Crystal Structure of the Human B-form Low Molecular Weight Phosphotyrosyl Phosphatase at 1.6-Å Resolution

The Journal of Biological Chemistry

Volume 281 • Number 10 • March 10, 2006

The crystal structure of HPTP-B, a human isoenzyme of the low molecular weight phosphotyrosyl phosphatase (LMW PTPase) is reported here at a resolution of 1.6 Å. This high resolution structure of the second human LMW PTPase isoenzyme provides the opportunity to examine the structural basis of different substrate and inhibitor/activator responses. The crystal packing of HPTP-B positions a normally surface-exposed arginine in a position equivalent to the tyrosyl substrate. A comparison of all deposited crystallographic coordinates of these PTPases reveals three atomic positions within the active site cavity occupied by hydrogen bond donor or acceptor atoms on bound molecules, suggesting useful design elements for synthetic inhibitors. A selection of inhibitor and activator molecules as well as small molecule and peptide substrates were tested against each human isoenzyme. These results along with the crystal packing seen in HPTP-B suggest relevant sequence elements in the currently unknown target sequence.

Tyrosine phosphorylation and dephosphorylation are critical components of eukaryotic signaling. The superfamly of protein-tyrosine phosphatases (PTPases) is defined in part by a canonical CX$_2$R(S/T) active site motif (P-loop), and can be divided into families based on substrate specificity and protein size (1–3). Following the classification scheme of Alonso et al. (1), the receptor-like PTPases such as CD45 (4) and non-receptor PTPases such as PTP1B (5) are subdivisions of the class I Cys-based family, multidomain proteins where the canonical phosphatase domain is roughly 30 kDa. The low molecular weight phosphatase domain is roughly 25 kDa. The low molecular weight phosphotyrosyl phosphatase; MES, 2-(bis(2-hydroxyethyl)amino)-2-(hydroxymethyl)ethanesulfonic acid; PLP, pyridoxal phosphate; pNPP, p-nitrophenyl phosphate; EGFP, epidermal growth factor protein; bis-Tris, 2-(bis(2-hydroxyethyl)aminomino)-2-(hydroxymethyl)propane-1,3-diol; STAT, signal transducer and activator of transcription. The abbreviations used are: PTPase, protein-tyrosine phosphatases; HPTP-A and -B, human red cell low molecular weight PTPase A and B isoenzymes; LMW PTPase, low molecular weight phosphotyrosyl phosphatase; MES, 2-(2-hydroxyethyl)ethanesulfonic acid; PLP, pyridoxal phosphate; pNPP, p-nitrophenyl phosphate; EGFP, epidermal growth factor protein; bis-Tris, 2-(bis(2-hydroxyethyl)aminomino)-2-(hydroxymethyl)propane-1,3-diol; STAT, signal transducer and activator of transcription.
**Materials and Methods**

Unless otherwise stated, all reagents were purchased from commercial suppliers and used without further purification.

**Wild-type Expression, Purification, and Crystallization—Cloning, expression, and purification of HPTP-A and HPTP-B were performed as previously described (7, 24). Analysis of purity by SDS-PAGE showed two impurities of lower molecular weight for the A isoform and three for the B isoform. These fragments are the result of Asp-Pro bond cleavage (25–27), and appear because of heating the sample prior to loading the gel. Unboiled samples did not show these bands (data not shown). HPTP-B was concentrated to 10 mg/ml and dialyzed into nanopurified water prior to all crystallization trials. Fine needle clusters appeared consistently across a relatively broad range of conditions, so variations in drop size and well-to-protein ratio were examined to produce diffraction quality crystals. Spontaneous overnight growth of small distinct plates (0.1 × 0.04 × 0.02 mm) was eventually obtained in a 6-μl hanging drop using 0.5 μl of the concentrated HPTP-B and 5.5 μl of a well solution consisting of 100 mM MES at pH 6.5 with 30% polyethylene glycol monomethyl ether (M₉ 5000) and 50 mM (NH₄)₂SO₄. Individual crystals were mounted in a cryoloop and flash frozen in liquid N₂ after a 10-s soak in a cryoprotectant solution of 90% well solution, 10% glycerol.

**Mutagenesis, Expression, and Purification of HPTP-B Mutant Proteins—**Mutant proteins were obtained by site-directed mutagenesis (28) of the wild-type HPTP-B coding sequence (7). Fragments of ~550 bp containing the HPTP-B gene were digested from the original pET vector and subcloned into the corresponding sites of the bacteriophage M13mp18. Mutagenesis was performed using a Mutagene M13 in vitro mutagenesis kit from Bio-Rad. A Sty1 restriction site was engineered into each mutagenic primer to facilitate initial screening. The mutant gene was digested using Ncol-BamHI restriction enzymes and subcloned into the pET-11d expression vector. The ligation mixture was transformed into the Escherichia coli strain DH5α, and individual colonies were sequenced to confirm the presence of the desired mutation. E. coli strain BL21(DE3) was used for overexpression of the mutant proteins.

Expression and purification of the mutant B-W49Y was accomplished by the standard LMW PTPase two-step purification scheme (24). For B-N50E, B-R53N, and B-W49Y/N50E the first step of the purification scheme used a hydrophobic interaction chromatography column. The lysate was loaded onto the hydrophobic interaction chromatography column using 1.5 M (NH₄)₂SO₄, after which the salt concentration was gradually reduced. The protein was eluted using 0.5 M (NH₄)₂SO₄ buffer. The protein was further purified using size exclusion chromatography on Sephadex G-50 as in the standard scheme. Yields for all mutants were at least 20 mg of protein per liter of culture.

**Data Collection and Structure Refinement—**Initial data collection was done on a Rigaku RU200 rotating anode x-ray generator with an R-AxisIV++ image plate detector. The preliminary refinement indicated a highly ordered crystal that diffracted to 2.0 Å. To maximize the number of observable reflections, high resolution data were collected on the 1D19D beamline at the Advanced Photon Source at 100 K and the frames were indexed and integrated using HKL2000 (29). HPTP-B crystallized in the monoclinic space group P2₁ with unit cell parameters a = 31.3 Å, b = 35.5 Å, c = 60.4 Å, and β = 100.0° and diffracted to 1.62 Å. The structure determination was performed with elements of the CCP4 package (30, 31). Molar replacement was performed using MOLREP (32) and the full coordinates for HPTP-A (Protein Data Bank code 5PNT (33)) as the search model to find initial phases. The R-factor for this initial solution was 40.2%. The final structure was established after eight cycles of manual adjustment followed by refinement using Refmac5.2 (34) with an Rwork of 15.8% and an Rfree of 21.5%. Water molecules were added using Refmac starting with the second round of refinement, deleting by hand those with high B factors or unassociated with the enzyme. A glycerol molecule was fit to a portion of density in the 2Fₒ – Fc map, not sufficiently explained by water. Starting with the sixth round of refinement, 10 residues were built with half-occupancy in two conformations. In each case, multiple rotamers were assigned only to those residues that exhibited reasonable density in each position as measured with a 2σ window. For B-N50E, B-R53N, and B-W49Y/N50E the first step of the purification scheme used a hydrophobic interaction chromatography column. The lysate was loaded onto the hydrophobic interaction chromatography column using 1.5 M (NH₄)₂SO₄, after which the salt concentration was gradually reduced. The protein was eluted using 0.5 M (NH₄)₂SO₄ buffer. The protein was further purified using size exclusion chromatography on Sephadex G-50 as in the standard scheme. Yields for all mutants were at least 20 mg of protein per liter of culture.

**Table 1**

| Crystal parameters | Z' (mol/asymmetric unit) | Matthews coefficient (Å³/Da) | Solvent (%) | Rmerge (%) |
|--------------------|--------------------------|-----------------------------|-------------|------------|
| Space group        | P₂₁                      | 1                           | 1.63        | 24.0       |
| a (Å)              | 31.3                     |                             |             |            |
| b (Å)              | 35.5                     |                             |             |            |
| c (Å)              | 60.4                     |                             |             |            |

**Data processing statistics**

- Temperature of measurement (K): 100
- Wavelength (Å): 1.0332
- Resolution (Å): 50–1.62 (1.68–1.62)
- Unique reflections: 15,887
- Redundancy: 2.71
- Completeness (%): 94.8 (79.5)
- Rmerge (%): 7.2 (42.4)
- (l/σ(l))²: 23.2 (2.49)

**Refinement statistics**

- Reflections with I > 1σ: 15,558
- Reflections, working set: 14,002
- Total atoms: 1,450
- Water molecules: 139
- R-factor, all reflections (%): 18.6
- Rwork (%): 15.8
- Rfree (%): 21.5
- Root mean square deviation, bond length (Å): 0.02
- Root mean square deviation, bond angle (°): 1.7
- Average protein B-factor (Å²): 17.8

**Ramachandran plot (%)**

- Most favored: 93.7
- Allowed: 6.3
**Crystal Structure of HPTP-B**

| Peptide*  | Sequence     | HPTP-A $k_{cat}/K_m$ (s⁻¹ μM⁻¹) | HPTP-B $k_{cat}/K_m$ (s⁻¹ μM⁻¹) |
|-----------|--------------|---------------------------------|---------------------------------|
| EGFR1092  | FLPVPEYMQRGS | $1.60 \times 10^2$             | $1.14 \times 10^2$             |
| EGFR1172  | SSLEFPQDDVF  | $8.88 \times 10^2$             | $8.28 \times 10^2$             |
| EGFR1197  | TAEYNMYQDDVF | $1.23 \times 10^2$             | $6.44 \times 10^2$             |
| B38       | QDVDYEDM     | $3.49 \times 10^1$             | $2.90 \times 10^1$             |
| Lck509    | TEGQYQQPOP   | $1.41 \times 10^0$             | $6.76 \times 10^0$             |
| Syk525    | DENTYFAQ     | $1.66 \times 10^0$             | $8.56 \times 10^0$             |
| Syk527    | LRSNYYDVN    | $2.79 \times 10^0$             | $4.04 \times 10^0$             |
| STAT1iso  | GPGKTVYKTELIS| $1.39 \times 10^{-1}$         | $2.18 \times 10^{-1}$         |
| STAT1iso  | QERKKLYHKRLLV | $9.45 \times 10^{-3}$        | $1.80 \times 10^{-3}$         |
| STAT2iso  | QERKKLYHKRLLV | $8.00 \times 10^{-3}$        | $9.00 \times 10^{-3}$         |
| AA01      | AAAAYA       | $3.10 \times 10^{-2}$         | $9.42 \times 10^{-2}$         |
| AA02      | AAAAYA       | $2.20 \times 10^{-2}$         | $8.76 \times 10^{-2}$         |
| AA03      | AAAAYA       | $1.22 \times 10^{-2}$         | $5.32 \times 10^{-2}$         |
| AA04      | DNTYAYA      | $7.88 \times 10^{-3}$         | $3.76 \times 10^{-3}$         |
| AA05      | AAYAADY      | $2.36 \times 10^{-2}$         | $1.05 \times 10^{-2}$         |

* Peptide names are based on the source protein: EGFR, epidermal growth factor receptor; B3, band 3; Lck, T-cell-specific protein-tyrosine kinase; Syk, spleen tyrosine kinase; STAT, signal transducer and activator of transcription; AA, poly-Ala based peptides. Subscript numbers refer to the tyrosine position in the source protein.

**Inhibition and Activation Studies**—Inhibition constants for HEPES, inorganic phosphate, pyridoxal phosphate (PLP), vanadate, and ZnCl₂ were determined at five different pNPP concentrations (0.1–5 $K_m$). At each substrate concentration, 10 different inhibitor concentrations were used to determine the initial velocity. A plot of $1/V$ versus (I) was made, and the $K_i$ (inhibition constant) value determined from the point where all lines intersected in the second quadrant (39).

The cGMP activation studies were performed under the same conditions as the inhibitor studies at a fixed concentration of pNPP (10 mM). Ten different cGMP concentrations were used, ranging from 0 to 2.4 mM. Reactions were quenched after 4 min and performed in duplicate. The data were fit to Equation 1 where $V_m'$ is the maximal activity at saturation, $M$ is the cGMP concentration, and $K_c$ is the apparent activation constant as shown in the equation.

$$V = 100 + \frac{V_m' [M]}{K_c' + [M]}$$  
(Eq. 1)

**Peptide Substrate Studies**—A selection of peptides (Table 2) was examined for substrate specificity against both isoenzyme forms by analogy to previous studies done with PTP1B (40, 41) and the rat LMW PTPase isoenzyme forms (42, 43). The PTP1B study showed that class I PTPs have a preference for sequences with anionic residues preceding the phosphotyrosine, whereas the rat isoenzyme study suggested the class II HPTP-B-like PTPs prefer hydrophobic residues to precede the phosphotyrosine. The AA series of phosphopeptides were synthesized at the Purdue University Protein Separation and Analysis Facility, whereas the remaining phosphopeptides were a generous gift from Dr. Chidambaran Ramachandran, Merck Sharp and Dohme. Phosphopeptide kinetics were conducted at 37 °C in 100 mM bis-Tris buffer, pH 7.0, with the ionic strength adjusted to 150 mM with NaCl. For each peptide, six concentrations ranging from 0.5 to 4.5 mM were prepared to a final volume of 45 μL. The reaction was initiated by the addition of 5 μL of either isoenzyme at a concentration of 0.045 mg/ml except for the peptide B38, which required an enzyme concentration of 0.7 mg/ml. Reactions were run for 6 min in all cases except for experiments with the signal transducer and activator of transcription (STAT) peptides, where the reactions were run for 10 min. The reaction was terminated by the addition of 450 μL of diluted Malachite Green assay reagent (see below) and the reaction was allowed to sit for at least 5 min before measuring absorbance at 620 nm. The phosphate concentration was determined by comparison to a standard curve for sodium phosphate.

Inorganic phosphate production was measured using a variation of the Malachite Green assay (44, 45) that was chosen over the standard method of Black and Jones (38) because the increased sensitivity made it possible to minimize peptide use. The Malachite Green assay reagent was prepared by combining 10 ml of Malachite Green dye concentrate (130 mM Malachite Green in 3.6 M sulfuric acid) with 2.5 ml of 7.5% ammonium molybdate and 0.2 ml of 11% Nonidet P-40. Diluted assay reagent was prepared by combining 100 μL of assay reagent with 350 μL of either water or bis-Tris buffer. The assay reagent was prepared daily and remained stable for several hours, whereas the diluted reagent was prepared immediately prior to use.

**RESULTS**

**Phosphatase Topology**—The overall fold of HPTP-B (Fig. 1) is essentially identical to other LMW PTPase crystal structures with an average Ca root mean square deviation of 0.79 Å. The HPTP-B structure consists of a pair of ββββ motifs, where the four β-strands form a continuous parallel β-sheet flanked on both sides by α-helices. The active site P-loop, residues 12–19, extends from the end of β1 to the start of α1. A 42-residue section between strands β2 and β3 contains two α-helices separated by relatively extended segments. This region forms one side of the active site and contains most of the variable region that distinguishes the HPTP isoenzymes. Following β4 is another extended loop that forms the other side of the active site and leads into a long 16-residue α-helix that ends just before the C terminus.

The active site forms a deep pocket in the surface of the enzyme, flanked by aromatic residues at the mouth and the three essential catalytic residues at the base. Two of the catalytic residues are part of the PTPase superfAMILY signature sequence, CX3C(R/S/T), in the characteristic P-loop conformation required to position the Cys and Arg for catalytic action. The cysteine accepts the phosphate group from the phosphotyrosine substrate to form the phosphoenzyme intermediate.
crystal structure of HPTP-B

The other side of the active site is capped by Trp49. Together these three aromatic residues help form a pocket deep enough to prevent catalytic His69 face each other in the crystal packing, but fail to form a plausible carboxyl rotamers and either of the histidine rotamers.

monomers are such that the side chains lie against the monomer surface that extends from the active site and is thought to contain the residues responsible for the differences in isoenzyme specificity. Residues with Partial Occupancy—At a resolution of 1.6 Å, the structure of HPTP-B affords us the opportunity to determine which rotamers, if any, could be exploited during inhibitor design. Ten residues were found that clearly exist in multiple rotamers. All but one of these residues are at least partially exposed to solvent. The buried residue, Ile77, is surrounded by hydrophobic residues Leu15, Val11, Phe82, Ile88, and Leu99, and the acyl portion of Lys102. Residues Leu29, Asn34, Glu37, and Leu125 are well separated from symmetry-related monomers and the multiple rotamers are likely the result of simple heterogeneity upon exposure to solvent. Other multiple positions are influenced by crystal packing. The two rotamers of Gln105 are both within hydrogen bonding distance of Asp129 on a symmetry related monomer. One rotamer of Lys131 has no intermolecular contacts, whereas the other rotamer forms a hydrogen bond to Ser7 O and Thr84 O of another monomer. Residues Glu92 and His59 face each other in the crystal packing, but fail to form a plausible hydrogen bonding pattern. The van der Waals contacts between the monomers are such that the side chains lie against the monomer surface, preventing an ideal hydrogen bond angle between either of the carboxyl rotamers and either of the histidine rotamers.

The most interesting of the multiple rotamers is Arg53 because it is the closest to the active site and has been suggested to play a significant role in determining isoenzyme specificity (33). This position is an Asn in HPTP-A, and mutation of this single amino acid in HPTP-B can be sufficient to change the effect of certain small molecule modulators, particularly if they are negatively charged (see below). Both rotamers of Arg53 are within a reasonable distance to form hydrogen bonds with at least one symmetry related molecule; one rotamer may form hydrogen bonds with two different monomers.

Crystal Packing at the Active Site—The electron density in and around the active site of HPTP-B demonstrated several interesting features. A well defined tetrahedral density above the P-loop was modeled as a bound sulfate because there was no phosphate in the crystallization solution (Fig. 2A). The backbone amide nitrogens of residues 13–18 sit within hydrogen bonding distance of one of the oxygen atoms in the sulfate, as do the Ne atom and one Nη atom of Arg49, to form a binding site compatible with what is needed for the terminal oxygen atoms of the phosphotyrosine substrate.

The most unique aspect of the crystal packing in HPTP-B is the full insertion of Arg101 from a symmetry-related monomer into the active site (Fig. 2B). Previously solved crystal structures of LMW PTPases contain a phosphate ion or buffer molecule (e.g. MES or HEPES) in the active site. The remainder of the active site is left open to solvent, or incorporates active site packing of the aromatic residues at the mouth of the active site. By comparison, Arg101 in HPTP-B is positioned directly above the sulfate with Nη-1 and Nη-2 hydrogen bonding to the oxygen that points away from the conserved CX5R(S/T) loop. The arginine also forms hydrogen bonds between Nη-2 and the side chain of the third catalytic residue, Asp129.

The sequential tyrosine residues Tyr131 and Tyr132 at one side of the active site interact with the Arg101 and Lys102 of their packed neighbor in a series of electrostatic π-based interactions (Fig. 2C). Consistent with previous structure determinations, the intramolecular interaction between Tyr131 and Tyr132 occurs in a T-shaped edge-to-face arrangement, perhaps the most common architecture for protein π-π interactions (49, 50). The insertion of Arg101 in the active site during crystal formation does not disrupt this architecture. Rather, it appears to be the core upon which a pair of cation-π interactions is formed. The first, an interaction between Lys102 and Tyr132, stacks the lysine Nζ atom almost directly above the aromatic ring of the tyrosine with a N-to-ring centroid distance of 3.5 Å. This is almost exactly the ideal theoretical value of 3.6 Å (51), and well within the 3.2–3.8 Å range of X-H/π hydrogen bond distances seen in other protein crystal structures (52). The second cation-π interaction is between Arg101 and Tyr131, and is shifted off-center with the closest contact between the arginine Ne and the tyrosine Cβ-2. The alignment of the two planes is roughly parallel, in agreement with other crystallographic examples of cation-π stacking (53), with a Cζ to ring centroid distance of 4.3 Å.

MARCH 10, 2006•VOLUME 281•NUMBER 10 JOURNAL OF BIOLOGICAL CHEMISTRY 6523
stable positions of the aromatic residue at position 49, the human isoenzymes might be expected to have differing sensitivities toward inhibitors with cyclic substituents. To probe this possibility, the $K_I$ values of HPTP-A and HPTP-B were determined at pH 5.0 using the heterocyclic inhibitors HEPES and PLP. Three ionic species known to be inhibitors in other PTPs were also tested: inorganic phosphate, vanadate, and ZnCl$_2$ (Table 4). Vanadate, a covalent inhibitor, was six times more effective against HPTP-A. More modest isozyme specificity was seen with PLP and HEPES, both of which were three to four times more effective against HPTP-A. Inorganic phosphate, HEPES, and ZnCl$_2$ were poor inhibitors, with $K_I$ values in the millimolar range, whereas vanadate and PLP had $K_I$ values in the micromolar range.

Activation studies of the wild-type isoenzymes and the HPTP-B mutants by cGMP were performed using the procedure outlined by Dissing et al. (19). The activity at saturation, $V_{	ext{max}}^*$, is presented as an increased percentage in activity relative to what is seen without cGMP. An approximate 1000% increase in activity was calculated for HPTP-B and B-R53N. For HPTP-A, B-W49Y/N50E, B-W49Y, and B-N50E, the activation at saturation was 44, 14, 467, and 429%, respectively (Fig. 3).

Peptide Substrates—Activities for both human isoenzymes were tested at pH 7.0 against five sets of synthetic phosphopeptides (Table 2). The first three sets are based on cellular protein sequences containing sites of tyrosine phosphorylation: autophosphorylation sites on the epidermal growth factor receptor (EGFR), regulatory sites on Band 3 or the dermal growth factor receptor (EGFR), and regulatory sites on STAT proteins 1 and 2. Increased positive charge to the carboxy-terminal side of the tyrosine was further examined by the insertion of a second Arg residue at the Y + 4 position in STAT2R. Peptides are numbered by the position of the tyrosine (the first tyrosine for the Syk kinase, followed by another Arg residue at the Y + 4 position in STAT2R).

The ratio $k_{cat}/K_m$ is an apparent second-order rate constant and describes the specificity of an enzyme for a given substrate. Using phosphotyrosine, the $k_{cat}/K_m$ values for HPTP-B and HPTP-A were $1.8 \times 10^3$ and $19 \times 10^3$ s$^{-1}$ M$^{-1}$, respectively. All four mutants showed altered specificity, with the specificity of B-R53N greater than the parent enzyme, $4.2 \times 10^5$ s$^{-1}$ M$^{-1}$. By comparison, the mutants B-W49Y, B-N50E, and B-W49Y/N50E showed increased specificity over the wild-type, $3.1 \times 10^5$, $6.9 \times 10^4$, and $2.4 \times 10^5$ s$^{-1}$ M$^{-1}$, respectively. The increased $k_{cat}/K_m$ values for B-N50E and B-W49Y over HPTP-B suggest these residues improve substrate specificity. Consistent with this, the B-W49Y/N50E double mutant has a $k_{cat}/K_m$ value on the same order as HPTP-A. The use of pNPP as the substrate achieves similar results with the single mutants. Interestingly, the B-W49Y/N50E double mutant shows specificity more like HPTP-B than either B-W49Y or B-N50E.

Inhibition and Activation Studies—The crystal structures of HPTP-A and HPTP-B show that residue 49 is involved in the formation of a hydrophobic wall of the active site, but Tyr$^{49}$ in HPTP-A has a different position than Trp$^{49}$ in HPTP-B (33). In HPTP-A the side chain of Tyr$^{49}$ flips back into the active site pocket and stacks face to edge with the morpholino ring of the bound MES, whereas the side chain of Trp$^{49}$ in HPTP-B does not interact with the arginine residue inserted into the active site from the symmetry related monomer. If these are the most

![Crystal Structure of HPTP-B](image-url)
In an attempt to predict which amino acid residues adjacent to the phosphotyrosine might provide favorable enzyme-substrate interactions, a baseline alanine hexamer was created along with four additional peptides that could potentially interact with surface residues on the enzyme.

To assess the peptides as substrates, $k_{cat}/K_m$ was determined at pH 7.0 for each peptide. The triple tyrosine peptide Syk629 was the best overall substrate at pH 7.0, and none of the tested peptides were noteworthy substrates. With the exception of Syk629, all peptides showed a preference for HPTP-A over HPTP-B, but only the STAT peptides demonstrated better than a 5-fold preference. Interestingly, the poly-Ala peptide AA01, designed to be a baseline from which the other rational peptides could be evaluated, was in fact the best substrate of that set for HPTP-A and only slightly less effective than AA05 for HPTP-B.

**DISCUSSION**

Although HPTP-B has the expected LMW PTPase protein fold, clear differences in charge distribution around the active site and the observation of multiple rotamers in HPTP-B make possible new inferences about isoenzyme specificity. Whereas there is minor variation in the shape of the human isoenzyme surfaces, the most obvious difference between HPTP-A and HPTP-B is the change in the surface charges near the active site (Fig. 4). Residues 50 and 53 (Glu and Asn in HPTP-A, Asn and Arg in HPTP-B) form the lip of the active site leading to the proposed specificity cleft where the peptide substrate is predicted to bind. The negative or neutral charge at position 50 coupled with the neutral or positive charge at position 53 should make a significant contribution to substrate specificity or efficacy of small molecule modulators directed against the two human isoenzymes. Because Arg53 is a surface-exposed cationic side chain, the two rotamers in the crystal structure emphasize the flexibility this residue can have. The multiple rotamers may be the result of an attempt to satisfy conflicting hydrogen bond partners, because one rotamer is extended away from the protein surface to contact two symmetry related molecules while the other folds along the protein surface. The fact that neither position is favored in the crystal suggests a highly mobile residue that may be important for drug design, offering related but independent surfaces for a given compound.

The other significant feature of the active site is the pair of sequential aromatic residues Tyr131 and Tyr132, which form the left side of the active site as shown in Fig. 4. The intramolecular, T-shaped $\pi-\pi$ interaction of the two aromatic residues is well maintained in all known crystal structures of the LMW PTPase family, even in the case of the yeast form of the enzyme where the second tyrosine is replaced by a tryptophan. The centroid-to-centroid distance for these aromatic residues in HPTP-B is 5.0 Å, identical to that seen for the lowest energy, T-shaped interaction between molecules in the benzene crystal structure (56). In the HPTP-B crystal packing, these two tyrosine residues are also involved in cation–$\pi$ stacking to Arg101 and Lys102. Interestingly, the intramolecular cation–$\pi$ interaction does not seem to affect the intramolecular $\pi-\pi$ interaction, because previous crystal structures retain the same angle and distance between the aromatic pairs. Assuming that Arg101 occupies the place of the tyrosyl substrate, the crystal packing interactions at the active site of HPTP-B may mirror the interactions of the natural substrate. This suggests that a positively charged residue occupying the same position as Lys102 should follow the phosphotyrosine. A similar analysis of crystal packing interactions based on a mutant bovine form of the protein (PDB code 1COE (57)) suggested the phosphotyrosine substrate would be preceded by an aromatic residue to establish the $\pi-\pi$ interactions.

For the competitive small inhibitors PO4 or ZnCl2 that should not extend to the mouth of the active site, we see modest to poor efficacy and little difference between the isoenzymes. The larger competitive inhibitors HEPES and PLP show a 3–6-fold degree of difference between isoenzymes, suggestive of the important role that residues at the top of the active site play in isoenzyme specificity. Three residues that deviate between the two isoenzymes were investigated for their possible role in substrate specificity: Trp49, Asn50, and Arg53 in the B isoenzyme, and Tyr49, Glu50, and Asn53 in the A isoenzyme. Mutations of HPTP-B were made at these residues to the corresponding ones in HPTP-A. These changes were sufficient to affect the specificity of small molecule substrate mimics, particularly with phosphotyrosine. Although B-R53N contained the mutation farthest removed from the active site, this mutant enzyme showed a decrease in specificity relative to HPTP-B. Because residues 50 and 53 hydrogen bond to each other in both wild-type forms, it is likely that the reduced chain length and modified charge distribution in the B-R53N mutation affects the orientation of Asn50, which in turn causes a negative effect on substrate specificity. The single mutants B-W49Y and B-N50E exhibited $k_{cat}/K_m$ values that were between those of HPTP-B and HPTP-A, whereas the double mutant B-W49Y/N50E exhibited a $k_{cat}/K_m$ value very close to that of HPTP-A, suggesting that only these two mutations are necessary to interconvert between A-type and B-type specificity. Activation by cGMP followed a similar pattern, where $V_{max}$ for B-R53N was virtually identical to HPTP-B and the maximal activities at saturation for B-W49Y and B-N50E were reduced to roughly half of the wild-type isoenzyme. The B-W49Y/N50E double mutant, like HPTP-A, exhibited little cGMP activation. These results are consistent with earlier structural data (33) that suggested residues at the mouth of the active site would be important for substrate recognition, and show that these residues are involved in tyrosine-specific recognition.

Previous work on the rat isoenzymes at pH 5.5 with a series of phosphotyrosine-containing peptides (42, 43) failed to find a good substrate...
for the ACP1 (HPTP-A like) isoform and provided only limited information about the ACP2 (HPTP-B like) isoform. To examine reasonable substrates at a physiological pH, we conducted our peptide survey at pH 7.0 against the human isoenzymes. Unfortunately, even our best peptide, Syk629, is 4–5 orders of magnitude poorer than the ideal peptides for class I PTPases (40, 41). Whereas this may be due in part to making the measurements at a pH above the enzymatic optimum, it is much more likely to be an indication that, as with the rat isoenzymes, none of the peptides match the sequence of the natural substrate.

The results with the synthetic peptide Syk629 coupled with the rational peptides AA01–AA05 suggest relevant elements for an ideal phosphopeptide substrate. Replacing either terminal residue of AA01 with an Asp resulted in a negligible change of activity versus HPTP-B but slightly decreased the activity versus HPTP-A, indicating a need to avoid a negative charge distant from the catalytic site of the A isoenzyme. HPTP-A has comparable catalytic activity for peptide Syk629 and AA01, whereas HPTP-B is roughly four times more active against Syk629, demonstrating that the change in sequence has little effect on the activity for the former but has a significant contribution for the latter. Although the sequential phosphorylated tyrosine residues of Syk629 are certainly the most noteworthy characteristics of this peptide, comparing the presence of both cationic and anionic residues in Syk629 with the presence of only anionic residues in AA02 and AA05 suggests that a cationic residue would increase activity with HPTP-B.

Recent studies have suggested that the ephrin receptor EphA2 may be the physiological substrate for the HPTP enzyme (22), and several of the 17 tyrosine residues of the cytoplasmic segment of EphA2 within the kinase domain, the juxtamembrane region, and the SAM domain have all been proposed to regulate the activity of the Eph family (58, 59). Based only on the structural evidence from HPTP-B and assuming Arg101 is acting as a phosphotyrosine mimic in the present structure, we propose the natural substrate for this isoenzyme would favor a cationic residue in either the −1 or +1 position. This results in only four possible substrate sites, all within the kinase domain of EphA2. Examination of the kinase domain structure (PDB code 1MQB (60)) shows the most accessible of these four residues to be Ty⁶⁸⁵, flanked on either side by a lysine and part of an extended loop on the N-terminal lobe of the kinase on the opposite face from the ATP binding site.

A comparison of all known LMW PTPase crystal structures reveals a useful triad of hydrogen bond donor or acceptor atoms that should be significant in the design of small molecule inhibitors. A subset of these structures (PDB codes 1C0E (57), 5PNT (33), and 1D2A (20)) is shown in Fig. 2D, with the three highly conserved atomic positions of a bound substrate circled in red and connected to their hydrogen bonding partner on the enzyme by yellow bars. Although there does not seem to be any specific preference for the small molecule to contain donor or acceptor atoms at these positions, every crystal structure fills at least one position by a potential hydrogen bonding atom. Previous inhibitor design in our laboratory has taken advantage of only one of these positions, a hydrogen bond donor intended to interact with Asp¹²⁹ and based on the orientation of adenine bound to the yeast form of the enzyme (20). These early inhibitors had relatively low binding affinity in the high micromolar to low millimolar range, but revealed interesting structural requirements for binding (23). These results, together with the structural insights from this crystal structure have led us to propose a new generation of inhibitors using similar molecular scaffolds, but incorporating a variation of the crystal contact interactions at the base, midpoint, and entrance to the active site. The synthesis, refinement, and binding specificity tests to exploit the structural differences between PTPase subfamilies of these second-generation inhibitors are currently under investigation and will be reported separately.

REFERENCES
1. Alonso, A., Sasin, J., Bottini, N., Friedberg, L., Osterman, A., Godzik, A., Hunter, T., Dixon, J., and Mustelin, T. (2004) Cell 117, 699–711
2. Zhang, Z.-Y. (2003) Acc. Chem. Res. 36, 385–392
3. Staufacher, C. V., and Charbonneau, H. (2000) in Principles of Molecular Regulation (Conn, P. M., and Means, A. R., eds) pp. 323–347, Humana Press, Totowa, NJ
4. Dahlke, M. H., Larsen, S. R., Rasko, J. E. J., and Schlitt, H. J. (2004) Leuk. Lymphoma 45, 229–236
5. Tonks, N. K. (2003) FEBS Lett. 546, 140–148
6. Rangei, G., Ramponi, G., and Chiarugi, P. (2002) Cell Mol. Life Sci. 59, 941–949
7. Wu, Y. Y. P., McCormack, A. L., Shabanowitz, J., Hunt, D. F., Davis, J. P., Mitchell, G. L., and Van Etten, R. L. (1992) J. Biol. Chem. 267, 10856–10865
8. Park, E. K., Warner, N., Mood, K., Pawson, T., and Daar, I. O. (2002) Mol. Cell. Biol. 22, 3404–3414
9. Miller, D. T., Read, R., Rusconi, J., and Cagan, R. L. (2000) Gene (Amst.) 243, 1–9
10. Saeed, A. (1999) J. Chem. Soc. Pakistan 21, 311–320
11. Manao, G., Pazzaglia, L., Cirri, P., Caselli, A., Camici, G., Cappugi, G., Saeed, A., and Ramponi, G. (1992) J. Protein Chem. 11, 333–345
12. Fujimoto, S., Murakami, K., Ishikawa, A., Himi, K., and Ohara, A. (1988) Chem. Pharm. Bull. (Tokyo) 36, 3020–3026
13. Baxter, J. H., and Suelter, C. H. (1985) Arch. Biochem. Biophys. 239, 29–37
14. Thomas, C. L., McKinnon, E., Granger, B. L., Harms, E., and Van Etten, R. L. (2002) Biochemistry 41, 15601–15609
15. Modesti, A., Cirri, P., Rangei, G., Carrarese, L., Magherini, F., Manao, G., Camici, G., and Ramponi, G. (1995) FEBS Lett. 375, 235–238
16. Ostain, K., Pokalsky, C., Wang, S., and Van Etten, R. L. (1995) J. Biol. Chem. 270, 18491–18499
17. Zhang, Z. Y., and Van Etten, R. L. (1990) Arch. Biochem. Biophys. 282, 39–49
18. Madhurantakam, C., Rajakumara, E., Mazumdar, P. A., Saha, B., Mitra, D., Wilker,
