YopH (28). There are strong hydrogen bonding and ionic interactions between the side chain of Asp231 and the main chain amides of the F$_2$Pmp moiety and the leucine residue at position +1 in the hexapeptide. Yet, the side chain of the C-terminal leucine residue makes very few favorable contacts with the protein. These observations are consistent with the finding that deleting residues C-terminal to the phosphotyrosine has more severe consequences than replacing them with alanine (28, 29).

Potential Function of the Noncatalytic Substrate Binding Site. The activity of many PTPases is regulated not only by the sequence context in which the phosphotyrosine residue resides, but also by subcellular localization and accessory modules that participate in substrate recognition (42). Such is the case for YopH. Persson and co-workers determined that the surface loop sequence 223−226 functions in localizing YopH to peripheral focal complexes (43). Moreover,

together with the C-terminal catalytic domain, the N-terminal domain of YopH plays a role in substrate recognition by independently binding to tyrosine-phosphorylated sites (*5, 11*). In principle, either or both of these factors could contribute to the specificity of YopH in vivo. Yet, whereas deleting the N-terminal domain of YopH or abolishing its ability to bind phosphotyrosyl peptides by site-directed mutagenesis reduced the rate of substrate dephosphorylation in vivo, it did not diminish the intrinsic specificity of the PTPase domain (*5*). It therefore seems likely that the role of the N-terminal substrate-binding domain is to anchor the enzyme to the target so that its catalytic machinery can operate processively on multiple phosphotyrosines. A similar phenomenon had been observed for the two-headed ATP-hydrolyzing motor protein that progresses along the length of its target in "end over end" mechanical steps (*44*). The fact that the known targets of YopH possess multiple phosphorylation sites in tandem is consistent with this hypothesis.

The unexpected discovery of a second substrate-binding site in the catalytic domain of YopH raises the possibility that it too may contribute to the processive action of the enzyme in vivo. It is interesting that there seem to be more specific interactions between the protein and ligand in the second site than the active site, especially outside of the F$_2$Pmp moiety. Unlike the way in which phosphopeptides adhere to a shallow groove on the surface of the N-terminal domain of YopH (*14*), the second substrate analogue in the structure described here binds in a very well-defined cleft (Figure 4B). Further work will be required to determine whether this binding site contributes to the activity of YopH in vivo, and if so, what its specificity is.

Prospects for the Development of YopH Inhibitors. The presence of multiple phosphotyrosine binding sites in YopH has the potential to complicate the development of inhibitors for this enzyme. Yet, because there is no direct way to measure the binding of inhibitors to these auxiliary sites, the active site of the enzyme remains the most practical target. The structure of the YopH PTPase domain in complex with a hexapeptide mimetic represents an important first step in this direction. The cocrystal structure suggests that the tetrapeptide analogue DE-F$_2$Pmp-L-NH$_2$, which encompasses the bulk of the interactions with the enzyme active site, would be a good starting point for further optimization. The side chain of the C-terminal leucine residue appears to be a particularly good prospect for modification because it projects into but does not make optimal contacts with the body of the enzyme. The crystal structure of the YopH/hexapeptide complex also suggests additional ways in which the inhibitor could be modified to take advantage of adjacent hydrophobic and polar pockets on the surface of the protein.

The challenges that lie ahead are formidable. As in any drug development project, specificity, toxicity, and bioavailability are major concerns. Yet, the specter of antibiotic-resistant plague, unleashed by terrorists or rogue nations, demands immediate attention and YopH is presently the most promising target for the development of antivirulence drugs. A noteworthy advantage of drugs that directly target essential virulence factors such as YopH is that it would probably be very difficult to engineer organisms that are resistant to antivirulence agents but that can still cause disease. With the structure of YopH in complex with a nonhydrolyzable substrate analogue now in hand, we are potentially one step closer to the eradication of this deadly bioterrorist threat.

REFERENCES

1. Galan, J. E., and Collmer, A. (1999) *Science 284*, 1322−1328.
2. Cheng, G. R., and Schneewind, O. (2000) *Trends Microbiol. 8*, 214−220.
3. Cornelis, G. R. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 8778−8783.
4. Persson, C., Carballeira, N., Wolf-Watz, H., and Fallman, M. (1997) *EMBO J. 16*, 2307−2318.
5. Black, D. S., Montagna, L. G., Zitsmann, S., and Bliska, J. B. (1998) *Mol. Microbiol. 29*, 1263−1274.
6. Black, D. S., and Bliska, J. B. (1997) *EMBO J. 16*, 2730−2744.
7. Hamid, N., Gustavsson, A., Anderson, K., McGee, K., Persson, C., Rudd, E. C., and Fallman, M. (1999) *Microb. Pathog. 26*, 231−242.
8. Yao, T., Mecas, J., Healy, J. I., Falkow, S., and Chien, Y.-H. (1999) *J. Exp. Med. 190*, 1343−1350.
9. Black, D. S., Marie-Cardine, A., Schraven, B., and Bliska, J. B. (2000) *Cell. Microbiol. 2*, 401−414.
10. Guan, K., and Dixon, J. E. (1990) *Science 249*, 553−556.
11. Montagna, L. G., Ivanov, M. I., and Bliska, J. B. (2001) *J. Biol. Chem. 276*, 5005−5011.
12. Evdokimov, A. G., Tropea, J. E., Routzan, K. M., Copeland, T. D., and Waugh, D. S. (2001) *Acta Crystallogr. D57*, 793−799.
13. Smith, C. L., Khandelwal, P., Keliikuli, K., Zuiderweg, E. R., and Saper, M. A. (2001) *Mol. Microbiol. 42*, 967−979.
14. Khandelwal, P., Keliikuli, K., Smith, C. L., Saper, M. A., and Zuiderweg, E. R. P. (2002) *Biochemistry 41*, 11425−11437.
15. Zhang, Z.-Y., Clemens, J. C., Schubert, H. L., Stuckey, J. A., Fischer, M. W. F., Hume, D. M., Saper, M. A., and Dixon, J. E. (1992) *J. Biol. Chem. 267*, 23759−23766.
16. DeVinney, R., Steele-Mortimer, O., and Finlay, B. B. (2000) *Trends Microbiol. 8*, 29−33.
17. Stuckey, J. A., Schubert, H. L., Fauman, E. B., Zhang, Z.-Y., Dixon, J. E., and Saper, M. A. (1994) *Nature 370*, 571−575.
18. Denu, J. M., Lohse, D. L., Vijayalakshmi, J., Saper, M. A., and Dixon, J. E. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 2493−2498.
19. Schubert, H. L., Fauman, E. B., Stuckey, J. A., Dixon, J. E., and Saper, M. A. (1995) *Protein Sci. 4*, 1904−1913.
20. Fauman, E. B., Yuvaniyama, C., Schubert, H. L., Stuckey, J. A., and Saper, M. A. (1996) *J. Biol. Chem. 271*, 18780−18788.
21. Burke, T. R., Kole, H. K., and Roller, P. P. (1994) *Biochem. Biophys. Res. Commun. 204*, 129−134.
22. Gordeev, M. F., Patel, D. V., Barker, P. L., and Gordon, E. M. (1994) *Tetrahedron Lett. 35*, 7585−7588.
23. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276A*, 307−326.
24. Matthews, B. W. (1968) *J. Mol. Biol. 33*, 491−497.
25. Navaza, J. (1994) *Acta Crystallogr. D50*, 157−163.
26. Brünger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D54*, 905−921.
27. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
28. Zhang, Z.-Y., Maclean, D., McNamara, D. J., Sawyer, T. K., and Dixon, J. E. (1994) *Biochemistry 33*, 2285−2290.
29. Zhang, Z.-Y., Thieme-Sefler, A. M., Maclean, D., McNamara, D. J., Dobrusin, E. M., Sawyer, T. K., and Dixon, J. E. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 4446−4450.
30. Burke, T. R., Jr., and Zhang, Z.-Y. (1998) *Biopolymers 47*, 225−241.
31. Jia Z., Barford, D., Flint A. J., and Tonks, N. K. (1995) *Science 268*, 1754−1758.
32. Aqvist, J. and Kolmodin, K. (2001) *FEBS Lett. 498*, 208−213.
33. Denu, J. M., and Dixon, J. E. (1998) *Curr. Opin. Chem. Biol. 2*, 633−641.
34. Zhang, Z.-Y. (1998) *Crit. Rev. Biochem. Mol. Biol. 33*, 1−52.
35. Huijsduijnen, R. H. V. (1998) *Gene 225*, 1−8.

36. Zhang, Z.-Y., Wang, Y., and Dixon, J. E. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 1624−1627.

37. Barford, D., Flint, A. J., and Tonks, N. K. (1994) *Science 263*, 1397−1404.

38. Groves, M. R., Yao, Z. J., Roller, P. P., Burke, T. R., Jr., and Barford, D. (1998) *Biochemistry 37*, 17773−17783.

39. Dunn, D., Chen, L, Lawrence, D. S., and Zhang, Z.-Y. (1996) *J. Biol. Chem. 271*, 168−173.

40. Andersson, K., Carballeira, N., Magnusson, K. E., Persson, C., Stendahl, O., Wolf-Watz, H., and Fallman, M. (1996) *Mol. Microbiol. 20*, 1057−1069.

41. Deleuil, F., Mogemark, L., Francis, M. S., Wolf-Watz, H., and Fallman, M. (2003) *Cell. Microbiol. 5*, 53−64.

42. Li, L., and Dixon, J. E. (2000) *Semin. Immunol. 12*, 75−84.

43. Persson, C., Nordfelth, R., Anderson, K., Forsberg, A., Wolf-Watz, H., and Fallman, M. (1999) *Mol. Microbiol. 33*, 828−838.

44. Visscher, K., Schnitzer, M. J., and Block, S. M. (1999) *Nature 400*, 184−189.

45. Carson, M. (1997) *Methods Enzymol. 277*, 493−505.

46. Esnouf, R. M. (1997) *J. Mol. Graph. 15*, 132−134.

47. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.

48. Nicholls, A., Sharp, K., and Honig, B. (1991) *Proteins 11*, 281−296.

49. Meritt, E. A., and Murphy, M. E. P. (1994) *Acta Crystallogr. D50*, 869−873.