High-resolution crystal structures of the D1 and D2 domains of protein tyrosine phosphatase epsilon for structure-based drug design

George T. Lountos,^{a,b} Sreejith Raran-Kurussi,^b Bryan M. Zhao,^{c,d} Beverly K. Dyas,^d Terrence R. Burke Jr,^e Robert G. Ulrich^d and David S. Waugh^{b,*}

Received 19 July 2018

Accepted 22 August 2018

Edited by G. Cingolani, Thomas Jefferson University, USA

Keywords: microarray assay; protein tyrosine phosphatase; PTP; receptor-like protein tyrosine phosphatase; RPTP; structure-based drug design.

PDB references: PTPε D1 domain, 6d4d; PTPε D2 domain, 6d3f; PTPε D2 domain, A455N/V457Y/E597D mutant, 6d4f

Supporting information: this article has supporting information at journals.iucr.org/d

^aBasic Science Program, Frederick National Laboratory for Cancer Research sponsored by the National Cancer Institute, Frederick, MD 21702, USA, ^bMacromolecular Crystallography Laboratory, Center for Cancer Research, National Cancer Institute, Frederick, MD 21702, USA, ^cThe Oak Ridge Institute for Science and Education, Oak Ridge, TN 37831, USA, ^dMolecular and Translational Sciences Division, US Army Medical Research Institute of Infectious Diseases, Frederick, MD 21702, USA, and ^eChemical Biology Laboratory, Center for Cancer Research, National Cancer Institute, Frederick, MD 21702, USA. *Correspondence e-mail: waughd@mail.nih.gov

Here, new crystal structures are presented of the isolated membrane-proximal D1 and distal D2 domains of protein tyrosine phosphatase epsilon (PTPε), a protein tyrosine phosphatase that has been shown to play a positive role in the survival of human breast cancer cells. A triple mutant of the PTPε D2 domain (A455N/V457Y/E597D) was also constructed to reconstitute the residues of the PTPε D1 catalytic domain that are important for phosphatase activity, resulting in only a slight increase in the phosphatase activity compared with the native D2 protein. The structures reported here are of sufficient resolution for structure-based drug design, and a microarray-based assay for high-throughput screening to identify small-molecule inhibitors of the PTPε D1 domain is also described.

1. Introduction

The reversible phosphorylation of tyrosine residues in proteins, which depends on the balanced coordination between a multitude of protein tyrosine kinases (PTKs) and protein tyrosine phosphatases (PTPs), is an essential step in many cellular signaling pathways (Hunter, 1995; Sun & Tonks, 1994; Denu & Dixon, 1998). While drug-design efforts targeting PTKs have resulted in the development of many inhibitors that are used in the clinic today, medicinal chemistry efforts targeting PTPs lag far behind and remain very challenging. Nevertheless, several PTPs have the potential to serve as important drug targets for a variety of diseases such as cancer, diabetes, neurological and autoimmune disorders, inflammation and numerous pathogenic infections (Blaskovich, 2009; Bialy & Waldmann, 2005; Tautz *et al.*, 2006; Alonso *et al.*, 2004; Mustelin, 2006; Bahta & Burke, 2012; Guan & Dixon, 1993; Zhou *et al.*, 2010; Böhmer *et al.*, 2013). PTPs are generally considered to be difficult targets for therapeutic intervention, in part because of their highly conserved active sites. Additionally, phosphotyrosine (pTyr) residues play critical roles in substrate recognition, and replicating these interactions with pTyr mimetics often entails the inclusion of anionic moieties, which impede cell permeability and reduce bioavailability (Barr, 2010). Even so, there continues to be considerable interest in discovering novel therapeutic agents to target PTPs (Ghattas *et al.*, 2016), and inhibitors of multiple PTPs are currently being developed (Stanford & Bottini, 2017).

The classical type I cysteine-based PTPs include both receptor and nonreceptor forms that catalyze the removal of

the phosphate moiety from a targeted pTyr residue (Alonso *et al.*, 2004; Denu & Dixon, 1998). The active site of these PTPs contains the consensus sequence (H/V)C(X₅)R(S/T) that harbors a conserved catalytic cysteine residue located at the base of the catalytic pocket, as well as a WPD loop that provides a conserved aspartic acid which functions as a general acid/base during catalysis (Zhang *et al.*, 1994; Zhang, 1998). The negatively charged pTyr substrate binds to the active site *via* hydrogen bonds to the residues forming the phosphate-binding loop (P-loop) and is further stabilized by the highly conserved arginine residue in the consensus sequence (Jia *et al.*, 1995; Sarmiento *et al.*, 2000). Once the substrate binds, the WPD loop undergoes a transition from an 'open' conformation to a 'closed' conformation, in which the invariant aspartic acid is properly positioned to form hydrogen bonds to the phenolic O atom of the phosphoryl ester (Fauman & Saper, 1996; Denu *et al.*, 1996). A cysteinyl-phosphate intermediate is formed as a result of nucleophilic attack by the conserved cysteine (Guan & Dixon, 1991). In a second mechanistic step, the covalent thiophosphoryl intermediate is attacked by a water molecule that is deprotonated by the aspartic acid. This results in disassociation of the phosphate, which frees the cysteine residue to regenerate the ground state of the enzyme (Zhang & Dixon, 1994).

Receptor-like protein tyrosine phosphatases (RPTPs) are integral membrane proteins that are composed of an extracellular receptor-like domain that varies in size and structure among RPTPs and a cytoplasmic region which contains the catalytic PTP domains (Andersen, Mortensen *et al.*, 2001). Most RPTPs contain two tandem cytoplasmic phosphatase domains. The catalytic activity is associated primarily with the membrane-proximal domain (D1), while the membrane-distal domain (D2) typically displays little or no activity (Itoh *et al.*, 1992). It has been proposed that the biological function of the D2 domain is regulatory; it may influence the binding of substrates or the activity of the D1 domain (Tsujikawa *et al.*, 2001; Wu *et al.*, 1997). However, it has been observed that the introduction of certain amino-acid substitutions into the inert D2 domains of LAR and PTP α can restore their catalytic activity, suggesting that these D2 domains might serve as auxiliary catalytic regions under certain circumstances (Lim *et al.*, 1997, 1998; Wu *et al.*, 1997; Nam *et al.*, 1999).

Several PTPs are potential molecular targets for inhibition in breast cancer therapy (Nunes-Xavier *et al.*, 2013; Hardy *et al.*, 2012). Among these, PTP ϵ has been viewed as a promising candidate because it is overexpressed in murine mammary tumors and experiments have demonstrated that PTP ϵ plays a positive role in cell growth and survival in MCF-7 cells and other breast cancer cell lines (Elson, 1999, 2017; Elson & Leder, 1995). Additional data have shown that PTP ϵ (and the closely related PTP α) are positive regulators of HER2-mediated and Src-mediated signaling in breast cancer cells (Parsons & Parsons, 2004; Berman-Golan *et al.*, 2008). Src is a major player in breast cancer, and dephosphorylation of its C-terminal regulatory phosphotyrosine (pTyr530) by PTP ϵ triggers a cascade of events that lead to the activation of Src, which favors a transformed phenotype (Espada & Martín-

Pérez, 2017). Additional studies have suggested that PTP ϵ may also be a valid molecular target for the treatment of osteoporosis (Granot-Attas *et al.*, 2007) and diabetes (Roussou-Noori *et al.*, 2011).

The receptor PTP ϵ (RPTP ϵ), which is located in the plasma membrane, possesses a small extracellular domain, a membrane-spanning domain and two tandem cytoplasmic domains, D1 and D2 (Nakamura *et al.*, 1996). Biochemical characterization has shown that catalytic activity is primarily found within the D1 domain, with only a low level of intrinsic activity observed within the D2 domain. A second form of PTP ϵ , referred to as cytosolic PTP ϵ or cytPTP ϵ , is transcribed from a separate promoter and lacks the extracellular and transmembrane domains (Andersen, Elson *et al.*, 2001; Wabakken *et al.*, 2002). It is possible that RPTP ϵ and cytPTP ϵ may have distinct physiological functions owing to their differing intracellular locations, but this has not been well studied. A 3.2 Å resolution crystal structure of cytPTP ϵ containing the tandem D1 and D2 domains has been deposited in the Protein Data Bank (PDB entry 2jjd; Barr *et al.*, 2009). However, higher resolution data could facilitate structure-based drug-design efforts (Almo *et al.*, 2007). Accordingly, we cloned, expressed and purified constructs of PTP ϵ corresponding to the individual D1 and D2 domains and determined their crystal structures at 1.76 and 2.27 Å resolution, respectively. We also developed a microarray assay that is suitable for the high-throughput screening of potential PTP ϵ inhibitors.

2. Materials and methods

2.1. Plasmid construction

Protein tyrosine phosphatase receptor type E (PTPRE) cDNA (Clone HsCD00438889) was obtained from the DNASU Plasmid Repository, Arizona State University, Arizona, USA (Supplementary Table S1). The D1 catalytic domain, spanning residues Ser107–Gly398, was amplified in a single PCR reaction using the primers PE-277, PE-2552 and PE-2589 (Supplementary Table S1) to add a Gateway *attB1* site, a *Tobacco etch virus* (TEV) protease cleavage site and a Gateway *attB2* site, yielding a PCR amplicon with the format *attB1*-TEV site-PTP ϵ (Ser107–Gly398)-*attB2*. The PCR product was recombined into pDONR221 (Thermo Fisher Scientific) by the Gateway BP reaction and verified by DNA sequencing. The gene was then recombined into the destination vector pDEST-HisMBP (Nallamsetty *et al.*, 2005) *via* the Gateway LR reaction to generate a vector (pBA2525) that directs the expression of a TEV protease-cleavable His₆-maltose-binding protein-PTP ϵ D1(Ser107–Gly398) fusion protein.

The D2 domain of PTP ϵ (residues Gly425–Lys700) was PCR-amplified from the above-mentioned cDNA clone in a single reaction using the primers PE-277, PE-2776 and PE-2777 (Supplementary Table S1). The PCR amplicon was recombined into the Gateway donor vector pDONR221 by Gateway cloning and the nucleotide sequence was confirmed experimentally. The D2 domain containing a recognition site

(ENLYFQ/G) for TEV protease at its N-terminus was moved into the destination vector pDEST-HisMBP by recombinational cloning to produce the expression vector pSRK2607. Because the TEV protease cleavage site was resistant to digestion, presumably owing to steric hindrance, we inserted a (Gly)₃ spacer between it and the N-terminus of the D2 domain in the vector pSRK2607 using the QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent Technologies) and primers PE-2806 and PE-2807 (Supplementary Table S1). The resulting expression vector (pSRK2657) was used to express and purify the wild-type PTP ϵ D2 domain. This vector directs the expression of the D2 domain as a fusion with the C-terminus of *Escherichia coli* maltose-binding protein (MBP) with an intervening TEV protease recognition site. The MBP contains an N-terminal His₆ tag for affinity purification by immobilized metal-affinity chromatography.

A triple mutant of the PTP ϵ D2 domain (A455N/V457Y/E597D) was also constructed to reconstitute the vital residues important for phosphatase activity in the PTP ϵ D1 active domain. This was performed in three steps. Initially, a double mutant, pSRK2608 (V457Y/E597D), was created with the primers PE-2784 and PE-2783 (Supplementary Table S1) using the QuikChange Lightning Multi Site-Directed Mutagenesis Kit. The expression vector pSRK2607 was used as the template in the mutagenesis reaction. In the second step, the resulting double mutant (pSRK2608) was further modified to incorporate the (Gly)₃ spacer at the N-terminus of the D2 domain, as described above, using the primers PE-2806 and PE-2807 (Supplementary Table S1) to generate the plasmid pSRK2610. In the final step, this TEV-cleavable construct was modified to generate the triple-mutant expression vector pSRK2612 (A455N/V457Y/E597D) using the QuikChange Lightning Site-Directed Mutagenesis Kit with the primers PE-2808 and PE-2809 (Supplementary Table S1). The mutations were confirmed experimentally. All mutagenesis reactions were performed as per the manufacturer's protocol.

2.2. Protein expression and purification

For protein expression, cells were grown at 310 K in Luria–Bertani broth containing the appropriate antibiotics and 0.2% glucose to mid-log phase, at which time overexpression of proteins was induced at 303 K for 4 h by the addition of 1 mM IPTG. The PTP ϵ D1 domain (Ser107–Gly398) was expressed in *E. coli* BL21(DE3) RIL cells. The PTP ϵ D2 domain and the D2 mutant (A455N/V457Y/E597D) were expressed in *E. coli* Rosetta 2 (DE3) cells (EMD Millipore).

For purification of the recombinant proteins, all procedures were performed at 277 K unless otherwise stated. For the purification of the D1 domain, 36 g of frozen cells were suspended in 300 ml ice-cold 50 mM Tris–HCl pH 8.0, 25 mM imidazole (buffer A). The cell suspension was lysed with an APV-1000 homogenizer (Invensys APV Products) at 69 MPa and centrifuged (30 000g, 30 min, 277 K). The supernatant was then filtered through a 0.45 μ m cellulose acetate membrane and applied onto a 20 ml HisPrep column (GE Healthcare

Bio-Sciences) equilibrated with buffer A. After loading, the column was washed to baseline and bound protein was eluted with a linear gradient from 25 to 250 mM imidazole using 50 mM Tris–HCl pH 8.0, 500 mM imidazole over five column volumes. Peak fractions corresponding to His₆-MBP-PTP ϵ (Ser107–Gly398) were pooled and concentrated. The concentrated protein was then diluted with 50 mM Tris–HCl pH 8.0 until the imidazole concentration was approximately 25 mM. 5 mg His₆-tagged *Tobacco etch virus* (TEV) protease (Kapust *et al.*, 2001) was added and incubated overnight at 277 K. The digest was applied onto a 25 ml HisPrep column equilibrated in buffer A and the flowthrough fractions were collected and concentrated to ~5 ml. Tris(2-carboxyethyl)-phosphine hydrochloride (TCEP) was added to this sample to a final concentration of 5 mM and allowed to stand at 277 K overnight. The protein was applied onto a 320 ml Sephacryl S100 HR gel-filtration column (GE Healthcare Bio-Sciences) equilibrated with 25 mM Tris–HCl pH 8.0, 2 mM TCEP, and peak fractions corresponding to PTP ϵ (Ser107–Gly398) were pooled and concentrated to 12.4 mg ml⁻¹ (extinction coefficient of 46 870 M⁻¹ cm⁻¹).

The PTP ϵ D2 domains (both wild-type and mutant) were purified as follows. Approximately 10 g of *E. coli* cell paste was suspended in 150 ml buffer B [50 mM HEPES–NaOH pH 7.5, 500 mM NaCl, 5%(v/v) glycerol] containing 25 mM imidazole, cOmplete EDTA-free protease-inhibitor cocktail tablets (Roche Diagnostics Corporation) and 1 mM benzamidine–HCl (Sigma–Aldrich). The cell suspension was lysed with an APV-1000 homogenizer at 69 MPa and centrifuged at 30 000g for 30 min at 277 K. The supernatant was filtered through a 0.2 μ m polyethersulfone membrane and applied onto a 5 ml HisTrap FF column (GE Healthcare Bio-Sciences) equilibrated with buffer B. The column was washed to baseline with buffer B and eluted with a linear gradient of imidazole to 250 mM in buffer B. The eluted fractions containing the His₆-MBP-tagged proteins were pooled and concentrated. The concentrated protein was diluted with buffer B to reduce the imidazole concentration to approximately 25 mM and digested overnight with His₆-tagged TEV protease at 277 K. The digest was applied onto a 5 ml HisTrap HP column (GE Healthcare Bio-Sciences) to capture the cleaved His₆-MBP tag and His₆-TEV. The column effluent contained the pure protein of interest. The effluent was concentrated and incubated with 10 mM dithiothreitol overnight at 277 K. The reduced sample was loaded onto a HiPrep 26/60 Sephacryl S-200 HR column (GE Healthcare Bio-Sciences) that had been equilibrated with 20 mM HEPES–NaOH pH 7.5, 150 mM NaCl, 2 mM TCEP, 5%(v/v) glycerol buffer. The peak fractions containing the protein of interest were pooled and concentrated to about 10–15 mg ml⁻¹ (as estimated at 280 nm using the molar extinction coefficient derived from the ExPASy ProtParam web tool; Gasteiger *et al.*, 2003). The final products were judged to be >95% pure by SDS–PAGE and the protein molecular weight was confirmed by electrospray ionization mass spectroscopy. Aliquots were flash-frozen using liquid nitrogen and stored at 193 K.

Table 1
Crystallization.

	PTP ϵ D1 domain	PTP ϵ D2 domain	PTP ϵ (A455N/V457Y/E597D) D2 domain
Method	Vapor diffusion, hanging drop	Vapor diffusion, hanging drop	Vapor diffusion, hanging drop
Plate type	EasyXtal 15-well plate, Qiagen	EasyXtal 15-well plate, Qiagen	EasyXtal 15-well plate, Qiagen
Temperature (K)	292	292	292
Protein concentration (mg ml ⁻¹)	13.9	5	9
Composition of reservoir solution	0.1 M Tris-HCl pH 8.5, 0.2 M sodium chloride, 22% (w/v) polyethylene glycol 3350	0.1 M Bicine/Trizma base pH 8.5, 0.03 M diethylene glycol, 0.03 M triethylene glycol, 0.03 M tetraethylene glycol, 0.03 M pentaethylene glycol, 10% (w/v) polyethylene glycol 4000, 20% (v/v) glycerol	0.1 M imidazole pH 8.0, 0.2 M calcium acetate hydrate, 20% (w/v) polyethylene glycol 1000
Volume and ratio of drop	4 μ l, 1:1 mixture of protein and reservoir solutions	4 μ l, 1:3 mixture of protein and reservoir solutions	4 μ l, 1:1 mixture of protein and reservoir solutions
Volume of reservoir (ml)	0.5	0.5	0.5

2.3. Crystallization

All purified PTP ϵ constructs were screened for crystals with several sparse-matrix screens from Hampton Research, Microlytic and Molecular Dimensions using a Gryphon crystallization robot (Art Robbins Instruments). Optimization trials of initial crystallization screening hits were conducted using the hanging-drop vapor-diffusion method in EasyXtal 15-well plates (Qiagen; Table 1). The PTP ϵ D1 domain was crystallized by mixing 2 μ l protein solution (13.9 mg ml⁻¹) with 2 μ l well solution [0.1 M Tris-HCl pH 8.5, 0.2 M sodium chloride, 22% (w/v) polyethylene glycol 3350] and incubating the trays at 292 K. A cluster of long thin rods was obtained after approximately 4 d. A single crystal was retrieved from a drop using a LithoLoop (Molecular Dimensions) and transferred to a solution consisting of 0.1 M Tris-HCl pH 8.5, 0.2 M sodium chloride, 22% (w/v) polyethylene glycol 3350, 20% (v/v) glycerol. After a 1 min soak, the crystal was then retrieved with a LithoLoop and immediately flash-cooled by plunging into liquid nitrogen. Crystals of the PTP ϵ D2 domain were grown by mixing 1 μ l protein solution (5 mg ml⁻¹) with 3 μ l well solution [100 mM Bicine/Trizma base pH 8.5, 30 mM diethylene glycol, 30 mM triethylene glycol, 30 mM tetraethylene glycol, 30 mM pentaethylene glycol, 10% (w/v) polyethylene glycol 4000, 20% (v/v) glycerol]. The drops were streak-seeded with a whisker and incubated at 292 K. A single crystal was looped from the drop with a LithoLoop and immediately flash-cooled by plunging into liquid nitrogen. The PTP ϵ D2 mutant (A455N/V457K/E597D) was crystallized by mixing 2 μ l protein solution (9 mg ml⁻¹) with 2 μ l well solution [100 mM imidazole-HCl pH 8.0, 200 mM calcium acetate hydrate, 20% (w/v) polyethylene glycol 1000]. A single crystal was retrieved with a LithoLoop, passed through a cryo-protectant solution consisting of 100 mM imidazole-HCl pH 8.0, 200 mM calcium acetate hydrate, 20% (w/v) polyethylene glycol 1000, 20% (v/v) glycerol and immediately flash-cooled with liquid nitrogen.

2.4. X-ray data collection, structure solution and refinement

X-ray diffraction data for crystals of the wild-type PTP ϵ D1 and D2 constructs were collected on the 22-ID beamline of the

SER-CAT facilities at the Advanced Photon Source, Argonne National Laboratory, Argonne, Illinois, USA. For PTP ϵ D1, 180 images were collected using a wavelength of 1.0000 Å, a crystal-to-detector distance of 150 mm, an oscillation angle of 1.0° and an exposure time of 2 s. Data were collected from the PTP ϵ D2 crystal using a wavelength of 1.0000 Å, a crystal-to-detector distance of 150 mm, an oscillation angle of 1.0° and an exposure time of 1 s. Diffraction data were collected from the PTP ϵ D2 mutant (A455N/V457Y/E597D) using a MAR345 detector mounted on a Rigaku MicroMax-007 HF high-intensity microfocus generator equipped with VariMax HF optics (Rigaku, The Woodlands, Texas, USA) and operated at 40 kV and 30 mA ($\lambda = 1.5418$ Å). Diffraction images were collected using a crystal-to-detector distance of 120 mm, an oscillation angle of 0.5° and an exposure time of 5 min. All X-ray diffraction images were processed with *HKL-3000* (Minor *et al.*, 2006). Data-collection and processing statistics are presented in Table 2.

The 1.76 Å resolution crystal structure of the PTP ϵ D1 domain was solved by molecular replacement with *Phaser* (McCoy *et al.*, 2007) using the coordinates of chain A extracted from the tandem PTP ϵ D1-D2 structure (PDB entry 2jjd, 100% sequence identity; Almo *et al.*, 2007) and searching for two molecules in the asymmetric unit. The structures of the PTP ϵ D2 wild type and D2 (A455N/V457Y/E597D) mutant were solved by molecular replacement with *Phaser* using the coordinates of chain B from the tandem PTP ϵ D1-D2 structure (PDB entry 2jjd, 100% sequence identity) and searching for two molecules and one molecule per asymmetric unit, respectively. *Coot* (Emsley *et al.*, 2010) was used for manual inspection and rebuilding of the models into electron-density maps and refinements were performed with *phenix.refine* (Afonine *et al.*, 2012). Water molecules were located with *Coot* and visually inspected and refined with *phenix.refine*. All structures were validated using the *MolProbity* server (Chen *et al.*, 2010). Refinement statistics and model validation are outlined in Table 3.

2.5. Phosphatase analysis by pTyr peptide microarrays

All peptides were synthesized to 95% purity with an N-terminal biotin-Ahx linkage and a C-terminal amide by

Table 2
Data collection and processing.

	PTP ϵ D1 domain	PTP ϵ D2 domain	PTP ϵ (A455N/V457Y/E597D) D2 domain
Diffraction source	SER-CAT 22-ID	SER-CAT 22-ID	Rigaku MicroMax-007 HF
Wavelength (Å)	1.0	1.0	1.5418
Temperature (K)	100	100	100
Detector	MAR CCD 300	MAR CCD 300	MAR345
Crystal-to-detector distance (mm)	150	150	120
Rotation range per image (°)	1.0	1.0	0.5
Total rotation range	180	180	180
Exposure per image (s)	2	1	300
Space group	$C2_1$	$P2_1$	$P2_12_12_1$
a, b, c (Å)	154.1, 35.9, 115.0	56.3, 74.3, 75.7	61.5, 73.4, 75.7
α, β, γ (°)	90.0, 115.7, 90.0	90.0, 91.98, 90.0	90.0, 90.0, 90.0
Mosaicity (°)	0.40	0.64	0.97
Resolution range (Å)	50–1.76 (1.79–1.76)	50–2.27 (2.31–2.27)	50–1.91 (1.94–1.91)
Total No. of reflections	192214	109450	168741
No. of unique reflections	56107	28821	26610
Completeness (%)	98.9 (92.2)	99.9 (100)	97.4 (71.8)
Multiplicity	3.4 (2.5)	3.8 (3.8)	6.3 (3.0)
$\langle I/\sigma(I) \rangle$	19.9 (2.1)	14.0 (2.1)	28.0 (2.0)
R_{merge}	0.082 (0.542)	0.128 (0.887)	0.064 (0.486)
Overall B factor from Wilson plot (Å ²)	18.3	27.8	32.5

Table 3
Structure solution and refinement.

	PTP ϵ D1 domain	PTP ϵ D2 domain	PTP ϵ (A455N/V457Y/E597D) D2 domain
Resolution range (Å)	38.2–1.76 (1.81–1.76)	44.9–2.27 (2.33–2.27)	40.02–1.91 (1.96–1.91)
Completeness (%)	98.5	99.4	97.2
No. of reflections			
Working set	55924 (3555)	28797 (1779)	26533 (1307)
Test set	1991 (130)	2016 (132)	1988 (101)
Final R_{cryst}	0.178 (0.266)	0.173 (0.205)	0.190 (0.297)
Final R_{free}	0.222 (0.306)	0.236 (0.289)	0.225 (0.313)
No. of non-H atoms			
Protein chain A	2313	2217	2233
Protein chain B	2261	2134	—
Polyethylene glycol	—	—	16
Water	532	198	140
R.m.s. deviations			
Bond lengths (Å)	0.006	0.008	0.007
Angles (°)	0.8	0.9	0.8
Average B factors (Å ²)			
Protein chain A	19.0	31.5	36.9
Protein chain B	22.6	33.4	—
Polyethylene glycol	—	—	50.5
Water	30.3	35.5	42.0
Ramachandran plot			
Favored (%)	98.0	94.8	95.6
Allowed (%)	1.6	4.8	3.7
Outliers (%)	0.4	0.4	0.7
PDB code	6d4d	6d3f	6d4f

Peptide 2.0 Inc., Chantilly, Virginia, USA. For peptide immobilization, JAK2, IRS1 and STAT3 peptides were first mixed with the carrier protein NeutrAvidin in a 4:1 ratio for at least 1 h and then printed onto 2×8 grids on nitrocellulose-coated FAST slides at six different concentrations (400, 267, 178, 120, 79 and 33 μM) using an Arrayjet Inkjet Microarrayer (Roslin, UK). Printed slides were dried in a desiccator overnight and stored at 253 K. To perform the dephosphorylation assay, printed peptide-microarray slides were blocked with

1 $\times$ FAST blocking buffer (Thermo Scientific, Rockford, Illinois, USA) for 30 min. 16-well gaskets were mounted on the slides to divide the surfaces into 16 wells. Purified recombinant PTP ϵ D1 (Ser107–Gly398) or PTP1B protein in citrate buffer (100 μl total volume) was added to the wells so that the final concentration of enzyme was 2.5 $\mu\text{g ml}^{-1}$ for PTP ϵ D1 and 1.5 $\mu\text{g ml}^{-1}$ for PTP1B. The inhibitor 6e was dissolved in DMSO and diluted to a final concentration of 100 μM in citrate buffer pH 6.4 and 5% (w/v) DMSO. Reference wells were treated with citrate buffer only. After 10 min, the gaskets were removed and the slides were washed with 1 $\times$ TBS + 0.1% Tween 20 buffer.

The slides were incubated with a pTyr mouse mAb (P-Tyr-100; Cell Signaling Technology, Danvers, Massachusetts, USA; 1:1000 dilution) for 1 h (295 K). Next, the slides were washed three times with 1 $\times$ TBS + 0.1% Tween 20 buffer. After the wash, Alexa 635-conjugated goat anti-mouse antibody (1:2000 dilution) was added to the slides and incubated for 1 h. Finally, the slides were washed three times with 1 $\times$ TBS + 0.1% Tween 20 buffer and three times with distilled water. All images were acquired by scanning dried slides at 635 nm wavelength with an Axon Instruments GenePix 4000B Microarray Scanner (Molecular Devices, Sunnyvale, California, USA). Spot intensity results were analyzed with the *GenePix Pro* 5.1 software (Molecular Devices) and the percentage of dephosphorylation was calculated using the formula

$$\text{percentage dephosphorylation (\%)} = \frac{\text{RFU}_{\text{reference}} - \text{RFU}_{\text{enzyme-treated}}}{\text{RFU}_{\text{reference}}} \times 100\%. \quad (1)$$

The average percentage of dephosphorylation was calculated using three concentrations at which the fluorescence signal measurements were in the linear range of spotted peptide concentrations. The percentage inhibition was calculated using the formula

$$\text{percentage inhibition (\%)} = \frac{\text{RFU}_{\text{DMSO}} - \text{RFU}_{\text{inhibitor-treated}}}{\text{RFU}_{\text{DMSO}}} \times 100\%. \quad (2)$$

K_m values were determined by the malachite green phosphatase method (Upstate Biotechnology) according to the kit instructions. Briefly, reactions were set up using citrate buffer (1 mM DTT, 1 mM EDTA pH 6.4) in a total volume of 25 μl in a 385-well black plate with a transparent bottom. The reactions were initiated by adding 0.25 μg purified phosphatase

Figure 1
Stereoviews of the three-dimensional structures of (a) the PTPε D1 domain and (b) the PTPε D2 domain. Secondary structures are illustrated by ribbons and the active-site cysteine residues are depicted as red spheres. (c) Stereoview of the superimposed structures of the PTPε D1 domain (gray ribbons) and the PTPε D2 domain (green ribbons).

Figure 2
Stereoview of the fit of the residues of the PTPε D1 active site and surrounding residues to the final $2F_o - F_c$ electron-density map (blue, 1.76 Å resolution, contoured at the 1σ level). C atoms are colored gray, N atoms blue, O atoms red and S atoms yellow.

proteins to wells containing various concentrations of peptide substrates (0–400 μ M), incubated (295 K, 15 min) and then stopped by the addition of 25 μ l malachite green reagent. The absorbance (650 nm) was measured with a Tecan Infinite M200 plate reader, and the readings were corrected for background absorbance without enzyme. The dephosphorylation of phosphopeptide substrates by PTPε and PTP1B was quantified by measuring the free inorganic phosphate in solution. The K_m and V_{max} values were obtained by fitting the data to the Michaelis–Menten equation using the *OriginPro* 2016 software, where v is the initial velocity, V_{max} is the maximal velocity, S is the substrate concentration and K_m is the Michaelis–Menten constant:

$$v = V_{max}S/(K_m + S). \quad (3)$$

3. Results and discussion

3.1. Structure of the PTPε D1 domain

The crystal structure of the isolated PTPε D1 domain was solved at 1.76 Å resolution with two molecules in the asymmetric unit (Fig. 1a, Table 3). The improved resolution compared with the 3.2 Å resolution structure of the tandem D1–D2 domain (PDB entry 2jjd) provides a more detailed view of the active-site environment, which allowed the fitting of several side-chain residues that were not seen in the original deposited structure. Additionally, the 2jjd structure lacked water molecules, many of which are readily visible and were modeled in the higher resolution data. Overall, the structure exhibits good geometry, with 98% of the residues in the favored region of the Ramachandran plot according to *MolProbity* analysis. Two outlier residues were found: Val378 in both chains A and B. However, the valine side chains fit into the electron-density map very well and are surrounded by a patch of hydrophobic side chains from Val347, Val363, Phe366 and the aryl portion of Tyr384.

The structure of the isolated PTPε D1 domain aligns well with the D1 molecule in the tandem D1–D2 domain construct (r.m.s.d. of 0.30 Å over 1720 atoms). The electron density is well defined for the residues in the phosphate-binding loop, consisting of the HCSAGVGR signature

motif, and the conserved Asp303 residue in the WPD loop (Fig. 2). The WPD loop is observed to exist in the 'open' conformation, as is typically observed in structures of classical PTPs without bound substrate, and is thus positioned appropriately to accept the incoming substrate (Choy *et al.*, 2017).

In comparison with the originally deposited structure of the tandem D1-D2 domains, new details regarding the residues neighboring the phosphate-binding loop are observed in the high-resolution structure of the D1 domain. In the 2jjd structure the side chains of several residues near the active site lacked electron density and thus were not included in the model. Most importantly, despite the presence of six molecules of the tandem D1-D2 structure in the asymmetric unit, the side-chain position of the conserved Asp303 residue in the WPD loop was not defined owing to a lack of electron density. Asp303 is a critical residue in the catalytic mechanism that serves as an acid/base during the dephosphorylation reaction (Denu *et al.*, 1995). The active-site cysteine, which serves as the catalytic nucleophile (Zhang & Dixon, 1993), forms water-mediated hydrogen-bond interactions with the backbone amides of Gly340 and Arg341. In the 2jjd model, a neighboring loop near the WPD loop and active-site phosphate-binding loop consisting of residues Lys237, Glu238, Lys239, Lys240, Glu241 and Glu242 lacked electron density for the side chains; therefore, these side chains are missing as well. On the other hand, these side chains are clearly defined in the 1.76 Å resolution structure of the D1 domain. Of particular note, it is observed that Glu238 forms a salt bridge with the conserved Arg341 in the phosphate-binding loop (Fig. 2). In summary, the new high-resolution structural data provide a more detailed model of the D1 domain. This should facilitate structure-based drug-design efforts.

3.2. Structure of the isolated PTP ϵ D2 domain

The isolated and purified PTP ϵ D2 domain was crystallized and its structure was solved at 2.27 Å resolution (Fig. 1*b*, Table 3). Approximately 95% of the residues are found to reside within the favored region of the Ramachandran plot, with only Val673 in both chains *A* and *B* being flagged as an outlier. However, this residue fits very

Figure 3

(a) Stereoview of the fit of the residues of the PTP ϵ D2 active site and surrounding residues to the final $2F_o - F_c$ electron-density map (blue, 2.27 Å resolution, contoured at the 1σ level). C atoms are colored green, N atoms blue, O atoms red and S atoms yellow. (b) Stereoview of the superimposed residues in the active-site environment of PTP ϵ D1 (C atoms in gray) and PTP ϵ D2 (C atoms in green). (c) The fit of the residues of the PTP ϵ D2 (A455N/V457Y/E597D) active site and surrounding residues to the final $2F_o - F_c$ electron-density map (blue, 1.97 Å resolution, contoured at the 1σ level). C atoms are colored cyan, N atoms blue, O atoms red and S atoms yellow. (d) Stereoview of the superimposed residues in the active-site environment of PTP ϵ D1 (C atoms in gray) and PTP ϵ D2 (A455N/V457Y/E597D) (C atoms in cyan).

Table 4
Enzymatic activities of PTPε D1 domain and PTP1B for pTyr peptide substrates.

Peptide	Substrate	PTPε D1 domain			PTP1B†		
		<i>K_m</i> (μ <i>M</i>)	<i>K_{cat}</i> (s ^{−1})	<i>K_{cat}/K_m</i> (M ^{−1} s ^{−1})	<i>K_m</i> (μ <i>M</i>)	<i>K_{cat}</i> (s ^{−1})	<i>K_{cat}/K_m</i> (M ^{−1} s ^{−1})
JAK2	PODKEY-pY-KVKEPG	119 ± 18	3.1	2.7 × 10 ⁴	72 ± 16	2.7	3.7 × 10 ⁴
IRS1	RKGS GD-pY-MPMSPK	124 ± 14	3.2	2.6 × 10 ⁴	106 ± 27	2.9	2.7 × 10 ⁴
STAT3	PGSAAP-pY-LKTKFI	127 ± 29	3.6	2.8 × 10 ⁴	147 ± 31	3.3	2.3 × 10 ⁴

† Sodium citrate buffer pH 6.4, 298 K.

	KNRY motif	WPD loop	Consensus sequence	Q-loop/α5 helix
PTPε D1	EKNRY P	SWP DE	VH C SAGVG R T	MVQ T DMQ Y T
PTPε D2	KKAR V L	GW P E I	VH C SAGAG R T	MVQ T LEQ Y E
PTPα D1	EKNRY V	SWP D F	VH C SAGVG R T	MVQ T DMQ Y V
PTPα D2	KKNR V L	GW P E V	VH C SAGAG R T	MVQ T LEQ Y E
LAR D1	PKNRY A	AWP D H	VH C SAGVG R T	MVQ T EDQ Y V
LAR D2	FKNR L V	DW P E Q	VH C SAGVG R T	MVQ T EDD Y Q

Figure 4
Sequence alignment of amino-acid sequences from RPTPs containing D1 and D2 domains, highlighting the residues (colored red) that are critical for catalytic activity in the D1 domain.

well into the electron-density map. The overall tertiary fold of the D2 domain is very similar to that of the D1 domain, and the D2 domain possesses essentially the same core architectural arrangement of secondary-structure elements as the D1 domain (Fig. 1c). The main structural features of both domains consist of a central, twisted β-sheet surrounded by α-helices on both sides. When superimposed, the structures exhibit an r.m.s.d of 0.71 Å over 1409 atoms. One significant difference in their tertiary structures is the absence of the N-terminal wedge motif that is found in the D1 domain but is absent in the D2 domain (Fig. 1c). However, the two domains drastically differ biochemically, in that virtually all of the catalytic activity of PTPε is found to reside within the D1 domain, while only low intrinsic activity is observed in the distal D2 domain (Lim *et al.*, 1997). The biological function of the PTPε D2 domain has yet to be elucidated, but it has been proposed that RPTP D2 domains exhibit a distinct function from that of their D1 domains, which may vary among RPTP family members. For instance, studies have shown that synchronized structural changes may regulate inter-domain crosstalk between the D1 and D2 domains that results in either inhibitory or activating effects (Madan *et al.*, 2011). Interaction of the D2 domain with other RPTPs has also been observed (Blanchetot & den Hertog, 2000). Studies have shown that the D2 domain in RPTPδ binds to and inhibits the D1 domain of RPTPσ (Wallace *et al.*, 1998). Therefore, a variety of functions that could be mediated by the D2 domains in different RPTP family members and the inactive D2 domains may have evolved from a common ancestor over time with accumulated mutations to define their specific functions (Pils & Schultz, 2004).

The active sites of the D1 and D2 domains are structurally very similar to each other as well as to other classical PTPs in

which the P-loop forms a cradle that serves as the phosphopeptide-binding site (Figs. 3a and 3b). The identities of the residues comprising the active-site HCSAGAGR motif are conserved between the two structures. The WPD loop is observed to adopt the ‘open’ conformation in both structures but exhibits one key difference. In the D1 structure the WPD loop harbors the Asp303 residue that serves as a general acid/base during the dephosphorylation reaction. However, in the D2 domain the equivalent position is occupied by the larger Glu597 residue. When a phosphopeptide binds to the active site of a PTP this loop undergoes a conformational shift and adopts a ‘closed’ conformation to properly position the aspartic acid for catalysis (Jia *et al.*, 1995). It is possible that the larger glutamate side chain in the D2 domain is sterically hindered, which may contribute to the lack of appreciable catalytic activity in the D2 domain. Other key differences between the active sites include a substitution of the conserved Tyr165 residue in the D1 domain by a valine residue (Val457) in the pTyr recognition loop (KNRY loop) of the D2 domain structure. This tyrosine residue has been shown to interact with the pTyr substrate in other phosphatases such as in the crystal structure of the PTP1B–phosphopeptide complex. It has been proposed that loss of this tyrosine side chain in the D2 domain would eliminate an important stabilizing interaction with the substrate (Jia *et al.*, 1995). Mutation of the catalytically important invariant aspartic acid and tyrosine residues is commonly found within other D2 domains (Fig. 4; Lim *et al.*, 1998). Another difference is observed in the KNRY loop, in which the Asn455 residue found in the D1 domain is replaced by an alanine residue (Ala455) in the D2 domain.

Earlier work by Lim and coworkers demonstrated the low intrinsic activity of the PTPε D2 domain, suggesting that it provides little, if any, direct catalytic contribution to the phosphatase activity of the intact PTPε protein (Lim *et al.*, 1997, 1999). The authors sought to restore catalytic activity to the D2 domain by introducing the V457Y/E597D mutations. Yet, while some increase in the activity of the double mutant was observed, the catalytic power of this construct was still substantially lower than that of the D1 domain. To examine the structural consequences of these mutations, we designed a PTPε D2 construct harboring the V457Y/E597D mutations along with an additional A455N mutation to restore the sequence identity of the KNRY loop and solved the crystal structure at 1.91 Å resolution (Fig. 3c, Table 3).

As reported by Lim *et al.* (1999), we also observed low intrinsic activity of this D2 domain triple mutant using a colorimetric *para*-nitrophenylphosphate (pNPP) assay (data

not shown). The crystal structure of the A455N/V457Y/E597D D2 domain reveals that the introduction of these mutations results in an active site that aligns well with the D1 active site, with only slight shifts in the Tyr165 (Tyr457) and Gln379 (Gln674) residues (Fig. 3*d*). Hence, although the replacement of the invariant residues in the WPD and KRNY loops results in structural conservation of the active site and an observable increase in catalytic activity compared with the native D2 domain, it is still not sufficient to endow it with similar catalytic power to the D1 domain (Lim *et al.*, 1999).

3.3. Identification of *in vitro* phosphopeptide substrates for PTP ϵ

To identify phosphopeptide substrates for the PTP ϵ D1 domain, recognition sites were experimentally determined by measuring the dephosphorylation of 6218 microarrayed pTyr peptides comprising confirmed and theoretical phosphorylation motifs from the cellular proteome. As previously described for other phosphatases (Zhao *et al.*, 2015), a broad continuum of dephosphorylation was observed across the microarrayed peptide substrates (data not shown). We subsequently selected and synthesized three pTyr peptide substrates for the PTP ϵ D1 active domain based on known physiological pathways that involve PTP ϵ and other PTPs (Table 4, Fig. 5*a*). We chose Janus kinase 2 (JAK2) and signal transducer and activator of transcription 3 (STAT3) peptides because many PTPs regulate the JAK/STAT pathway (Xu & Qu, 2008). Moreover, a recent study demonstrated that PTP ϵ negatively regulates hypothalamic leptin signaling by directly dephosphorylating JAK2 (Rousso-Noori *et al.*, 2011). Further,

an earlier study established that many PTPs (for example PTP1B, SHP-2, LAR and LRP) dephosphorylate insulin receptor substrate-1 (IRS1), with PTP1B exhibiting the highest specific activity (Goldstein *et al.*, 2000). As the structure of PTP ϵ is closely related to that of PTP1B, we anticipated that the IRS1 peptide may also be a good substrate for PTP ϵ . To ascertain whether or not the selected phosphopeptides are *in vitro* substrates of PTP ϵ , we immobilized the phosphopeptides on nitrocellulose-coated slides *via* biotin–NeutrAvidin linkers and conducted the dephosphorylation experiments in a microarray format. The chief benefit of the microarray dephosphorylation assay is that the dephosphorylation of multiple peptide substrates can be measured simultaneously within a single well. As shown in Fig. 5(*b*), all three pTyr peptides were readily dephosphorylated by both the PTP ϵ D1 domain and PTP1B. The relative order of pTyr peptide dephosphorylation was JAK2 > IRS1 > STAT3 (Fig. 5*c*). The kinetic data for PTP1B and the PTP ϵ D1 domain are summarized in Table 4. These results confirmed that three synthetic pTyr peptides are active *in vitro* substrates for PTP ϵ .

3.4. Evaluation of small-molecule inhibitors

Next, we investigated whether the pTyr peptide microarray could be used as a high-throughput method for screening potential PTP ϵ inhibitors. Previously, we reported that the potent YopH inhibitor 6e (Fig. 6*a*) also inhibited PTP1B with an IC₅₀ of 2.23 μ M (Bahta *et al.*, 2011). We evaluated inhibitor 6e in the microarray PTP ϵ D1 domain dephosphorylation assay and observed a significant increase in fluorescence signals for both PTP ϵ -treated and PTP1B-treated pTyr peptides (Fig. 6*b*), suggesting that the inhibitor 6e inhibits the catalytic activity of both PTP ϵ and PTP1B. Furthermore, we observed an overall higher inhibition of the activity of PTP ϵ compared with PTP1B for all three pTyr peptides (Fig. 6*c*). In contrast to the results obtained with the peptide substrates, which demonstrated equivalent inhibition for both PTP1B and

Figure 5 Relative dephosphorylation of JAK2, IRS1 and STAT3 pTyr peptide substrates. (*a*) Amino-acid sequences of the pTyr peptide substrates. (*b*) Fluorescence images of PTP ϵ D1 domain- and PTP1B-treated microarrays. (*c*) Relative peptide dephosphorylation compared with control wells treated with assay buffer only. The JAK2, IRS1 and STAT3 pTyr peptides were immobilized on nitrocellulose-coated slides.

PTP ϵ , only inhibition of PTP1B was apparent with the small-molecule substrate pNPP (Fig. 6*d*). While the greater number of molecular contacts expected between the PTPase and

peptide substrate represents a higher barrier for inhibition, we speculate that the difference in activity between PTP1B and PTP ϵ in the pNPP assay is owing to better competition of 6e

Figure 6

A high-throughput microarray assay for screening PTP ϵ D1 domain inhibitors. (a) Structure of the phosphatase inhibitor 6e. (b) Scanned images of the pTyr peptides to evaluate the inhibition of PTP ϵ D1 and PTP1B by inhibitor 6e. (c) Quantitation of the inhibition represented in (b). (d) Comparison of the inhibition by inhibitor 6e of the hydrolysis of *para*-nitrophenylphosphate (pNPP) by PTP ϵ D1 and PTP1B.

for the small-molecule substrate with PTP1B in comparison with the peptide.

Acknowledgements

We thank the Biophysics Resource in the Structural Biophysics Laboratory, NCI at Frederick for use of the LC/ESMS instrument, and Brian P. Austin for technical assistance. X-ray diffraction data were collected on the Southeast Regional Collaborative Access Team (SER-CAT) beamline 22-ID at the Advanced Photon Source, Argonne National Laboratory. Supporting institutions may be found at <http://www.ser-cat.org/members.html>. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services or the US Army, nor does the mention of trade names, commercial products or organizations imply endorsement by the US Government.

Funding information

This project has been funded in whole or in part with Federal funds from the Frederick National Laboratory for Cancer Research, National Institutes of Health under contract HHSN261200800001E and the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research. Use of the Advanced Photon Source was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under contract No. W-31-109-Eng-38.

References

- Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H. & Adams, P. D. (2012). *Acta Cryst. D* **68**, 352–367.
- Almo, S. C. *et al.* (2007). *J. Struct. Funct. Genomics*, **8**, 121–140.
- Alonso, A., Sasin, J., Bottini, N., Friedberg, I., Friedberg, I., Osterman, A., Godzik, A., Hunter, T., Dixon, J. & Mustelin, T. (2004). *Cell*, **117**, 699–711.
- Andersen, J. N., Elson, A., Lammers, R., Rømer, J., Clausen, J. T., Møller, K. B. & Møller, N. P. (2001). *Biochem. J.* **354**, 581–590.
- 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. & Møller, N. P. (2001). *Mol. Cell. Biol.* **21**, 7117–7136.
- Bahta, M. & Burke, T. R. Jr (2012). *Curr. Med. Chem.* **19**, 5726–5734.
- Bahta, M., Lountos, G. T., Dyas, B., Kim, S. E., Ulrich, R. G., Waugh, D. S. & Burke, T. R. Jr (2011). *J. Med. Chem.* **54**, 2933–2943.
- Barr, A. J. (2010). *Future Med. Chem.* **2**, 1563–1576.
- Barr, A. J., Ugochukwu, E., Lee, W. H., King, O. N. F., Filippakopoulos, P., Alfano, I., Savitsky, P., Burgess-Brown, N. A., Müller, S. & Knapp, S. (2009). *Cell*, **136**, 352–363.
- Berman-Golan, D., Granot-Attas, S. & Elson, A. (2008). *Cancer Metastasis Rev.* **27**, 193–203.
- Bialy, L. & Waldmann, H. (2005). *Angew. Chem. Int. Ed.* **44**, 3814–3839.
- Blanchetot, C. & den Hertog, J. (2000). *J. Biol. Chem.* **275**, 12446–12452.
- Blaskovich, M. A. (2009). *Curr. Med. Chem.* **16**, 2095–2176.
- Böhmer, F., Szedlacsek, S., Tabernero, L., Östman, A. & den Hertog, J. (2013). *FEBS J.* **280**, 413–431.
- Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst. D* **66**, 12–21.
- Choy, M. S., Li, Y., Machado, L. E. S. F., Kunze, M. B. A., Connors, C. R., Wei, X., Lindorff-Larsen, K., Page, R. & Peti, W. (2017). *Mol. Cell*, **65**, 644–658.e5.
- Denu, J. M. & Dixon, J. E. (1998). *Curr. Opin. Chem. Biol.* **2**, 633–641.
- Denu, J. M., Lohse, D. L., Vijayalakshmi, J., Saper, M. A. & Dixon, J. E. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 2493–2498.
- Denu, J. M., Zhou, G., Guo, Y. & Dixon, J. E. (1995). *Biochemistry*, **34**, 3396–3403.
- Elson, A. (1999). *Oncogene*, **18**, 7535–7542.
- Elson, A. (2017). *Int. J. Biochem. Cell Biol.* **96**, 135–147.
- Elson, A. & Leder, P. (1995). *J. Biol. Chem.* **270**, 26116–26122.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst. D* **66**, 486–501.
- Espada, J. & Martín-Pérez, J. (2017). *Int. Rev. Cell. Mol. Biol.* **331**, 83–122.
- Fauman, E. B. & Saper, M. A. (1996). *Trends Biochem. Sci.* **21**, 413–417.
- Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D. & Bairoch, A. (2003). *Nucleic Acids Res.* **31**, 3784–3788.
- Ghattas, M. A., Raslan, N., Sadeq, A., Al Sorkhy, M. & Atatreh, N. (2016). *Drug. Des. Dev. Ther.* **10**, 3197–3209.
- Goldstein, B. J., Bittner-Kowalczyk, A., White, M. F. & Harbeck, M. (2000). *J. Biol. Chem.* **275**, 4283–4289.
- Granot-Attas, S., Knobler, H. & Elson, A. (2007). *Crit. Rev. Eukaryot. Gene Expr.* **17**, 49–71.
- Guan, K. L. & Dixon, J. E. (1991). *J. Biol. Chem.* **266**, 17026–17030.
- Guan, K. L. & Dixon, J. E. (1993). *Semin. Cell Biol.* **4**, 389–396.
- Hardy, S., Julien, S. G. & Tremblay, M. L. (2012). *Anticancer Agents Med. Chem.* **12**, 4–18.
- Hunter, T. (1995). *Cell*, **80**, 225–236.
- Itoh, M., Streuli, M., Krueger, N. X. & Saito, H. (1992). *J. Biol. Chem.* **267**, 12356–12363.
- Jia, Z., Barford, D., Flint, A. J. & Tonks, N. K. (1995). *Science*, **268**, 1754–1758.
- Kapust, R. B., Tözsér, J., Fox, J. D., Anderson, D. E., Cherry, S., Copeland, T. D. & Waugh, D. S. (2001). *Protein Eng.* **14**, 993–1000.
- Lim, K. L., Kolatkar, P. R., Ng, K. P., Ng, C. H. & Pallen, C. J. (1998). *J. Biol. Chem.* **273**, 28986–28993.
- Lim, K. L., Lai, D. S., Kalousek, M. B., Wang, Y. & Pallen, C. J. (1997). *Eur. J. Biochem.* **245**, 693–700.
- Lim, K. L., Ng, C. H. & Pallen, C. J. (1999). *Biochim. Biophys. Acta*, **1434**, 275–283.
- Madan, L. L., Veeranna, S., Shameer, K., Reddy, C. C., Sowdhamini, R. & Gopal, B. (2011). *PLoS One*, **6**, e24766.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- Minor, W., Cymborowski, M., Otwinowski, Z. & Chruszcz, M. (2006). *Acta Cryst. D* **62**, 859–866.
- Mustelin, T. (2006). *Adv. Exp. Med. Biol.* **584**, 53–72.
- Nakamura, K., Mizuno, Y. & Kikuchi, K. (1996). *Biochem. Biophys. Res. Commun.* **218**, 726–732.
- Nallamsetty, S., Austin, B. P., Penrose, K. J. & Waugh, D. S. (2005). *Protein Sci.* **14**, 2964–2971.
- Nam, H. J., Poy, F., Krueger, N. X., Saito, H. & Frederick, C. A. (1999). *Cell*, **97**, 449–457.
- Nunes-Xavier, C. E., Martín-Pérez, J., Elson, A. & Pulido, R. (2013). *Biochim. Biophys. Acta*, **1836**, 211–226.
- Parsons, S. J. & Parsons, J. T. (2004). *Oncogene*, **23**, 7906–7909.
- Pils, B. & Schultz, J. (2004). *Mol. Biol. Evol.* **21**, 625–631.
- Rousso-Noori, L., Knobler, H., Levy-Apter, E., Kuperman, Y., Neufeld-Cohen, A., Keshet, Y., Akepati, V. R., Klinghoffer, R. A., Chen, A. & Elson, A. (2011). *Cell Metab.* **13**, 562–572.
- Sarmiento, M., Puius, Y. A., Vetter, S. W., Keng, Y.-F., Wu, L., Zhao, Y., Lawrence, D. S., Almo, S. C. & Zhang, Z.-Y. (2000). *Biochemistry*, **39**, 8171–8179.
- Stanford, S. M. & Bottini, N. (2017). *Trends Pharmacol. Sci.* **38**, 524–540.
- Sun, H. & Tonks, N. K. (1994). *Trends Biochem. Sci.* **19**, 480–485.

- Tautz, L., Pellecchia, M. & Mustelin, T. (2006). *Expert Opin. Ther. Targets*, **10**, 157–177.
- Tsujikawa, K., Kawakami, N., Uchino, Y., Ichijo, T., Furukawa, T., Saito, H. & Yamamoto, H. (2001). *Mol. Endocrinol.* **15**, 271–280.
- Wabakken, T., Hauge, H., Funderud, S. & Aasheim, H. C. (2002). *Scand. J. Immunol.* **56**, 276–285.
- Wallace, M. J., Fladd, C., Batt, J. & Rotin, D. (1998). *Mol. Cell. Biol.* **18**, 2608–2616.
- Wu, L., Buist, A., den Hertog, J. & Zhang, Z.-Y. (1997). *J. Biol. Chem.* **272**, 6994–7002.
- Xu, D. & Qu, C.-K. (2008). *Front. Biosci.* **13**, 4925–4932.
- Zhang, Z.-Y. (1998). *Crit. Rev. Biochem. Mol. Biol.* **33**, 1–52.
- Zhang, Z.-Y. & Dixon, J. E. (1993). *Biochemistry*, **32**, 9340–9345.
- Zhang, Z.-Y. & Dixon, J. E. (1994). *Adv. Enzymol. Relat. Areas Mol. Biol.* **68**, 1–36.
- Zhang, Z.-Y., Wang, Y., Wu, L., Fauman, E. B., Stuckey, J. A., Schubert, H. L., Saper, M. A. & Dixon, J. E. (1994). *Biochemistry*, **33**, 15266–15270.
- Zhao, B. M., Keasey, S. L., Tropea, J. E., Lountos, G. T., Dyas, B. K., Cherry, S., Raran-Kurussi, S., Waugh, D. S. & Ulrich, R. G. (2015). *PLoS One*, **10**, e0134984.
- Zhou, B., He, Y., Zhang, X., Xu, J., Luo, Y., Wang, Y., Franzblau, S. G., Yang, Z., Chan, R. J., Liu, Y., Zheng, J. & Zhang, Z.-Y. (2010). *Proc. Natl Acad. Sci. USA*, **107**, 4573–4578.