Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Water-molecule network and active-site flexibility of apo protein tyrosine phosphatase 1B

Anja K. Pedersen,[a,b,c] Günther H. Peters,[d] Karin B. Møller,[b] Lars F. Iversen[a] and Jette S. Kastrup[c]*

[a]Protein Science, Novo Nordisk, DK-2880 Bagsværd, Denmark, [b]Signal Transduction, Novo Nordisk, DK-2880 Bagsværd, Denmark, [c]Department of Medicinal Chemistry, The Danish University of Pharmaceutical Sciences, DK-2100 Copenhagen, Denmark, and [d]Department of Chemistry, MEMPHYS–Center for Biomembrane Physics, Technical University of Denmark, DK-2880 Lyngby, Denmark

Correspondence e-mail: jsk@dfuni.dk

Received 26 March 2004
Accepted 21 June 2004

PDB Reference: apo protein tyrosine phosphatase 1B, 1sug, r1sugsf.

Protein tyrosine phosphatase 1B (PTP1B) plays a key role as a negative regulator of insulin and leptin signalling and is therefore considered to be an important molecular target for the treatment of type 2 diabetes and obesity. Detailed structural information about the structure of PTP1B, including the conformation and flexibility of active-site residues as well as the water-molecule network, is a key issue in understanding ligand binding and enzyme kinetics and in structure-based drug design. A 1.95 Å apo PTP1B structure has been obtained, showing four highly coordinated water molecules in the active-site pocket of the enzyme; hence, the active site is highly solvated in the apo state. Three of the water molecules are located at positions that approximately correspond to the positions of the phosphate O atoms of the natural substrate phosphotyrosine and form a similar network of hydrogen bonds. The active-site WPD-loop was found to be in the closed conformation, in contrast to previous observations of wild-type PTPs in the apo state, in which the WPD-loop is open. The closed conformation is stabilized by a network of hydrogen bonds. These results provide new insights into and understanding of the active site of PTP1B and form a novel basis for structure-based inhibitor design.

1. Introduction

Reversible protein tyrosine phosphorylation reactions play important roles in the regulation of many biological processes in the eukaryotic cell. Tyrosine phosphorylation is regulated by protein tyrosine kinases, which catalyze the addition of phosphate to tyrosine residues, and by protein tyrosine phosphatases (PTPs), which remove the phosphate group from phosphotyrosine (pTyr). Owing to their key roles as both positive and negative regulators of signal-transduction pathways, *e.g.* insulin and leptin receptor signalling (Cheng *et al.*, 2002; Kenner *et al.*, 1996; Zabolotny *et al.*, 2002), the PTPs have received much attention during the last decade as attractive drug targets in several diseases such as cancer, diabetes, obesity and inflammation (Johnson *et al.*, 2002; Kennedy & Ramachandran, 2000; Møller *et al.*, 2000; Tobin & Tam, 2002; Zhang, 1998).

The catalytic domain of PTPs contains the characteristic active-site sequence known as the PTP signature motif, HC*XX*G*X*GR(S/T), with *X* representing varying amino acids. This sequence constitutes the phosphate-binding loop, denoted the P-loop. The cysteine (Cys215 in PTP1B) is required for catalysis, making a nucleophilic attack on the P atom of the substrate (Guan & Dixon, 1991), and the arginine (Arg221) is important for substrate binding and transition-state stabilization (Zhang *et al.*, 1994). Upon substrate or inhibitor binding, the WPD-loop is believed to move from an

© 2004 International Union of Crystallography
Printed in Denmark – all rights reserved

'open' to a 'closed' conformation. This reorientation moves the WPD-loop by about 10 Å to close over the active site; hence, the WPD-loop possesses inherent flexibility. The aspartic acid (Asp181) situated in the WPD-loop is directly involved in PTP-mediated catalysis and serves as a general acid in the first step of the catalytic mechanism (formation of the phosphoenzyme complex) by protonating the leaving group. In the second step (release of inorganic phosphate), it acts as a general base by abstracting a proton from a catalytic water molecule held in the required position by a conserved glutamine (Gln262; Pannifer *et al.*, 1998). The catalytic mechanism of PTPs is reviewed in Barford *et al.* (1998) and Zhang (1998).

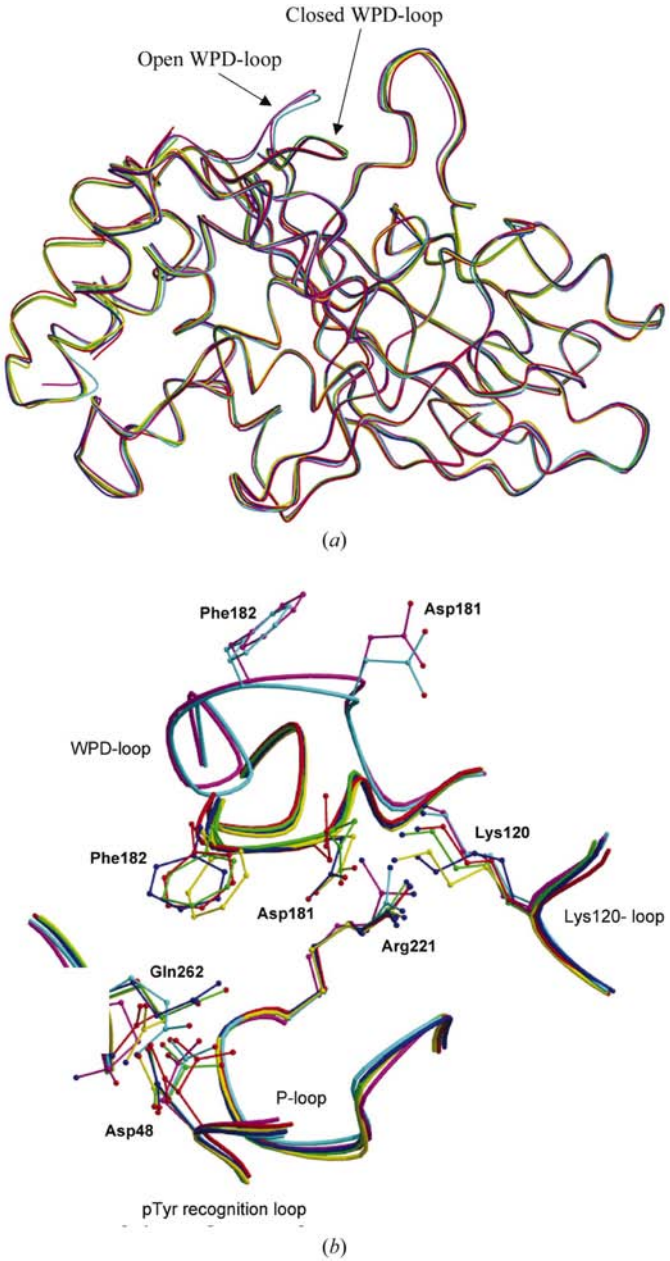

Figure 1
Comparison of PTP1B structures. (*a*) Superimposition of the structure of apo PTP1B (in green) with the 2.85 Å structure of apo PTP1B (2hnp, in magenta) and the structures of PTP1B in complex with tungstate (2hnq, in cyan), with pTyr (1pty, in red), with vanadate (in blue; Pedersen, Guo *et al.*, 2004) and with the inhibitor 6-(oxalylamino)-1*H*-indole-5-carboxylic acid (1c83, in yellow). The r.m.s.d.s on C^{α} atoms between apo PTP1B and the superimposed structures are 0.20 Å for the vanadate structure, 0.22 Å for 1c83, 0.27 Å for 1pty, 0.45 Å for 2hnp and 0.54 Å for 2hnq. The structures are shown as tube representations using the programs *MOLSCRIPT* (Kraulis, 1991) and *RASTER3D* (Merritt & Bacon, 1997). (*b*) Close-up of the active-site residues. Residues 47–49 of the pTyr-recognition loop, 119–121 of the Lys120-loop, 179–187 of the WPD-loop, 215–221 of the P-loop and 261–263 of the Q-loop are shown in tube representations. The side chains of residues Asp48, Lys120, Asp181, Phe182, Arg221 and Gln262 are shown.

Several structures of PTPs with the WPD-loop in either the open or closed conformations have been determined by X-ray crystallography. Generally, the loop is found in the open conformation when no ligand is bound in the active site (Barford *et al.*, 1994; Bilwes *et al.*, 1996; Hof *et al.*, 1998; Hoffmann *et al.*, 1997; Iversen *et al.*, 2002; Nam *et al.*, 1999; Stuckey *et al.*, 1994; Szedlacsek *et al.*, 2001; Yang *et al.*, 1998). In substrate and inhibitor complexes, the WPD-loop is normally found in the closed conformation (see, for example, Andersen *et al.*, 2000; Bleasdale *et al.*, 2001; Sun *et al.*, 2003), but in some cases an open WPD-loop has also been detected in structures with bound ligands (Barford *et al.*, 1994; Bleasdale *et al.*, 2001; Groves *et al.*, 1998; Larsen *et al.*, 2002).

In general, ligand binding (substrate or inhibitor) to an enzyme consists of several steps, all of which must be correctly evaluated in order to fully understand the binding. In this evaluation it is important to realise that (i) binding takes place in an aqueous environment and the solvated ligand–protein complex will be in equilibrium with unbound solvated ligand and solvated protein and (ii) both ligand and protein are flexible entities with multiple conformations; however, they have only one conformation in the ligand–protein complex. Thus, if the active site of an enzyme is heavily solvated and possesses multiple flexible loops, a low binding affinity for a ligand might reflect a solvation and conformation penalty rather than poor atomic interactions between the ligand and the enzyme (Boström *et al.*, 1998). Up to 50 enzyme complexes of PTP1B with substrates or inhibitors have been solved by X-ray crystallography. With Arg221 at the bottom of the active site, Gln262 at one side, a lysine (Lys120, which hydrogen bonds to some inhibitors) at the other side and Asp181 adjacent to this, the active-site pocket of PTP1B is highly polar. Owing to this polar surface in the active-site pocket, it is likely that water molecules are present in the apo state. However, this has not been investigated; hence, the impact of the water-molecule architecture on substrate and inhibitor binding has yet to be accounted for. Here, we present the 1.95 Å X-ray structure of apo PTP1B, allowing a detailed study of the active-site water-molecule architecture and of active-site flexibility of apo PTP1B in comparison with other PTP structures.

2. Materials and methods

2.1. Expression and purification

Human PTP1B 1–321 was cloned as described previously (Andersen *et al.*, 2000; Iversen *et al.*, 2000) and expressed in

the *Escherichia coli* BL21(DE3) strain. An overnight culture was diluted 1:160 into a total volume of 2 l SOB medium [2%(*w*/*v*) Bacto tryptone, 0.5%(*w*/*v*) Bacto yeast extract, 10 m*M* NaCl, 10 m*M* $MgCl_2$ and 10 m*M* $MgSO_4$] and grown at 310 K until A_{600} reached 0.6. Isopropyl-β-D-thiogalactoside was added to a final concentration of 0.1 m*M* and the incubation was continued at room temperature for 3 h. The cells were harvested by centrifugation and lysed by sonication. Remaining cell debris was removed by centrifugation and the supernatant was purified by affinity chromatography using Sepharose 6B coupled with a selective PTP1B inhibitor (Pedersen, Branner *et al.*, 2004). The affinity-purified PTP1B was then loaded onto a Superdex 200 column (2.5 × 60 cm; Amersham Biosciences) equilibrated in 10 m*M* Tris pH 7.5, 25 m*M* NaCl, 0.2 m*M* EDTA and 3 m*M* DTT. Fractions containing PTP1B (measured by an enzyme-activity assay and SDS–PAGE) were pooled and concentrated to 20 mg ml^{-1} using Centriprep and Centricon YM-10 Centrifugal Filter Devices (Millipore).

2.2. Crystallization

Crystals of apo PTP1B were produced by microseeding with seeds from crystals obtained with PTP1B in complex with the inhibitor 2-(oxalylamino)-4,5,6,7-tetrahydro-thieno[2,3-*c*]-pyridine-3-carboxylic acid (OTP), as seeding proved to be necessary for crystal growth. Crystallization of PTP1B in complex with OTP was performed as described previously (Iversen *et al.*, 2000). Stock seed solutions were prepared by crushing a single crystal of PTP1B complexed with OTP in a 2 µl drop of reservoir solution. 1 µl of the drop was transferred to 9 µl reservoir solution and this solution was then tenfold-diluted eight times. The eight solutions were used for micro-seeding into overnight pre-equilibrated drops containing PTP1B by streaking a whisker through the drops. The seeds were invisible in a light microscope. Crystals were grown by the hanging-drop vapour-diffusion method at room temperature. Drops contained 1 µl protein solution and 1 µl reservoir solution consisting of 0.1 *M* HEPES pH 7.5, 0.2 *M* magnesium acetate, 10–12%(*w*/*v*) polyethylene glycol 8000 and 0.1%(*v*/*v*) β-mercaptoethanol. The reservoir volume was 1 ml. Single crystals appeared within 3–4 d of seeding. The morphology and form of the apo PTP1B crystals are similar to those of PTP1B in complex with OTP; however, the apo crystals are smaller than the crystals of the complex.

2.3. Data collection

Diffraction data were collected at 100 K at the EMBL beamline X11 (wavelength 0.8110 Å) at the DORIS storage ring, DESY, Hamburg using a MAR CCD detector. 1.5 µl of a cryoprotectant solution consisting of 50%(*v*/*v*) glycerol was added to the hanging drop and the crystals were transferred through a drop of cryoprotectant solution and flash-frozen in liquid nitrogen prior to data collection. Data were auto-indexed and processed using *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997) and the *CCP*4 suite of programs (Collaborative Computational Project, Number 4, 1994).

Table 1
Crystallographic data and statistics of structure refinement.

Values in parentheses correspond to the outermost resolution shell.

Space group	$P3_121$
Unit-cell parameters (Å)	$a = b = 88.12$, $c = 103.90$
Resolution range (Å)	45–1.95
Unique reflections	34528
Average redundancy	7.4
Completeness (%)	100 (100)
$R_{\rm merge}$ (%)	7.3 (51.5)
$\langle I/\sigma(I)\rangle$	26.4 (4.4)
Atoms in structure	2720
No. of water/glycerol/Tris molecules	251/4/1
R factor† (%)	18.8
$R_{\rm free}$‡ (%)	20.4
R.m.s.d. bond lengths (Å)/angles (°)	0.005/1.2
Residues in allowed regions of Ramachandran plot§ (%)	98.6
Average B factors of main-chain atoms (Å^2)	25.2
Average B factors of side-chain atoms (Å^2)	28.8

† R factors were calculated using all data. Crystallographic R factor = $\sum_{(hkl)} \big||F_o| - |F_c|\big| / \sum_{(hkl)} |F_o|$. ‡ $R_{\rm free} = \sum_{(hkl)\in T} \big||F_o| - |F_c|\big| / \sum_{(hkl)\in T} |F_o|$, where T is a test set containing a random 5% of the observations omitted from the refinement process. § The Ramachandran plots were calculated according to Kleywegt & Jones (1996).

2.4. Structure determination

The crystal structure of apo PTP1B was solved by molecular replacement using *AMoRe* (Navaza, 1994) from *CCP*4. The coordinates of a high-resolution structure of PTP1B (PDB code 1c83) were used as search model (water molecules and the inhibitor were omitted from the model). Initial model building was performed using *ARP*/*wARP* (Perrakis *et al.*, 1999) and the structure was inspected and corrected using the program *O* (Jones *et al.*, 1991). Gradually, water, glycerol and Tris molecules were inserted. Refinements were performed in *CNS* (Brünger *et al.*, 1998). 5% of the data were used for $R_{\rm free}$ calculation. The N-terminal methionine and the C-terminal 22 residues could not be located in the electron-density maps. To ensure correct interpretation of the electron density for the WPD-loop, the coordinates of the WPD-loop atoms from the search model were omitted upon automated model building in *ARP*/*wARP* and different omit maps were generated using *CNS*. In addition, refinements were carried out, including refinements of occupancies, with the WPD-loop in both the open and closed conformations. The crystallographic data and statistics of structure refinement are given in Table 1.

2.5. Molecular-dynamics simulations

The coordinates from apo PTP1B were used for molecular-dynamics simulations. The initial structure was solvated using the program *SOLVATE* (http://www.mpibpc.gwdg.de/abteilungen/071/solvate/docu.html). Water molecules located outside the initial rectangular box dimensions of 79.8 × 68.6 × 58.3 Å were removed by using appropriate Tcl scripts executed within *VMD* (Humphrey *et al.*, 1996). Six water molecules were replaced with sodium ions in order to neutralize the system. For the simulations, the molecular-dynamics program *NAMD* (Kale *et al.*, 1999) was used with the Charmm27 all-H atoms parameter set and with the TIP3

water model (Jorgensen *et al.*, 1983). The simulations were carried out in an *NPT* ensemble in which the number of atoms *N*, the pressure *P* and the temperature *T* were held constant. The system was initially energy-minimized for 500 steps, which was followed by 100 ps of heating the system to $T = 300$ K. The simulations were performed for 6 ns. A time step of 1 fs was used throughout all simulations. A constant ambient pressure of 101.3 Pa was imposed using the Langevin piston method of Feller *et al.* (1995) with a damping coefficient of 5 ps^{-1}, a piston period of 200 fs and a decay of 100 fs. The Particle Mesh Ewald method was used for computation of the electrostatic forces (Darden *et al.*, 1993; Essmann *et al.*, 1995). The grid spacing applied was approximately 1.0 Å and a fourth-order spline was used for the interpolation. The long-range part of the electrostatic forces was evaluated every fourth femtosecond. The van der Waals interactions were cut off at 12 Å using a switching function starting at 10 Å. Periodic boundary conditions were imposed in all directions. The analyses of the trajectory were performed using appropriate Tcl scripts executed within *VMD* (Humphrey *et al.*, 1996).

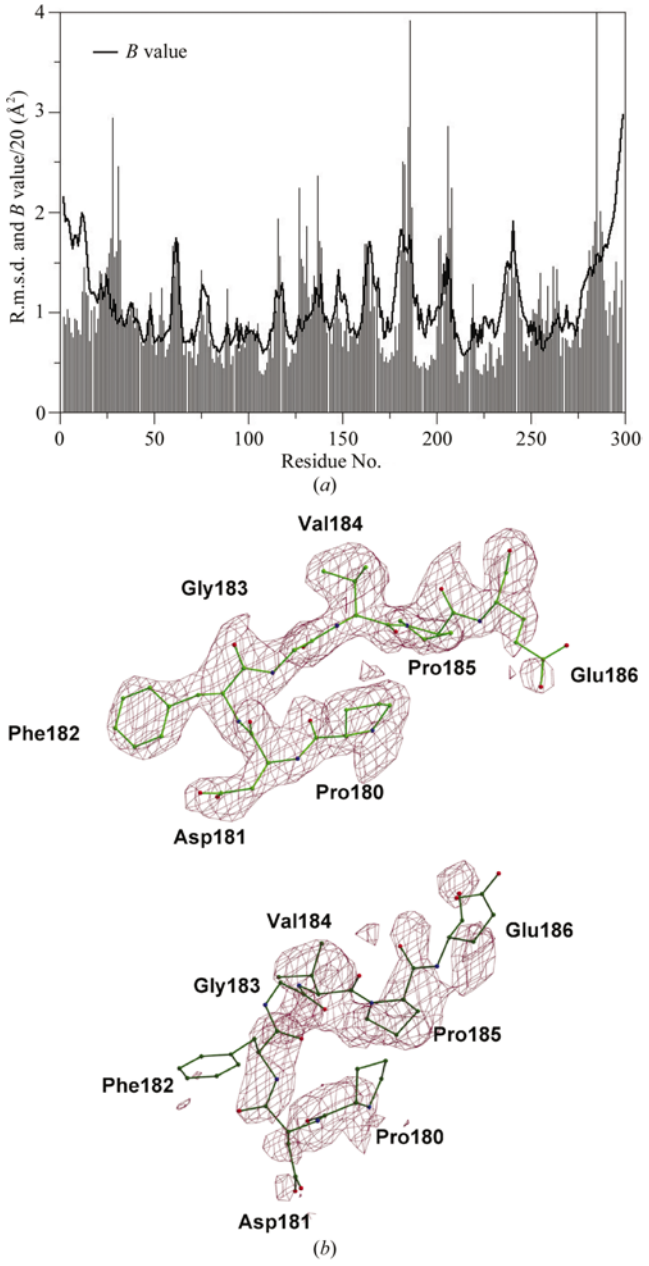

Figure 2
Flexibility of apo PTP1B. (*a*) *B* values (arbitrarily divided by a factor of 20 and shown as a continuous line) and average value of the r.m.s.d. (shown as columns) extracted from the molecular-dynamics simulations for C^{α} atoms as a function of apo PTP1B residue number. The r.m.s.d. data are calculated with respect to the structure extracted after 1 ns simulation. (*b*) WPD-loop $2F_o - F_c$ omit electron density for the closed and open conformations of the loop is shown in magenta in the upper and lower part of the figure, respectively. All atoms of the WPD-loop are shown in ball-and-stick representation in light green for the closed conformation and in dark green for the open conformation. N atoms are shown in blue and O atoms in red. The figure was prepared using the programs *MOLSCRIPT* (Kraulis, 1991) and *RASTER3D* (Merritt & Bacon, 1997).

3. Results

3.1. The 1.95 Å X-ray structure of apo PTP1B

The apo structure of PTP1B was refined to an *R* value of 18.8% (R_{free} = 20.4%) at 1.95 Å resolution; statistics are listed in Table 1. The overall structure of apo PTP1B is similar to other reported PTP1B structures crystallized in the same space group but representing different binding modes in the protein (see Fig. 1*a*) (as reflected by the r.m.s.d. on C^{α} atoms, which ranges between 0.20 and 0.54 Å). The only area in the enzyme where significant conformational differences are observed between apo PTP1B and other PTP1B structures is the active site. The most striking difference between apo PTP1B and the previously determined 2.85 Å apo PTP1B structure (termed 2hnp apo PTP1B; Barford *et al.*, 1994) is found in the conformation of the WPD-loop, which is closed in apo PTP1B in contrast to the open conformation seen in 2hnp apo PTP1B (Figs. 1*a* and 1*b*). Furthermore, water molecules were not determined in 2hnp apo PTP1B.

3.2. Flexibility of apo PTP1B

Analysis of the *B*-value distribution as a function of residue number, as well as the results from molecular-dynamics (MD) simulations, revealed that the overall structure of the protein is relatively rigid (Fig. 2*a*). The variations displayed in the r.m.s.d. data extracted from a 6 ns simulation are in good agreement with the experimentally determined *B* values. Both the experimental data and the r.m.s.d.s calculated from the MD simulations reveal similar localized regions of relatively high flexibility. The flexible regions are localized around residues 29, 63, 117, 128, 137, 166, 183, 208, 243 and 287. These regions include the WPD-loop and loop regions which are part of a turn connecting secondary-structural elements. For the regions located around residues 63, 208, 243 and 287, the r.m.s.d. data and low standard deviations from the MD simulation reflect fluctuations of main-chain atoms around an average position with low amplitude, whereas the data for residues 29, 117 and 183 correspond to a combination of structural changes and fluctuations (data not shown). An

interesting observation is that the loop region around residues Lys116–Lys120, termed the Lys120-loop, appears to move concurrently with the WPD-loop.

3.3. Interactions of active-site residues

The electron-density map for the WPD-loop in apo PTP1B is well defined for the closed conformation; however, as an indication of the dynamics of this loop, non-continuous electron density is observed for the residues of the WPD-loop in the open form (see Fig. 2*b*). It has not been possible to build the open form of the WPD-loop in the non-continuous electron-density map owing to the low occupancy of this conformation. The *B* values of apo PTP1B also point towards the high mobility of atoms within the WPD-loop (Fig. 2*a*).

In apo PTP1B, the side chain of Phe182 in the WPD-loop forms van der Waals contacts with the aliphatic part of the side chains of residues Gln262 and Thr263, as is often the case when the WPD-loop is in the closed conformation. Also stabilizing the loop in the closed conformation is a salt bridge between the side chains of Arg221 and Asp181. In addition, Arg221 interacts with the carbonyl O atom of Pro180 and the side chain of Trp179, as previously described for *Yersinia* PTP (Stuckey *et al.*, 1994). Furthermore, a glycerol molecule is positioned between the side chains of Tyr46 and Phe182, which might provide further stabilization of the WPD-loop in the absence of a stabilizing substrate (Fig. 3*a*). The glycerol molecule forms hydrogen bonds to the side chains of Asp48 and Gln262. Glycerol was also used as cryoprotectant for the data collection of 2hnp apo PTP1B (Barford *et al.*, 1994). As only protein atoms are included in this structure, it is not possible to unravel whether glycerol might be the inducer of the closed conformation of the WPD-loop. The PTP1B used to determine the two apo structures was crystallized under the same conditions, except that the present structure was crystallized using a seeding technique and at 293 K instead of 277 K. The same crystal packing is observed, which allows the WPD-loop to be present in both the open and closed conformations.

Another pronounced difference between the two PTP1B apo structures is the conformation of Gln262, which is important in positioning the catalytic water molecule. Gln262 points towards the active site in apo PTP1B but away from the active site in 2hnp apo PTP1B (Fig. 1*b*). Normally, when a ligand is bound at the active site, Gln262 points away from the active site and only in PTP1B structures with the transition-state analogue vanadate bound is Gln262 found pointing towards the active site. By analysis of all PTP apo structures deposited in the PDB, the residue equivalent to Gln262 is found to point towards the active site in six out of eight structures, reflecting the flexibility of residues involved in PTP-mediated catalysis. In general, when Gln262 points away from the active site, it is mainly stabilized by hydrogen-bond formation to the main-chain carbonyl O atom of Gly259 (PTP1B numbering), whereas the conformation facing the active site is stabilized by interactions with active-site water molecules.

3.4. The active-site water-molecule network

Four water molecules (W1–W4) are found in the active site of apo PTP1B (Fig. 3*b*). Two of these (W1 and W2) are very well defined in the electron-density maps, whereas the other two (W3 and W4) are situated in the same elongated density, being only 2.3 Å apart. This might indicate that a disordered water molecule is present in this region, moving between two positions (Ko *et al.*, 2003). However, additional positive electron density in the $F_o - F_c$ map was observed upon refinement with only one water molecule (data not shown). W1 and W3 are both situated 3.2 Å from the S atom of the catalytic Cys215 residue. W1 has the lowest *B* value (19 Å^2) of the four water molecules in the active site of apo PTP1B, whereas larger *B* values are observed for W2 (26 Å^2), W3 (31 Å^2) and W4 (32 Å^2).

The hydrogen-bonding pattern of the active-site water molecules can be seen in Fig. 3(*c*), displaying an array of potential hydrogen bonds. W1–W3 form hydrogen bonds with each other and with the backbone N atoms of the P-loop residues in the same way as observed in structures with bound substrates and inhibitors. In addition, W2 hydrogen bonds to Asp181 O$^{\delta1}$, thereby stabilizing the loop in the closed conformation. Together with W3, W2 also forms hydrogen bonds to the side chain of Arg221. Gln262 makes hydrogen bonds to both W1 and W4 and these contacts keep Gln262 pointing towards the active site.

4. Discussion

Protein hydration is known to play a crucial role in biological processes. Depending on the hydrogen-bonding environment of the water molecules in the active site of an enzyme, the water molecules may be used in two ways in structure-based drug design: (i) loosely bound water molecules may be displaced with an acceptable low energy penalty or (ii) tightly bound molecules may be used directly as pre-existing 'building' blocks in the formation of water-mediated hydrogen bonds from substrate (or inhibitor) to the enzyme (Lam *et al.*, 1994; Mikol *et al.*, 1995). The current apo PTP1B structure has demonstrated that the active site is highly solvated. Three extensively coordinated water molecules, W1–W3, were found in the vicinity of the catalytic cysteine and are situated at positions that approximately correspond to the positions of the phosphate O atoms of the substrate pTyr when bound to PTP1B (PDB code 1pty; see Fig. 3*d*). Therefore, these water molecules have to be displaced in all substrate complexes in order to form the required atomic interactions between substrate and enzyme.

Numerous inhibitors of PTP1B have been reported. Although much progress has been made in this field, with the development of highly selective and low-nanomolar PTP1B inhibitors (Shen *et al.*, 2001; Sun *et al.*, 2003), it seems difficult to also obtain cellular activity. This might be because of the fact that most active-site-directed PTP inhibitors synthesized so far mimic the interaction of the natural substrate pTyr and are thus highly charged. More than 30 X-ray structures of such PTP1B–inhibitor complexes have been deposited within the

Protein Data Bank and these structures show that in most cases the inhibitors displace all three water molecules (W1–W3). Of note, the basic low-molecular-weight scaffolds used for the design of, for example, oxamic acid-based inhibitors are of very low potency despite the fact that they utilize most of the interaction points of the natural substrate. Thus, as an example, the K_i value for 2-oxalylamino benzoic acid against PTP1B is only 23 μ*M* (Andersen *et al.*, 2000). It is tempting to speculate that the low affinity arises from the payment of a significant energy penalty to overcome displacement of the water molecules in the active site. The information obtained on the presence and location of these water molecules might be used to design a novel class of inhibitors using water molecules as interaction points rather than the P-loop itself and thereby circumvent the need for highly charged molecules.

One additional water molecule is observed in the active site of apo PTP1B, namely W4, which corresponds to the highly conserved water molecule seen in almost all PTP1B structures with closed WPD-loops, *i.e.* the water molecule coordinating to the side chain of Gln266, to the backbone N atom of Phe182 and to the substrate or inhibitor bound (Fig. 3). Inhibitors that displace all water molecules including W4 can be found in the PDB (Groves *et al.*, 1998). However, W4 is not displaced by the inhibitors as often as W1–W3 (only in about 50% of the structures deposited in the PDB). In most cases, these inhibitors give rise to an open WPD-loop. Thus, this water molecule might be important for stabilizing the WPD-loop of PTP1B in the closed conformation by the formation of water-mediated hydrogen bonds. However, the relatively high *B* value of W4 combined with a reduced coordination by hydrogen bonding compared with W1 suggests that from a

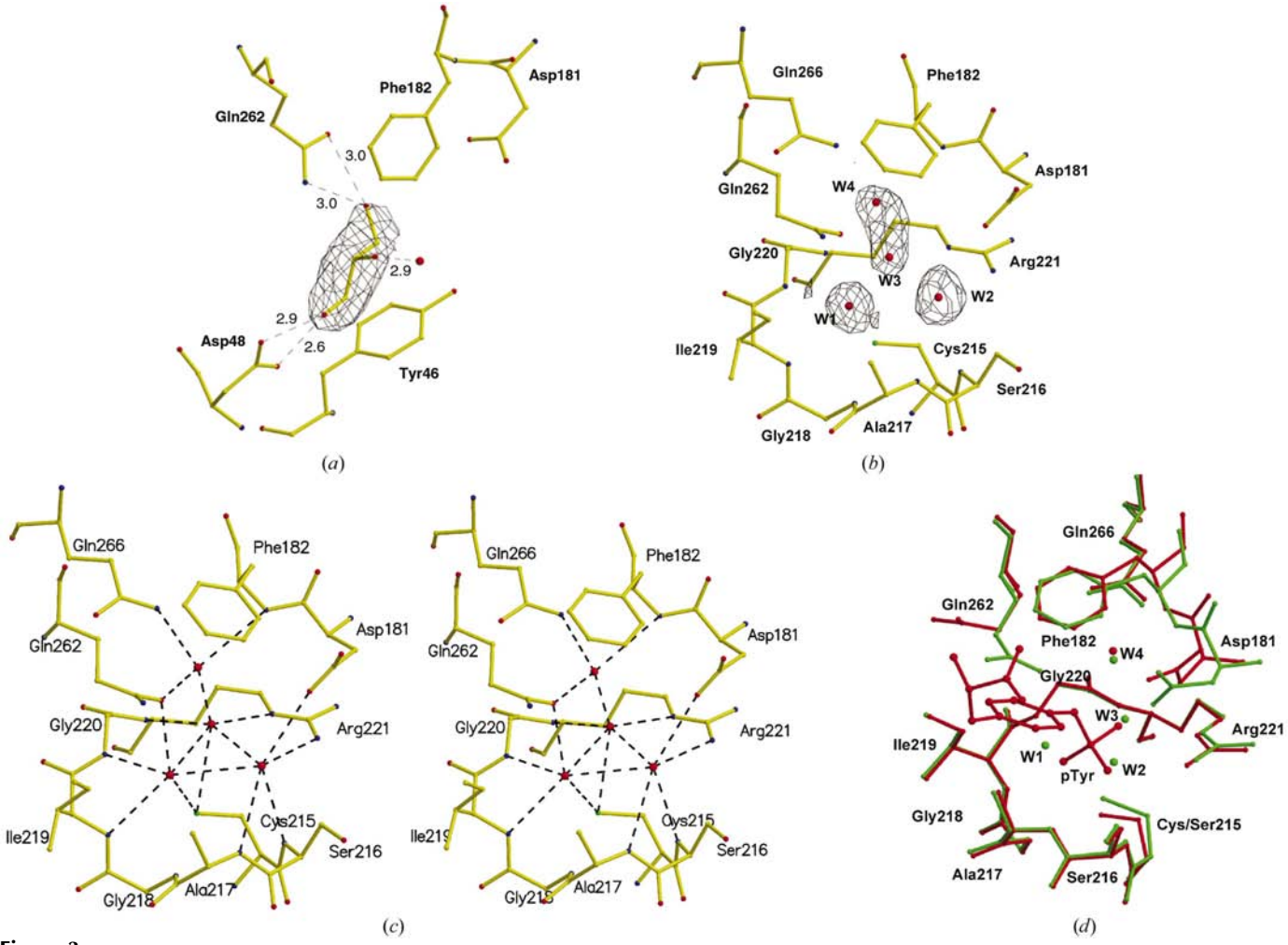

Figure 3
Water-molecule network in the active site of apo PTP1B. (*a*) A glycerol molecule is located between the side chains of Tyr46 and Phe182, close to the active site. Atoms are coloured according to type: carbon is yellow, oxygen is red, nitrogen is blue and sulfur is green. The $2F_o - F_c$ electron density is shown in black at the 1.5σ level. (*b*) Four water molecules (shown as red spheres) situated in the active site of apo PTP1B. The $2F_o - F_c$ electron density is shown in black at the 1.5σ level. Atoms are coloured as in (*a*). (*c*) Potential hydrogen bonds (dashed lines) within 3.4 Å from the water molecules (W1 to Cys215 S^{γ}, 3.2 Å, Ile219 N, 3.1 Å, Gly220 N, 2.7 Å, Gln262 $O^{\varepsilon 1}$, 2.9 Å, W2, 3.3 Å and W3, 3.1 Å; W2 to Asp181 $O^{\delta 1}$, 3.2 Å, Ser216 N, 3.0 Å, Ala217 N, 3.4 Å, Arg221 $N^{\eta 2}$, 3.0 Å, W1, 3.3 Å and W3, 2.8 Å; W3 to Cys215 S^{γ}, 3.3 Å, Arg221 N, 2.9 Å, Arg221 N^{ε}, 2.9 Å, W1, 3.1 Å, W2, 2.8 Å, and W3, 2.3 Å; W4 to Phe182 N, 3.2 Å, Gln262 $O^{\varepsilon 1}$, 2.8 Å, Gln266 $N^{\varepsilon 2}$, 3.2 Å and W3, 2.3 Å). The figure is shown in stereo and individual atoms are coloured as in (*a*). (*d*) Superimposition of apo PTP1B (in green) with PTP1B in complex with pTyr (1pty, in red). W1–W3 are found at positions corresponding to the positions of the phosphate O atoms of pTyr. The figure was prepared with *MOLSCRIPT* (Kraulis, 1991) and *RASTER3D* (Merritt & Bacon, 1997).

drug-design point of view it may be more favourable to displace W4 than W1.

In other apo PTP structures, *i.e.* SHP-2, PTPμ, PTP LAR and PTP-SL/BR7, water molecules are situated at positions similar to W1–W3 observed in apo PTP1B even though the WPD-loop is in the open conformation in these structures. In TC-PTP, two water molecules are found at positions corresponding to W1 and W3. In contrast, water molecules corresponding to W4 are not seen in any of the other PTPs. When analyzing the hydrogen-bonding pattern of the three water molecules, several interactions are found to be the same as in apo PTP1B. Interestingly, the hydrogen bonds from W2 and W3 to the side chain of Arg221 seen in our structure are not present in SHP-2, PTP LAR and PTP-SL/BR7, illustrating the influence of Arg221 on the conformation of the WPD-loop. In contrast, PTPα has no water molecules in the active site owing to the packing of residues from another PTPα molecule into the active site.

The present work demonstrates that, in addition to the expected open conformation, the catalytically important WPD-loop of PTP1B can also attain the closed conformation without bound substrates or inhibitors. This structure is the first example of a wild-type PTP in the apo state in which the WPD-loop is closed (Fig. 1). Interestingly, very recently Romsicki *et al.* (2003) reported the 1.6 Å structure of an apo PTP1B C215D mutant structure in which the WPD-loop also adopts the closed conformation. There is increasing evidence that the WPD-loops of PTPs are relatively flexible and might fluctuate between the open and closed conformation. This has been experimentally monitored by Juszczak *et al.* (1997), who determined the frequency of WPD-loop movement in *Yersinia*. Indirect evidence is provided by several crystal structures of PTPs in which the WPD-loop has been found in either open or closed conformations. The molecular-dynamics simulations confirm the flexibility within the WPD-loop of apo PTP1B, which is also in agreement with the present X-ray structure of apo PTP1B. Comparing the *B* values of residues situated in the WPD-loops in the PTP1B structures superimposed in Fig. 1 reveals the high mobility of the atoms within the WPD-loop in all structures, except in the inhibitor complex 1c83 (average *B* value 33.6 Å^2 on all WPD-loop atoms for apo PTP1B compared with 27.0 Å^2 for all atoms in the protein; for the inhibitor complex, these values are 11.8 Å^2 compared with 16.1 Å^2 for all atoms). This indicates that upon binding of certain inhibitors to PTP1B, a stabilization of the WPD-loop in the closed conformation occurs. However, it seems likely that flexibility of this loop is necessary for enzyme catalysis, allowing substrates to enter and products to leave the active site.

The water molecules W1–W3 have similar positions both in the case of the closed and open conformation of the WPD-loop. The water molecules directly involved in hydrogen bonding to the loop are W2 and W4. As only W4 is specific for the closed conformation, it can be inferred that the major contribution to the stabilization of the closed conformation of the WPD-loop arises from the unique hydrogen bond of the W4 water molecule. However, other forces than hydrogen bonding of the active-site water molecules might be of importance for stabilizing the closed conformation of this loop. In the structure of apo PTP1B C215D, the water-molecule network is comparable to the network seen in apo PTP1B (Romsicki *et al.*, 2003).

In PTP1B structures, the Lys120-loop is always seen to be pointing towards the active site, irrespective of the WPD-loop conformation. The molecular-dynamics simulations reveal concurrent movements of the WPD-loop and the neighbouring Lys120-loop. We therefore speculate that not only does the WPD-loop change conformation upon substrate binding, but also the Lys120-loop. In order for the WPD-loop to move from the open to the closed conformation, either the side chain of Asp181 has to move between Glu115 and Lys116 or the Lys120-loop region has to move (at least Lys116 should change its conformation). Glu115 forms a bidentate salt bridge to Arg221 and the conformation of these residues varies depending on the conformation of the WPD-loop. The movement of Arg221 is believed to be coupled to the movement of the WPD-loop (Stuckey *et al.*, 1994). It may indicate that the interaction between Arg221 and Glu115 is altered upon ligand binding at the active site. This in turn will allow the WPD-loop to close without interfering with Lys116. Following WPD-loop closure, the Lys120-loop could move to its original position and the interaction between Glu115 and Arg221 could be restored. It is interesting to note that the *Yersinia* PTP, the most efficient PTP described to date, has a very different conformation of the Lys120-loop region compared with most other PTPs. The Glu–Arg interaction is also present in *Yersinia* PTP, but the residues following Glu115 (Glu290 in *Yersinia*) adopt conformations similar to those of the corresponding residues in TC-PTP. Here, the loop is disordered and found to be in an 'open' conformation, pointing away from the active-site pocket (Iversen *et al.*, 2002). It might be speculated that the efficiency of the *Yersinia* PTP arises from the conformation of the Lys120-loop, which does not prevent free WPD-loop movement.

The active-site residues forming contacts to substrates (or inhibitors) are situated in the P-loop (residues 215–221) and in four other loop regions defining the boundaries of the active site, including the pTyr-recognition loop (residues 46 and 48–49), the Lys120-loop (residue 120), the WPD-loop (residue 181–182) and the Q-loop (residue 262). As seen in Fig. 1(*b*), the conformation of the side chains of Asp48, Asp181, Phe182, Gln262 and probably also of Lys120 varies depending on the bound state of the enzyme, suggesting an induced fit of these residues. For example, an outwards-facing side-chain conformation of Gln262 is a requirement for substrate or inhibitor binding. Furthermore, the conformation of Arg221 is coupled to the conformation of the WPD-loop. This range of conformations of active-site residues in PTP1B reflects the flexibility of the enzyme and suggests an important role for the catalytic efficiency of the enzyme. This inherent flexibility, combined with the obtained insight into active-site hydration, should be taken into account in the next generation of protein tyrosine phosphatase inhibitors.

Annette S. Petersen is thanked for providing the seeding technique for producing crystals of apo PTP1B and Osman Mirza is thanked for helping with data collection. JSK gratefully acknowledges support from The Danish Medical Research Council, DANSYNC (Danish Center for Synchrotron Based Research), Novo Nordisk Fonden, Apotekerfonden of 1991 and the European Community Access to Research Infrastructure Action of the Improving Human Potential Programme to the EMBL Hamburg Outstation, contract No. HPRI-CT-1999-00017. GHP gratefully acknowledges financial support from the Danish National Research Foundation *via* a grant to the MEMPHYS–Center for Biomembrane Physics. Simulations were performed at the Danish Center for Scientific Computing at the University of Southern Denmark.

References

Andersen, H. S., Iversen, L. F., Jeppesen, C. B., Branner, S., Norris, K., Rasmussen, H. B., Møller, K. B. & Møller, N. P. H. (2000). *J. Biol. Chem.* **275**, 7101–7108.

Barford, D., Das, A. K. & Egloff, M. P. (1998). *Annu. Rev. Biophys. Biomol. Struct.* **27**, 133–164.

Barford, D., Flint, A. J. & Tonks, N. K. (1994). *Science*, **263**, 1397–1404.

Bilwes, A. M., den Hertog, J., Hunter, T. & Noel, J. P. (1996). *Nature (London)*, **382**, 555–559.

Bleasdale, J. E. *et al.* (2001). *Biochemistry*, **40**, 5642–5654.

Boström, J., Norrby, P. O. & Liljefors, T. (1998). *J. Comput.-Aided Mol. Des.* **12**, 383–396.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Cheng, A., Uetani, N., Simoncic, P. D., Chaubey, V. P., Lee-Loy, A., McGlade, C. J., Kennedy, B. P. & Tremblay, M. L. (2002). *Dev. Cell*, **2**, 497–503.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Darden, T., York, D. & Pedersen, L. (1993). *J. Chem. Phys.* **98**, 10089–10092.

Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H. & Pedersen, L. G. (1995). *J. Chem. Phys.* **103**, 8577–8593.

Feller, S. E., Zhang, Y. H., Pastor, R. W. & Brooks, B. R. (1995). *J. Chem. Phys.* **103**, 4613–4621.

Groves, M. R., Yao, Z. J., Roller, P. P., Burke, T. R. & Barford, D. (1998). *Biochemistry*, **37**, 17773–17783.

Guan, K. L. & Dixon, J. E. (1991). *J. Biol. Chem.* **266**, 17026–17030.

Hof, P., Pluskey, S., Dhe-Paganon, S., Eck, M. J. & Shoelson, S. E. (1998). *Cell*, **92**, 441–450.

Hoffmann, K. M. V., Tonks, N. K. & Barford, D. (1997). *J. Biol. Chem.* **272**, 27505–27508.

Humphrey, W., Dalke, A. & Schulten, K. (1996). *J. Mol. Graph.* **14**, 33–38.

Iversen, L. F., Andersen, H. S., Branner, S., Mortensen, S. B., Peters, G. H., Norris, K., Olsen, O. H., Jeppesen, C. B., Lundt, B. F., Ripka, W., Møller, K. B. & Møller, N. P. H. (2000). *J. Biol. Chem.* **275**, 10300–10307.

Iversen, L. F., Møller, K. B., Pedersen, A. K., Peters, G. H., Petersen, A. S., Andersen, H. S., Branner, S., Mortensen, S. B. & Møller, N. P. H. (2002). *J. Biol. Chem.* **277**, 19982–19990.

Johnson, T. O., Ermolieff, J. & Jirousek, M. R. (2002). *Nature Rev. Drug Discov.* **1**, 696–709.

Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Jorgensen, W. L., Chandrasekhar, J., Medura, J. D., Impey, R. W. & Klein, M. L. (1983). *J. Chem. Phys.* **79**, 926–935.

Juszczak, L. J., Zhang, Z. Y., Wu, L., Gottfried, D. S. & Eads, D. D. (1997). *Biochemistry*, **36**, 2227–2236.

Kale, L., Skeel, R., Bhandarkar, M., Brunner, R., Gursoy, A., Krawetz, N., Phillips, J., Shinozaki, A., Varadarajan, K. & Schulten, K. (1999). *J. Comput. Phys.* **151**, 283–312.

Kennedy, B. P. & Ramachandran, C. (2000). *Biochem. Pharmacol.* **60**, 877–883.

Kenner, K. A., Anyanwu, E., Olefsky, J. M. & Kusari, J. (1996). *J. Biol. Chem.* **271**, 19810–19816.

Kleywegt, G. J. & Jones, T. A. (1996). *Structure*, **4**, 1395–1400.

Ko, T., Robinson, H., Gao, Y., Cheng, C. C., DeVries, A. L. & Wang, A. H. (2003). *Biophys. J.* **84**, 1228–1237.

Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.

Lam, P. Y., Jadhav, P. K., Eyermann, C. J., Hodge, C. N., Ru, Y., Bacheler, L. T., Meek, J. L., Otto, M. J., Rayner, M. M. & Wong, Y. N. (1994). *Science*, **263**, 380–384.

Larsen, S. D., Barf, T., Liljebris, C., May, P. D., Ogg, D., O'Sullivan, T. J., Palazuk, B. J., Schostarez, H. J., Stevens, F. C. & Bleasdale, J. E. (2002). *J. Med. Chem.* **45**, 598–622.

Merritt, E. A. & Bacon, D. J. (1997). *Methods Enzymol.* **277**, 505–524.

Mikol, V., Papageorgiou, C. & Borer, X. (1995). *J. Med. Chem.* **38**, 3361–3367.

Møller, N. P. H., Iversen, L. F., Andersen, H. S. & McCormack, J. G. (2000). *Curr. Opin. Drug Discov. Dev.* **3**, 527–540.

Nam, H. J., Poy, F., Krueger, N. X., Saito, H. & Frederick, C. A. (1999). *Cell*, **97**, 449–457.

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Pannifer, A. D. B., Flint, A. J., Tonks, N. K. & Barford, D. (1998). *J. Biol. Chem.* **273**, 10454–10462.

Pedersen, A. K., Branner, S., Mortensen, S. B., Andersen, H. S., Klausen, K. M., Møller, K. B., Møller, N. P. H. & Iversen, L. F. (2004). *J. Chromatogr. B*, **799**, 1–8.

Pedersen, A. K., Guo, X.-L., Møller, K. B., Peters, G. H., Andersen, H. S., Kastrup, J. S., Mortensen, S. B., Iversen, L. F., Zhang, Z.-Y. & Møller, N. P. H. (2004). *Biochem. J.* **378**, 421–433.

Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.

Romsicki, Y., Scapin, G., Beaulieu-Audy, V., Patel, S., Becker, J. W., Kennedy, B. P. & Asante-Appiah, E. (2003). *J. Biol. Chem.* **31**, 29009–29015.

Shen, K., Keng, Y.-F., Wu, L., Guo, Z.-L., Lawrence, D. S. & Zhang, Z.-Y. (2001). *J. Biol. Chem.* **276**, 47311–47319.

Stuckey, J. A., Schubert, H. L., Fauman, E. B., Zhang, Z. Y., Dixon, J. E. & Saper, M. A. (1994). *Nature (London)*, **370**, 571–575.

Sun, J. P., Fedorov, A. A., Lee, S. Y., Guo, X.-L., Shen, K., Lawrence, D. S., Almo, S. C. & Zhang, Z.-Y. (2003). *J. Biol. Chem.* **278**, 12406–12414.

Szedlacsek, S. E., Aricescu, A. R., Fulga, T. A., Renault, L. & Scheidig, A. J. (2001). *J. Mol. Biol.* **311**, 557–568.

Tobin, J. F. & Tam, S. (2002). *Curr. Opin. Drug Discov. Dev.* **5**, 500–512.

Yang, J., Liang, X., Niu, T., Meng, W., Zhao, Z. & Zhou, G. W. (1998). *J. Biol. Chem.* **273**, 28199–28207.

Zabolotny, J. M., Bence-Hanulec, K. K., Stricker-Krongrad, A., Haj, F., Wang, Y. P., Minokoshi, Y., Kim, Y. B., Elmquist, J. K., Tartaglia, L. A., Kahn, B. B. & Neel, B. G. (2002). *Dev. Cell*, **2**, 489–495.

Zhang, Z. Y. (1998). *Crit. Rev. Biochem. Mol. Biol.* **33**, 1–52.

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.