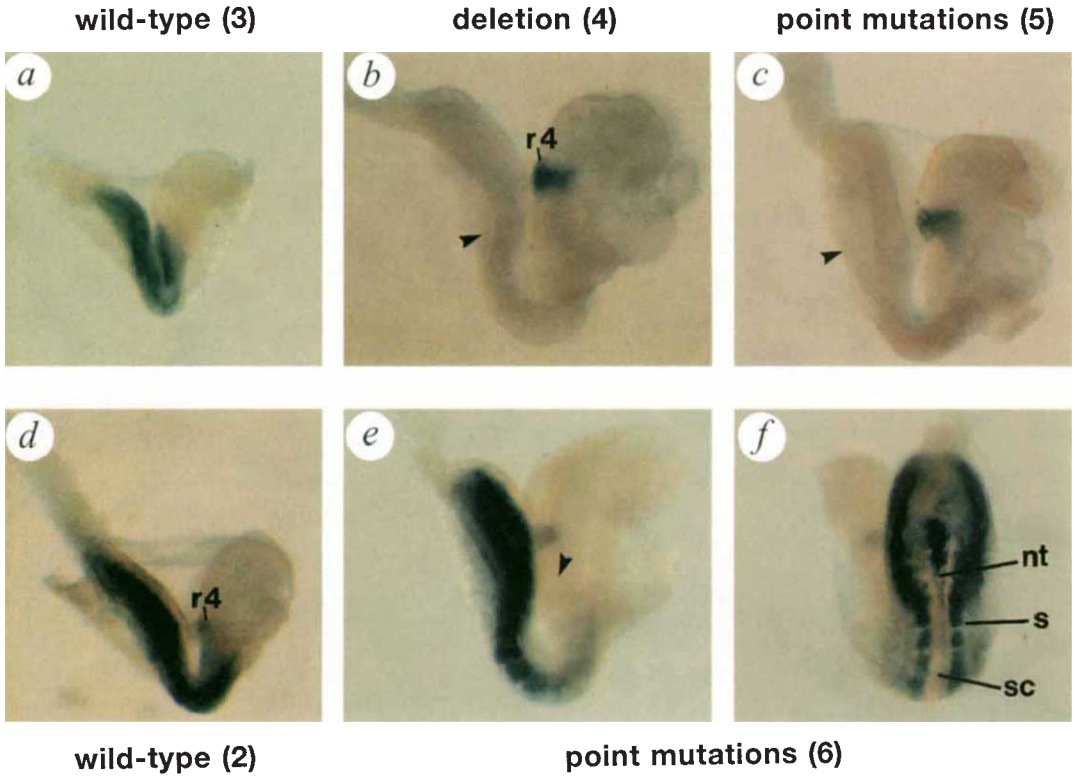


FIG. 4 The 3′ RARE of the *Hoxb-1* gene is necessary for establishing early neural expression. Numbers in parentheses indicate construct number. *a*, *d*, Control staining patterns of transgenic embryos with the wild-type *Eco*RV-*Hin*dIII fragment (*a*) and the *Eco*RV–*Sal* frament (*d*) linked to a construct directing expression in rhombomere 4 (constructs 3 and 2, Fig. 1*a*). Lateral views of 8.25-d.p.c. embryos containing a deletion (*b*) or point mutations (*c*) in the 3′ RARE in the *Eco*RV–*Hin*dIII enhancer (constructs 4 and 5). At all stages examined (7.25–9.0 d.p.c.) the only expression observed arises in rhombomere 4 after 8.25 d.p.c, indicating that both mutations completely eliminate the early expression in mesoderm and ectoderm. *e*, *f*, Lateral and posterior views, respectively, of an embryo carrying point mutations in the 3′ RARE in the *Eco*RV–*Sal* enhancer fragment (construct 6). Posterior expression is present in somitic mesoderm (s) and notochord (nt) but is absent in the spinal cord (sc) and neural tube. Arrowheads indicate domains of expression missing in RARE mutant constructs.

3′ region which is necessary for the response to retinoic acid in tissue culture cells[27]. By analogy to our results with *Hoxb-1*, this RARE could be important for establishing *Hoxa-1* expression, and could reflect a common role for retinoids in regulating vertebrate *labial*-related genes. □

Received 18 April; accepted 24 June 1994.

1. Wilkinson, D. *et al. Nature* **341,** 405–409 (1989).
2. Murphy, P., Davidson, D. & Hill, R. *Nature* **341,** 156–159 (1989).
3. Frohman, M., Boyle, M. & Martin, G. *Development* **110,** 589–607 (1990).
4. Sundin, O. & Eichele, G. *Genes Dev.* **4,** 1267–1276 (1990).
5. Maden, M. *et al. Development* **111,** 35–44 (1991).
6. Guthrie, S. *et al. Nature* **356,** 157–159 (1992).
7. Murphy, P. & Hill, R. *Development* **111,** 61–74 (1991).
8. Conlon, R. & Rossant, J. *Development* **116,** 357–368 (1992).
9. Marshall, H. *et al. Nature* **360,** 737–741 (1992).
10. Sundin, O. & Eichele, G. *Development* **114,** 841–852 (1992).
11. Kessel, M. *Neuron* **10,** 379–393 (1993).
12. Morriss-Kay, G. *et al. EMBO J.* **10,** 2985–2996 (1991).
13. Kessel, M. & Gruss, P. *Cell* **67,** 89–104 (1991).
14. Dolle, P. *et al. Nature* **342,** 767–772 (1989).
15. Izpisua-Belmonte, J.-C. *et al. Nature* **350,** 585–589 (1991).
16. Brenner, S. *et al. Nature* **366,** 265–268 (1993).
17. Leid, M., Kastner, P. & Chambon, P. *Trends biochem. Sci.* **17,** 427–433 (1992).
18. Stunnenberg, H. *Bioessays* **15,** 309–315 (1993).
19. Kliewer, S. A. *et al. Nature* **355,** 446–449 (1992).
20. Durston, A. *et al. Nature* **340,** 140–144 (1989).
21. Ruiz i Altaba, A. & Jessell, T. *Genes Dev.* **5,** 175–187 (1991).
22. Ruiz i Altaba, A. & Jessell, T. *Development* **112,** 945–958 (1991).
23. Papalopulu, N. *et al. Development* **113,** 1145–1159 (1991).
24. Sive, H. & Cheng, P. *Genes Dev.* **5,** 1321–1332 (1991).
25. Sive, H. *et al. Genes Dev.* **4,** 932–942 (1990).
26. Holder, N. & Hill, J. *Development* **113,** 1159–1170 (1991).
27. Langston, A. W. & Gudas, L. J. *Mechanisms of Development* **38,** 217–228 (1992).
28. Whiting, J. *et al. Genes Dev.* **5,** 2048–2059 (1991).
29. Yee, S.-P. & Rigby, P. W. R. *Genes Dev.* **7,** 1277–1289 (1993).
30. Popperl, H. & Featherstone, M. S. *Molec. cell. Biol.* **13,** 257–265 (1993).

ACKNOWLEDGEMENTS. We thank A. Gould for help in generating the point mutations; W. Hatton for help in sectioning embryos; Z. Webster and L. Jones for animal husbandry; P. Chambon and H. Gronemeyer for providing the RAR and RXR proteins for *in vitro* bindings assays; and members of the laboratories for comments and discussion. R.K., H.M. and A.K. were supported by an HFSPO collaborative grant, M.S. by an SNF and EMBO fellowship, H.P. by an EMBO fellowship, and S.A. by an MRC student traineeship.

Crystal structure of *Yersinia* protein tyrosine phosphatase at 2.5 Å and the complex with tungstate

Jeanne A. Stuckey*, Heidi L. Schubert*†‡, Eric B. Fauman*, Zhong-Yin Zhang†‡§, Jack E. Dixon†‡ & Mark A. Saper*†¶

* Biophysics Research Division, † Department of Biological Chemistry and ‡ Walther Cancer Institute, The University of Michigan, Ann Arbor, Michigan 48109-1055, USA

PROTEIN tyrosine phosphatases (PTPases) and kinases coregulate the critical levels of phosphorylation necessary for intracellular signalling, cell growth and differentiation[1,2]. *Yersinia*, the causative bacteria of the bubonic plague and other enteric diseases, secrete an active PTPase[3], Yop51, that enters and suppresses host immune cells[4,5]. Though the catalytic domain is only ~20% identical to human PTP1B[6], the *Yersinia* PTPase contains all of the invariant residues present in eukaryotic PTPases[7], including the nucleophilic Cys 403 which forms a phosphocysteine intermediate during catalysis[3,8–10]. We present here structures of the unliganded (2.5 Å resolution) and tungstate-bound (2.6 Å) crystal forms which reveal that Cys 403 is positioned at the centre of a distinctive phosphate-binding loop. This loop is at the hub of several hydrogen-bond arrays that not only stabilize a bound oxyanion, but may activate Cys 403 as a reactive thiolate. Binding of tungstate triggers a conformational change that traps the oxyanion and swings Asp 356, an important catalytic residue[7], by ~6 Å into the active

§ Present address: Department of Biochemistry, Albert Einstein College of Medicine, Bronx, New York 10461, USA.
¶ To whom correspondence should be addressed.

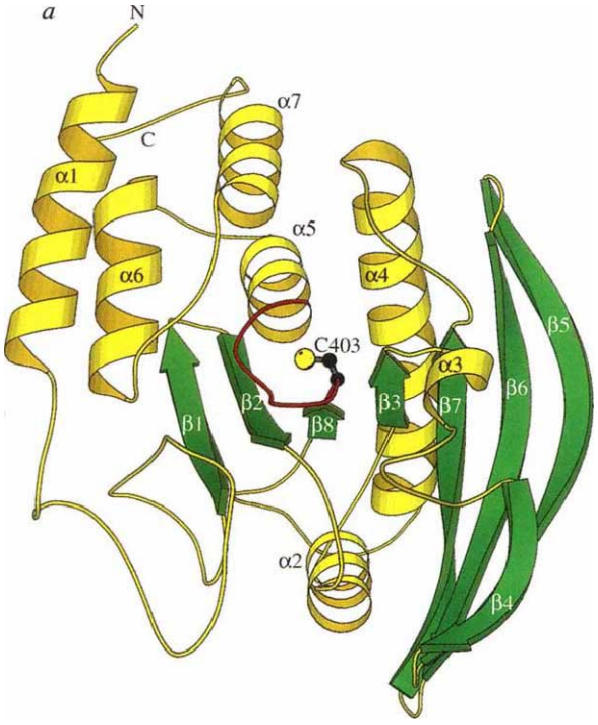
site. The same anion-binding loop in PTPases is also found in the enzyme rhodanese[11].

The central feature of the *Yersinia* PTPase tertiary fold, as in the structurally homologous PTP1B[12], is a highly twisted, eight-stranded, mixed β-sheet flanked by five α-helices on one side and two α-helices (α2 and α3) on the other (Table 1 and Fig. 1*a*). At the centre of the β-sheet are two split $\beta\alpha\beta$ motifs[13] that interdigitate to form four adjacent parallel strands (β2, β8, β3 and β7). A prominent depression forming the substrate-binding site is located at the carboxy end of these strands (Fig. 1*b*). Centred within the site, and located on the β-turn following β8 (residues 403–406) and the first turn of helix α5 (407–410), is the PTPase signature sequence (I/V)HCXAGXGR(S/T)(G/A)[14] (single-letter amino-acid code). Residues 403–410 form the PTPase phosphate-binding loop or P-loop with the invariant Cys 403 thiol[3,9] centred within the loop and close to the α5 helix axis (Figs 1 and 2*a*).

Although sulphydryls in proteins normally have a dissociation constant (pK_a) of ~8.5, Cys 403 in *Yersinia* PTPase has an apparent pK_a of 4.7 suggesting that it may be stabilized as a negatively charged thiolate at neutral pH[9]. Active-site thiolates in other hydrolases, such as papain[15] and dienelactone hydrolase[16], are stabilized by ion-pairing to nearby histidines. The only invariant and positively charged residue in the vicinity of Cys 403 is Arg 409, a residue that is important for catalysis[17] (Z.-Y.Z. *et al.*, manuscript in preparation). In the unliganded *Yersinia* PTPase structure, the Arg 409 side chain projects from the α5 amino terminus towards helix α3 and the active-site opening (Figs 1 and 2*a*). The position of its guanidinium group is anchored by hydrogen bonds to the Leu 285 carbonyl and Thr 358 Oγ1, and by a bidentate salt bridge to the carboxylate of Glu 290, the putative catalytic base[7] (Fig. 2*c*). Cys 403 is clearly in the $g^+(\chi_1 \approx +60°)$ rotamer (Fig. 2*a*), and Sγ is 5–6 Å from the Arg 409 guanidinium group. In the unliganded structure of PTP1B solved at 2.9 Å resolution[12], the equivalent Cys 215 Sγ is in a different rotamer conformation that places Sγ within 3.0 Å of the Arg 221 guanidinium group. This was proposed to partially stabilize the negatively charged thiolate[12]. On the basis of our electron density maps at 2.5 Å resolution (Fig. 2*a*), we conclude that Arg 409 does not play a major role in stabilizing the thiolate in *Yersinia* PTPase.

In *Yersinia*, the thiol anion of Cys 403 appears to be stabilized by an extensive network of hydrogen bonds[18] (Fig. 2*c*). The closest proton donor is the Oγ1 hydroxyl of conserved Thr 410 (Thr or Ser in all PTPases). The corresponding Ser 222 in the PTP1B structure[12] does not make this interaction, but hydrogen bonds to the amide of Gly 218. This is probably due to the different cysteine orientation in PTP1B mentioned above.

Cys 403 Sγ is also located between 3.0 and 4.2 Å from every amide nitrogen of the P-loop (residues 403–410). At least five of these are shorter than 4.0 Å and should make reasonable S–HN hydrogen bonds[19]. Alternatively, the P-loop amides may be considered individual microdipoles with their δ^+ ends[20] all oriented towards the thiolate. The P-loop peptide nitrogens are coupled electrostatically through their carbonyl oxygens[18] to several hydrogen-bonding arrays radiating out from the P-loop (Fig. 2*c*). The invariant (and mostly buried) residues His 402, Tyr 301, Arg 437, Arg 440 and Asn 245 participate in three of these arrays. The first array begins with the Cys 403 carbonyl, and continues through the His 402 imidazole ring, the Tyr 301 hydroxyl, and finally the His 270 side chain (Fig. 2*c*). Site-directed mutations of the invariant His 402 residue to Asn or Ala raise the apparent pK_a of Cys 403[9], probably by partially disrupting this hydrogen-bond network. In the second array, peptide bonds following residues 405 and 406 mediate ionic interactions between the thiolate and the buried guanidinium group of Arg 440. Mutation of Arg 440 to Ala or Lys severely impairs catalytic efficiency and binding (Z.-Y.Z. *et al.*, manuscript in preparation). The third array couples Val 407 O to the Arg 437 side chain, but is mediated through an additional



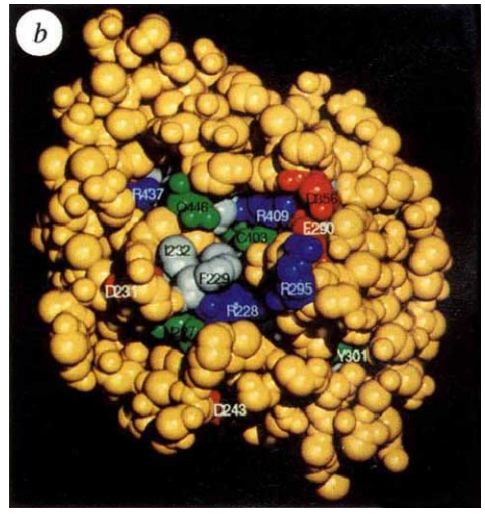


FIG. 1 *a*, Ribbon diagram of the unliganded *Yersinia* PTPase structure (drawn with MOLSCRIPT[30]). Molecule is oriented with the C-terminal end of the β8 strand towards the viewer (α-helices and turns in yellow, β-strands in green). The catalytic site is centred around the anion-binding loop (P-loop) formed between β8 and α5 (residues 403–410, red; Cys 403, yellow). Surrounding the binding site are residues contributed by turns α1–β1, α6–α7, β2–α2, α3–β4 and α4–β7, and helix α3. *b*, CPK model of unliganded PTPase in the same orientation as *a*, highlighting the conserved and invariant residues[7] that are solvent accessible. Polar uncharged residues, green; basic, blue; acidic red; and hydrophobic, grey. Positions of unlabelled conserved residues: Gly 408, grey, upper-left of Cys 403; Gln 450, green, behind Gly 408 and Arg 409; Val 433, Met 444, Val 445, grey, in pocket adjacent to Arg 437. Only the main-chain of Tyr 301 is accessible; the side chain is buried and interacts with His 402 and His 270 in the hydrogen-bonding network (Fig. 2*c*). As noted for the PTP1B structure[12], most of the residues clustered within or around the catalytic depression are invariant of structurally similar within the PTPase family[7]. This suggests that all PTPases are likely to recognize phosphotyrosine in a similar way and have similar catalytic mechanisms.

peptide bond (444–445). Because mutations of the invariant Arg 437 have little effect on catalysis or oxyanion binding (Z.-Y.Z. *et al.*, unpublished data), we suggest that this residue might interact with conserved functional groups (such as main-chain or anionic side-chain atoms) of a phosphorylated peptide substrate. The Arg 437 guanidinium group is located ~14 Å from the active site, but is partially exposed in a predominantly hydrophobic pocket between $\alpha 1$ and $\alpha 7$ (Fig. 1*b*). Finally, the amides of P-loop residues 408 and 409 on the first turn of the $\alpha 5$ helix are coupled by hydrogen bonds to peptide bonds on

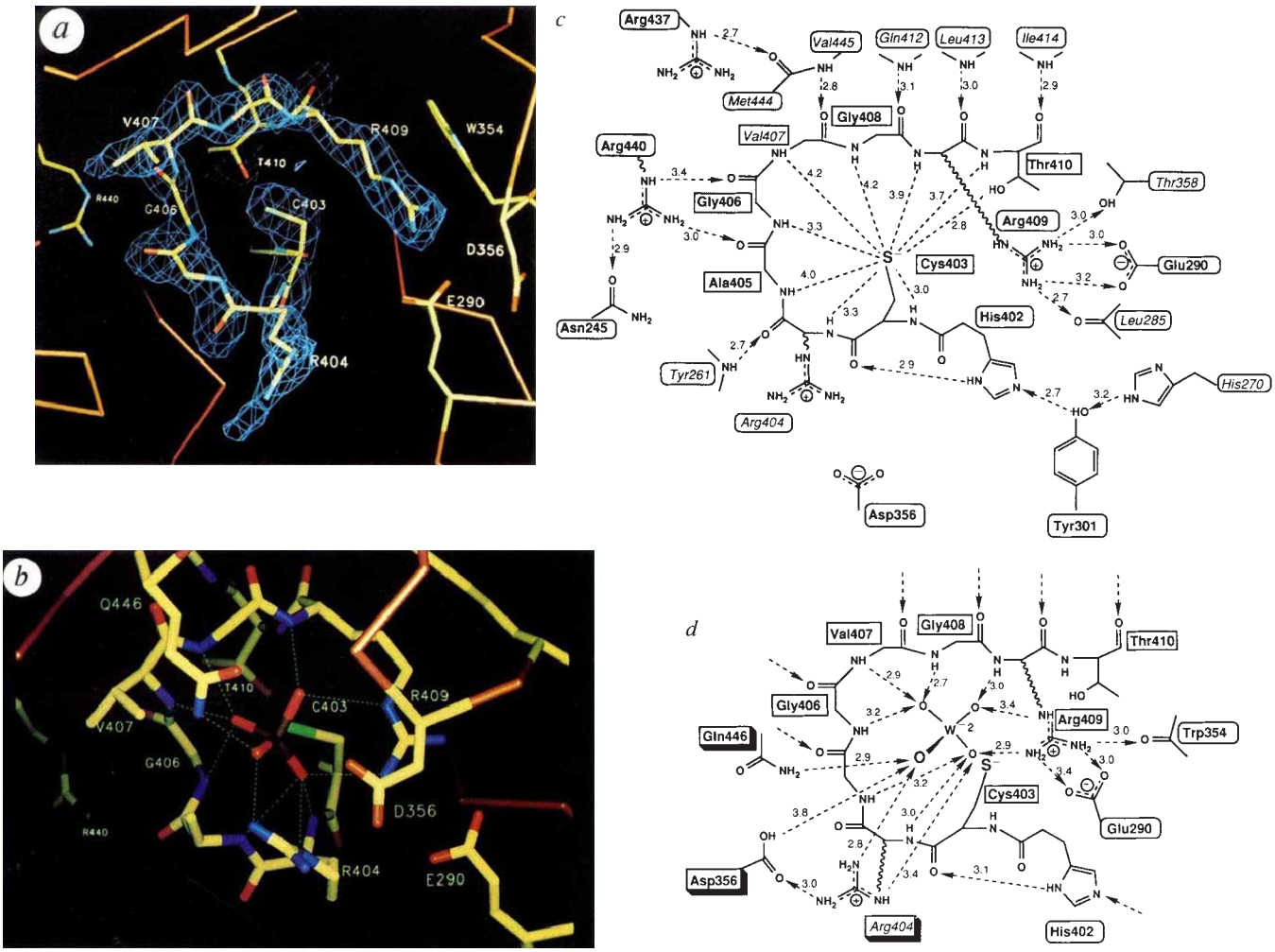


FIG. 2 The *Yersinia* PTPase P-loop (residues 403–410) and hydrogen-bonding arrays. *a* and *b*, Carbons, nitrogens, oxygens, sulphurs and the tungsten are shown in yellow, blue, red, green and violet, respectively. *a*, Unliganded *Yersinia* PTPase structure. Superimposed on the P-loop are electron density contours (2.5 σ) from an 'omit' map calculated with coefficients ($F_{obs} - F_{calc}$), α_{calc}. *b*, PTPase–tungstate structure. Hydrogen bonds (3.5 Å or less) between the tungstate oxygens and the enzyme are shown as dashed green lines. As described in the text, the $\beta 7$–$\alpha 4$ loop containing Asp 356 has moved to interact with the bound oxyanion. The corresponding $\beta 11$–$\alpha 3$ loop in both unliganded and tungstate-bound structures of PTP1B[12] appears to be in a position similar to the *Yersinia* PTPase unliganded state with the analogous Asp 181 quite distant from the bound tungstate. Crystal lattice restrictions in PTP1B may have inhibited a conformational change because the PTP1B–tungstate complex was prepared by soaking unliganded crystals[12], whereas the *Yersinia* PTPase–WO_4^{2-} crystals were co-crystallized with tungstate. Alternatively, sequence differences in PTP1B may make the loop movement energetically less favourable, and account for the enzyme's lower catalytic rate when compared to *Yersinia* PTPase[31]. *c*, *d*, Active-site hydrogen-bonding arrays in the *Yersinia* PTPase structures. Arrows point from hydrogen donor to acceptor. Residues involved in primary hydrogen bonds with the Cys 403 thiolate or bound tungstate are boxed. Conserved and invariant residues are shown in bold and non-conserved residues in italics. All distances in Å. *c*, Cys 403 is stabilized as an anion by several hydrogen-bonding arrays radiating from the P-loop (see text). Sulphur–nitrogen interactions under 4.0 Å are expected to form reasonable S–HN hydrogen bonds[19]. Cys 403 Sγ is ~10 Å from Asp 356, and 6 Å from Arg 409 Nε. *d*, PTPase–WO_4^{2-} structure showing hydrogen bonds to the tungstate oxygens. All secondary hydrogen-bonding arrays shown in *c* also exist in *d*, but have been omitted for clarity. The oxygen of WO_4^{2-} indicated with an enlarged O is expected to correspond to the oxygen of the scissile O–P bond in a phosphotyrosine substrate. Glu 290 Oε2 is 5.5 Å from the nearest tungstate oxygen; Cys 403 Sγ is 3.6 Å from the tungsten.

METHODS. Calculated phases for the difference map in *a* were obtained from a simulated-annealing X-PLOR refinement with residues 403–410 omitted from the structure factor calculations. *b* and *d*, Crystals of the PTPase–tungstate complex ($P2_12_12_1$, $a = 56.4$ Å, $b = 49.8$ Å, $c = 100.8$ Å; size $0.8 \times 0.2 \times 0.02$ mm^3) were grown by mixing equal volumes of a PTPase solution containing 1 mM tungstate, with precipitant (22% (w/v) polyethylene glycol 4000, 200 mM Li_2SO_4, 10% isopropanol, 0.1% β-mercaptoethanol, 100 mM Tris-HCl, pH 8.5). The crystal diffracted to at least 2.2 Å, and 81% of the reflections between 7.0 Å and 2.6 Å were recorded. These amplitudes (and those from a Hg derivative) were used to phase data from isomorphous crystals of the *Yersinia* PTPase Cys 403→Ser mutant containing bound sulphate (H.L.S. *et al.*, manuscript in preparation). After most of the Cys 403→Ser structure had been built, a tetrahedral tungstate ion (net −2 charge, W–O bond length = 1.77 Å) was fit into a difference map, and the entire model further refined with X-PLOR[29] against the PTPase–tungstate data. Force-field interactions between the tungstate and Cys 403 were turned off to allow close-approach of these two groups if necessary. The PTPase–WO_4^{2-} structure currently has a crystallographic *R*-factor of 19.5% to 2.6 Å ($F/\sigma_F \geq 2.0$). The Cα r.m.s. deviation from the unliganded structure is 0.9 Å. A detailed description of the structure refined to 2.5 Å will be published elsewhere (E.B.F. *et al.*, manuscript in preparation).

TABLE 1 Diffraction data, phasing, density averaging and refinement statistics for unliganded *Yersinia* PTPase

Compound*	Resolution (Å)	No of unique reflections	Completeness (%)	R_{merge}† (%)	R_{iso}† (%)	No. of sites	Phasing power§ (max. resolution used in phasing, Å)
Native	2.5	19,476	91	5.9			NA
$K_2Pt(CN)_4$ ‖	2.3	19,501	72	9.6	15.7	1	1.42 (2.8)
$K_2Pt(CN)_4$ ‖	2.3	12,793	46	10.0	13.9	1	1.52 (2.8)
$PtCl_4$	2.8	11,106	75	8.0	16.2	1	1.26 (3.5)
K_2PtCl_4	3.0	9,133	77	10.6	20.2	2	1.38 (3.5)

Overall mean figure of merit (20–2.8 Å) = 0.51

The catalytic domain (residues 163–468) of Yop51, a PTPase from *Yersinia enterocolitica*, was expressed, purified, and crystallized as described previously[24]. Trigonal crystals (space group $P3_2$; $a = b = 71.53$ Å, $c = 107.48$ Å, $\alpha = \beta = 90°$, $\gamma = 120°$; two molecules per asymmetric unit) grew only in the absence of oxyanion-containing salts. All diffraction data were collected on a SDMS multiwire area detector system mounted on a Rigaku RU-200 rotating anode generator (50 kV, 100 mA), and processed by the accompanying software[25]. Single isomorphous replacement phases, derived from the first $K_2Pt(CN)_4$ derivative, were improved by solvent flattening using PHASES[26]. Sites for the other heavy-atom derivatives were located in difference Fourier maps. Multiple isomorphous replacement (MIR) phases (20–3.5 Å resolution; mean, figure of merit, 0.56) were improved by iterative solvent flattening, and an initial backbone trace of one molecule was fitted to the resulting density map with the program O[27]. The location of the non-crystallographic 2-fold axis was found by a rotation/translation search of the MIR map (PROTEIN[28]), and refined by PHASES (density correlation value, 0.32). Over 60% of the model (mainly polyalanine) was fit to a 20–3.0 Å MIR map (mean figure of merit = 0.51) that had been averaged without a mask. Masks derived from this model were used to improve the MIR map by iterative density averaging and solvent flattening[26]. PTPase models were refined with the 'slow-cool', simulated annealing (3,000 K) protocol of X-PLOR[29] (including non-crystallographic restraints), and refit to 3.0 Å electron density maps calculated with MIR-combined phase probabilities, and later at 2.8 Å and 2.5 Å with $(2F_{obs} - F_{calc})$, α_{calc} maps. The *R*-factor is currently 18.4% (free *R*-factor[29] = 27.0%) for 17,095 unique reflections in the 7–2.5 Å shell ($F/\sigma_F \geq 2$, 84% complete). The model includes residues 189–468 and 191–468 for molecules A and B, respectively, a total of 4,290 non-hydrogen atoms. There was no interpretable electron density for the N-terminal residues 163–188 in either molecule and they are assumed to be disordered. The r.m.s. deviations for bond distances and angles were 0.015 Å and 3.3°, respectively. Refinement to higher resolution, relaxation of the non-crystallographic restraints, and inclusion of solvent molecules is in progress. Coordinates will be submitted to the Brookhaven Protein Data Bank.

* Derivative crystals were soaked in solutions containing 0.05–0.1 mM heavy-atom reagent in 22% polyethylene glycol 1500, 10% 2-methyl-2,4-pentanediol, 1 mM imidazole, pH 7.2, for up to 20 h before data collection.

† $R_{merge} = \Sigma_{hkl} \Sigma_i |I_i - \langle I \rangle| / \Sigma \langle I \rangle$ where I_i is the intensity measurement of a particular symmetry-related reflection.

‡ $R_{iso} = \Sigma(||F_{PH}| - |F_P||) / \Sigma |F_P|$, where F_P and F_{PH} are the structure factor amplitudes of the native protein and protein complexed with heavy atoms, respectively.

§ Phasing power is defined as r.m.s. F_h/(r.m.s. lack of closure error), where F_h is the heavy-atom structure factor.

‖ Data from two $K_2Pt(CN)_4$ soaked crystals were treated as separate heavy-atom derivatives for multiple isomorphous replacement phasing.

the second turn of the helix. Though previously attributed to a helix macrodipole effect[21] in the PTP1B structure[12], the polarized microdipoles of peptide bonds on the first two turns of an α-helix (such as in the P-loop) have been shown to confer the majority of the charge stabilization energy[20].

To examine how phosphoryl groups of substrates would bind to PTPases, we transferred the unliganded PTPase crystals to solutions containing 0.01–1.0 mM sodium tungstate, a competitive inhibitor (dissociation constant $K_D = 61$ μM at pH 7; Z.-Y. Z. and J. E. D., unpublished data). The crystals cracked within hours. By adding 1 mM tungstate to the enzyme before crystallization, we obtained a different crystal form containing the PTPase–WO_4^{2-} complex and have independently solved its structure to 2.6 Å resolution (Fig. 2*b*; E.B.F. *et al.*, manuscript in preparation). The P-loop orients the bound tungstate directly over the nucleophilic Cys 403 Sγ (Sγ–W distance = 3.6 Å) by anchoring three of the anion's oxygens to the main-chain amides of residues 404–409 through hydrogen bonds (Fig. 2*d*; mean distance, 3.0 Å). The −2 charge on the tungstate, or a −3 charge of the pentacoordinated transition state expected during

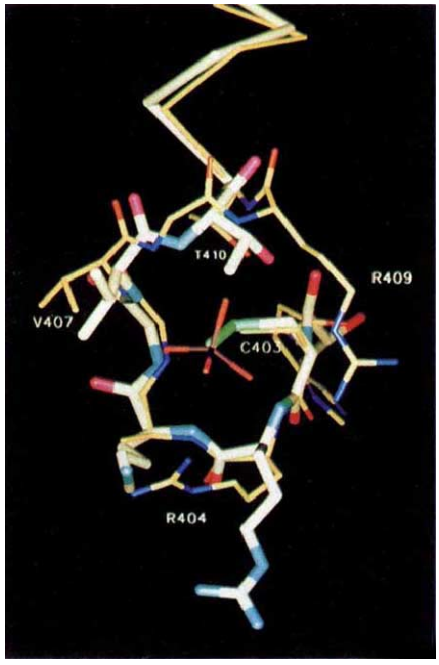


FIG. 3 Superposition of the *Yersinia* PTPase P-loop structure onto the anion-binding loop of rhodanese. The PTPase–tungstate complex is in yellow with thin bonds, and rhodanese is white with thick bonds. Residue numbers are for the PTPase structure. The corresponding sequence alignment for the anion-binding loops is:

PTPase–WO_4^{2-} 403 **C R** A **G V** G R T A 411

Rhodanese 247 **C R** K **G V** T . A C 254

Arg 409, catalytically important in all PTPases, is an insertion relative to rhodanese, which in effect widens the first turn of α5 in the PTPase. In both structures, the main-chain NH groups from 403–409 are all oriented towards the centre of the anion-binding loop. Though functionally unrelated to PTPases, the catalytic mechanism is similar. Rhodanese transfers a sulphur from thiosulphate ($S_2O_3^{2-}$) to a nucleophilic thiolate to form a charged, persulphide (Cys 247–Sγ–Sδ^-) enzyme intermediate as depicted here. When superimposed on the PTPase P-loop, the covalently-bound sulphur anion (Sδ^-) is just 1.6 Å from the bound tungsten. In rhodanese, cyanide (or possibly a non-haem iron protein) binds and accepts S to form thiocyanate (SCN^-)[11]. In PTPases, a hydroxyl ion is the acceptor forming HPO_4^{2-}.

METHODS. Rhodanese coordinates[11] (PDB code 1RHD) were superimposed on the PTPase–WO_4^{2-} structure by minimizing the Cα r.m.s. deviation between residues 398–407, 410–421 and 254–260 in PTPase–WO_4^{2-}, and 242–251, 253–264 and 266–272 in rhodanese. The r.m.s. deviation for these 29 pairs of α-carbons was 1.4 Å.

phosphate hydrolysis, will be delocalized by the same hydrogen-bond arrays discussed above in the unliganded structure. The invariant Arg 409 guanidinium group moves ~2 Å to form a bifurcated salt bridge with the tungstate oxygens while maintaining its salt bridge with Glu 290 (Fig. 2*d*). Because mutation of Arg 409 to Lys severely impairs catalysis, we suspect that this bifurcated interaction is essential to bind preferentially the transition state that occurs during phosphate hydrolysis (Z.-Y.Z. *et al.*, manuscript in preparation).

The major difference between the PTPase–WO_4^{2-} complex and the unliganded structure is that the $\beta7$–$\alpha4$ loop (residues 350–360) has moved an average of 3.3 Å and folds over the active site to sequester the bound oxyanion (Fig. 2*a*, *b*; H.L.S. *et al.*, manuscript in preparation). In the unliganded PTPase structure, the putative catalytic general acid[7] Asp 356 is ~10 Å from the phosphate-binding site, and interacts with Ser 287 and Ser 289. In the tungstate complex, Asp 356 Cα moves towards the active site by 6 Å, hydrogen-bonds with Arg 404, and positions its carboxylate 3.5–3.8 Å away from the tungstate oxygen that is not ligated to the P-loop (large O of tungstate in Fig. 2*d*). Because this oxygen is directly opposite the Cys 403 nucleophile it probably corresponds to the atom of the scissile ester bond in a phosphotyrosine residue. Assuming this conformational change also occurs on binding of a phosphotyrosyl substrate, Asp 356 would be ideally positioned to aid proton transfer during one or both of the proposed hydrolysis steps[14]. The 2.5 Å structure of the inactive Cys 403→Ser PTPase mutant has a sulphate anion bound to the P-loop and displays the identical conformational change (H.L.S. *et al.*, manuscript in preparation). The reported structures of tungstate-free and tungstate-bound forms of the PTP1B structure[12] do not exhibit the above conformational change (discussed in Fig. 2).

Unexpectedly, we discovered a striking similarity between the active-site structures of *Yersinia* PTPase and that of rhodanese, a sulphur transferase found in a variety of organisms[11,22,23]. Not only are both anion-binding loop conformations very similar (Fig. 3) but they also have an identically positioned cysteine thiolate which is stabilized by hydrogen bonds to amides of the P-loop, and by the amino-end of an α-helix[11]. Each enzyme also forms a charged covalent intermediate during catalysis, Cys 403–Sγ–PO_3^{2-} in PTPase and Cys 247–Sγ–Sδ^- in rhodanese, with the anion positioned at the centre of the P-loop. Additionally, rhodanese Arg 110 (conserved in all rhodanese sequences) is buried and forms hydrogen bonds to the same P-loop amides as the buried Arg 440 in PTPases. We might expect other unrelated enzyme structures to also use the P-loop motif for anion binding and transfer. □

Received 5 April; accepted 28 June 1994.

1. Fischer, E. H., Charbonneau, H. & Tonks, N. K. *Science* **253,** 401–406 (1991).
2. Walton, K. M. & Dixon, J. E. *A. Rev. Biochem.* **62,** 101–120 (1993).
3. Guan, K. L. & Dixon, J. E. *Science* **249,** 553 –556 (1990).
4. Bolin, I. & Wolf-Watz, H. *Molec. Microbiol.* **2,** 237–245 (1988).
5. Bliska, J. B., Guan, K. L., Dixon, J. E. & Falkow, S. *Proc. natn. Acad. Sci. U.S.A.* **88,** 1187–1191 (1991).
6. Tonks, N. K., Diltz, C. D. & Fischer, E. H. *J. biol. Chem.* **263,** 6722–6730 (1988).
7. Zhang, Z. Y., Wang, Y. & Dixon, J. E. *Proc. natn. Acad. Sci. U.S.A.* **91,** 1624–1627 (1994).
8. Guan, K. L. & Dixon, J. E. *J. biol. Chem.* **266,** 17026–17030 (1991).
9. Zhang, Z. Y. & Dixon, J. E. *Biochemistry* **32,** 9340–9345 (1993).
10. Cho, H. *et al. J. Am. chem. Soc.* **114,** 7296–7298 (1992).
11. Ploegman, J. H. *et al. Nature* **273,** 124–129 (1978).
12. Barford, D., Flint, A. J. & Tonks, N. K. *Science* **263,** 1397–1404 (1994).
13. Orengo, C. A. & Thornton, J. M. *Structure* **1,** 105–120 (1993).
14. Zhang, Z. Y. & Dixon, J. E. *Adv. Enzym. Relat. Areas molec. Biol.* **68,** 1–36 (1994).
15. Lewis, S. D., Johnson, F. A. & Shafer, J. A. *Biochemistry* **20,** 48–51 (1981).
16. Pathak, D. & Ollis, D. *J. molec. Biol.* **214,** 497–525 (1990).
17. Streuli, M., Krueger, N. X., Thai, T., Tang, M. & Saito, H. *EMBO J.* **9,** 2399–2407 (1990).
18. Quiocho, F. A., Sack, J. S. & Vyas, N. K. *Nature* **329,** 561–564 (1987).
19. Gregoret, L. M., Rader, S. D., Fletterick, R. J. & Cohen, F. E. *Proteins* **9**, 99–107 (1991).
20. Åqvist, J., Luecke, H., Quiocho, F. A. & Warshel, A. *Proc. natn. Acad. Sci. U.S.A.* **88,** 2026–2030 (1991).
21. Hol, W. G., van Duijnen, P. T. & Berendsen, H. J. *Nature* **273,** 443–446 (1978).
22. Ploegman, J. H., Drent, G., Kalk, K. H. & Hol, W. G. *J. molec. Biol.* **127,** 149–162 (1979).
23. Wierenga, R. K. *Trans. Am. Crystallogr. Ass.* **22,** 49–62 (1986).
24. Zhang, Z. Y. *et al. J. biol. Chem.* **267,** 23759–23766 (1992).
25. Howard, A. J., Nielsen, C. & Xuong, N. H. *Meth. Enzym.* **114,** 452–472 (1985).
26. Furey, W. & Swaminathan, S. *Am. Crystallogr. Ass. Meeting Abstr.* **18,** 73 (1990).
27. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. *Acta crystallogr.* **A47,** 110–119 (1991).
28. Steigemann, W. thesis, Technische Universität, Munich (1974).
29. Brünger, T. A. *XPLOR Version 3.1 Manual* (Yale Univ. Press, New Haven, 1993).
30. Kraulis, P. J. *J. appl. Crystallogr.* **24,** 946–950 (1991).
31. Zhang, Z. Y., Maclean, D., McNamara, D. J., Sawyer, T. K. & Dixon, J. E. *Biochemistry* **33,** 2285–2290 (1994).

ACKNOWLEDGEMENTS. E.B.F. and H.L.S. contributed equally to these results. We thank members of M. Ludwig's and J. Dixon's laboratories, and K. Shaw for their expertise; D. Hume, J. Clement and J. Clemens for their contributions; F. E. Cohen and J. F. Bazan for discussions; and W. Furey, J. Pflugrath, W. Steigemann, T. A. Jones and P. Kraulis for providing software. Financial support provided by the American Cancer Society (J.A.S.), U-M Multipurpose Arthritis Center (M.A.S.), U-M Program in Protein Structure and Design (M.A.S.), Parke-Davis/Warner Lambert (M.A.S.), the NIH (J.E.D., M.A.S.) and Walther Cancer Institute (J.E.D.). H.L.S. is a Joe Dawson Fellow. M.A.S. is a Pew Scholar in the Biomedical Sciences.

The crystal structure of a low-molecular-weight phosphotyrosine protein phosphatase

Xiao-Dong Su*, Niccolo' Taddei†, Massimo Stefani†, Giampietro Ramponi† & Pär Nordlund*

* Department of Molecular Biology, University of Stockholm, Stockholm, Sweden
† Department of Biochemical Sciences, University of Florence, Florence, Italy

PROTEIN tyrosine phosphorylation and dephosphorylation are central reactions for control of cellular division, differentiation and development[1]. Here we describe the crystal structure of a low-molecular-weight phosphotyrosine protein phosphatase (PTPase)[2], a cytosolic phosphatase present in many mammalian cells. The enzyme catalyses the dephosphorylation of phosphotyrosine-containing substrates[3–6], and overexpression of the protein in normal and transformed cells inhibits cell proliferation[7,8]. The structure of the low-molecular-weight PTPase reveals an α/β protein containing a phosphate-binding loop motif at the amino end of helix α1. This motif includes the essential active-site residues Cys 12 and Arg 18 and bears striking similarities to the active-site motif recently described in the structure of human PTP1B[9]. The structure of the low-molecular-weight PTPase supports a reaction mechanism involving the conserved Cys 12 as an attacking nucleophile in an in-line associative mechanism. The structure also suggests a catalytic role for Asp 129 in the reaction cycle.

The bovine liver low-molecular-weight PTPase contains 157 amino-acid residues (M_r 17,953)[2]; it has been crystallized in 0.1 M acetate buffer, pH 5.5, in the presence of 40% saturated ammonium sulphate[10]. The structure of the enzyme was solved using the multiple isomorphous replacement method combined with density modification techniques (Table 1). A model containing residues 4 to 157, one sulphate ion and 83 water molecules has been built and subsequently refined at 2.1 Å resolution. The present model has a crystallographic *R*-factor of 15.6% and shows good stereochemistry (Table 1).

The structure of the protein is composed of a single α/β domain containing a central four-stranded, parallel open-twisted β-sheet surrounded by connecting α-helices on both sides (Fig. 1*a*). The secondary structure topology is shown in Fig. 1*b*. Residues 13–17 form a smooth loop connecting strand $\beta1$ and helix $\alpha1$. A strong density at the amino end of helix $\alpha1$ has been interpreted as a sulphate ion (Fig. 2).

In the refined structure, this sulphate ion forms an extensive hydrogen-bonding network to the amide peptide groups of the