Crystal structures and inhibitor identification for PTPN5, PTPRR and PTPN7: a family of human MAPK-specific protein tyrosine phosphatases

Jeyanthy ESWARAN*, Jens Peter von KRIES†, Brian MARSDEN*, Emma LONGMAN*, Judit É. DEBRECZENI*, Emilie UGOCHUKWU*, Andrew TURNBULL*, Wen Hwa LEE*, Stefan KNAPP*[1] and Alastair J. BARR*

*Structural Genomics Consortium, University of Oxford, Botnar Research Centre, Oxford OX3 7LD, U.K., and †Screening Unit, Forschunginstitut für Molekulare Pharmakologie, Robert-Roessle Strasse 10, 13125 Berlin, Germany

Protein tyrosine phosphatases PTPN5, PTPRR and PTPN7 comprise a family of phosphatases that specifically inactivate MAPKs (mitogen-activated protein kinases). We have determined high-resolution structures of all of the human family members, screened them against a library of 24 000 compounds and identified two classes of inhibitors, cyclopenta[c]quinolinecarboxylic acids and 2,5-dimethylpyrrolyl benzoic acids. Comparative structural analysis revealed significant differences within this conserved family that could be explored for the design of selective inhibitors. PTPN5 crystallized, in two distinct crystal forms, with a sulphate ion in close proximity to the active site and the WPD (Trp-Pro-Asp) loop in a unique conformation, not seen in other PTPs, ending in a 3_{10}-helix. In the PTPN7 structure, the WPD loop was in the closed conformation and part of the KIM (kinase-interaction motif) was visible, which forms an N-terminal aliphatic helix with the phosphorylation site Thr[66] in an accessible position. The WPD loop of PTPRR was open; however, in contrast with the structure of its mouse homologue, PTPSL, a salt bridge between the conserved lysine and aspartate residues, which has been postulated to confer a more rigid loop structure, thereby modulating activity in PTPSL, does not form in PTPRR. One of the identified inhibitor scaffolds, cyclopenta[c]quinoline, was docked successfully into PTPRR, suggesting several possibilities for hit expansion. The determined structures together with the established SAR (structure–activity relationship) propose new avenues for the development of selective inhibitors that may have therapeutic potential for treating neurodegenerative diseases in the case of PTPRR or acute myeloblastic leukaemia targeting PTPN7.

Key words: crystal structure, phosphatase inhibitor, protein tyrosine phosphatase, PTPN5, PTPN7, PTPRR.

INTRODUCTION

Reversible phosphorylation of tyrosine residues is a key regulatory mechanism for numerous cellular events such as proliferation, differentiation, gene expression and migration [1]. There are at least 107 genes coding for PTPs (protein tyrosine phosphatases) in the human genome, which cluster into four evolutionarily distinct families: the class I, II, and III cysteine-based PTPs, and the aspartate-based phosphatases [2,3]. The class I cysteine-based PTPs form, with 99 members, the largest PTP gene family. To date, about half of all PTP genes in the human genome have been implicated in human disease or are recognized as potential drug targets [4,5]. The phenotype of the PTP1B-knockout mouse, which demonstrated an increased response to insulin and resistance to high-fat-diet-induced obesity [6], stimulated many inhibitor-development programmes both in academia as well as in private industry. Since then the increasing number of human diseases associated with deregulation of phosphatases has led to a growing interest in PTP inhibitor development [7].

PTPN5, PTPRR and PTPN7 comprise a family of PTPs that are distinguished by the presence of a 16-amino-acid KIM (kinase-interaction motif) that binds specifically to members of the MAPK (mitogen-activated protein kinase) family. These PTPs are highly specific in their substrate recognition preferring the MAPKs ERK1/2 (extracellular-signal-regulated kinase 1/2) and p38, over JNK (c-Jun N-terminal kinase) [8]. Other reported substrates include ERK5 for the mouse homologue of PTPRR [9] and the Src family kinase Fyn for PTPN5 [10]. Tyrosine dephosphorylation in the activation loop of the MAPK causes inactivation of the kinase and blocks its nuclear translocation [8]. In a reciprocal manner, KIM domain phosphatases are also substrates of ERK1/2 and are phosphorylated on a threonine residue in the KIM. The activity of these three phosphatases is regulated further by phosphorylation of the KIM by protein kinase A, which reduces the affinity for MAPK substrates [11,12].

PTPN5 (also named STEP for striatal-enriched PTP) is preferentially expressed in neurons of the central nervous system [13,14], where it regulates dopaminergic and glutaminergic neurotransmission [15,16]. The related phosphatase PTPRR (also named PCPTP1, PTPSL or PTPBR7), is also expressed predominantly in brain, but has, in addition, been detected in cartilage and may play a role in bone morphogenesis [17]. Alternative splicing produces transmembrane and cytosolic variants of both PTPN5 and PTPRR. Studies in PC12 (pheochromocytoma) cells have indicated the involvement of PTPRR in nerve growth factor signalling, suggesting that it is a target for the treatment of neurodegenerative processes such as Alzheimer's disease [18,19]. In contrast, PTPN7 (also named HePTP for haematopoietic PTP) is expressed mainly in thymus, spleen and leucocytes. This phosphatase plays a negative role in TCR (T-cell antigen receptor) signalling by down-regulating MAPK activity [20–22]. Interestingly, PTPN7 is found at chromosome locus 1q32.1, a site

---

of frequent abnormalities in preleukaemic myelodysplastic syndrome, as well as in other haematopoietic malignancies [23]. PTPN7 has been shown to be overexpressed in some patients with acute myeloblastic leukaemia, suggesting a linkage of increased PTPN7 activity to this disease [24].

To date, the X-ray crystallographic structure of catalytic domains from several non-receptor PTPs, such as PTP1B [25], and receptor PTPs have been determined, including the mouse homologue of PTPRR (PTPSL) [26]. Recently, the structure of PTPN7 has been reported [27]. Structural studies, together with enzyme kinetic studies, of the PTP family prototype PTP1B have provided important insights into the mechanism of substrate recognition and catalysis (reviewed in [28]). The PTPs have a signature motif, (I/V)HCXAGXGR(S/T), containing the catalytically essential cysteine and arginine residues that form a rigid cradle-like structure that co-ordinates the phosphate moiety of the substrate. This motif lies at the base of a cleft surrounded by four loops. One of these loops, the WPD (Trp-Pro-Asp) loop, undergoes a large conformational change closing over the active site on substrate binding and determines the size of the active-site cavity. Another loop surrounding the PTP1B active site, which has the YRD (Tyr-Arg-Asp) motif, is important for substrate recognition; however, in the KIM-containing PTPs, this sequence is YKT (Tyr-Lys-Thr).

PTP inhibitor screening has focused predominantly on the identification of PTP1B-specific compounds and inhibitors for the CDC25 family; however, to date, there is a lack of information on inhibitors of other PTPs, even though many of these enzymes have been implicated as therapeutic targets. In addition, specific inhibitors would also have the potential to be used as pharmacological tools to elucidate functional roles of the targeted enzymes. In order to provide insight into the mechanism of action of KIM-containing PTPs, and facilitate identification of selective inhibitors, we have determined the high-resolution structures of all catalytic domains of the human KIM-containing PTP family, PTPN5, PTPRR and PTPN7, and compared the three dimensional structures and catalytic activities of these enzymes. Furthermore, we screened the human KIM-containing PTP family and PTP1B against a library of 24 000 compounds and identified a number of novel inhibitors. Among those, we discuss here two classes of inhibitors: cyclopenta[c]quinolinecarboxylic acids and 2,5-dimethylpyrrolyl benzoic acids. The generated screening data, together with the high-resolution structural information and docking experiments, allowed us to rationalize the binding mode and establish the SARs (structure–activity relationships) for these two inhibitor classes.

MATERIALS AND METHODS

Cloning

A cDNA encoding the human full-length PTPN5 (NCBI accession number NP_116170) was obtained from Origene and was used as a template to amplify a sequence encompassing residues 282–563. This sequence corresponds to residues 258–539 of Swiss-Prot entry P54829 with the changes $Asp^{289} \rightarrow Val$, $Leu^{298} \rightarrow Arg$, $Val^{299} \rightarrow Cys$ and $Thr^{517} \rightarrow His$, which have been documented. The PCR product was subcloned into a pET-21a-derived vector, pLIC-SGC, by ligation-independent cloning [29]. The vector includes a TEV (tobacco etch virus)-cleavable N-terminal His$_6$ tag (MHHHHHHSSGVDLGTENLYFQ*SM). Plasmids with the above PTPN5 sequence; a sequence encompassing the catalytic domain of PTPRR (residues 375–655; NCBI accession number NP_002840), and a sequence encompassing the catalytic domain of PTPN7 (residues 79–358; NCBI accession number NP_002823) subcloned into the SmaI site of pGEX-6P2 were

obtained from Purely Proteins. The PTPN7 clone was used as a template to amplify a sequence encompassing residues 65–358 and was subcloned into a pLIC-SGC-derived kanamycin-resistant vector, pNIC-Bsa4, using primers that extended the existing boundary.

Expression and purification

Escherichia coli BL21(DE3) cells transformed with the expression constructs were grown at 37 °C in Luria–Bertani medium containing 100 μg/ml ampicillin or kanamycin until the D_{600} reached $\sim$ 0.6 and were then transferred to 18 °C. Protein expression was induced at a D_{600} of 0.8 using 1 mM IPTG (isopropyl β-D-thiogalactoside). Cells were harvested after either 4 h or overnight growth.

Cells expressing His$_6$-tagged PTPN5 were lysed in 50 mM Hepes, pH 7.5, 500 mM NaCl, 1 mM PMSF and 0.5 mM TCEP [tris-(2-carboxyethyl)phosphine] using an EmulsiFlex high-pressure homogenizer, and the cell extract was centrifuged at 20 000 **g** for 30 min at 4 °C. The supernatant was loaded on to a 1 ml Ni^{2+}–Sepharose affinity HiTrap HP column on an ÄKTAxpress system (GE Healthcare). The column was washed with 10 vol. of wash buffer (50 mM Hepes, pH 7.5, 500 mM NaCl, 30 mM imidazole, 0.5 mM TCEP and 5 % glycerol), then eluted with elution buffer (50 mM Hepes, pH 7.5, 500 mM NaCl, 250 mM imidazole, 0.5 mM TCEP and 5 % glycerol). The eluted peak at A_{280} was applied to a Superdex 200 16/60 gel-filtration column equilibrated in 50 mM Hepes, pH 7.5, 500 mM NaCl, 0.5 mM TCEP and 5 % glycerol. The protein eluted from the gel-filtration column was diluted 10-fold in buffer A [50 mM Hepes, pH 7.5, and 10 mM DTT (dithiothreitol)] and was purified further by ion-exchange chromatography using a 1 ml HiTrapQ column. The protein was eluted with a 20 min linear gradient to 50 % buffer B (50 mM Hepes, pH 7.5, 1.0 M NaCl and 10 mM DTT). His$_6$-tagged PTPN7 was purified using a method similar to that used for PTPN5 with some changes. The cell lysate was passed over a DEAE-cellulose resin before applying to a Ni^{2+}–Sepharose (Qiagen) gravity flow column. The protein was eluted with a range of imidazole step elutions. Samples containing PTPN7 were digested with TEV protease overnight at 4 °C. After exchanging the buffer with gel-filtration buffer (50 mM Hepes, pH 7.5, 150 mM NaCl and 0.5 mM TCEP) and re-applying the digested protein to Ni^{2+}–Sepharose, the cleaved protein was purified further by gel filtration on an S200 16/60 HiLoad column equilibrated in 50 mM Hepes, pH 7.5, 150 mM NaCl and 0.5 mM TCEP.

Cells expressing GST (glutathione S-transferase) fusions of PTPN5 or PTPRR were lysed as above in 50 mM Tris/HCl, pH 7.5, 150 mM NaCl and 5 mM DTT, the supernatant was loaded on to 5 ml of glutathione–Sepharose, and the washed resin was mixed gently overnight at 4 °C with GST-tagged PreScission protease ($\sim$ 50 μg per mg of protein bound to the glutathione–Sepharose). The cleaved PTPN5 and PTPRR proteins were eluted and purified further by gel-filtration chromatography using the above buffer.

The purified proteins were homogeneous as assessed by SDS/PAGE and electrospray MS, which also confirmed the predicted mass of the proteins. Proteins were concentrated to 7–10 mg/ml using a 10 kDa cut-off concentrator (Vivascience).

Crystallization

Crystallization was performed using the sitting-drop method, mixing protein and precipitant solutions in 2:1, 1:1 and 1:2 ratios at 4 °C. Crystals of PTPN5 were obtained at 4 °C in conditions containing 25 % PEG [poly(ethylene glycol)]-3350, 0.2 M LiSO$_4$ and 100 mM Bis-Tris, pH 5.5. The crystal forms obtained with

Table 1 Crystallographic data and refinement statistics

Rmsd, root mean square deviation.

(a)

Data collection	PTPN5 (1)	PTPN5 (2)	PTPRR	PTPN7
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12$	$P2_12_12_1$
Cell dimensions (Å)	51.81, 64.32, 107.08	39.96, 64.01, 136.15	63.19, 74.10, 62.33	39.1, 81.0, 100.4
Resolution (Å)	2.0	1.8	2.3	2.5
Total observations (unique, redundancy)	103 306 (20 825, 4.71)	165 645 (32 877, 4.97)	46 787 (13 031, 3.6)	38 503 (10 922, 3.5)
Completeness (outer shell)	95.0 % (97.4 %)	98.7 % (95.4 %)	97 % (98.4 %)	99.4 % (96.6 %)
R_{merge}	0.086	0.082	0.105	0.093
I/σ (outer shell)	11.4 (3.05)	12.9 (2.5)	5.2 (2.2)	10.5 (2.0)

(b)

Refinement	PTPN5 (1)	PTPN5 (2)	PTPRR	PTPN7
R_{work} (R_{free}) (%)	21.2 (26.4)	16.4 (20.1)	19.2 (25.6)	21.9 (27.5)
Protein atoms (water)	2272 (103)	2314 (271)	2353 (155)	2134 (32)
Hetero groups:	Sulphate	Glycerol, sulphate	Chlorine	Phosphate
Rmsd bond length (Å)	0.015	0.011	0.011	0.011
Rmsd bond angle (°)	1.484	1.322	1.317	1.270
Ramachandran (allowed, disallowed) (%)	99.6, 0.4	99.6. 0.4	99.6, 0.4	98.5, 1.5

the His$_6$-tagged PTPN5 are referred to as PTPN5(1) and those obtained with the cleaved GST fusion protein are referred to as PTPN5(2). Crystals of PTPRR and PTPN7 were obtained in precipitant solutions containing 0.15 M malic acid, pH 7.0, 20 % PEG-3350 and 2.0 M (NH$_4$)H$_2$PO$_4$, 0.1 M Tris/HCl, pH 8.5, respectively at 20 °C.

Data collection and processing

PTPN5(1) data were collected using a Rigaku FRE rotating anode equipped with Varimax multilayer mirrors and a Rigaku HTC detector to 2.05 Å (1 Å = 0.1 nm). PTPN5(2) data were measured at the BL1 beamline at BESSY (Berliner Elektronenspeicherring-Gesellschaft für Synchrotronstrahlung m.b.H., Berlin, Germany). Data collection was performed on flash-frozen crystals at 100 K, with 15 % glycerol used as cryoprotectant. Diffraction images of both PTPN5(1) and PTPN5(2) were processed with HKL2000 [30]. The PTPRR dataset was collected at the beamline X10 at SLS (Swiss Light Source, Paul Scherrer Institut, Villigen, Switzerland) to a resolution of 2.3 Å. Images were indexed and integrated using MOSFLM [31], and scaled using SCALA [32] in the CCP4 [33] suite of programs. PTPN7 data were collected at beamline X10 at the SLS to a maximum resolution of 2.5 Å, and these data were reduced using HKL2000. Data collection statistics and cell parameters are listed in Table 1.

Structure solution and refinement

All structures were solved with molecular replacement using Phaser [34] with the mouse homologue of PTPRR (PDB code 1JLN) as a search model. Iterative rounds of rigid-body refinement and restrained refinement using TLS tensors, against maximum likelihood targets, were interspersed by manual rebuilding of the model using Coot and Xfit/XtalView [35].

Docking procedure

The ICM docking method [36] was used for all docking experiments. This approach performs a flexible ligand/grid receptor docking process using potential grids which are pre-calculated (including van der Waals, electrostatic, hydrogen bond and hydrophilic potentials). These are then applied to the molecule to be docked and the side chains of the receptor site using a global optimization method [37,38]. The ability of protein side chains within the receptor site to rotate helps to relieve bad contacts and also promote positive interactions between the docked ligand and protein. The docking process involves two main steps: first with only the ligand flexible, and, secondly, with the ligand and receptor side chains flexible (refinement). The structure of PTPRR (PDB code 2A8B) was used as the template to which compounds were docked. Docked conformations of the compounds were assessed by their occupancy of the active site and interactions with active-site residues, in particular Cys[588] of PTPRR.

High-throughput screening and compound library

A library of 24 000 compounds was screened as described previously in a 384-well format using pNPP (*p*-nitrophenyl phosphate) as the substrate [39]. The compound collection is a fragment-based library selected because of its diverse representation of reportedly bioactive scaffold elements and compliance with physicochemical boundaries, which include the Lipinski rules.

RESULTS AND DISCUSSION

Multiple phosphatases are involved in regulating the degree and duration of MAPK activation, which are critical parameters in determining the cellular response to a signal. The family of PTPs (PTPN5, PTPRR and PTPN7) is responsible for the rapid inactivation of MAPKs in the cytosol, thereby blocking kinase nuclear translocation, which is crucial for certain cellular functions, such as learning processes in neurons as well as in the haematopoietic system in response to inflammation. While all PTP family members share a high degree of sequence homology within their phosphatase domains, differences in their functions and substrate-specificity prompted us to compare all three enzymes of this subfamily on a structural level and initiate a high-throughput small-molecule screen to identify inhibitors.
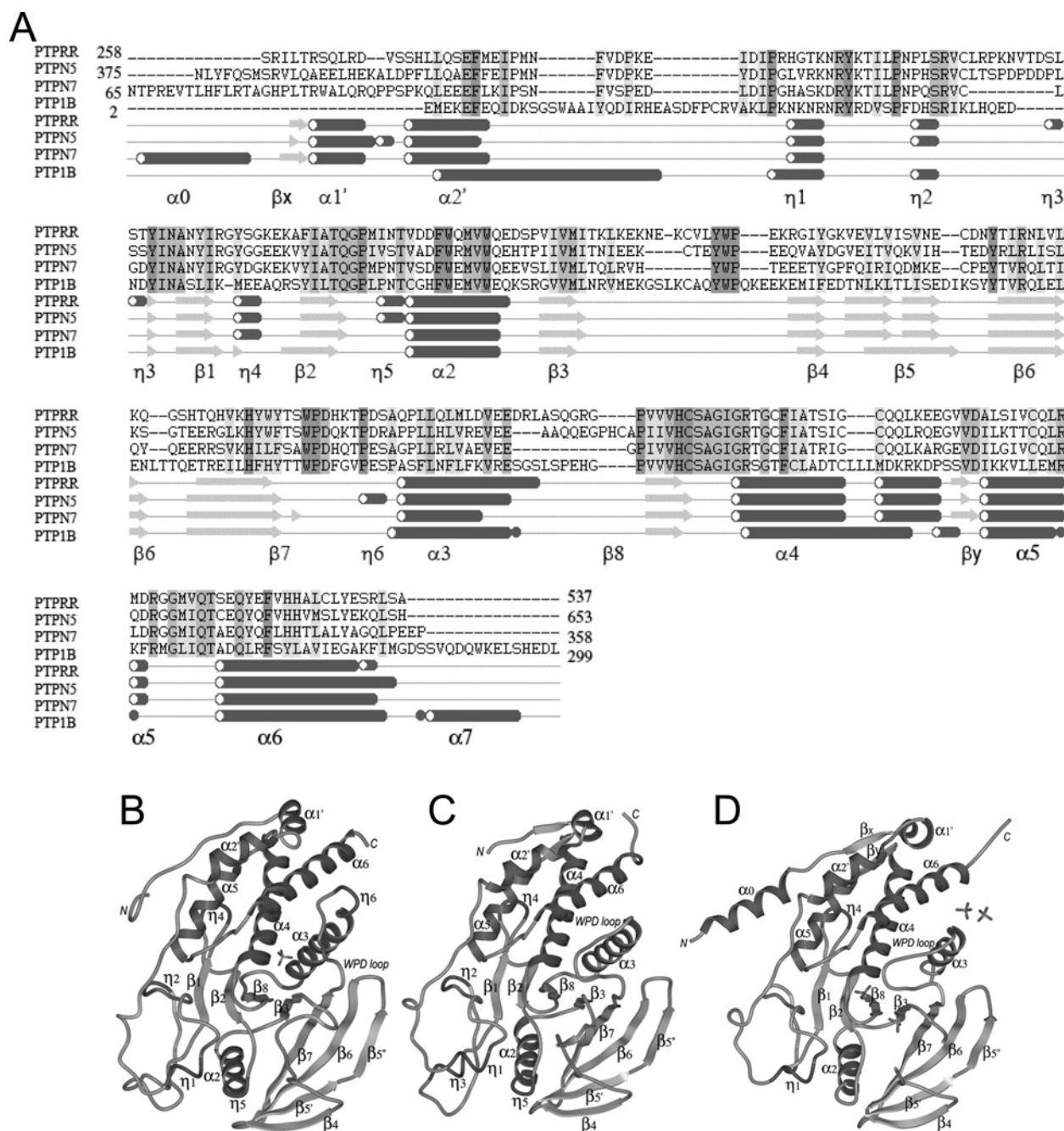
Figure 1   Sequence alignment of PTP catalytic domains and ribbon diagram of their three-dimensional structures

(**A**) PTPN5, PTPRR, PTPN7 and PTP1B were aligned using the program ICM Pro 3.3 (Molsoft LLC). Secondary-structure elements described in the text are labelled in the same way as in the ribbon diagrams showing the three-dimensional structure of (**B**) PTPN5(1) in complex with a sulphate ion shown; (**C**) PTPRR and (**D**) PTPN7 in complex with phosphate ions.

Quality of the structural models

The crystal structures of all three family members have been refined to high-resolution, low R-values and satisfactory geometry (Table 1). Residues 258–539 of PTPN5 crystallized with either the N-terminal His$_6$ tag intact or after removal of the GST tag of a second expression construct giving two different crystal forms, referred to as PTPN5(1) and PTPN5(2) respectively (Table 1). The entire structure of PTPRR (residues 375–655) was well defined and ordered except for the last two C-terminal residues. The structure of PTPN7 containing the catalytic domain and part of the KIM domain (residues 65–356) has the loop regions Arg[140]–

Gly[147], Glu[199]–Cys[204], Ser[277]–Glu[285] as well as the two C-terminal residues Glu[357]–Pro[358] disordered.

Overall structure

The overall structures of the PTPN5, PTPRR and PTPN7 phosphatase domains closely resembles the classical architecture of PTPs [25] consisting of highly twisted mixed β-sheets flanked by α-helices (Figure 1). In addition to the catalytic domain, the structure of PTPN7 contains part of the KIM domain which forms an N-terminal helix (α0) as observed in PTPSL and recently reported by Mustelin et al. [27]. At the secondary-structure level

there are a number of 3_{10} helices unique to KIM-domain containing PTPs ($\eta3$, $\eta4$, $\eta5$; Figure 1) which have not been observed in other tyrosine phosphatase structures so far [26]. Most notably we observed a PTPN5-specific 3_{10} helix ($\eta6$; Figure 1) towards the end of the WPD loop, contributing to the unique conformation of this catalytically important loop region. In contrast with the single $\beta5$ sheet of PTP1B, this structural element in PTPN5, PTPRR and PTPN7 is split into two smaller β-sheets annotated as $\beta5'$ and $\beta5''$ (Figure 1).

Structural comparison with other PTPs and phosphatase activity

The structures of PTPN5, PTPRR and PTPN7 are monomeric with an unhindered catalytic site. This conformation contrasts with the structure reported for the related receptor PTP, PTPRα [40], which forms homodimers with the N-terminal segment of one molecule inserted into the active site of the other molecule. Although this was originally proposed as a common mechanism for regulating the activity of receptor PTPs, it has subsequently been shown that the D1 catalytic domain of other receptor PTPs, such as RPTPμ, have an open and uninhibited conformation, similar to PTPN5, PTPRR and PTPN7 [41]. Similarly, crystal structures of the PTPs containing two catalytic domains (D1 and D2), such as CD45 [42], suggest that homodimerization would be impossible given the D1–D2 intramolecular domain orientation, indicating that dimerization is an exceptional, rather than a general, mechanism of phosphatase regulation.

The structures considered in the present study contain a number of structural elements that are different from other PTP structures reported so far. For example, the sequence following the $\alpha2'$ helix in PTP1B forms a secondary phosphotyrosine-binding pocket [43], which has been exploited by several high-affinity inhibitors. The corresponding sequences of PTPN5, PTPRR and PTPN7 show only a weak sequence identity with PTP1B, are shorter and take a different path in the structure which results in closure of this pocket. Also at the C-terminus of PTP1B, a helix $\alpha7$ is present which is thought to stabilize closure of the WPD loop [44]; however, this structural element is absent in PTPN5, PTPRR and PTPN7.

Measurement of the kinetic parameters of phosphatase activity for the catalytic domains of PTP1B, PTPN5, PTPRR and PTPN7 (see the Supplementary Data at http://www.BiochemJ.org/bj/395/bj3950483add.htm) using the fluorescent substrate DiFMUP (6,8-difluoro-4-methylumbelliferyl phosphate), indicated that all phosphatases were active. The catalytic efficiency (k_{cat}/K_m) was not significantly different among the four phosphatases, as would be expected when using bulkier physiological phosphoprotein substrates [45].

The N-terminal helix $\alpha0$

All members of this phosphatase family are characterized by the presence of a KIM domain which can be subdivided into a highly conserved MAPK-interaction motif and a protein-specific sequence which plays a role determining specificity for certain members of the MAPK family [8]. The complete KIM region is not present in the structures examined in the present study, but residues 65–78, which form the target specific interaction motif, were very well defined in the electron density of PTPN7. As described for PTPSL, these residues form an α-helix ($\alpha0$) stabilized by mainly hydrophobic interactions with a groove formed by the extended region following helix $\alpha2'$ and helix $\alpha5$. We found that all hydrophobic interactions that stabilize this aliphatic helix are conserved between PTPSL and PTPRR. Interestingly, the $\alpha0$ helix of PTPN7 contains Thr[66], which has been described to be phosphorylated by the kinases ERK2 and p38. In the structure considered in the present study, this residue is accessible and solvent-exposed, which is in contrast with PTPSL and a PTPN7 structure published recently in which phosphorylation of this residue would require significant structural changes [26,27].

Ion binding

Ions from the crystallization mother liquor are present in the PTPN5 and PTPN7 structures (Figures 1 and 2). A sulphate ion (SO_4^{2-}) is present in close proximity to the catalytic site in both PTPN5 structures. The sulphate oxygen atoms form hydrogen bonds with Trp[435] and Lys[439] of the WPD loop, Arg[478] of the PTP 'signature motif' and Gln[520]. Four phosphate (PO_4^{3-}) ions are present in the PTPN7 structure, of which two are found in the active-site pocket. One phosphate ion mimics the phosphate moiety of a phosphotyrosine substrate molecule and forms hydrogen bonds with the catalytic cysteine residue, Cys[291], as well as with Arg[297], and main-chain amide groups (Figure 2D). The second phosphate ion found in the proximity of the active site also forms two hydrogen bonds with Arg[297]. Interestingly, the phosphate moiety of the threonine residue of the substrate sequence pTEpY found in the activation loop of ERK2 may be mimicked by this phosphate ion, suggesting that phosphorylation of both ERK activation loop phosphates is necessary for substrate recognition.
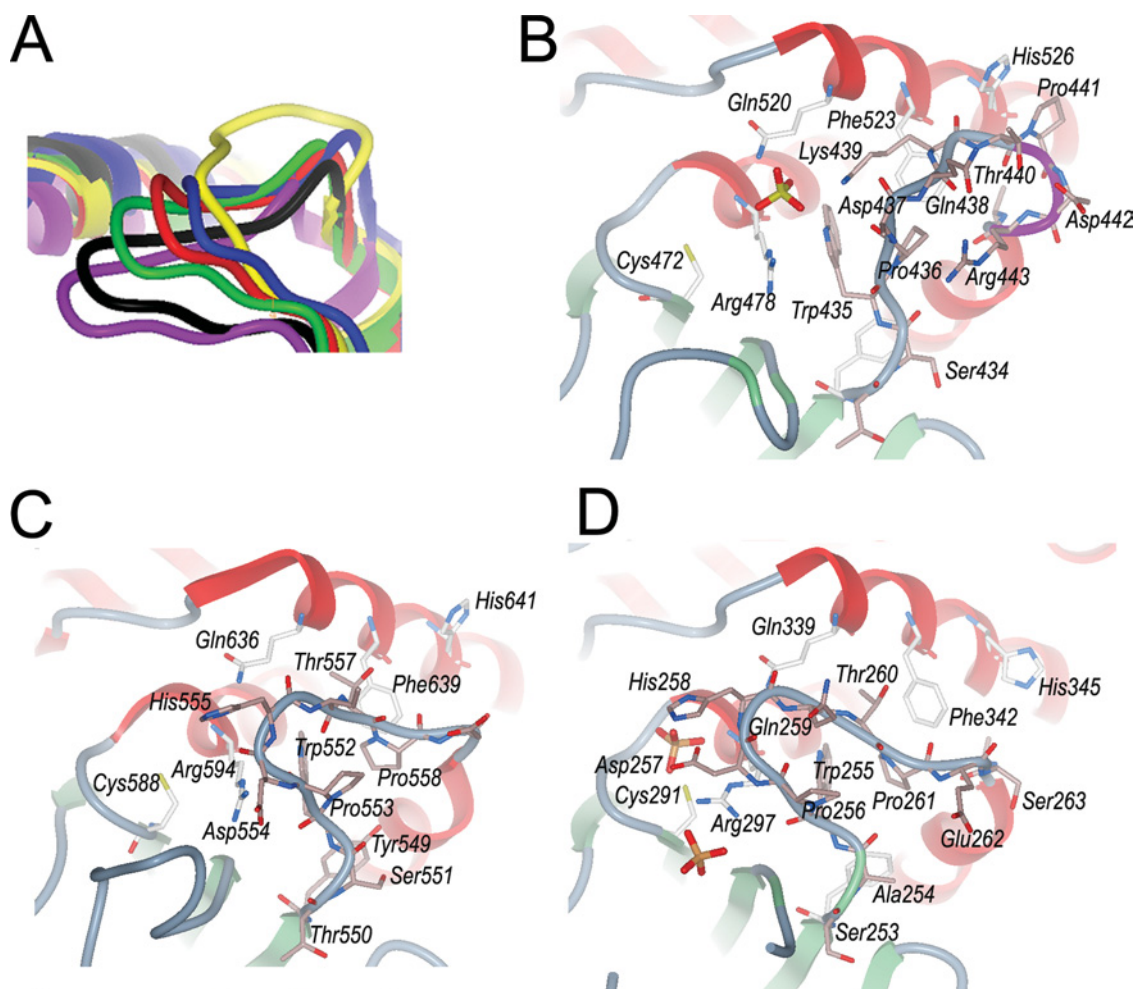
WPD loop

The WPD loop is a key element in the mechanism of PTP catalysis (reviewed in [46]). In PTP1B, this loop undergoes a large conformational change, of approx. 8–12 Å, on substrate binding that allows Asp[181] of the WPD loop to act as a general acid in the catalytic region. Arg[221] of the conserved PTP 'signature motif' reorients into a position to optimize salt-bridge interactions with the substrate phosphate. In this closed conformation, Cys[215] is in a position to commence a nucleophilic attack on the phosphotyrosine phosphorus atom. In addition, the flexibility of the WPD loop affects the size and shape of ligands that bind to the active site and therefore has major implications on any drug-design strategy.

Comparison of the WPD loop conformation of the PTPN5, PTPRR and PTPN7 structures with that of PTP1B (PDB codes 1SUG for the closed conformation and 2HNP for the open conformation) suggests that PTPN7 is in the closed conformation, while PTPRR and PTPN5 structures are in an open conformation (Figure 2A). Comparison of the PTPRR structure with that of its mouse homologue, PTPSL [26], indicates that both WPD loop conformations are open, with PTPSL more closely resembling the backbone conformation of PTP1B. All WPD loops have relatively low B-factors, indicating that the loops are not flexible. The WPD loop conformation in PTPN5 is additionally stabilized in an open conformation by the presence of an unusual 3_{10} helix not observed in other phosphatase structures determined so far, in which hydrogen bonds form between main-chain atoms of Thr[440] and Arg[443] as well as Pro[441] and Ala[444] (Figure 2B).

Crystal contacts are present in this region of PTPN5(1) (Gln[438] contacts the main-chain oxygen of Gln[403] in another molecule), but are unlikely to account for the unusual conformation, since crystal contacts occur elsewhere in PTPN5(2) and both loop conformations are similar. In PTPN5(2) Gln[438] is free to make a hydrogen bond with Asp[437], which is disordered in PTPN5(1), thus stabilizing this region.

Alignments of KIM-containing PTPs with other tyrosine phosphatases revealed differences in sequences that are largely conserved in tyrosine phosphatases: the WPD loop motif Gly-Val-Pro is replaced by Lys-Thr-Pro in PTPN5 and PTPRR and

Figure 2   Conformation of the WPD loops

(**A**) Structural overlay of WPD loops found in PTP1B (PDB code 1SUG; purple) in the closed conformation; PTPN7 (black); PTPRR (green); PTPSL (PDB code 1JLN; red); PTP1B (PDB code 2HNP; blue) in the open conformation and PTPN5 (yellow). (**B**) Structural details of the WPD loop of PTPN5. In order to differentiate the WPD loop from other residues, carbon atoms of WPD loop residues are coloured pink. The 3_{10} helix $\eta 6$ is coloured magenta. Side-chain conformations were identical in both PTPN5 structures, except for Arg[443] which has a slightly different conformation in PTPN5(1). Asp[437] is largely disordered in PTPN5(1), but was visible in the second structure determined. The sulphate moiety found in both crystal structures is shown in a ball-and-stick representation. (**C**) Structural details of the WPD loop of PTPRR. (**D**) Structural details of the WPD loop of PTPN7. The phosphate moieties are shown in a ball-and-stick representation.

Gln-Thr-Pro in PTPN7. In PTPSL, Lys[448] within this motif forms a salt bridge with the carboxy group of Asp[451], and it has been postulated that this interaction confers a more rigid loop structure, thereby hindering catalytic activity and contributing to substrate specificity [26]. A key difference is evident in this region of the PTPRR and PTPN5 structures in that the conserved lysine residue does not contact the conserved aspartate, suggesting that this interaction is specific to PTPSL. At the same position in PTPN7, a glutamine residue (Gln[259]) is present, and Huang et al. [47] reported that no significant changes in PTPN7 activity were observed when this sequence was replaced by a glycine residue. Together, these observations suggest that interactions made by Lys[448] (PTPSL) are not a critical determinant limiting catalytic activity and determining substrate specificity for KIM-containing PTPs.

The substrate-recognition loop of PTP1B with the YRD motif (Tyr-Arg-Asp) is replaced by Tyr-Lys-Thr in the KIM-containing PTPs. The tyrosine residue is invariant among the PTPs and is engaged in hydrophobic packing with phosphotyrosine substrates while the aspartate residue makes hydrogen bonds with the main-chain nitrogen atoms of the phosphotyrosine and the +1 residue
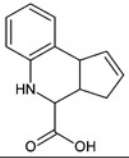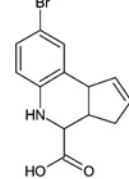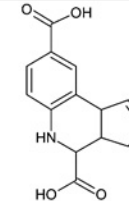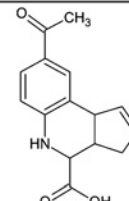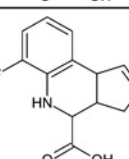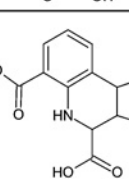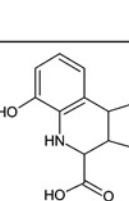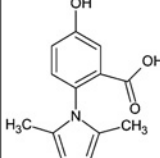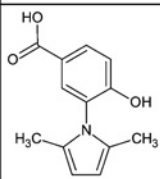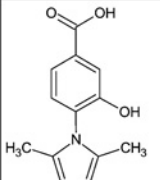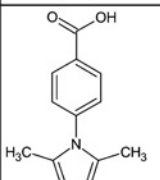
in the peptide substrate [48]. It is evident from the structures presented here that the threonine residue, corresponding to aspartate, is unable to make such interactions, thereby reducing substrate affinity and acting as a negative determinant limiting catalytic ac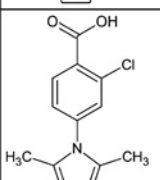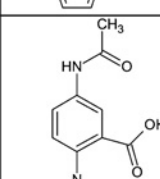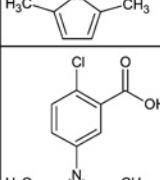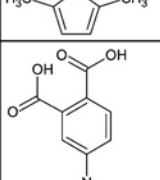tivity. In support of this, Huang et al. [47] reported that mutation of this threonine residue in PTPN7 to aspartate enhances phospho-peptide hydrolysis more than 10-fold.

Comparison of the surface-exposed electrostatic potential of the three KIM-containing PTPs with PTP1B reveals marked differences in surface charge in the proximity of the active site. It is interesting to note that PTPN7 has a mainly negatively charged surface close to the active site. It is very likely that the electrostatic potential contributes to the rather weak affinity for negatively charged substrates as indicated by the 10-fold higher K_{m} value for DiFMUP that has been determined (see the Supplementary Data at http://www.BiochemJ.org/bj/395/bj3950483add.htm).

Inhibitor identification

High-throughput screening of 24 000 compounds in a phosphatase assay with pNPP as a substrate using the enzymes PTP1B,

Table 2   Chemical stuctures and activity profiles of compounds

*Data are expressed as a fraction of control (i.e. 1.0 = no inhibition). Compounds were tested at 100 μM.

Inhibitor structure	Chemical name	PTP1B	PTPN5	PTPN7	PTPRR
cyclopenta[c]quinoline-carboxylic acids		Phosphatase Activity*			
	1a: 3a,4,5,9b-tetrahydro-3H-cyclopenta[c]quinoline-4-carboxylic acid	0.23	0.8	0.75	0.09
	1b: 8-bromo-3a,4,5,9b-tetrahydro-3H-cyclopenta[c]quinoline-4 carboxylic acid	0.33	0.9	0.41	0.08
	1c: 3a,4,5,9b-tetrahydro-3H-cyclopenta[c]quinoline-4,8-dicarboxylic acid	0.3	0.79	0.63	0.53
	1d: 8-acetyl-3a,4,5,9b-tetrahydro-3H-cyclopenta[c]quinoline-4-carboxylic acid	0.48	0.81	0.83	0.66
	1e: 6-fluoro-3a,4,5,9b-tetrahydro-3H-cyclopenta[c]quinoline-4-carboxylic acid	0.63	0.72	0.75	0.35
	1f: 3a,4,5,9b-tetrahydro-3H-cyclopenta[c]quinoline-4,6-dicarboxylic acid	0.25	0.73	0.65	0.39
	1g: 6-hydroxy-3a,4,5,9b-tetrahydro-3H-cyclopenta[c]quinoline-4-carboxylic acid	0.5	0.73	0.8	0.38

Inhibitor structure	Chemical name	PTP1B	PTPN5	PTPN7	PTPRR
2,5-dimethyl-pyrrolyl benzoic acids		Phosphatase Activity*			
	2a: 2-(2,5-dimethyl-1H-pyrrol-1-yl)-5-hydroxybenzoic acid	0.09	0.36	0.07	0.03
	2b: 3-(2,5-dimethyl-1H-pyrrol-1-yl)-4-hydroxybenzoic acid	0.07	0.52	0.2	0.03
	2c: 4-(2,5-dimethyl-1H-pyrrol-1-yl)-3-hydroxybenzoic acid	0.21	0.72	0.67	0.09
	2d: 4-(2,5-dimethyl-1H-pyrrol-1-yl)benzoic acid	0.06	0.88	0.63	0.09
	2e: 2-chloro-4-(2,5-dimethyl-1H-pyrrol-1-yl)benzoic acid	0.36	0.95	0.98	0.46
	2f: 5-(acetylamino)-2-(2,5-dimethyl-1H-pyrrol-1-yl)benzoic acid	0.35	0.94	0.82	0.46
	2g: 2-chloro-5-(2,5-dimethyl-1H-pyrrol-1-yl)benzoic acid	0.47	0.88	0.91	0.77
	2h: 4-(2,5-dimethyl-1H-pyrrol-1-yl)phthalic acid	0.18	0.8	0.78	0.54

PTPN5, PTPN7 and PTPRR led to the identification of two main inhibitor scaffolds that displayed significant inhibition at a compound concentration of 100 μM. Inhibition data for the first scaffold (cyclopenta[c]quinoline-4-carboxylic acids) are compiled in Table 2. Compound 1a inhibited PTPRR strongly and showed approx. 10-fold selectivity against PTPN5 and PTPN7. Furthermore,
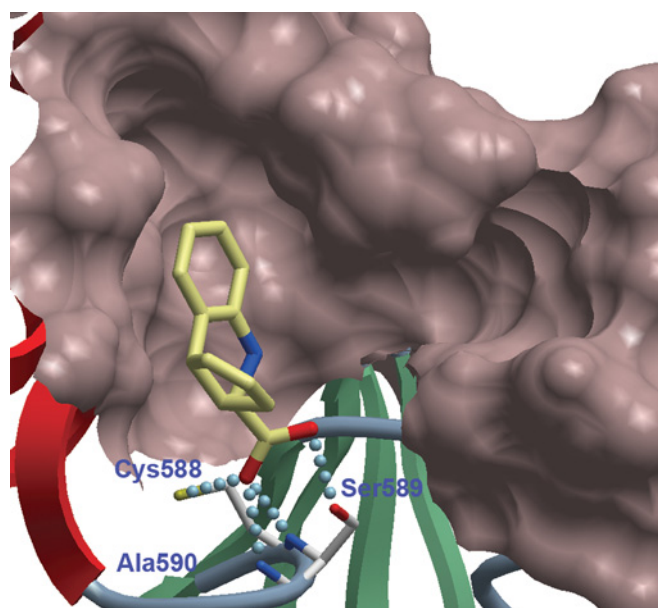
Figure 3   Docking of compound 1a in the active site of PTPRR

Representation of the active-site surface is shown as well as predicted hydrogen bonds formed by the carboxy acid group of compound 1a to PTPRR residues. Unoccupied regions in the PTPRR active-site cavity offer multiple possibilities for expansion of this compound class.

MS showed that none of the inhibitors studied oxidized the phosphatase or attached covalently to the target. Substitution at position 8 on the quinoline ring other than bromine is unfavourable and makes these inhibitors non-selective and weak. Also, substitution in position 6 reduced activity substantially for PTPRR and generates non-selective inhibitors. Compound 1a was successfully docked in the active site of PTPRR and shows good shape complementarity with the active-site surface (Figure 3). The docking model predicts a binding mode in which hydrogen bonds form between the carboxylic acid group of compound 1a and the active-site Cys[588], backbone nitrogens of the phosphate-binding loop and the hydroxy side chain of Ser[589] (Figure 3). The SAR observed in position 8 can be also rationalized using this model: a bromine atom at this position fits nicely into a hydrophilic pocket in that binding site, but substituents that are larger than that moiety (such as carboxy or acetyl) cannot be accommodated by PTPRR. Also, large substituents in position 6 are not favourable for sterical reasons. The docking model of compound 1a into the active site of PTPRR suggests also that a large portion of the binding pocket is not occupied by this inhibitor class, which allows for expansion of this compound class and would lead to more potent and selective inhibitors. On the basis of our docking model, the best position to target unoccupied areas in the binding site would be at the nitrogen atom of the quinoline ring.

The parent compound of the dimethyl-1*H*-pyrrolyl benzoic acid class of phosphatase inhibitors, 2d, is quite potent and selective for PTPRR over the closely related PTPN5 and PTPN7; however, PTP1B is also inhibited (Table 2). Substitution of the pyrrolyl moiety in position 2 rather than 4 of the benzoic acid ring and addition of a hydroxy group in position 5 (compound 2a) is favourable only for inhibitor development targeting PTPN7 and has no effect on inhibitor potency for PTPRR, for which it had an IC$_{50}$ value of 2 μM. Moving the pyrrolyl moiety to position 3 and the hydroxy group to position 4 (compound 2b) results in some selectivity for PTPRR inhibitors against PTPN7, but does not affect potency for PTPRR (IC$_{50}$ 4 μM). Other substitution patterns

shown in Table 2 resulted in a significant reduction of inhibitory activity. Unfortunately, we were not able to unambiguously dock this compound class into the active site of PTPRR, but it is reasonable to assume that the carboxy group in this inhibitor class interacts in a similar way with the active-site cysteine residue as observed for the cyclopenta[*c*]quinoline-carboxylic acid class. Hit expansion and co-crystal structures with the identified inhibitors of this class will facilitate the development of more potent and selective inhibitors and is an ongoing activity in our laboratories.

Conclusions

A large number of phosphatases have been linked to a wide spectrum of human disease and have been suggested as potential drug targets; however, to date, there has been a lack of information on selective inhibitors of individual phosphatase family members. In the present study, we have determined the structure of all human KIM-containing PTPs (PTPN5, PTPN7 and PTPRR) that negatively regulate MAPK signalling. A comparison of the determined high-resolution structures revealed important significant differences in the catalytic domain and in the WPD loop conformations. Two inhibitor classes (cyclopenta[*c*]quinolinecarboxylic acids and 2,5-dimethylpyrrolyl benzoic acids) for the KIM-containing phosphatase family were identified in a high-throughput screen which may provide new means of pharmacological intervention for MAPK signalling as well as interesting tools for cell biology research. SAR determined on the initial screening hits indicated that specificity may be obtained using these scaffolds. Furthermore, owing to the tissue-specific expression pattern of KIM-containing PTPs, selective PTP inhibitors would allow, in contrast with kinase inhibitors, the study of tissue-specific inhibition of MAPK signalling. Docking studies suggested routes for chemical modification of the identified cyclopenta[*c*]quinoline-carboxylic acid scaffold through which increased potency and selectivity may be achieved. Potent and selective inhibitors may have therapeutic potential for the treatment of neurodegenerative diseases such as Alzheimer's or acute myeloblastic leukaemia.

REFERENCES

1   Neel, B. G. and Tonks, N. K. (1997) Protein tyrosine phosphatases in signal transduction. Curr. Opin. Cell Biol. **9**, 193–204

2   Andersen, J. N., Mortensen, O. H., Peters, G. H., Drake, P. G., Iversen, L. F., Olsen, O. H., Jansen, P. G., Andersen, H. S., Tonks, N. K. and Moller, N. P. (2001) Structural and evolutionary relationships among protein tyrosine phosphatase domains. Mol. Cell. Biol. **21**, 7117–7136

3   Alonso, A., Sasin, J., Bottini, N., Friedberg, I., Friedberg, I., Osterman, A., Godzik, A., Hunter, T., Dixon, J. and Mustelin, T. (2004) Protein tyrosine phosphatases in the human genome. Cell **117**, 699–711

4   Hooft van Huijsduijnen, R., Bombrun, A. and Swinnen, D. (2002) Selecting protein tyrosine phosphatases as drug targets. Drug Discovery Today **7**, 1013–1019

5   Tautz, L., Pellecchia, M. and Mustelin, T. (2006) Targeting the PTPome in human disease. Expert Opin. Ther. Targets **10**, 157–177

6   Elchebly, M., Payette, P., Michaliszyn, E., Cromlish, W., Collins, S., Loy, A. L., Normandin, D., Cheng, A., Himms-Hagen, J., Chan, C. C. et al. (1999) Increased insulin sensitivity and obesity resistance in mice lacking the protein tyrosine phosphatase-1B gene. Science **283**, 1544–1548

7   Bialy, L. and Waldmann, H. (2005) Inhibitors of protein tyrosine phosphatases: next-generation drugs? Angew. Chem. Int. Ed. Engl. **44**, 3814–3839

8   Munoz, J. J., Tarrega, C., Blanco-Aparicio, C. and Pulido, R. (2003) Differential interaction of the tyrosine phosphatases PTP-SL, STEP and HePTP with the mitogen-activated protein kinases ERK1/2 and p38α is determined by a kinase specificity sequence and influenced by reducing agents. Biochem. J. **372**, 193–201

9   Buschbeck, M., Eickhoff, J., Sommer, M. N. and Ullrich, A. (2002) Phosphotyrosine-specific phosphatase PTP-SL regulates the ERK5 signaling pathway. J. Biol. Chem. **277**, 29503–29509

10   Nguyen, T. H., Liu, J. and Lombroso, P. J. (2002) Striatal enriched phosphatase 61 dephosphorylates Fyn at phosphotyrosine 420. J. Biol. Chem. **277**, 24274–24279

11   Pulido, R., Zuniga, A. and Ullrich, A. (1998) PTP-SL and STEP protein tyrosine phosphatases regulate the activation of the extracellular signal-regulated kinases ERK1 and ERK2 by association through a kinase interaction motif. EMBO J. **17**, 7337–7350

12   Saxena, M., Williams, S., Tasken, K. and Mustelin, T. (1999) Crosstalk between cAMP-dependent kinase and MAP kinase through a protein tyrosine phosphatase. Nat. Cell Biol. **1**, 305–311

13   Lombroso, P. J., Naegele, J. R., Sharma, E. and Lerner, M. (1993) A protein tyrosine phosphatase expressed within dopaminoceptive neurons of the basal ganglia and related structures. J. Neurosci. **13**, 3064–3074

14   Boulanger, L. M., Lombroso, P. J., Raghunathan, A., During, M. J., Wahle, P. and Naegele, J. R. (1995) Cellular and molecular characterization of a brain-enriched protein tyrosine phosphatase. J. Neurosci. **15**, 1532–1544

15   Paul, S., Nairn, A. C., Wang, P. and Lombroso, P. J. (2003) NMDA-mediated activation of the tyrosine phosphatase STEP regulates the duration of ERK signaling. Nat. Neurosci. **6**, 34–42

16   Valjent, E., Pascoli, V., Svenningsson, P., Paul, S., Enslen, H., Corvol, J. C., Stipanovich, A., Caboche, J., Lombroso, P. J., Nairn, A. C. et al. (2005) Regulation of a protein phosphatase cascade allows convergent dopamine and glutamate signals to activate ERK in the striatum. Proc. Natl. Acad. Sci. U.S.A. **102**, 491–496

17   Augustine, K. A., Rossi, R. M., Silbiger, S. M., Bucay, N., Duryea, D., Marshall, W. S. and Medlock, E. S. (2000) Evidence that the protein tyrosine phosphatase (PC12,Br7,Sl) gamma (−) isoform modulates chondrogenic patterning and growth. Int. J. Dev. Biol. **44**, 361–371

18   Sharma, E. and Lombroso, P. J. (1995) A neuronal protein tyrosine phosphatase induced by nerve growth factor. J. Biol. Chem. **270**, 49–53

19   Shiozuka, K., Watanabe, Y., Ikeda, T., Hashimoto, S. and Kawashima, H. (1995) Cloning and expression of PCPTP1 encoding protein tyrosine phosphatase. Gene **162**, 279–284

20   Saxena, M., Williams, S., Brockdorff, J., Gilman, J. and Mustelin, T. (1999) Inhibition of T cell signaling by mitogen-activated protein kinase-targeted hematopoietic tyrosine phosphatase (HePTP). J. Biol. Chem. **274**, 11693–11700

21   Saxena, M., Williams, S., Gilman, J. and Mustelin, T. (1998) Negative regulation of T cell antigen receptor signal transduction by hematopoietic tyrosine phosphatase (HePTP). J. Biol. Chem. **273**, 15340–15344

22   Gronda, M., Arab, S., Iafrate, B., Suzuki, H. and Zanke, B. W. (2001) Hematopoietic protein tyrosine phosphatase suppresses extracellular stimulus-regulated kinase activation. Mol. Cell. Biol. **21**, 6851–6858

23   Zanke, B., Suzuki, H., Kishihara, K., Mizzen, L., Minden, M., Pawson, A. and Mak, T. W. (1992) Cloning and expression of an inducible lymphoid-specific, protein tyrosine phosphatase (HePTPase). Eur. J. Immunol. **22**, 235–239

24   Zanke, B., Squire, J., Griesser, H., Henry, M., Suzuki, H., Patterson, B., Minden, M. and Mak, T. W. (1994) A hematopoietic protein tyrosine phosphatase (HePTP) gene that is amplified and overexpressed in myeloid malignancies maps to chromosome 1q32.1. Leukemia **8**, 236–244

25   Barford, D., Flint, A. J. and Tonks, N. K. (1994) Crystal structure of human protein tyrosine phosphatase 1B. Science **263**, 1397–1404

26   Szedlacsek, S. E., Aricescu, A. R., Fulga, T. A., Renault, L. and Scheidig, A. J. (2001) Crystal structure of PTP-SL/PTPBR7 catalytic domain: implications for MAP kinase regulation. J. Mol. Biol. **311**, 557–568

27   Mustelin, T., Tautz, L. and Page, R. (2005) Structure of the hematopoietic tyrosine phosphatase (HePTP) catalytic domain: structure of a KIM phosphatase with phosphate bound at the active site. J. Mol. Biol. **354**, 150–163

28   Zhang, Z. Y. (2002) Protein tyrosine phosphatases: structure and function, substrate specificity, and inhibitor development. Annu. Rev. Pharmacol. Toxicol. **42**, 209–234

29   Stols, L., Gu, M., Dieckman, L., Raffen, R., Collart, F. R. and Donnelly, M. I. (2002) A new vector for high-throughput, ligation-independent cloning encoding a tobacco etch virus protease cleavage site. Protein Expression Purif. **25**, 8–15

30   Otwinowski, Z. and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. **276**, 307–326

31   Leslie, A. G. W. (1992) Recent changes to the MOSFLM package for processing film and image plate data. Joint CCP4 and ESF-EAMCB Newsletter on Protein Crystallography **26**

32   Evans, P. R. (1993) Data reduction. Proceedings of CCP4 Study Weekend on Data Collection & Processing, pp. 114–122

33   Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography. Acta Crystallogr. Sect. D Biol. Crystallogr. **50**, 760–763

34   Storoni, L. C., McCoy, A. J. and Read, R. J. (2004) Likelihood-enhanced fast rotation functions. Acta Crystallogr. Sect. D Biol. Crystallogr. **60**, 432–438

35   McRee, D. E. (1999) XtalView/Xfit – a versatile program for manipulating atomic coordinates and electron density. J. Struct. Biol. **125**, 156–165

36   Totrov, M. and Abagyan, R. (1997) Flexible protein–ligand docking by global energy optimization in internal coordinates. Proteins Suppl. **1**, 215–220

37   Abagyan, R. and Totrov, M. (1994) Biased probability Monte Carlo conformational searches and electrostatic calculations for peptides and proteins. J. Mol. Biol. **235**, 983–1002

38   Abagyan, R., Totrov, M. and Kuznetsov, D. (1994) ICM – a new method for protein modeling and design: applications to docking and structure prediction from the distorted native conformation. J. Comput. Chem. **15**, 488–506

39   Manger, M., Scheck, M., Prinz, H., von Kries, J. P., Langer, T., Saxena, K., Schwalbe, H., Furstner, A., Rademann, J. and Waldmann, H. (2005) Discovery of *Mycobacterium tuberculosis* protein tyrosine phosphatase A (MptpA) inhibitors based on natural products and a fragment-based approach. ChemBioChem **6**, 1749–1753

40   Bilwes, A. M., den Hertog, J., Hunter, T. and Noel, J. P. (1996) Structural basis for inhibition of receptor protein-tyrosine phosphatase-α by dimerization. Nature (London) **382**, 555–559

41   Hoffmann, K. M., Tonks, N. K. and Barford, D. (1997) The crystal structure of domain 1 of receptor protein-tyrosine phosphatase μ. J. Biol. Chem. **272**, 27505–27508

42   Nam, H. J., Poy, F., Saito, H. and Frederick, C. A. (2005) Structural basis for the function and regulation of the receptor protein tyrosine phosphatase CD45. J. Exp. Med. **201**, 441–452

43   Puius, Y. A., Zhao, Y., Sullivan, M., Lawrence, D. S., Almo, S. C. and Zhang, Z. Y. (1997) Identification of a second aryl phosphate-binding site in protein-tyrosine phosphatase 1B: a paradigm for inhibitor design. Proc. Natl. Acad. Sci. U.S.A. **94**, 13420–13425

44   Wiesmann, C., Barr, K. J., Kung, J., Zhu, J., Erlanson, D. A., Shen, W., Fahr, B. J., Zhong, M., Taylor, L., Randal, M. et al. (2004) Allosteric inhibition of protein tyrosine phosphatase 1B. Nat. Struct. Mol. Biol. **11**, 730–737

45   Denu, J. M., Lohse, D. L., Vijayalakshmi, J., Saper, M. A. and Dixon, J. E. (1996) Visualization of intermediate and transition-state structures in protein-tyrosine phosphatase catalysis. Proc Natl. Acad. Sci. U.S.A. **93**, 2493–2498

46   Barford, D., Das, A. K. and Egloff, M. P. (1998) The structure and mechanism of protein phosphatases: insights into catalysis and regulation. Annu. Rev. Biophys. Biomol. Struct. **27**, 133–164

47   Huang, Z., Zhou, B. and Zhang, Z. Y. (2004) Molecular determinants of substrate recognition in hematopoietic protein-tyrosine phosphatase. J. Biol. Chem. **279**, 52150–52159

48   Salmeen, A., Andersen, J. N., Myers, M. P., Tonks, N. K. and Barford, D. (2000) Molecular basis for the dephosphorylation of the activation segment of the insulin receptor by protein tyrosine phosphatase 1B. Mol. Cell **6**, 1401–1412