






ARTICLE

DOI: 10.1038/s41467-017-01226-8

OPEN

# The molecular basis of phosphite and hypophosphite recognition by ABC-transporters

Claudine Bisson <sup>1,3</sup>, Nathan B.P. Adams <sup>1</sup>, Ben Stevenson<sup>1</sup>, Amanda A. Brindley<sup>1</sup>, Despo Polyviou<sup>2</sup>, Thomas S. Bibby<sup>2</sup>, Patrick J. Baker <sup>1</sup>, C. Neil Hunter <sup>1</sup> & Andrew Hitchcock <sup>1</sup>

Inorganic phosphate is the major bioavailable form of the essential nutrient phosphorus. However, the concentration of phosphate in most natural habitats is low enough to limit microbial growth. Under phosphate-depleted conditions some bacteria utilise phosphite and hypophosphite as alternative sources of phosphorus, but the molecular basis of reduced phosphorus acquisition from the environment is not fully understood. Here, we present crystal structures and ligand binding affinities of periplasmic binding proteins from bacterial phosphite and hypophosphite ATP-binding cassette transporters. We reveal that phosphite and hypophosphite specificity results from a combination of steric selection and the presence of a P-H... $\pi$  interaction between the ligand and a conserved aromatic residue in the ligand-binding pocket. The characterisation of high affinity and specific transporters has implications for the marine phosphorus redox cycle, and might aid the use of phosphite as an alternative phosphorus source in biotechnological, industrial and agricultural applications.

<sup>1</sup>Department of Molecular Biology and Biotechnology, University of Sheffield, Firth Court, Western Bank, Sheffield S10 2TN, UK. <sup>2</sup>Ocean and Earth Sciences, National Oceanography Centre Southampton, University of Southampton, Waterfront Campus, European Way, Southampton SO14 3ZH, UK. <sup>3</sup>Present address: Department of Biological Sciences, Crystallography ISMB, Birkbeck, University of London, Malet Street, London WC1E 7HX, UK. Claudine Bisson and Nathan B.P. Adams contributed equally to this work. Correspondence and requests for materials should be addressed to A.H. (email: [a.hitchcock@sheffield.ac.uk](mailto:a.hitchcock@sheffield.ac.uk))

Phosphorus (P), a component of nucleic acids, nucleotide cofactors and phospholipids, is essential to cell structure, metabolism and signalling. In nature it is mainly present in its most oxidised state (P valence = +5) predominantly as inorganic phosphate, which is the only form that cells can utilise directly for incorporation into biological molecules<sup>1–3</sup>. However, phosphate availability can vary significantly in natural environments and despite the presence of high-affinity phosphate uptake systems, in many ecosystems is sufficiently depleted to limit microbial growth. For example, phosphate concentrations in the oligotrophic surface waters of the vast ocean gyres can be extremely low, influencing the biogeochemistry and productivity of marine microorganisms<sup>4, 5</sup>. When phosphate is unavailable, some bacteria have evolved mechanisms to metabolise reduced phosphorus compounds to acquire P<sup>2, 3, 6–10</sup>. ATP-binding cassette (ABC) transporters for the active uptake of these compounds are encoded in the genomes of a range of microorganisms, with the transporter genes usually arranged in an operon with genes encoding enzymes that catalyse oxidation of the imported species to phosphate<sup>11, 12</sup> (Supplementary Fig. 1). Phosphonates (P valence = +3) are a diverse range of organic C-P bond containing compounds ubiquitous in natural habitats<sup>9</sup>. There are multiple mechanisms for their breakdown, the most common of which is the carbon-phosphorus (C-P) lyase, which converts phosphonates to a hydrocarbon and inorganic phosphate<sup>6, 8, 13</sup>. The crystal structure and ligand binding properties of the periplasmic binding protein (PBP) subunit from an *Escherichia coli* C-P lyase linked organophosphonate transporter have been reported previously<sup>14, 15</sup>; here we focus on the uptake of the alternative inorganic reduced phosphorus compounds phosphite ( $\text{HPO}_3\text{H}^-/\text{HPO}_3^{2-}$ ; P valence = +3) and hypophosphite ( $\text{H}_2\text{PO}_2^-$ ; P valence = +1).

The concentration of phosphite and hypophosphite in many natural environments is unknown, but they have been detected in up to micromolar amounts in some aquatic systems, where they can account for up to a third of the total dissolved P<sup>16, 17</sup>. The major sources of phosphite and hypophosphite are likely anthropogenic in origin (e.g. fungicides in agriculture), but they may also be derived from geothermal and biological processes<sup>16, 17</sup>. Many bacterial species are capable of utilising phosphite as a P source by oxidising it to phosphate (assimilatory phosphite oxidation)<sup>12, 17–20</sup>. The deltaproteobacteria *Desulfotomum phosphitoxidans* can also use phosphite as an electron donor and energy source by coupling its oxidation to the reduction of sulphate, CO<sub>2</sub> or nitrate (dissimilatory phosphite oxidation)<sup>21, 22</sup>. Phosphite is highly water-soluble and kinetically stable, and its oxidation by the NAD:phosphite oxidoreductase/phosphite dehydrogenase (PtxD) is highly exergonic (the reduction potential of phosphate/phosphite couple at pH 7 = –650 mV resulting in a  $\Delta G = 63.7 \text{ kJ mol}^{-1}$ ) and regenerates reductant in the form of NAD(P)H<sup>2, 23</sup>. Phosphite is also non-toxic and

inexpensive and, when provided as the sole source of P, it allows PtxD to be used as a selective marker, which has resulted in its use in numerous biotechnological and agricultural applications<sup>24–29</sup>. Hypophosphite can also be used as a sole P source by some microorganisms<sup>18, 19, 30, 31</sup>. The hypophosphite/2-oxoglutarate dioxygenase (HtxA) oxidises hypophosphite to phosphite (reduction potential at pH 7 = –740 mV) coupled to the oxidative decarboxylation of 2-oxoglutarate to succinate<sup>32</sup>; the phosphite produced is subsequently further oxidised to phosphate by PtxD.

The energetically favourable oxidation of phosphite and hypophosphite, yielding phosphate and NADPH, makes them attractive P sources to microorganisms under phosphate-limited conditions; however, the potentially low environmental concentration of these compounds necessitates the need for high-affinity transporters. While the genetic and biochemical basis of phosphite and hypophosphite utilisation and the PtxD and HtxA enzymes have been characterised previously<sup>18, 23, 32, 33</sup>, the mechanism of high-affinity ligand binding by the cognate ABC transporters is not well studied. The use of non-physiologically high concentrations (up to 1 mM) of P ligands in previous in vivo studies has hindered clear interpretation of the physiologically relevant specificity of reduced P compound transporters<sup>18, 19, 30, 34</sup>. Here we elucidate the molecular basis of high-affinity phosphite and hypophosphite acquisition by determining the ligand-binding affinity and structure of the PBPs of phosphite and hypophosphite ABC transporters from environmental microorganisms. Our results reveal how a ligand-protein P-H... $\pi$  interaction coupled with steric selection forms the basis of ligand specificity, explaining how these transporters allow bacteria to scavenge these alternative sources of P with ultra-high affinity.

## Results

**Ligand specificity of recombinant binding proteins.** We over-produced and purified PBPs of three putative phosphite transporters; PtxB from the globally important marine diazotroph *Trichodesmium erythraeum* IMS101 (Te\_PtxB)<sup>20</sup>, PhnD from the oceanic picocyanobacteria *Prochlorococcus marinus* MIT9301 (Pm\_Phnd)<sup>35</sup> and PtxB from the soil bacterium *Pseudomonas stutzeri* WM88 (Ps\_PtxB)<sup>18</sup>, as well as a PBP from a putative hypophosphite transporter, HtxB, also from *P. stutzeri* WM88 (HtxB)<sup>18</sup> (Supplementary Fig. 2). The binding affinity of each protein for a range of P ligands was measured by microscale thermophoresis (MST)<sup>36</sup>, using a RED-tris-NTA label (Nanotemper Technologies GmbH) via the C-terminal His-tag (Table 1; Supplementary Fig. 3). Dyes conjugated to the tris-NTA moiety provide a stable and very selective method to non-covalently fluorescently label proteins<sup>37</sup>. Dissociation constants ( $K_d$ ) are derived from three independent MST experiments and errors

**Table 1** Microscale thermophoresis determined dissociation constants ( $K_d$ ) of Te\_PtxB, Ps\_PtxB, Pm\_Phnd and HtxB for selected P-ligands

Ligand	$K_d$ ( $\mu\text{M}$ )			
	Te_PtxB	Ps_PtxB	Pm_Phnd	HtxB
Phosphate	46.80 $\pm$ 3.45	159.5 $\pm$ 36.0	182.2 $\pm$ 26.3	NBD
Phosphite	0.17 $\pm$ 0.07	0.24 $\pm$ 0.03	0.051 $\pm$ 0.004	>10 $\times$ 10 <sup>3</sup>
Methylphosphonate	30.63 $\pm$ 2.48	40.64 $\pm$ 5.70	108.8 $\pm$ 8.3	NBD
Ethylphosphonate	>60 $\times$ 10 <sup>3</sup>	—	111.9 $\pm$ 8.7	—
2-Aminoethylphosphonate	NBD	NBD	NBD	—
Hypophosphite	2.21 $\pm$ 0.16	—	—	0.56 $\pm$ 0.11

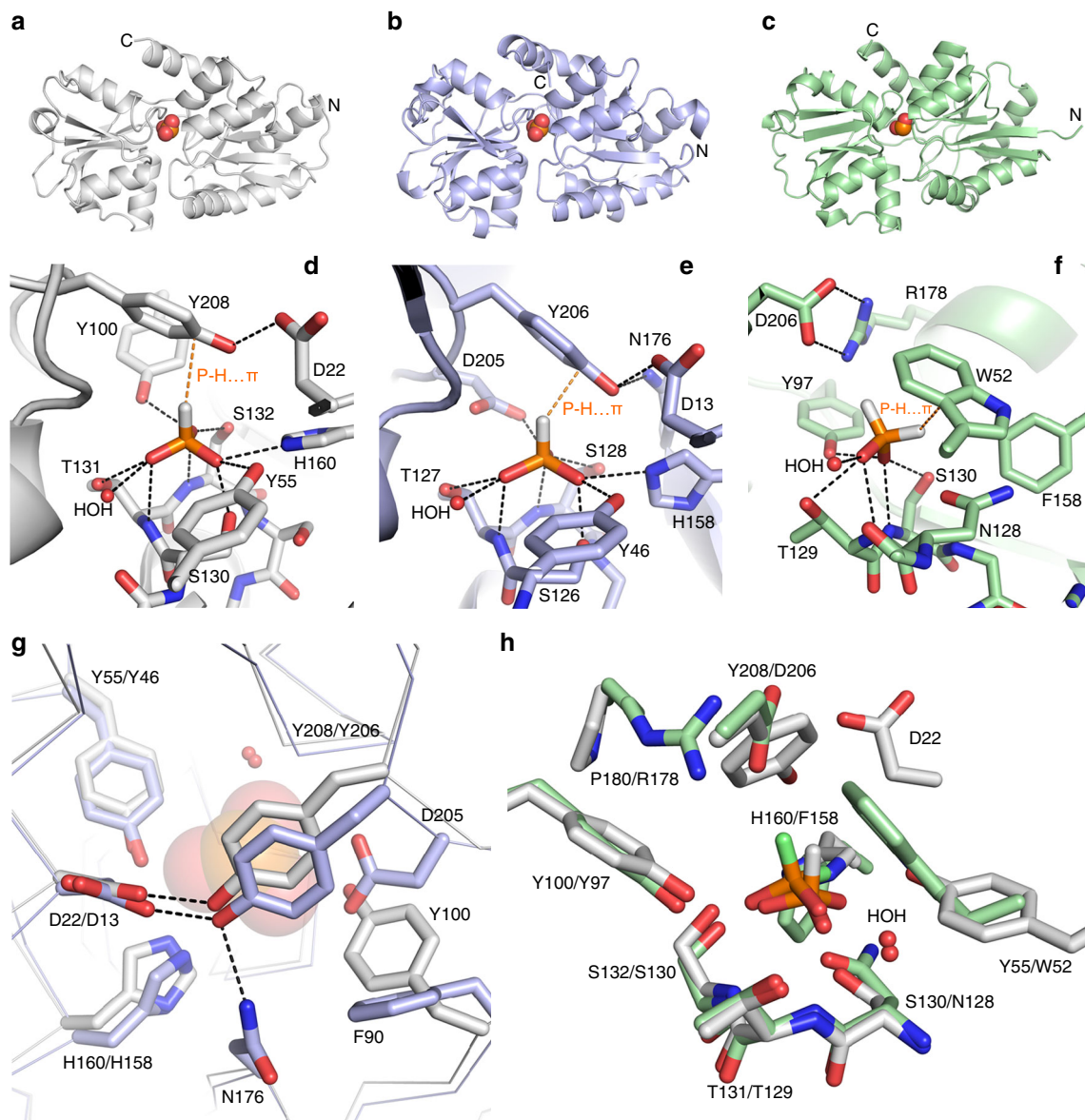
$K_d$  values presented are derived from fitting three independent experiments to Eq. (1) and errors reported as  $\pm$  the estimated standard deviation of the  $K_d$ . NBD indicates no binding was detected up to a ligand concentration of 10 mM; — indicates  $K_d$  not measured

**Table 2** Data collection and refinement statistics part 1

	Te_PtxB/phosphite (PDB: 5JVB)	Te_PtxB/MPn (PDB:5LQ1)	Pm_Phnd/phosphite (PDB:5LQ5)	Pm_Phnd/MPn (PDB:5LQ8)	Ps_PtxB apo (PDB: 5O2K)
<i>Data collection</i>					
Wavelength (Å)	0.96861	0.97629	0.97951	0.97951	0.97951
Beamline	i24	i03	i24	i24	i04
Resolution (Å)	47.96-1.95 (2.00-1.95)	65.05-1.41 (1.43-1.41)	44.27-1.46 (1.49-1.46)	44.0-1.52 (1.55-1.52)	79.75-2.1 (2.14-2.1)
Space group	P2 <sub>1</sub>	P1	P2 <sub>1</sub>	P2 <sub>1</sub>	P2 <sub>1</sub>
Unit cell (a, b, c, α, β, γ)	54.34, 69.31, 66.53, 90, 93.07, 90	39.48, 53.70, 66.39, 84.82, 93.07, 69.74	46.18, 57.41, 54.57, 90, 107.1, 90	45.86, 57.08, 54.17, 90, 106.56, 90	87.947, 136.640, 90.844, 90, 115.022, 90
Total reflections <sup>a</sup>	125,055 (9783)	302,564 (10,007)	162,732 (4664)	88,958 (3373)	762,654 (37,446)
Unique reflections <sup>a</sup>	35,004 (2628)	91,920 (3601)	46,137 (1859)	40,089 (1854)	113,030 (5565)
Multiplicity <sup>a</sup>	3.6 (3.7)	3.3 (2.8)	3.5 (2.5)	2.2 (1.8)	6.7 (6.7)
Completeness (%) <sup>a</sup>	97.2 (98.9)	95.2 (75.4)	97.3 (78.4)	97 (90.4)	99.9 (98.7)
Mean I/σ (I) <sup>a</sup>	7.3 (1.3)	9.1 (1.1)	15.8 (1.1)	16.2 (1.2)	6.3 (1.2)
CC <sub>half</sub> <sup>a</sup>	0.994 (0.554)	0.999 (0.48)	0.999 (0.569)	0.997 (0.679)	0.992 (0.535)
R <sub>merge</sub> <sup>a,b</sup>	0.081 (0.825)	0.044 (0.813)	0.035 (0.934)	0.021 (0.434)	0.174 (1.408)
R <sub>pim</sub> <sup>a,c</sup>	0.074 (0.725)	0.032 (0.664)	0.021 (0.670)	0.021 (0.431)	0.078 (0.625)
<i>Refinement</i>					
R <sub>factor</sub> /R <sub>free</sub>	0.215/0.267	0.180/0.230	0.146/0.180	0.158/0.207	0.236/0.283
RMSD bonds	0.0106	0.0125	0.0127	0.0135	0.0121
RMSD angles	1.48	1.51	1.47	1.52	1.53
No. of non-H atoms					
Protein	3925	4016	2225	2192	12,071
Ligands	8	18	4	10	0
Water	41	213	129	91	300
Protein residues	504 (2 chains)	514 (2 chains)	273 (1 chain)	271 (1 chain)	1562 (6 chains)
Average B-factors					
Main chain/side chain	40.5/45.8	19.1/28.6	25.6/38.0	29.7/40.2	34.9/38.6
Ligands/solvent	26.9/36.6	20.8/26.8	16.4/33.8	25.2/34.2	- /26.7
Ramachandran	97.42/100	97.07/100	98.81/100	97.78/100	90.56/100
favoured/allowed (%)					
Molprobrity score	1.17 (100th percentile N = 13,349, 1.95 ± 0.25 Å)	1.01 (99th percentile N = 3343, 1.41 ± 0.25 Å)	1.08 (99th percentile N = 4372, 1.46 ± 0.25 Å)	1.05 (99th percentile N = 4870, 1.52 ± 0.25 Å)	1.47 (98th percentile N = 11,758, 2.1 ± 0.25 Å)
MPn methylphosphonate					
<sup>a</sup> Values in parenthesis are for data in the high-resolution shell					
<sup>b</sup> R <sub>merge</sub> = $\sum_{hkl} \sum_i  I_i - I_m  / \sum_{hkl} \sum_i I_i$					
<sup>c</sup> R <sub>pim</sub> = $\sum_{hkl} \sqrt{1/n - 1/\sum_{i=1}^n  I_i - I_m  / \sum_{hkl} \sum_i I_i}$ , where I <sub>i</sub> and I <sub>m</sub> are the observed intensity and mean intensity of related reflections, respectively					

reported as ± the estimated standard deviation of the  $K_d$ . Te\_PtxB, Pm\_Phnd and Ps\_PtxB all bound phosphite with a nanomolar  $K_d$  in the range of ~50–240 nM (Table 1; Supplementary Fig. 3). In contrast, all three proteins bound methylphosphonate (MPn) and phosphate less tightly, with dissociation constants between ~30–100 and ~50–180 μM, respectively (Table 1; Supplementary Fig. 3); these values are an order of magnitude higher than the environmental concentration of these P compounds<sup>35</sup>. None of the proteins showed any evidence of binding the environmentally abundant<sup>15</sup> 2-aminoethylphosphonate (2AEPn). As an independent verification of the binding affinities we performed isothermal titration calorimetry (ITC) with Te\_PtxB with phosphite ( $K_d = 289 \pm 64$  nM) and phosphate ( $K_d = 15.9 \pm 0.36$  μM), which gave dissociation constants in good agreement with those determined by MST (Supplementary Fig. 4). Furthermore, our MST-derived binding affinities for Pm\_Phnd were similar to those reported previously by Feingersch et al.<sup>35</sup>. The proposed hypophosphite transporter, HtxB, bound hypophosphite with high affinity ( $K_d = 560 \pm 11$  nM), with some evidence of phosphite binding in the millimolar range but no binding detected for phosphate (Supplementary Fig. 5) or methylphosphonate up to a concentration of 10 mM (Table 1). These data show that at environmentally relevant concentrations Te\_PtxB, Pm\_Phnd and Ps\_PtxB clearly favour phosphite as a ligand, and HtxB appears to be specific for hypophosphite.

**Phosphite binding involves a P-H...π interaction.** To determine the structural mechanism of phosphite selectivity we co-crystallised Te\_PtxB and Pm\_Phnd with phosphite resulting in high-resolution structures (1.95 and 1.46 Å respectively) of the closed, ligand-bound complexes (Table 2). Te\_PtxB and Pm\_Phnd are both typical type II PBPs<sup>38</sup>, with two lobes separated by a hinge region surrounding a buried, enclosed central cavity that forms the ligand-binding pocket (Fig. 1a, b; Supplementary Figs. 6–9) (RMSD  $\alpha$ : ~1.7 Å). In the closed form, the amino acids that form the walls of the binding pocket pack tightly against the phosphite moiety and a single water molecule that is also buried in the binding pocket. Each oxygen atom of the phosphite contributes to an extensive hydrogen bonding network to the main chain and side chains of a cluster of conserved residues from lobe 2 (Te\_PtxB; Y55, Y100, S130, T131, S132 and H160; Pm\_Phnd; Y46, S126, T127, S128, H158 and D205, Figs. 1d, e and 2e, f). In Pm\_Phnd, the interaction of the D205 carboxylate with the ligand suggests that, under the crystallisation conditions used (pH 8), mono-anionic ( $\text{HPO}_3\text{H}^-$ ) rather than di-anionic ( $\text{HPO}_3^{2-}$ ) phosphite is the predominantly bound form. In Te\_PtxB, it is not possible to determine the protonation state of the phosphite from the pattern of hydrogen bonding, but given the pH of the crystallisation experiment (pH 4.2), it is also likely that the mono-anionic form of phosphite is bound (phosphite  $\text{pK}_a$  1.5, 6.7<sup>39</sup>). Both proteins contain a conserved histidine as one of the phosphite

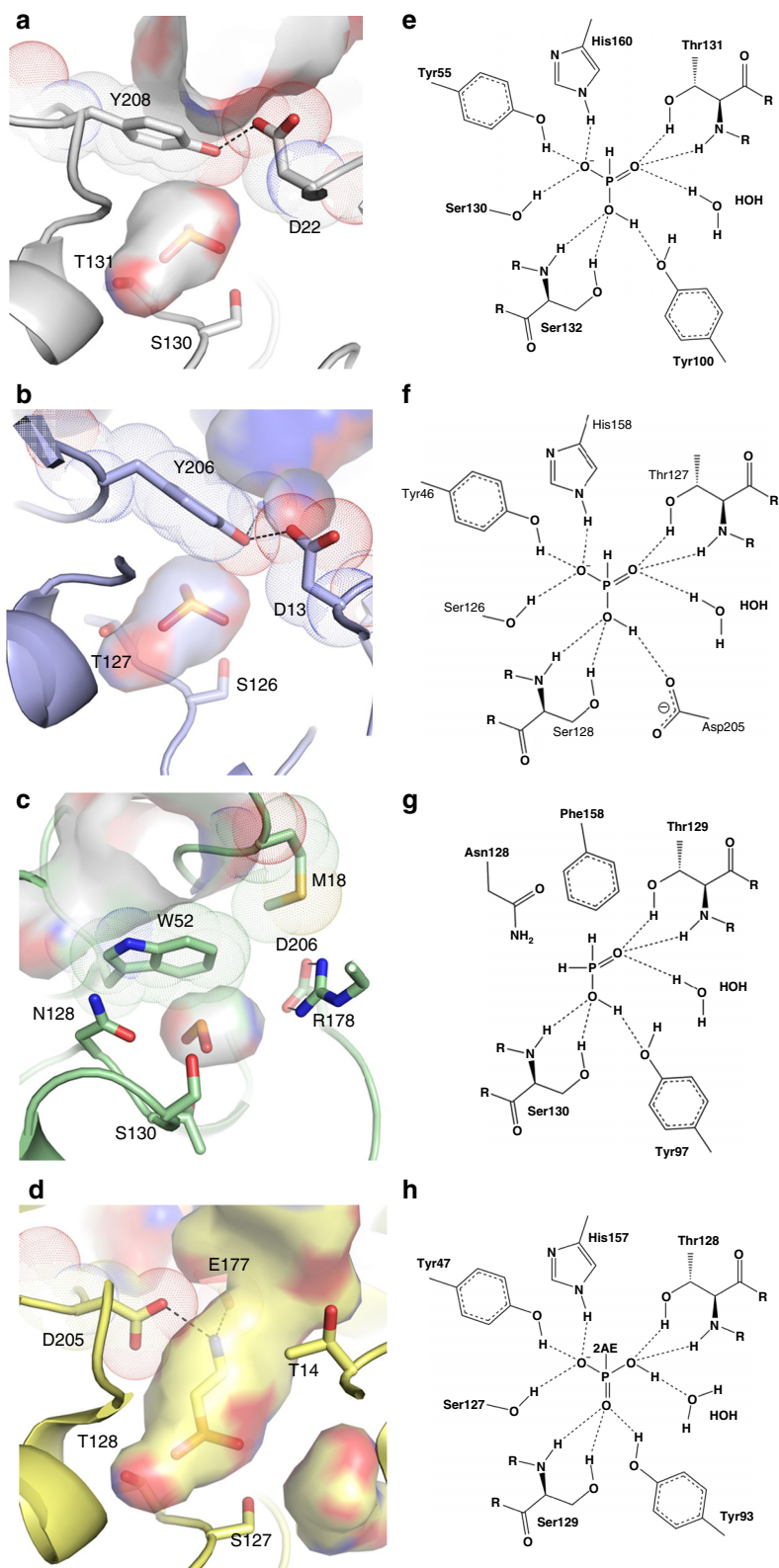


**Fig. 1** The crystal structures of Te\_PtxB and Pm\_Phnd in complex with phosphite and HtxB in complex with hypophosphite. **a–c** A cartoon representation of the overall fold of Te\_PtxB (grey, **a**), Pm\_Phnd (blue, **b**) and HtxB (green, **c**), with the ligand drawn as spheres and coloured in atom colours (N = blue, P = orange, O = red, S = yellow, H = white). The termini are labelled. See also domain topology diagram in Supplementary Fig. 6. **d–f** A detailed view of the ligand-binding pocket, highlighting the interactions made between the ligand and the protein for Te\_PtxB (**d**), Pm\_Phnd (**e**) and HtxB (**f**). The protein backbones are drawn as a cartoon, with the bound ligand and surrounding sidechains drawn as sticks, and coloured as in parts (**a–c**). The single buried water molecule (HOH) is drawn as a red sphere. Hydrogen bonds are drawn as dashed lines and the P–H... $\pi$  interaction is drawn as an orange dashed line. **g** A superposition of Te\_PtxB and Pm\_Phnd, viewed from above the phosphite binding site, shows how exchanging Y100 in Te\_PtxB for a phenylalanine (F90) in Pm\_Phnd requires D205 to be recruited to complete the network of hydrogen bonds around the phosphite ligand. The phosphite is drawn as partially transparent spheres and the buried water molecule is indicated as a red sphere. **h** A comparison of the binding pockets of Te\_PtxB and HtxB showing how the interactions around two of the P-oxygen binding sites are largely spatially conserved, but the third site is modified in HtxB (H160 to F158) to essentially block binding of a third oxygen moiety. The hydrogen atoms of the ligand are highlighted in green for HtxB and in white for Te\_PtxB, with most of the main chain atoms omitted for clarity

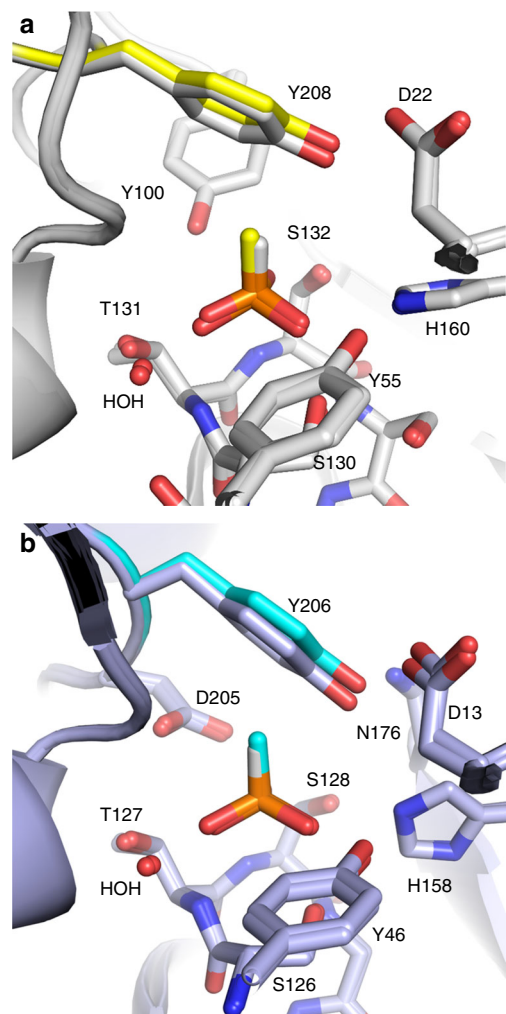
ligands (Te\_PtxB; H160 and Pm\_Phnd; H158), which could help stabilise the negative charge on the phosphite. However, in Te\_PtxB, one imidazole nitrogen atom acts as a hydrogen bond acceptor to a mainchain N–H (S178), suggesting that the histidine is neutral. In Pm\_Phnd the equivalent interaction is made to a buried water molecule outside of the binding pocket and the hydrogen bonding network around this water molecule is not sufficient to determine the charge on the histidine.

In both structures, the hydrogen atom at the R1 position of the phosphite points towards lobe 1, packing against the face of the

aromatic ring of a tyrosine residue (Te\_PtxB; Y208, Pm\_Phnd; Y206). This tyrosine sidechain forms a cap that is stabilised by intramolecular hydrogen bonds to residues on lobe 1 (Te\_PtxB; D22, Pm\_Phnd; D13 and N176), which in the closed conformation, contributes to the network of interactions that connect both lobes of the protein and bury the binding pocket from solvent (Fig. 2a, b, e, f). The distance between the R1 hydrogen and the plane of the aromatic ring of this tyrosine is  $\sim 2.6$  Å, which is less than the sum of the expected van der Waals radii and, as the P–H bond also impinges upon the  $\pi$  system at an angle of  $\sim 140^\circ$ , this



**Fig. 2** A detailed comparison of the spatial arrangement of the binding pocket and network of interactions around the ligand. The interior surface (partially transparent, atom colours) of the binding pocket and surrounding residues (sticks) in **a** *Te\_PtxB*; **b** *Pm\_PhxD*; **c** *HtxB* and **d** *E. coli PhnD* (PDB:3P7I). In each case, the capping residues are highlighted with a dotted surface that represents the van der Waals radii of the atoms. The external surface is also shown in each case, demonstrating how the binding pocket in reduced phosphorus binding PBP (a–c) is buried from the external solvent and in each case is much smaller than that of the phosphonate binding PhnD from *E. coli* (d). The protein backbone is drawn as a cartoon, with sidechains drawn in sticks and hydrogen bonds between the capping residues indicated by black dashes. Each figure (a–d) is accompanied by a schematic (e–h) that displays the hydrogen-bonding network between the protein and the oxygen atoms of the ligand. The 2-aminoethyl group of 2AEPn is abbreviated to 2AE in **h**



**Fig. 3** A comparison of phosphite and methylphosphonate binding in Te\_PtxB and Pm\_Phnd. **a** Superposition of the complexes of Te\_PtxB (grey) with phosphite (white hydrogen) and methylphosphonate (MPn; yellow methyl) showing that the residues around the binding pocket are in consistent positions in each structure apart from the capping tyrosine which moves by  $\sim 0.5$  Å in the MPn complex (Y208, yellow) in response to the larger van der Waals radii of the methyl group of the ligand. **b** Superposition of the phosphite (white hydrogen) and MPn (cyan methyl) complexes with Pm\_Phnd (blue) showing a similar movement of the capping tyrosine (Y206, cyan)

suggests a P-H... $\pi$  interaction occurs<sup>40, 41</sup> (Fig. 1d, e). To test the importance of the capping tyrosine for phosphite binding we generated mutants of Te\_PtxB, exchanging the tyrosine for a phenylalanine or alanine; circular dichroism (CD) spectra showed that neither substitution changed the secondary structure of the protein (Supplementary Fig. 10). Both mutants show reduced capacity to bind phosphite, with  $\sim 2$  and  $\sim 3$  orders of magnitude weaker binding for the Y208A and Y208F mutants, respectively (Table 3; Supplementary Fig. 11). The reduced affinity of the Y208F mutant suggests that simply positioning a  $\pi$  electron system above the R1-hydrogen of the phosphite ligand is not enough to achieve stable phosphite binding. The tyrosine cap therefore appears to play a dual role, both providing a platform for the P-H... $\pi$  interaction with phosphite and also stabilising the closed state by providing a hydrogen bond network between the two domains, centred on the tyrosine hydroxyl. We also tested the affinity of the Y208A and Y208F mutants for phosphate. Interestingly, the tyrosine to alanine mutation made little

difference to binding affinity, suggesting that phosphate may interact weakly with the protein in a partially open state, with few, if any interactions with the tyrosine side chain. The Y208F mutation essentially abolished phosphate binding, again showing the importance of the inter domain hydrogen bond network. Overall, this suggests that the tyrosine cap is a pre-requisite to achieve a fully closed state of the protein that engulfs the phosphite within the binding pocket and acts as a steric barrier for the binding of ligands with bulkier R groups, providing an explanation for the observed specificity of Te\_PtxB and Pm\_Phnd for phosphite.

Te\_PtxB and Pm\_Phnd have 2–4 orders of magnitude higher binding affinity for phosphite than phosphate and phosphonates (Table 1), but whilst attempts to crystallise complexes of either protein with ethylphosphonate (EPn), 2AEPn or phosphate were unsuccessful, structures of both proteins in complex with methylphosphonate (MPn) were obtained (Table 2; Fig. 3). The overall fold of the Te\_PtxB and Pm\_Phnd complexes with MPn were very similar to the comparative complexes with phosphite (RMSD  $\alpha$ :  $\sim 0.3$  Å) and since both MPn complexes were determined at high resolution (1.41–1.52 Å), the electron density for the ligand could be confidently interpreted within the binding pocket, with the MPn refining unequivocally in only one orientation (Supplementary Figs. 12 and 13). The only difference in chemical structure between phosphite and MPn is an exchange of the R1 hydrogen in phosphite for a methyl group in MPn. Given this similarity, it was unsurprising that the three oxygen atoms and the phosphorus atom of the MPn bind in the same position as those of the phosphite, with the methyl group of the MPn pointing in the same direction as the R1 hydrogen of the phosphite, towards the capping tyrosine residue (Y208/Y206). To accommodate the methyl group of the MPn the tyrosine moves in both structures by  $\sim 0.5$  Å, compared to the complexes with phosphite, resulting in an  $\sim 16\%$  expansion in the volume of the binding pocket (Supplementary Table 1). Despite this movement, the methyl-carbon atom of the MPn lies  $\sim 3.0$  Å from the plane of the tyrosine ring in both structures. This might suggest that a C-H... $\pi$  interaction is present, however the relative geometry of the two partners is sub-optimal for this type of interaction<sup>41</sup>. Thermal ellipsoids were calculated from the anisotropic B-factors for the three structures that were better than 1.5 Å resolution (Te\_PtxB and Pm\_Phnd with MPn and Pm\_Phnd with phosphite). These show an elongation of the ellipsoids for the tyrosine atoms in the MPn complexes, compared to those of in the phosphite complex, suggesting a degree of instability in the position of the tyrosine next to the methyl group of the MPn (Supplementary Fig. 14). Since Te\_PtxB and Pm\_Phnd have a significantly lower binding affinity for MPn than for phosphite (100- and 2000-fold, respectively), this suggests that the  $\sim 0.5$  Å movement of the capping tyrosine may alter the network of hydrogen bonds that stabilise the closed, ligand-bound conformation. To explore this, we measured the binding affinity of the Y208A and Y208F mutants of Te\_PtxB for MPn. The Ala substitution weakened the binding of MPn by a factor of 10, whilst the Phe mutation essentially abolished binding of MPn (Table 3 and Supplementary Fig. 11). This suggests that by swapping the bulky tyrosine for an alanine, the steric barrier provided by the Tyr is removed, providing space in the binding pocket for the methyl group of MPn. However, as the alanine is unable to form bridging hydrogen bonds to stabilise the closed conformation, the binding of MPn is weaker. CD shows that Y208F mutant is folded (Supplementary Fig. 10), however, we could not detect binding of MPn, suggesting that this amino-acid substitution blocks the closed ligand-bound complex from forming. Taken together, the exquisite selectivity of PtxB for phosphite can be explained not only by the favourable P-H... $\pi$

**Table 3** Microscale thermophoresis determined dissociation constants ( $K_d$ ) of Te\_PtxB and HtxB binding pocket mutants

Ligand	$K_d$ ( $\mu$ M)						
	Te_PtxB	Te_PtxB Y208A	Te_PtxB Y208F	HtxB	HtxB W52A	HtxB W52F	HtxB W52Y
Phosphate	46.80 $\pm$ 3.45	51.27 $\pm$ 1.76	>60 $\times 10^3$	NBD	—	—	—
Phosphite	0.17 $\pm$ 0.07	134.5 $\pm$ 6.4	294.81 $\pm$ 4.64	>10 $\times 10^3$	—	—	—
Methylphosphonate	30.63 $\pm$ 2.48	255.4 $\pm$ 20.2	>60 $\times 10^3$	NBD	—	—	—
Hypophosphite	2.21 $\pm$ 0.16	—	—	0.56 $\pm$ 0.11	NBD	28.49 $\pm$ 6.10	147.54 $\pm$ 14.48

$K_d$  values presented are derived from fitting three independent experiments to Eq. (1) and errors reported as  $\pm$  the estimated standard deviation of the  $K_d$ . NBD indicates no binding was detected up to a ligand concentration of 10 mM; - indicates  $K_d$  not measured

**Table 4** Data collection and refinement statistics part 2

	Ps_PtxB/phosphite (PDB:5O2J)	Ps_PtxB/MPn (PDB:5O37)	HtxB/hypophosphite (PDB: 5ME4)	Pm_PtxB/phosphite (PDB: 5LV1)
<i>Data collection</i>				
Wavelength ( $\text{\AA}$ )	0.92819	0.92819	0.97951	0.92819
Beamline	i04 - 1	i04 - 1	i24	i04 - 1
Resolution ( $\text{\AA}$ )	33.25-1.52 (1.55-1.52)	37.61-1.37 (1.39-1.37)	22.81-1.52 (1.55-1.52)	131.79-2.12 (2.16-2.12)
Space group	P2 <sub>1</sub> 2 <sub>1</sub> 2	P2 <sub>1</sub> 2 <sub>1</sub> 2	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P3 <sub>1</sub> 21
Unit cell ( <i>a</i> , <i>b</i> , <i>c</i> , $\alpha$ , $\beta$ , $\gamma$ )	113.21, 39.03, 63.52, 90, 90, 90	112.8, 39.12, 63.3, 90, 90, 90	40.24, 55.21, 125.2, 90, 90, 90	152.175, 152.175, 67.915, 90, 90, 120
Total reflections <sup>a</sup>	548,602 (22,254)	672,723 (21,736)	303,161 (11,860)	982,237 (39,470)
Unique reflections <sup>a</sup>	44,234 (2169)	59,851 (2940)	42,787 (1986)	51,504 (2522)
Multiplicity <sup>a</sup>	12.4 (10.3)	11.2 (7.4)	7.1 (6.0)	19.1 (15.7)
Completeness (%) <sup>a</sup>	100 (100)	99.9 (98.8)	97.7 (91.3)	100 (100)
Mean <i>I</i> / $\sigma$ ( <i>I</i> ) <sup>a</sup>	11.9 (1.2)	16.8 (1.2)	9.8 (1.1)	7.7 (1.4)
CC <sub>half</sub> <sup>a</sup>	0.999 (0.510)	1 (0.485)	0.998 (0.525)	0.996 (0.648)
$R_{\text{merge}}$ <sup>a,b</sup>	0.127 (1.774)	0.069 (1.463)	0.117 (1.209)	0.36 (2.82)
$R_{\text{pim}}$ <sup>a,c</sup>	0.039 (0.607)	0.022 (0.615)	0.05 (0.553)	0.086 (0.733)
<i>Refinement</i>				
$R_{\text{factor}}/R_{\text{free}}$	0.159/0.189	0.138/0.177	0.167/0.198	0.218/0.264
RMSD bonds	0.0131	0.015	0.0108	0.0112
RMSD angles	1.48	1.5	1.45	1.47
No. of non-H atoms				
Protein	1998	2038	2034	6010
Ligands	8	5	6	12
Water	193	225	175	123
Protein residues	259 (1 chain)	261 (1 chain)	255 (1 chain)	789 (3 chains)
Average B-factors				
Main chain/side chain	17.8/26.0	17.1/24.1	18.1/24.4	28.7/32.2
Ligands/solvent	21.1/32.1	11.0/32.0	31.8/31.6	19.3/25.9
Ramachandran favoured/allowed (%)	97.28/100	97.3/100	96.48/100	98.09/100
Molprobrity score	0.73 (100th percentile N = 4870, 1.52 $\pm$ 0.25 $\text{\AA}$ )	1.05 (99th percentile N = 3026, 1.37 $\pm$ 0.25 $\text{\AA}$ )	1.18 (97th percentile N = 4870, 1.52 $\pm$ 0.25 $\text{\AA}$ )	0.78 (100th percentile N = 11300, 2.12 $\pm$ 0.25 $\text{\AA}$ )

MPn methylphosphonate

<sup>a</sup>Values in parenthesis are for data in the high-resolution shell

<sup>b</sup> $R_{\text{merge}} = \sum_{hkl} \sum_i |I_i - I_m| / \sum_{hkl} \sum_i I_i$

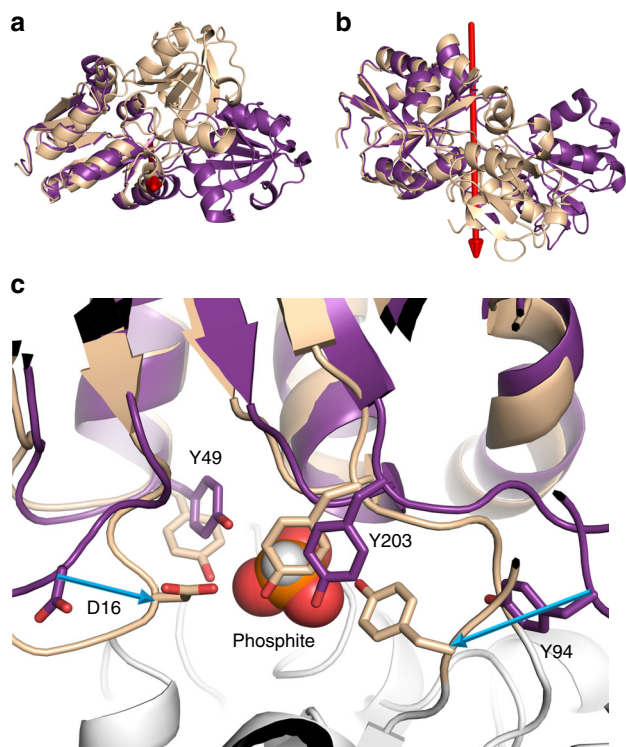
<sup>c</sup> $R_{\text{pim}} = \sum_{hkl} \sqrt{1/n - 1/\sum_{i=1}^n |I_i - I_m|} / \sum_{hkl} \sum_i I_i$ , where  $I_i$  and  $I_m$  are the observed intensity and mean intensity of related reflections, respectively

interaction between the phosphite and the protein, but also by a combination of a steric requirement for a hydrogen at the R1 position of the ligand and a strengthening network of hydrogen bonds around the binding pocket that stabilise the closed, phosphite-bound state.

#### Ligand binding induces a large conformational change in PtxB.

Determining an apo-structure of a PBP from this family of ABC transporters is difficult due to the inherent flexibility of the hinge region between the two domains, and it has only been achieved previously with PhnD from *E. coli* (PDB:3S4U), which was locked in an open conformation by substituting H157 with an alanine<sup>14</sup>. In

our study, numerous attempts to crystallise apo forms of Te\_PtxB and Pm\_PhnD were unsuccessful, but a 2.1  $\text{\AA}$  resolution structure of *Pseudomonas stutzeri* PtxB (Ps\_PtxB) in an open conformation was determined (Table 2). We also obtained structures of phosphite and MPn complexed Ps\_PtxB (Table 4; Fig. 4) and determined the associated ligand affinities (Table 1); both the structures and binding data were similar to the other phosphite-binding proteins discussed previously. Analysis of the conformational change between the open and closed, phosphite-bound structure of Ps\_PtxB using DynDom<sup>42</sup> revealed that domain closure is induced by a  $\sim 60^\circ$  rotation around an axis that lies between the two lobes (Fig. 4). This is facilitated by flexible loops that form the hinge region between the lobes (residues A85-T93 and Y200-Q202), as



**Fig. 4** The conformational change between the open and closed structures of Ps\_PtxB. **a, b** A comparison of the open (purple) and closed (beige) structure of Ps\_PtxB, superimposed on lobe 2, showing the 60° rotation of one domain relative to the other around an axis that lies between the two domains (red arrow). **c** Superimposing the open and closed structures on lobe 2 (lobe 1 is shown in grey) shows that the spatial arrangement of Y49 and Y203 is conserved in both the open and closed states, whilst Y94 and D16, which both sit on flexible loop regions, undergo large conformational movements (blue arrows) of 7 and 4.5 Å ( $\alpha$ - $\alpha$ ), respectively. The phosphite is shown as spheres

well as residues from the extreme end of the C-terminal helix (G255-L258). These changes are similar to those observed between the open and closed structures of *E. coli* PhnD<sup>14</sup>.

Superposition of the individual lobes of Ps\_PtxB in the open and closed conformations shows that lobe 2 (residues 96–196, 0.35 Å RMSD  $\alpha$ ) adopts a very similar structure, with the conserved STS-H motif (S124, T125, S126 and H155) spatially conserved in both conformations. In contrast, lobe 1 (residues 1–84 and 202–255, 1.55 Å RMSD  $\alpha$ ) shows some differences between the open and ligand-bound states. The positions of Y94, which hydrogen bonds to one of the phosphite oxygen atoms in the closed structure, and D16, which hydrogen bonds to and stabilises the capping tyrosine in the closed structure, move ~7 Å and ~4.5 Å ( $\alpha$ - $\alpha$ ), respectively, on ligand binding. Y94 resides on the hinge region, the conformation of which alters substantially during domain closure, whilst D16 sits on a flexible loop that latches over the binding pocket in the closed conformation. In addition, the B-factors of the core residues in lobe 2 are much lower in the closed, ligand bound, conformation compared to those in the open structure, implying lower overall flexibility when phosphite is bound to the protein (Supplementary Fig. 15).

#### HtxB binds hypophosphite using a similar P-H... $\pi$ interaction.

The structure of HtxB in complex with hypophosphite (Table 4) shows that the protein adopts the same closed conformation as seen in the complex of Te\_PtxB with phosphite (RMSD  $\alpha$  1.86 Å) (Fig. 1c, f). The ligand-binding pocket is very similar, with

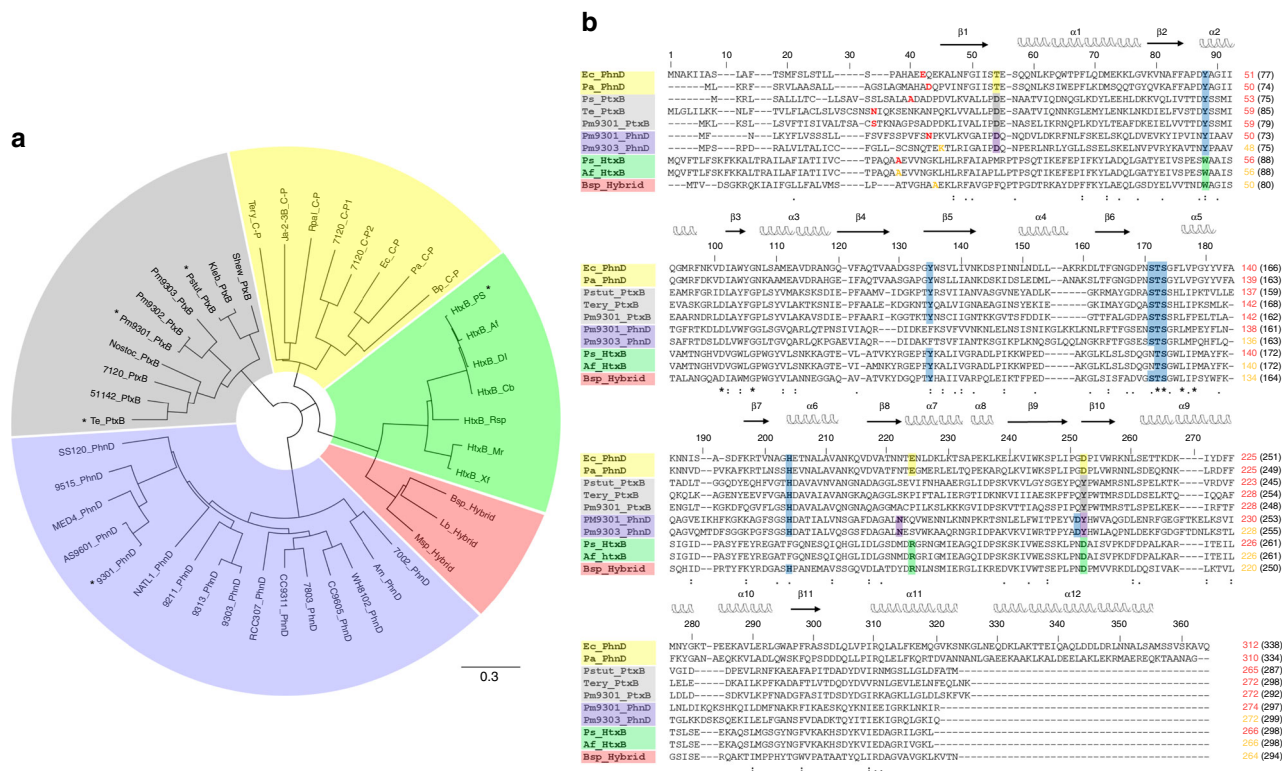
two of the hypophosphite oxygen atoms (O1 and O2) adopting the same position and making the same interaction as the equivalent oxygen atoms in the Te\_PtxB/phosphite complex (Fig. 1f, h). However, H160, Y55 and S130, which all make interactions with O3 of the phosphite in Te\_PtxB are replaced in HtxB with F158, W52 and N128, respectively. This has the effect of removing the hydrogen bonding potential around the O3 site, reducing the volume of the binding pocket by ~36 % (Fig. 2c, g; Supplementary Table 1) and providing steric selectivity for hypophosphite (two oxygen substituents) over phosphite (three oxygen substituents). Consistent with this, we observed only negligible levels of phosphite binding to HtxB in our MST assays (Table 1). Conversely, hypophosphite binds to Te\_PtxB with low micromolar affinity ( $K_d$  ~2  $\mu$ M), indicating that presumably its smaller size can be accommodated within the Te\_PtxB binding pocket, albeit with suboptimal interactions.

The changes in the ligand-binding pocket also extend to the capping residue, Y208, which in Te\_PtxB provides the  $\pi$  system for the P-H... $\pi$  interaction. In HtxB, a buried ion pair occurs in this region of the structure (D206 from lobe 1 and R178 from lobe 2), which connects the two lobes of the protein when it is in the closed conformation (Fig. 1f). Nevertheless, in HtxB the P-H... $\pi$  interaction between the ligand and the protein is conserved by the presence of a tryptophan residue, W52 (Fig. 1h). The aromatic ring system of this tryptophan is positioned adjacent to the hydrogen at the R2 position of hypophosphite and is in the appropriate orientation for a P-H... $\pi$  interaction to occur (~2.6 Å, 140°), in the same way as seen in the phosphite-binding proteins (Fig. 1f). Replacing this tryptophan with a phenylalanine or tyrosine substantially reduces the binding affinity of HtxB for hypophosphite by 50-fold and 250-fold, respectively (Table 3; Supplementary Fig. 11). The smaller size of both sidechains would clearly alter the packing within this region of the binding pocket and may not provide a large enough platform for the R2 hydrogen of the hypophosphite to pack against. The lower binding affinity in the tyrosine mutant may be due to its hydroxyl group interfering with Phe158 on the adjacent side of the binding pocket, possibly blocking closure of the protein around its substrate. Changing the tryptophan to an alanine results in the complete abolition of binding, suggesting the packing interactions between the tryptophan, nearby residues and the ligand are critical for high-affinity binding of hypophosphite. Therefore, we have demonstrated that in both phosphite and hypophosphite binding proteins, positioning the  $\pi$  electron system of an aromatic amino acid adjacent to the P-H atom of the ligand is a conserved molecular mechanism governing specificity of the PBPs for reduced phosphorus ligands.

#### Structural comparisons with phosphonate binding PBPs.

A phylogenetic tree of members of the phosphonate/phosphite/phosphate periplasmic substrate binding protein family (KEGG orthology K02044; InterPro Family IPR005770; Pfam12974) shows they cluster into five main sub-groups based on primary sequence (Fig. 5a). These groups are: (1) C-P lyase-linked phosphonate transporters (PhnD); (2) phosphite dehydrogenase (PtxD)-linked phosphite transporters (PtxB); (3) the PhnD PBP of the marine PhnDCE phosphite transporter; (4) hypophosphite transporters (HtxB); and (5) transporters we have referred to as ‘hybrid’ as they share sequence features with both PtxB and HtxB. We have determined the structure of three proteins from group 2 (Te\_PtxB, Pm\_PtxB and Ps\_PtxB) and proteins from groups 3 (Pm\_Phnd) and 4 (HtxB). From structure-based sequence alignments of representative members of these groups it is clear that the residues imparting phosphite and hypophosphite specificity are strictly conserved (Fig. 5b). The only



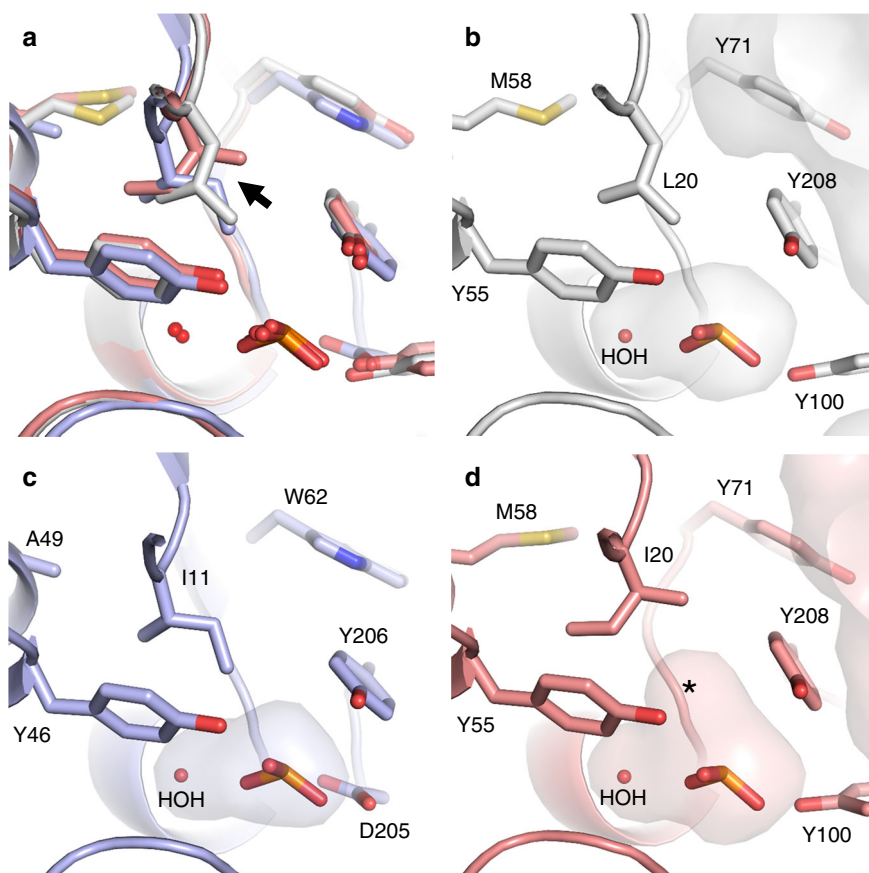


**Fig. 5** Protein phylogeny and sequence alignments of reduced phosphorus compound PBPs. **a** Unrooted phylogenetic tree of reduced P compound transporter PBP amino-acid sequences. Phosphite binding PtxB (grey) and PhnD (purple) homologues cluster separately from each other and from C-P lyase-linked phosphonate binding PhnDs (yellow), HtXB homologues (green) and a putative ‘hybrid’ class of transporters (red). Proteins characterised in this study are marked with an asterisk and the scale bar represents the number of substitutions per site. **b** Sequence alignment of reduced phosphorus compound binding PBPs, coloured as in **a**. 2-Aminoethylphosphonate-specific cap residues are highlighted yellow, phosphite-specific cap residues are highlighted grey (PtxB) or purple (PhnD) and hypophosphite-specific cap residues are highlighted green. Conserved residues that form H-bonds with ligand oxygen atoms are highlighted in blue. The position in the alignment is shown above the sequence and residue numbering for each protein at row ends. For residue numbering, the red (or orange, see below) value corresponds to the residue numbering of the structures, as the predicted N-terminal signal peptides were not present in recombinant proteins, and the black number in parentheses is the residue number of the unprocessed proteins. The second residue (following the initiator Met1) in the recombinant proteins used in this study is shown in red and bold in the sequence; in the case of proteins not directly studied here, the equivalent residues/residue numbers are shown in orange. For full details of the proteins included in this figure see Supplementary Table 3

other proteins from this family with known structures are C-P lyase-linked PhnDs, namely the *E. coli* 2AEPn binding protein (PDB:3P7I) (Ec\_PhnD)<sup>14, 15</sup> and its close homologue from *Pseudomonas aeruginosa* (PDB:3N5L). A structural comparison between Te\_PtxB and Ec\_PhnD (RMSD  $\alpha$ : 1.56 Å) shows that whilst the residues that interact with the ligand oxygen atoms are conserved, the capping residues that confer phosphite specificity in Te\_PtxB (Y208 and D22) are replaced with D205 and T14 in Ec\_PhnD (Fig. 2d, h). These changes result in Ec\_PhnD having a larger binding pocket required for phosphonate binding, with D205 and the adjacent E177 also providing hydrogen bond acceptors to the amino group of the 2AEPn (Fig. 2d), contributing to the low nanomolar binding affinity for this ligand<sup>14, 15, 35</sup>. This D/T/E cap is conserved in many other group 1 members, including the *P. aeruginosa* PhnD, indicating high-affinity 2AEPn binding, however, there are also PhnD homologues in other bacteria with alternative residues at these positions, reflecting the diversity of R1 side chains of potential environmental phosphonate ligands<sup>14</sup>.

**Prochlorococcus** ecotypes with two phosphite transporters. *Prochlorococcus* is abundant in low phosphate ocean regions but cannot grow on phosphonates<sup>12, 35, 43</sup>, despite some ecotypes having homologues of PhnYZ, which have been shown to oxidise

2AEPn *in vitro*<sup>44, 45</sup>; therefore, phosphite could be an important P source to this globally significant oceanic bacteria. The gene encoding Pm\_PhnD in *P. marinus* MIT9301 is not responsive to P-limitation and is not genomically linked to known phosphite or phosphonate utilisation genes<sup>12, 43, 46</sup>, nevertheless, we and others<sup>35</sup> have shown that it binds phosphite with high affinity. Some *P. marinus* ecotypes also have a *ptxABCD* operon, encoding a PtxABC-transporter and a phosphite dehydrogenase<sup>12</sup>, indicating that these bacteria have a second phosphite uptake system. Based on protein phylogeny (Fig. 5a and ref. <sup>12</sup>), these two types of phosphite binding protein not only branch separately from C-P lyase-linked PhnD and HtXB homologues, but also from each other, indicating that they represent phylogenetically separate phosphite transporters. Conversely to the phosphite-specific PtxB proteins discussed above, the *P. marinus* MIT9301 PtxB (Pm\_PtxB) has previously been reported to have similar low micromolar binding affinities for phosphite, MPn and EPn, measured by ITC<sup>35</sup>. To try to explain this, we determined the structure of the Pm\_PtxB/phosphite complex (Table 4), which shows that Pm\_PtxB has the same overall fold as Te\_PtxB, with complete conservation of the amino acids that coordinate the ligand, including the P-H... $\pi$  interaction with Y206 (Supplementary Fig. 16). We calculated the volume of the binding pocket in Pm\_PtxB and found that it is 23–29% larger than those of the



**Fig. 6** The volume of the binding pocket in Pm\_PtxB is ~25% bigger than that of Te\_PtxB and Pm\_PhnD. Superposition of the Pm\_PtxB/phosphite structure (pink) with those of Te\_PtxB (grey) and Pm\_PhnD (blue) (**a**) shows how a sequence difference (black arrow) in the binding pocket increases the volume of the cavity in Pm\_PtxB. In Te\_PtxB (**b**) and Pm\_PhnD (**c**) Leu20 and Ile11 (respectively) provide a hydrophobic surface for the ligand to pack against within the binding pocket. In Pm\_PtxB (**d**) Ile20 adopts an alternative rotamer making the hydrophobic region of the pocket slightly larger, as indicated by the asterisk (\*)

phosphite-specific PBPs (Supplementary Table 1). Comparison of the Pm\_PtxB binding pocket with that of Te\_PtxB (RMSD  $\alpha$ : 0.75 Å) and Pm\_PhnD (RMSD  $\alpha$ : 1.52 Å) shows that this expansion is due to Ile20, which is conserved in Pm\_PhnD (Ile11) and a leucine (Leu20) in Te\_PtxB, as it adopts a different rotamer that forms a small hydrophobic cavity next to the capping tyrosine that is absent in the other proteins (Fig. 6). This seems to be due to several sequence changes in the second shell of amino acids within lobe 1 that create space for the Ile to rotate (Fig. 6). We were unable to measure the dissociation constant of Pm\_PtxB for P-ligands here as the protein showed evidence of aggregation in the MST assay. Thus, from the structural similarities with Te\_PtxB we can speculate that Pm\_PtxB is most likely specific for phosphite, however, due to the extra space in the binding pocket could also bind simple phosphonates, in agreement with the previously reported ITC binding data<sup>35</sup>. The unknown substrate of the *Prochlorococcus* PhnYZ system, encoded downstream of the *ptxABCD* genes<sup>12, 44</sup>, may also be a ligand of Pm\_PtxB. Therefore, some *Prochlorococcus* ecotypes appear to have two independent phosphite transporter systems, and may also be able to take up some simple phosphonate species. The PhnDCE transporter is ubiquitous in *Prochlorococcus* and is also present in numerous marine *Synechococcus* strains, but its *in vivo* function is unknown as only ecotypes that have PtxABCD can grow on phosphite as a sole P source<sup>12</sup>. Further studies are therefore required to ascertain the biological role of phosphite uptake by the PhnDCE transporter and the *in situ* substrates of the PtxABC transporter in this important primary producer.

## Discussion

We have determined structures of PBPs from reduced phosphorus compound ABC transporter systems in complex with phosphite and hypophosphite. The structures reveal that ligand specificity, measured here by thermophoresis, arises from a combination of steric considerations and the interaction between a ligand P-H hydrogen atom and an aromatic  $\pi$  system in the binding pocket. In phosphite binding PBPs a strictly conserved tyrosine residue provides the  $\pi$  system, and in HtxB an invariant tryptophan makes the equivalent interaction with hypophosphite. We identified similar P-H... $\pi$  interactions in 12 small-molecule crystal structures within the Cambridge Structural Database (Supplementary Table 2; Supplementary Fig. 17), however, such interactions with reduced phosphorus species could not be identified in protein structures in the PDB. It thus appears that, although C-H... $\pi$ , N-H... $\pi$  and cation-X... $\pi$  bonds are well characterised in proteins<sup>40, 41, 47–50</sup>, the structures reported here represent the first observations of P-H... $\pi$  interactions in protein/ligand complexes.

Phosphite is a useful P source for biotechnological, industrial and agricultural applications. PtxD has been introduced to *Arabidopsis*, tobacco and rice, allowing the use of phosphite as a dual fertilisation and weed control system<sup>28</sup>. PtxD and phosphite are also used as selectable markers in bacteria<sup>20, 34</sup>, yeast<sup>24</sup>, plants<sup>25</sup> and algae<sup>26</sup>; to provide competitive advantage in industrial fermentation processes<sup>27</sup>; and in NADH-dependent biocatalysis for *in situ* cofactor regeneration<sup>29</sup>. So far these applications have only exploited PtxD, relying on non-specific phosphite uptake

that requires media phosphite concentrations of  $\geq 0.1$  mM. As phosphite is cheap it can be provided at high concentrations, but our binding data suggest that introducing a high-affinity transporter would reduce the phosphite requirement in these applications. Furthermore, some technologies require that phosphite uptake is highly specific. For example, a recent study by Hirota et al.<sup>34</sup> described the use of phosphite-dependent growth as a biocontainment strategy, which requires a phosphite transporter that does not transport phosphate. The authors introduced PtxD and the *Ralstonia* sp. 4506 PtxABC transporter into an *E. coli* strain devoid of its native phosphate acquisition systems, resulting in a strain that grew on phosphite, but also phosphate when the ligand was supplied at 1 mM. This concentration is several orders of magnitude higher than environmental phosphate concentrations, and based on the micromolar  $K_d$  for phosphate binding to PtxB, determined in our study (Table 1), it is not surprising that phosphate supported growth. In the same study, addition of the *P. stutzeri* WM88 HtxBCDE hypophosphite transporter to the same *E. coli* strain did not allow growth on phosphate, as predicted by our binding studies which show that HtxB does not bind phosphate (Supplementary Fig. 5), but did allow growth on phosphite, which we would not have anticipated. Hirota et al.<sup>34</sup> show that the affinity of the HtxBCDE hypophosphite transporter is clearly high enough to support uptake in vivo when phosphite was supplied at 1 mM, but not when lowered to 5  $\mu$ M, which still exceeds the highest known environmental phosphite concentrations. We did not detect significant phosphite binding by recombinant HtxB, which, from our structural evidence, is consistent with the architecture (Fig. 1f) and size (Supplementary Table 1) of the HtxB binding pocket. This suggests that for HtxB to bind phosphite the protein must undergo a conformational change or the mode of binding of the ligand would have to be different to that with hypophosphite.

In summary, we have determined the molecular basis for high affinity and specific phosphite and hypophosphite uptake in transporters from globally important and abundant terrestrial and marine microorganisms. The structures of two distinct types of phosphite-binding protein and a hypophosphite-binding protein reveal that a P-H... $\pi$  interaction ensures high-affinity binding of specific ligands, representing a universal mechanism in these PBPs for recognising ligands with a P-H bond. The initial encounter with phosphite or hypophosphite in the bacterial periplasm is the first step in assimilating these potentially important P species, which are increasingly recognised as important constituents of the global phosphorus redox cycle<sup>51</sup> that may influence the biogeochemistry and productivity of the oligotrophic oceans and terrestrial ecosphere<sup>1–12</sup>.

## Methods

**DNA manipulation.** Primers and plasmids used in this study are provided in Supplementary Tables 4 and 5 respectively. *Escherichia coli* JM109 competent cells (Promega, UK) were used for all standard cloning steps. The *Trichodesmium ptxB* gene (Tery\_0366, *Te\_ptxB*) minus predicted N-terminal signal peptide coding sequence (nucleotides 1–81 = amino acids 1–27, SignalP 4.1 Server<sup>52</sup>) and stop codon was amplified from *Trichodesmium erythraeum* IMS101 (Prufert-Bebout et al.<sup>53</sup>, obtained from NCMA, USA) genomic DNA using Q5 DNA polymerase (New England Biolabs, UK) and primers *Te\_ptxB-F* and *Te\_ptxB-R* and cloned into the *NdeI* and *XhoI* sites of pET21a(+) (Novagen, UK). The *Prochlorococcus marinus* MIT 9301 *phnD* (P9301\_07261, *Pm\_phnD*) and *ptxB* (P9301\_12511, *Pm\_ptxB*) and *Pseudomonas stutzeri* WM88 (Ps) *ptxB* and *htxB* genes with their N-terminal signal peptide coding sequences<sup>52</sup> (*Pm\_phnD* = nucleotides 1–72, amino acids 1–24; *Pm\_ptxB* = nucleotides 1–63, amino acids 1–21; *Ps\_ptxB* = nucleotides 1–69, amino acids 1–23; *htxB* = nucleotides 1–99, amino acids 1–33) and stop codons removed were synthesised with codons optimised for expression in *E. coli* (Integrated DNA Technologies, Inc., USA) and sub-cloned into pET21a(+) as above. The Quik-Change II Site-Directed Mutagenesis Kit (Agilent Technologies, USA) was used to generate genes encoding point mutants. All genes were sequence verified (GATC Biotech) and confirmed to be in-frame with the C-terminal His<sub>6</sub>-tag.

**Protein production and purification.** Plasmids were introduced into *E. coli* BL21 (DE3) (Invitrogen, UK) for protein production. Cultures were grown in LB broth with 100  $\mu$ g ml<sup>-1</sup> ampicillin shaking (250 rpm) at 37 °C. At an optical density at 600 nm of ~0.6, isopropyl- $\beta$ -D-thiogalactopyranoside was added to a final concentration of 0.4 mM to induce expression and the cultures were incubated for a further 16 h at 18–20 °C. Cells were harvested by centrifugation at 12,000  $\times$  g for 20 min at 4 °C and suspended in binding buffer (25 mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) pH 7.5, 0.5 M NaCl, 5 mM imidazole, EDTA-free protease inhibitor (Roche, UK)). Cell suspensions were sonicated on ice (6  $\times$  30 s bursts with 30 s intervals between) and insoluble debris removed by centrifugation at 45,000  $\times$  g for 20 min at 4 °C. Proteins were purified by immobilised Ni-affinity chromatography on a 5 ml Chelating Sepharose™ Fast Flow resin column (GE Healthcare, UK). Bound protein was washed with binding buffer containing 20 mM imidazole and then eluted into 25 mM HEPES pH 7.5, 400 mM imidazole, 0.1–0.2 M NaCl. Proteins were further purified on a 22 ml Superdex S200 Increase column (GE Healthcare) in 25 mM Tris/HCl, 0.2 M NaCl pH 7.4 and eluted as single peaks. Peak fractions were pooled and concentrated to 2.5 ml and protein was buffer exchanged into 25 mM Tris/HCl pH 7.4 with 0–0.2 M NaCl using a PD-10 desalting column (GE Healthcare). Proteins were used immediately or stored at 4 °C.

**Thermophoresis.** The binding affinity of proteins for P ligands was determined using Microscale Thermophoresis (MST) with a Monolith NT.115 instrument (NanoTemper Technologies, Germany). Protein (100 nM) was labelled with RED-tris-NTA dye (Nanotemper Technologies) according to the manufacturer's instructions in 50 mM HEPES pH 7.4, 250 mM NaCl, 0.05% Tween-20. A volume of 10  $\mu$ l of 100 nM labelled protein was mixed with 10  $\mu$ l of ligand in 50 mM HEPES pH 7.4, 250 mM NaCl, 0.05% Tween-20. 4  $\mu$ l of protein–ligand mixture was loaded into Premium grade capillaries (Nanotemper Technologies) and thermophoresis was measured at 22 °C for 22 s with 20% LED power and 40% infrared laser power. Data from three independent measurements were combined and analysed using the MO.Affinity Analysis software version 2.1 (Nanotemper Technologies), fitted to a single binding site model (Eq. 1) where  $[l]$  is the concentration of ligand and data plotted using Igor Pro version 7 (Wavemetrics Inc., USA).

$$f(l) = \text{Unbound} + \frac{(\text{Bound} - \text{Unbound}) \times \left( [l] + [\text{protein}] + K_d - \sqrt{([l] + [\text{protein}] + K_d)^2 - 4 \times [l] \times [\text{protein}]} \right)}{2 \times [\text{protein}]} \quad (1)$$

Equation 1 shows the single binding site model used to determine dissociation constants.

**Isothermal titration calorimetry.** ITC experiments were performed on a Nano ITC (TA Instruments, USA) at 22 °C. Purified *Te\_ptxB* (200  $\mu$ M) was exchanged into a buffer containing 50 mM HEPES, 250 mM NaCl, 0.05% Tween-20, pH 7.4 using a PD-10 desalting column. The ligand was dissolved in the same buffer, and then diluted to 2 mM. Each titration series consisted of 32 injections of 1.49  $\mu$ l ligand into a 160  $\mu$ l protein sample. Control experiments were carried out by injecting ligand into the buffer, and the resulting heat of dilution was subtracted from the binding isotherm data. The first injection was ignored in the final analysis. The raw ITC data were processed using NanoAnalyse Data Analysis V 3.7.0 (TA Instruments) and fitted to a single-site binding model. Data were plotted in Igor Pro version 7.04 (Wavemetrics Inc.).

**Circular dichroism spectroscopy.** Spectra were recorded with a JASCO-810 spectrometer (JASCO, UK). Protein (0.1 mg ml<sup>-1</sup>) was in 5 mM sodium phosphate buffer, pH 7.4 at 25 °C. Spectra were recorded in a cuvette with a 0.1-cm path length. Spectra were recorded continuously, from 250 to 190 nm (50 nm s<sup>-1</sup>, 1 nm increments, 4 s response, 6 accumulations).

**Crystallisation and structure determination.** *Te\_ptxB* (10 mg ml<sup>-1</sup>) was co-crystallised with 5 mM sodium phosphite or methylphosphonic acid (MPn) (pH 7.0) by sitting-drop vapor diffusion (200 nl: 200 nl, 290 K) in 0.2 M NaCl, 0.1 M phosphate-citrate buffer pH 4.2 and 20% (w/v) PEG 8000 or 0.2 M magnesium chloride, 0.1 M sodium acetate pH 5 and 20% (w/v) PEG 6000, respectively. *Pm\_phnD* (7 mg ml<sup>-1</sup>) was co-crystallised using the same method, but in 0.1 M SPG (succinic acid, sodium dihydrogen phosphate and glycine) buffer pH 8 and 25% (w/v) PEG 1500 or 0.1 M MMT (DL-malic acid, MES and Tris base) buffer pH 8 and 25% (w/v) PEG 1500, for the phosphite or MPn complexes, respectively. Apo *Ps\_ptxB* (11 mg ml<sup>-1</sup>) was crystallised using the same method, but in conditions containing 0.1 tri-sodium citrate pH 5.5 and 20% (w/v) PEG 3000 and in the presence of 5 mM ethylphosphonate (EPn), no electron density for which was observed in the resulting structure. Crystals of *Ps\_ptxB* (11 mg ml<sup>-1</sup>) in complex with phosphite and MPn were grown at pH 5 and pH 4, respectively, from 0.1 M MMT buffer (DL-malic acid, MES and Tris base) and 25% (w/v) PEG 1500. *HtxB* (7 mg ml<sup>-1</sup>) was co-crystallised with 3 mM sodium hypophosphite using the same method, but in 0.2 M sodium/potassium tartrate, 0.1 M bis-tris propane pH 8.5 and 20% (w/v) PEG 3350. For all the above proteins, thin, rod-shaped

crystals grew within 48 h (290 K). Conversely, long rod-shaped crystals of Pm\_PtxB (5 mg ml<sup>-1</sup>) in complex with phosphite (5 mM) took a number of weeks to grow at 290 K using the same method as the other proteins in conditions containing 0.1 M magnesium chloride, 0.1 M sodium citrate pH 5.0 and 15% PEG 4000. All crystals were harvested and cryoprotected in their mother-liquor with an additional 25% ethylene glycol or glycerol before plunge cooling in liquid nitrogen.

Crystals were mounted on a 100 K cryostream at the Diamond Light Source and for each structure data were collected from a single crystal at a single wavelength (for details see Tables 2 and 4). Data were autoprocessed using Xia2<sup>54</sup> and a resolution cut-off was applied using CC<sub>half</sub> (for processing statistics see Tables 2 and 4). The structure of the Te\_PtxB/phosphite complex was determined using the automated molecular replacement pipeline, Mr. Bump<sup>55</sup>. The best solution was produced by Phaser<sup>56</sup> using a homology model generated by MolRep<sup>57</sup> from the structure of *E. coli* PhnD (PDB:3P7I)<sup>14</sup>. Automated model building with Buccaneer<sup>58</sup> produced an initial model comprising two copies of Te\_PtxB in a back-to-back arrangement. Several rounds of manual model building and refinement were carried out in COOT<sup>59</sup> and Refmac<sup>60</sup>, before validating the final model with Molprobity<sup>61</sup>. The structures of the other complexes were determined by molecular replacement with Phaser<sup>56</sup> using a monomer of the highest-resolution structure of Te\_PtxB or Pm\_PhnD available at the time. The models were built, refined and validated using the same methods as those used with the Te\_PtxB/phosphite complex, but the three structures that were sub-1.5 Å resolution (Te\_PtxB/MPn, Pm\_PhnD/phosphite and Ps\_PtxB/MPn) were refined with anisotropic B-factors. The apo structure of Ps\_PtxB was determined by molecular replacement using a model of the open structure of the protein, the conformation of which was based on the structure of the H157A mutant of *E. coli* PhnD. Phaser successfully found the six subunits that were contained within the asymmetric unit, which were subjected to ridged body refinement in PHENIX<sup>62</sup> to improve the fit to the electron density, prior to the same process of model building and refinement that was applied to the other structures (for refinement statistics see Tables 2 and 4).

In the structure of each ligand complex, clear electron density for the ligand was present in the binding pocket. However, as the three P ligands under study have similar chemical structures to each other and to phosphate, careful assignment of the ligand to the difference density was required. For each structure the relevant ligand was refined in the density in the orientation that made most chemical sense, considering van der Waals contacts and hydrogen bonds. These refinements were compared to controls, in which, for example, the phosphite was replaced with phosphate, or the methylphosphonate was refined with the methyl group in different orientations. In each case the controls confirmed that the ligand had been modelled correctly as evidenced by the B-factors, difference maps and relevant overall omit maps (Supplementary Figs. 12 and 13), which were calculated using SFCHECK (CCP4<sup>63</sup>). Ligand coordinate files and dictionaries were all generated in JLigand<sup>64</sup>. Superpositions were carried out in Superpose<sup>65</sup> by secondary structure matching of the C $\alpha$  atoms. The volume of the ligand-binding pocket in each protein complex was analysed using the CASTp server using a probe radius of 1.4 Å<sup>66</sup> and UCSF Chimera<sup>67</sup>. Analysis of the crystal packing in each structure using the PISA server<sup>68</sup> showed that no more than 5% of the total surface area of each protein was buried between neighbouring chains in the crystals, suggesting that all the proteins are monomeric, which agrees with the elution volume from size exclusion chromatography.

**Bioinformatics.** Homologues of PhnD/PtxB/HtxB proteins were identified by BLAST searches<sup>69</sup>. Sequence alignments were performed using ClustalW<sup>70</sup>. The phylogenetic tree was made with Geneious version 10.0.2 (<http://www.geneious.com>)<sup>71</sup>. For details of proteins used for these analyses see Supplementary Table 3.

**Data availability.** Atomic coordinates and structure factors for the reported crystal structures have been deposited in the RCSB Protein Data Bank ([www.pdb.org](http://www.pdb.org)) with accession codes: 5JVB (Te\_PtxB/phosphite), 5LQ1 (Te\_PtxB/MPn), 5LQ5 (Pm\_PhnD/phosphite), 5LQ8 (Pm\_PhnD/MPn), 5O2K (apo Ps\_PtxB), 5O2J (Ps\_PtxB/phosphite), 5O37 (Ps\_PtxB/MPn), 5ME4 (HtxB/hypophosphite) and 5LV1 (Pm\_PtxB/phosphite). Uniprot accession codes of proteins described in this study are given in Supplementary Table 3. The authors declare that the data supporting the findings of this study are available from the corresponding author upon reasonable request.

Received: 27 January 2017 Accepted: 30 August 2017

Published online: 23 November 2017

## References

- Karl, D. M. Aquatic ecology: phosphorus, the staff of life. *Nature* **406**, 31–33 (2000).
- White, A. K. & Metcalf, W. W. Microbial metabolism of reduced phosphorus compounds. *Annu. Rev. Microbiol.* **61**, 379–400 (2007).
- Karl, D. M. Microbially mediated transformations of phosphorus in the sea: new views of an old cycle. *Ann. Rev. Mar. Sci.* **6**, 279–337 (2014).
- Wu, J., Sunda, W., Boyle, E. A. & Karl, D. M. Phosphate depletion in the western North Atlantic Ocean. *Science* **289**, 759–762 (2000).
- Sañudo-Wilhelmy, S. A. et al. Phosphorus limitation of nitrogen fixation by Trichodesmium in the central Atlantic Ocean. *Nature* **411**, 66–69 (2001).
- Quinn, J. P., Kulakova, A. N., Cooley, N. A. & McGrath, J. W. New ways to break an old bond: the bacterial carbon-phosphorus hydrolases and their role in biogeochemical phosphorus cycling. *Environ. Microbiol.* **9**, 2392–2400 (2007).
- Stone, B. L. & White, A. K. Most probable number quantification of hypophosphite and phosphite oxidizing bacteria in natural aquatic and terrestrial environments. *Arch. Microbiol.* **194**, 223–228 (2012).
- Villarreal-Chiu, J. F., Quinn, J. P. & McGrath, J. W. The genes and enzymes of phosphonate metabolism by bacteria, and their distribution in the marine environment. *Front. Microbiol.* **3**, 19 (2012).
- McGrath, J. W., Chin, J. P. & Quinn, J. P. Organophosphonates revealed: new insights into the microbial metabolism of ancient molecules. *Nat. Rev. Microbiol.* **11**, 412–419 (2013).
- Horsman, G. P. & Zechel, D. L. Phosphonate biochemistry. *Chem. Rev.* **117**, 5704–5783 (2017).
- Dyhrman, S. T. et al. Phosphonate utilization by the globally important marine diazotroph Trichodesmium. *Nature* **439**, 68–71 (2006).
- Martínez, A., Osburne, M. S., Sharma, A. K., DeLong, E. F. & Chisholm, S. W. Phosphite utilization by the marine picocyanobacterium Prochlorococcus MIT9301. *Environ. Microbiol.* **14**, 1363–1377 (2012).
- Seweryn, P. et al. Structural insights into the bacterial carbon-phosphorus lyase machinery. *Nature* **525**, 68–72 (2015).
- Alicia, I. et al. Structure of the *Escherichia coli* phosphonate binding protein PhnD and rationally optimized phosphonate biosensors. *J. Mol. Biol.* **414**, 356–369 (2012).
- Rizk, S. S., Cuneo, M. J. & Hellinga, H. W. Identification of cognate ligands for the *Escherichia coli* phnD protein product and engineering of a reagentless fluorescent biosensor for phosphonates. *Protein Sci.* **15**, 1745–1751 (2006).
- Pasek, M. A., Sampson, J. M. & Atlas, Z. Redox chemistry in the phosphorus biogeochemical cycle. *Proc. Natl Acad. Sci. USA* **111**, 15468–15473 (2014).
- Figueroa, I. A. & Coates, J. D. Microbial phosphite oxidation and its potential role in the global phosphorus and carbon cycles. *Adv. Appl. Microbiol.* **98**, 93–117 (2016).
- Metcalf, W. W. & Wolfe, R. S. Molecular genetic analysis of phosphite and hypophosphite oxidation by *Pseudomonas stutzeri* WM88. *J. Bacteriol.* **180**, 5547–5558 (1998).
- Wilson, M. M. & Metcalf, W. W. Genetic diversity and horizontal transfer of genes involved in oxidation of reduced phosphorus compounds by *Alcaligenes faecalis* WM2072. *Appl. Environ. Microbiol.* **71**, 290–296 (2005).
- Polyviou, D., Hitchcock, A., Baylay, A. J., Moore, C. M. & Bibby, T. S. Phosphite utilization by the globally important marine diazotroph Trichodesmium. *Environ. Microbiol. Rep.* **7**, 824–830 (2015).
- Schink, B. & Friedrich, M. Bacterial metabolism: phosphite oxidation by sulphate reduction. *Nature* **406**, 37 (2000).
- Simeonova, D. D., Wilson, M. M., Metcalf, W. W. & Schink, B. Identification and heterologous expression of genes involved in anaerobic dissimilatory phosphite oxidation by *Desulfotignum phosphitoxidans*. *J. Bacteriol.* **192**, 5237–5244 (2010).
- Costas, A. M., White, A. K. & Metcalf, W. W. Purification and characterization of a novel phosphorus-oxidizing enzyme from *Pseudomonas stutzeri* WM88. *J. Biol. Chem.* **276**, 429–436 (2001).
- Kanda, K. et al. Application of a phosphite dehydrogenase gene as a novel dominant selection marker for yeasts. *J. Biotechnol.* **20**, 68–73 (2014).
- López-Arredondo, D. L. & Herrera-Estrella, L. A novel dominant selectable system for the selection of transgenic plants under in vitro and greenhouse conditions based on phosphite metabolism. *Plant Biotechnol. J.* **11**, 516–525 (2013).
- Loera-Quezada, M. M. et al. A novel genetic engineering platform for the effective management of biological contaminants for the production of microalgae. *Plant Biotechnol. J.* **10**, 2066–2076 (2016).
- Shaw, A. J. et al. Metabolic engineering of microbial competitive advantage for industrial fermentation processes. *Science* **353**, 583–586 (2016).
- López-Arredondo, D. L. & Herrera-Estrella, L. Engineering phosphorus metabolism in plants to produce a dual fertilization and weed control system. *Nat. Biotechnol.* **30**, 889–893 (2012).
- Relyea, H. A. & van der Donk, W. A. Mechanism and applications of phosphite dehydrogenase. *Bioorg. Chem.* **33**, 171–189 (2005).
- White, A. K. & Metcalf, W. W. Two C-P lyase operons in *Pseudomonas stutzeri* and their roles in the oxidation of phosphonates, phosphite, and hypophosphite. *J. Bacteriol.* **186**, 4730–4739 (2004).
- Wilson, M. M. *Characterization of Reduced Phosphorus Metabolism in Alcaligenes faecalis WM2072 and Xanthobacter flavus WM2814*. PhD thesis, University of Illinois, Urbana-Champaign, IL (2006).
- White, A. K. & Metcalf, W. W. Isolation and biochemical characterization of hypophosphite/2-oxoglutarate dioxygenase. A novel phosphorus-oxidizing enzyme from *Pseudomonas stutzeri* WM88. *J. Biol. Chem.* **277**, 38262–38271 (2002).

33. Vrtis, J. M., White, A. K., Metcalf, W. W. & van der Donk, W. A. Phosphite dehydrogenase: an unusual phosphoryl transfer reaction. *J. Am. Chem. Soc.* **123**, 2672–2673 (2001).
34. Hirota, R. et al. A novel biocontainment strategy makes bacterial growth and survival dependent on phosphite. *Sci. Rep.* **7**, 44748 (2017).
35. Feingersch, R. et al. Potential for phosphite and phosphonate utilization by *Prochlorococcus*. *ISME J.* **6**, 827–834 (2012).
36. Wienken, C. J., Baaske, P., Rothbauer, U., Braun, D. & Duhr, S. Protein-binding assays in biological liquids using microscale thermophoresis. *Nat. Commun.* **1**, 100 (2010).
37. Lata, S., Reichel, A., Brock, R., Tampé, R. & Piehler, J. High-affinity adaptors for switchable recognition of histidine-tagged proteins. *J. Am. Chem. Soc.* **127**, 10205–10215 (2005).
38. Fukami-Kobayashi, K., Tateno, Y. & Nishikawa, K. Domain dislocation: a change of core structure in periplasmic binding proteins in their evolutionary history. *J. Mol. Biol.* **286**, 279–290 (1999).
39. Troev, K. D. *Chemistry and Application of H-Phosphonates* (Elsevier, 2006).
40. Meyer, E. A., Castellano, R. K. & Diederich, F. Interactions with aromatic rings in chemical and biological recognition. *Angew. Chem. Int. Ed. Engl.* **42**, 1210–1250 (2003).
41. Nishiro, M. CH/π hydrogen bonds in crystals. *Cryst. Eng. Comm.* **6**, 130–158 (2004).
42. Hayward, S. & Berendsen, H. J. C. Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme. *Proteins* **30**, 144–154 (1998).
43. Martiny, A. C., Coleman, M. L. & Chisholm, S. W. Phosphate acquisition genes in *Prochlorococcus* ecotypes: evidence for genome-wide adaptation. *Proc. Natl Acad. Sci. USA* **103**, 12552–12557 (2006).
44. Martínez, A., Tyson, G. W. & Delong, E. F. Widespread known and novel phosphonate utilization pathways in marine bacteria revealed by functional screening and metagenomic analyses. *Environ. Microbiol.* **12**, 222–238 (2010).
45. McSorley, F. R. et al. PhnY and PhnZ comprise a new oxidative pathway for enzymatic cleavage of a carbon-phosphorus bond. *J. Am. Chem. Soc.* **134**, 8364–8367 (2012).
46. Ilikchyan, I. N., McKay, R. M., Zehr, J. P., Dyhrman, S. T. & Bullerjahn, G. S. Detection and expression of the phosphonate transporter gene *phnD* in marine and freshwater picocyanobacteria. *Environ. Microbiol.* **11**, 1314–1324 (2009).
47. Steiner, T. & Koellner, G. Hydrogen bonds with pi-acceptors in proteins: frequencies and role in stabilizing local 3D structures. *J. Mol. Biol.* **305**, 535–557 (2001).
48. Ma, J. C. & Dougherty, D. A. The cation–π interaction. *Chem. Rev.* **97**, 1303–1324 (1997).
49. Brandl, M., Weiss, M. S., Jabs, A., Sühnel, J. & Hilgenfeld, R. C-H...π-interactions in proteins. *J. Mol. Biol.* **307**, 357–377 (2001).
50. Baker, E. N. & Hubbard, R. E. Hydrogen bonding in globular proteins. *Prog. Biophys. Mol. Biol.* **44**, 97–179 (1984).
51. Van Mooy, B. A. et al. Phosphorus cycling. Major role of planktonic phosphate reduction in the marine phosphorus redox cycle. *Science* **348**, 783–785 (2015).
52. Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* **8**, 785–786 (2011).
53. Prufert-Bebout, L., Paerl, H. W. & Lassen, C. Growth, nitrogen fixation and spectral attenuation in cultivated *Trichodesmium* species. *Appl. Environ. Microbiol.* **59**, 1367–1375 (1993).
54. Winter, G. xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* **43**, 186–190 (2010).
55. Keegan, R. M. & Winn, M. D. Automated search-model discovery and preparation for structure solution by molecular replacement. *Acta Crystallogr. D* **63**, 447–457 (2007).
56. McCoy, A. J. et al. *Phaser* crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
57. Vagin, A. A. & Isupov, M. N. Spherically averaged phased translation function and its application to the search for molecules and fragments in electron-density maps. *Acta Crystallogr. D* **57**, 1451–1456 (2001).
58. Cowtan, K. The Buccaneer software for automated model building. *Acta Crystallogr. D* **62**, 1002–1011 (2006).
59. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
60. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**, 240–255 (1997).
61. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010).
62. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
63. Vaguine, A. A., Richelle, J. & Wodak, S. J. SFCHECK: a unified set of procedure for evaluating the quality of macromolecular structure-factor data and their agreement with atomic model. *Acta Crystallogr. D* **55**, 191–205 (1999).
64. Lebedev, A. A. et al. JLigand: a graphical tool for the CCP4 template-restraint library. *Acta Crystallogr. D* **68**, 431–440 (2012).
65. Krissinel, E. & Henrick, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D* **60**, 2256–2268 (2004).
66. Dundas, J. et al. CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Res.* **34**, 116–W118 (2006).
67. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
68. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).
69. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
70. Larkin, M. A. et al. Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).
71. Kearse, M. et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647–1649 (2012).

## Acknowledgements

This work was supported by grants BB/M000265/1 (to C.N.H.) and BB/M011305/1 (to T.S.B.) from the Biotechnology and Biological Sciences Research Council UK and grant 338895 (to C.N.H.) from the European Research Council. This work was also supported in part by the European Community's Seventh Framework Program (FP7/2007–2013) under Grant Agreement 283570 (BioStruct-X) and the Photosynthetic Antenna Research Center (PARC), an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences (Award number DE-SC0001035). PARC provided partial support for C.N.H. and C.B. D.P. acknowledges a PhD studentship jointly funded by the Natural Environment Research Council and the A.G. Leventis Foundation. We thank the Diamond Light Source for beam time under BAG applications MX-8987 and MX-12788, and the beamline scientists on i03, i04, i04-1 and i24 for assistance with data collection. The authors also thank Professor G. Michael Blackburn and Dr Jack W. Chidgey (University of Sheffield), Dr Russell Viner (Syngenta), Professor C. Mark Moore (University of Southampton) and Dr Benjamin Van Mooy (Woods Hole Oceanographic Institution) for useful discussions relating to this work.

## Author contributions

A.H. and T.S.B. conceived the study. C.B., N.B.P.A., C.N.H. and A.H. designed the experiments. D.P. and A.H. generated expression plasmids. N.B.P.A., B.S., A.A.B. and A.H. purified proteins. N.B.P.A. acquired and interpreted binding data and did CD spectroscopy. C.B., B.S. and A.H. prepared crystals. C.B. collected and processed diffraction data, determined structures and performed structural analysis. B.S. assisted with binding data acquisition, diffraction data processing and structure determination for Ps\_PtxB. P.J.B. performed structural analysis. A.H. performed bioinformatics. C.B. and A.H. wrote the paper. N.B.P.A., T.S.B., P.J.B. and C.N.H. edited and proofread the manuscript. All authors read and approved the final paper.

## Additional information

**Supplementary Information** accompanies this paper at doi:10.1038/s41467-017-01226-8.

**Competing interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017