Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Refined structures of placental alkaline phosphatase show a consistent pattern of interactions at the peripheral site

Boguslaw Stec,[a]* Anton Cheltsov[b] and José Luis Millán[a]

[a]Sanford–Burnham Medical Research Institute, La Jolla, CA 92037, USA, and [b]Q-MOL LLC, San Diego, CA 92105, USA

Correspondence e-mail: bstec@burnham.org

Received 21 March 2010
Accepted 26 May 2010

PDB References: placental alkaline phosphatase complexed with three putative regulators, 3mk0; 3mk1; 3mk2.

In order to gain deeper insights into the functional sites of human placental alkaline phosphatase, the structures of the enzyme with the putative regulators L-Phe, pNPP and 5′-AMP [Llinas *et al.* (2005), *J. Mol. Biol.* **350**, 441–451] were re-refined. Significant variations in ligand positioning and identity were found compared with the previous report. The multiple corrections to the model improved the phases and the electron-density maps, allowing the modeling of omitted side chains and multiple disordered residues. These improvements led to a change in the position of L-Phe at the peripheral binding site, which appeared to be reversed. The structure with pNPP contained only *p*-nitrophenol in three distinct sites, while the structure with 5′-AMP contained the *p*-nitrophenyl group in two of the sites instead of 5′-AMP. Comparison of the re-refined models shows a consistent pattern of interactions at the peripheral site.

1. Introduction

Alkaline phosphatases (APs; EC 3.1.3.1) are one of the most ubiquitous enzyme families processing phosphate groups (Millán, 2006*a*) and are present in many species, including humans (Millán, 2006*b*). They catalyze the hydrolysis of phosphomonoesters, with the release of inorganic phosphate and an alcohol group. Those APs whose structures have been determined showed a dimeric organization of these metalloenzymes. Enzymes from several species have been solved, such as those from *Escherichia coli*, shrimp, human *etc*. All of them share the same general folding motif of the catalytic core (Stec *et al.*, 2000; de Backer *et al.*, 2002; Le Du *et al.*, 2001).

The three-dimensional structure of the core of the protein is strongly conserved among different species. The active site, which contains two zinc ions and one magnesium ion, is also strongly conserved, suggesting a very similar mode of catalysis. Extensive structural studies of *E. coli* AP have elucidated the role of the catalytic site residues and details of the catalytic mechanism (Stec *et al.*, 2000; Kozlenkov *et al.*, 2002). Thus, the overall catalytic mechanism is conserved between the *E. coli* enzyme and the mammalian enzymes. The catalysis involves the sequential activation of the catalytic serine by a Zn atom, formation of a covalent phosphoserine intermediate, hydrolysis of the phosphoserine by a water molecule activated by the second Zn atom and release of the phosphate product or its transfer to a phosphate acceptor. A recent structure–function study of residues that are preserved between *E. coli* and the mammalian enzymes revealed a conserved function for the residues that stabilize the active-site Zn and Mg ions as well as other preserved residues (Kozlenkov *et al.*, 2002; Hoylaerts *et al.*, 1997).

In humans, there are four AP isozymes (Le Du & Millán, 2002). Three of these, *i.e.* placental AP (PLAP), germ cell AP (GCAP) and intestinal AP (IAP), are tissue-specific enzymes and have relative homologies in the 90–98% range; their genes are clustered on chromosome 2. The fourth (tissue-nonspecific AP; TNAP) is found in bone, liver and kidney. TNAP is only 50% identical to the other three isozymes and its gene is located on chromosome 1. As mentioned above, the catalytic cores of all these enzymes are very similar;

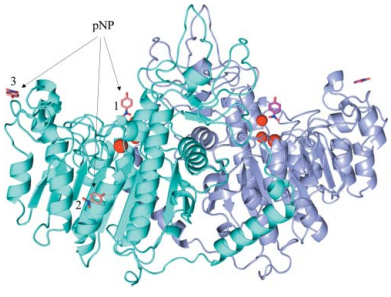


Table 1
Data and refinement statistics for human placental alkaline phosphatase.

Values in parentheses are for the last shell.

Data set (ligand)	L-Phe		PNPP		5′-AMP	
Original PDB code	1zef		1zed		1zeb	
Space group	$C222_1$		$C222_1$		$C222_1$	
Unit-cell parameters						
a (Å)	87.79		88.35		89.09	
b (Å)	114.94		114.57		115.30	
c (Å)	106.34		106.57		107.32	
Data†						
Diffraction limit (Å)	1.89		1.57		1.90	
Resolution (Å)	99.0–1.89 (1.93–1.89)		99.0–1.57 (1.61–1.57)		70.0–1.90 (1.94–1.90)	
No. of reflections	43024		74477		37407	
$\langle I/\sigma(I)\rangle$	25.4 (4.6)		17.4 (1.6)		8.47 (1.4)	
Completeness (%)	98.5 (98.5)		97.7 (80.8)		77.3 (76.8)	
R_{merge}‡	0.066 (0.183)		0.045 (0.411)		0.083 (0.405)	
Refinement	New	Old	New	Old	New	Old
Resolution (Å)	40.0–1.886		40.0–1.573		40.0–1.90	
R factor§	0.125	0.129	0.141	0.151	0.137	0.153
R_{free}¶	0.163	0.173	0.178	0.188	0.189	0.193
Refined elements						
Protein atoms	3692	3699	3840	3673	3850	3677
Missing residues	0	9	0	7	0	7
Disordered residues	0	0	23	0	25	0
Water molecules	559	702	570	761	572	710
Ligands	2 Phe	2 Phe	3 PNP	2 PNPP	2 PNP	—
B factors (Å^2)						
Protein	18.2	20.6	17.2	18.0	18.2	20.5
Water	48.6	41.4	36.7	37.5	46.4	38.8
Active-site ligands	44.7	45.2	53.4	44.8	—	—
Peripheral-site ligand	48.2	48.8	28.7	34.7	35.6	—
R.m.s. deviation from ideal						
Bond lengths (Å)	0.010	0.017	0.009	0.010	0.010	0.010
Bond angles (°)	1.27	1.72	1.23	1.42	1.25	1.34
Chiral angles (°)	0.081	0.175	0.078	0.105	0.082	0.082

† The data section of this table contains the originally reported statistics from Llinas *et al.* (2005). ‡ $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the *i*th observation of reflection *hkl* and $\langle I(hkl)\rangle$ is the mean intensity of all observations of reflection *hkl*. § Crystallographic *R* factor = $\sum_{hkl}||F_{obs}| - |F_{calc}||/\sum_{hkl}|F_{obs}|$, where F_{obs} and F_{calc} are the observed structure-factor amplitudes and the structure-factor amplitudes calculated from the model, respectively. ¶ The free *R* factor was monitored using a randomly selected 5% of data that were excluded from refinement.

however, unique features exist in some of these APs, such as their allosteric properties and their susceptibility to inhibition by L-amino acids, *e.g.* L-Phe, L-Trp, L-homoarginine, L-Leu and levamisole (Hoylaerts *et al.*, 1992; Kozlenkov *et al.*, 2004).

To study the allosteric binding properties (Hoylaerts *et al.*, 1997) of L-Phe, structures of wild-type PLAP as well as of PLAP in complex with L-Phe, pNPP and 5′-AMP were determined (Le Du & Millán, 2002; Llinas *et al.*, 2005). The results of these structural studies showed the existence of a peripheral noncatalytic binding site located ∼30 Å away from the active site. The structures suggest a mode of interaction at this peripheral site that may play a role in stabilizing the vicinity of the Ca^{2+}-binding site and in so doing affect the catalytic site. In support of this hypothesis, mutations in this peripheral site have been linked to hypophosphatasia, a disease of bone mineralization (Whyte, 1995).

2. Experimental

The deposited structure coordinates were retrieved from the PDB (PDB entries 1zeb, 1zed and 1zef; Llinas *et al.*, 2005). These structures were superposed and compared. Some differences were noted, for instance in the region of residue 470. Subsequently, the files containing the structure factors corresponding to the deposited models were retrieved. Both the coordinates as well as the data were imported into the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). The process produced new test sets for R_{free} calculations conforming to the default of 5% of reflections. We used the R_{free} to monitor the progress of our refinement and to provide a guiding direction in trying to avoid overfitting by introducing an excessive number of refinement parameters. The initial rounds of refinement starting from the deposited models were performed in *REFMAC*5 (Murshudov *et al.*, 1999). The numerical values of *R* and R_{free} for the refined models were not significantly different from those for the deposited structures. However, inspection of the difference density maps in *Coot* (Emsley *et al.*, 2010) suggested modestly over-refined structures. A significant number of water molecules were not within any substantial density; additionally, some of the residues produced significant negative difference maps.

In particular, in every model the electron-density maps suggested that around 50 residues would require manual intervention. In the model 1zeb 58 water molecules were located below the 0.8σ contour level in the $2F_o - F_c$ electron density. Similarly, in models 1zed and 1zeb 112 and 48 water molecules were below this level, respectively. Several of the ligand sites as well as solvent sites showed excessive negative electron density. As is well known, errors in positioning the protein atoms are reflected in the positioning of the water peaks; therefore, we removed all solvent before the next round of refinement.

Motivated by these findings and pursuing our goal of investigating the peripheral site in detail, we stripped the entire model of water molecules and solute molecules and initiated a new refinement. The electron-density maps indicated that a number of protein side chains needed some adjustment (particularly Arg and Leu). However, two fragments of the main chain (around residues 60 and 470) indicated a different conformation for the backbone. The introduction of an

initial set of water molecules (∼300) improved the maps so that the multiple side chains that were not present in the previously refined models were localized (∼7). Also, the improved maps allowed modeling of the disordered side chains (∼25) and improvement in the positioning and identification of the bound ligands. Introduction and refinement of the remaining protein side chains led to improved maps that showed clearly visible (in the difference electron densities) water molecules (∼200), finalizing the refinement. The final R factors as well as $R_{\rm free}$ were improved, together with the geometry of the model. The fit to the electron density was improved despite a significantly smaller number of water molecules being present in the final model than in the original models (around 150 less). The final statistics for the refinements and a summary of the data taken from Llinas *et al.* (2005) are presented in Table 1.

3. Results and discussion

We have analyzed the deposited models and analyzed the electron-density maps computed from the retrieved data. The general quality of the models was good, as was the fit to the electron density. We focused our attention on the refined ligands as they showed some anomalies in their temperature factors. The level of temperature factors was higher than in the direct protein vicinity. Closer inspection of the $2F_{\rm o} - F_{\rm c}$ and $F_{\rm o} - F_{\rm c}$ electron-density maps suggested numerous small improvements to the structures that were carried out by modeling multiple disordered residues and correcting conformational errors as described above. An example of such corrections is presented in Fig. 1. The figure shows a difference electron-density map calculated from the original model (1zed). The negative contour showed several carbonyl O atoms requiring corrections. One of the required peptide-bond flips (Glu469-Pro470) resulted in Pro470 attaining the *cis* conformation of the peptide bond. More than 5% of residues in all three structures were modeled as disordered (Table 1).

In the model 1zeb a total of 57 side chains were corrected, including three rearrangements of the backbone geometry. The corrections constituted three classes of amendments. The first class constituted the modeling of residues that were not present in the original model. Seven residues were added and modeled belonging to this category, as well as two units of NAG that were added to single units of already modeled sugars. The second category of amendments constituted the modeling of residues in disordered conformations. We limited ourselves to modeling two alternative positions for 25 amino acids that were mostly charged and located on the surface of the protein. The third category was composed of residues that were modeled with an incorrect conformer. These were mostly Leu, Val and Arg residues. Closer inspection of the water structure in *Coot* indicated that 40 water molecules were located in positions characterized by electron density at less than the 0.8σ level.

The refinement showed improved stereochemical characteristics of the model, with improved electron density. The global change was not detected by the r.m.s.d. measures. The initial and final models showed an r.m.s.d. of 0.05 Å^2 calculated for all C$^\alpha$ atoms and of less than 1 Å^2 for corrected residues. Finally, around 70% of the water molecules refined in our model were located less than 1 Å from their original positions.

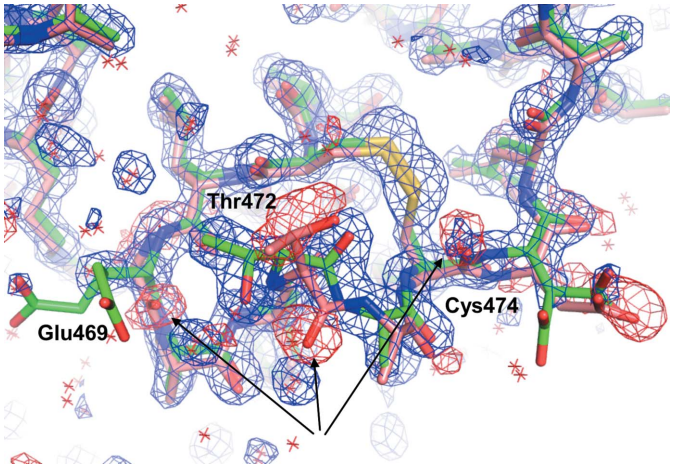
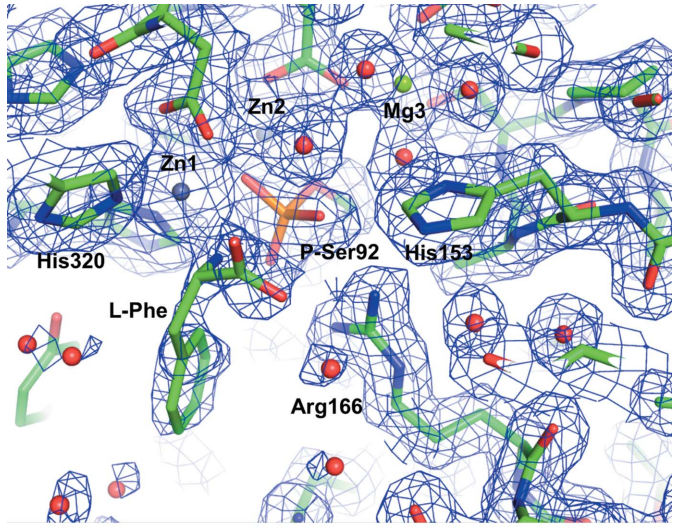
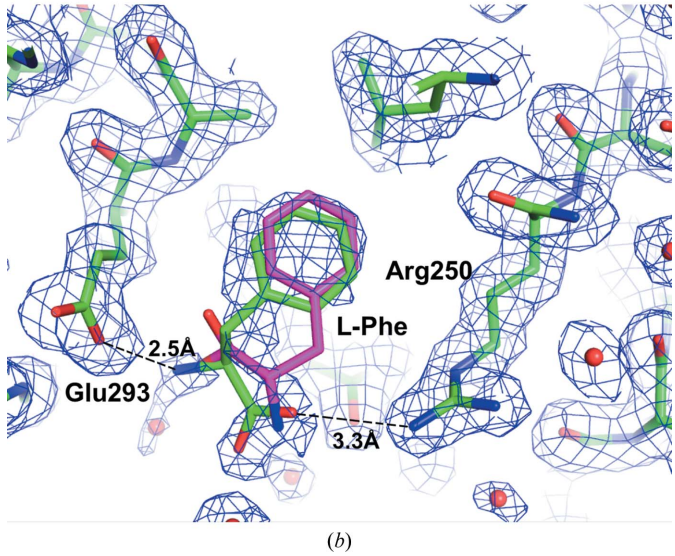


Figure 1
The $2F_{\rm o} - F_{\rm c}$ electron density for both models of PLAP with pNPP bound (1zed): original model in mauve and the model presented in this paper (newly refined) in green. The difference electron density contoured at the -3σ level is depicted by red contours and indicated that three peptide bonds should be flipped (369, 372 and 374). The reversed peptide 369 creates a *cis*-Pro residue. The arrows indicate the three peptide bonds that were flipped in the process of refinement.



(*a*)



(*b*)

Figure 2
The $2F_{\rm o} - F_{\rm c}$ electron-density map contoured at the 1σ level phased with the refined model (originating from 1zef). (*a*) The L-Phe bound at the active site is in a position very close to the original position present in 1zef. (*b*) The L-Phe bound at peripheral site 2. The map shows that the L-Phe is reversed compared with its original position (in magenta) as suggested by the pattern of donor–acceptor relationships (Arg250 and Glu293).

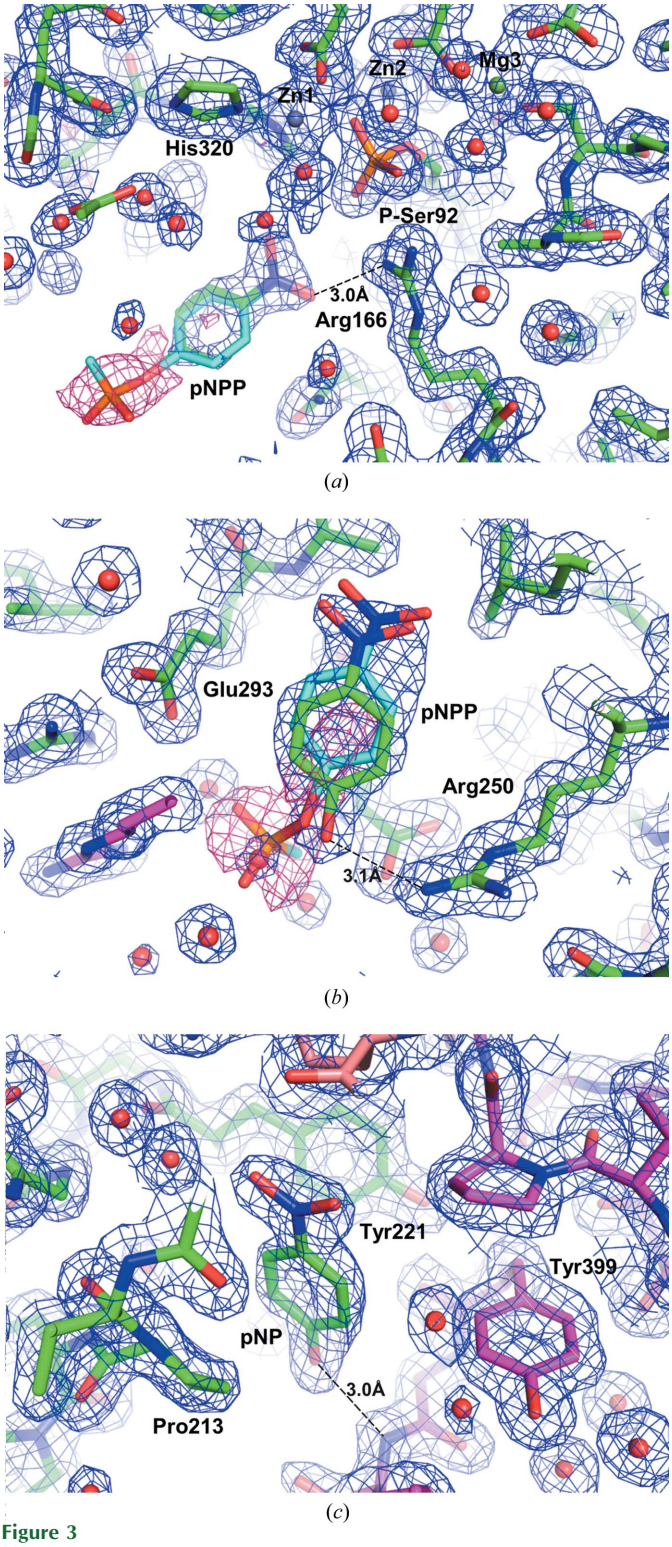


Figure 3
The final refined model covered by a $2F_o - F_c$ electron-density map contoured at the 1σ level phased with the presently refined model (originating from 1zed). The difference electron density contoured at the -4σ level phased with the original model is shown in red. The map shows that the phosphate group of pNPP has been hydrolyzed and the bound species is a *p*-nitrophenol molecule (pNP). The blue model represents the original position of pNPP as present in 1zed. (*a*) The model of the active site with pNPP bound. The gray spheres represent Zn^{2+} ions, while the green sphere represents the Mg^{2+} ion. Ser104 has a covalently attached phosphate representing an active intermediate in the catalysis. (*b*) The peripheral site 2 near Arg250 with bound pNP molecule. (*c*) The third *p*-nitrophenol-binding site near the interface between three molecules.

Similar results were obtained with the two remaining models. Somewhat unexpected were the errors encountered in the 1zeb structure. In this structure the conformers of several His-coordinating metal ions or water molecules were incorrectly modeled with C atoms in coordinating positions. The global improvement in all the models was associated with the correction of a limited number of residues that subsequently influenced a substantial number of neighboring residues and atoms, leading to a globally improved stereochemistry, and resulted in improvements in electron density as well as in the crystallographic *R* factor.

Analysis of the active site of all the complexes confirmed the conclusion of the original study regarding the mode of binding by pNPP and L-Phe. The model of L-Phe at the active site was practically isostructural with the original structure (Fig. 2*a*). The most interesting observation was that the L-Phe molecule bound at the peripheral site did not refine well. The presence of multiple difference peaks suggested that while the phenyl ring was positioned correctly, the amino and carboxy groups were not. After removing the original L-Phe from the model using the difference electron density, we were able to properly position the L-Phe into the resulting $2F_o - F_c$ electron-density map. The temperature factors of both the moieties bound at the active site and at the peripheral site suggested that the L-Phe was bound with fractional rather than full occupancy. The occupancy was estimated to be no higher than 80%.

As mentioned above, the general location of the pNPP molecules in the 1zed model was similar. However, the strongly negative difference electron density covering the phosphate group indicated that instead of pNPP only the product *p*-nitrophenol was bound at the active site (Fig. 3*a*). The same observation was made for the peripheral binding site 2. Here also only a *p*-nitrophenol molecule was bound. Encouraged by these improvements, we inspected the entire molecule and found an additional molecule of *p*-nitrophenol bound at the interface of three symmetry-related molecules (Figs. 3*c* and 4).

The largest discrepancy we encountered was in the model with 5′-AMP. Our attempts at refining the ligand were unsuccessful. In every attempt we obtained non-optimal AMP conformations with close contacts to the protein and extensive difference maps. Addi-

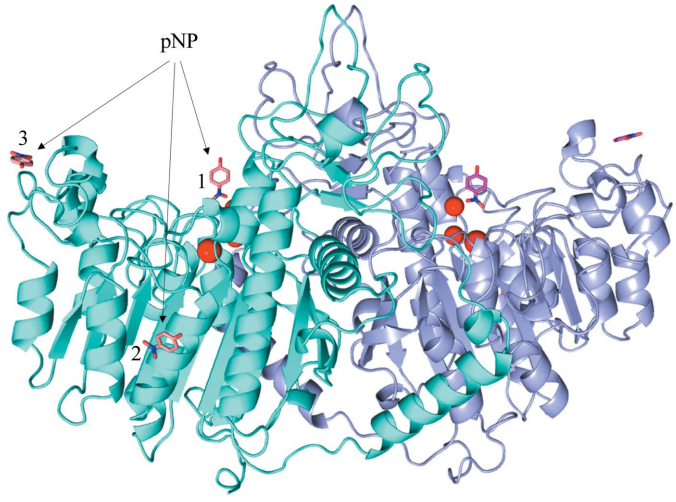


Figure 4
Structure of the PLAP dimer (derived from the original 1zed structure). Red spheres indicate metal ions bound at the active site. Three molecules of *p*-nitrophenol are bound at each subunit. The active site (site 1) and the peripheral site 2 have been described previously. An additional peripheral site 3 is found at the edge of the molecule, wedged between the symmetry-related molecules, as seen in Fig. 3(*c*).

tionally, strong negative difference densities suggested that the phosphate group as well as the sugar moieties were not present. After computing the complete OMIT map, we concluded that the most likely molecule to be bound at this site was *p*-nitrophenol (pNP). A model with this ligand at the peripheral site 2 and the additional peripheral site 3 located previously in the pNP model refined well (Figs. 3 and 4). The structure was very similar to that with pNP, with the exception that the pNP molecule located at the active site had a much lower occupancy than the remainder of the sites and was not refined.

The presence of the pNP molecule at the peripheral site 2 as well as a reversed model for the L-Phe fully explain and conform to the mode of the ionic interactions observed at this site. L-Phe forms proper hydrogen bonds from its carboxy group to the guanidine group of Arg250 and from its amino group to the carboxy group of Glu293. The hydroxyl group of the pNP molecule at the peripheral site 2 forms a similar bond to Arg250 as the L-Phe. Both moieties (pNP and L-Phe) occupy a similar volume and suggest that the site is important for the regulation of wild-type enzyme activity. It suggests that this might be the position of protein–protein interaction and that the most likely amino acid interacting with this site might be a large aromatic residue such as Phe or Trp.

In order to test our hypothesis, we carried out an en masse docking experiment testing the entire NCI database of small molecules. The result of this experiment confirmed the preference for this site to receive larger hydrophobic entities such as the five compounds designated by NSC codes 34412, 633364, 716295, 719716 and 732317 (for details, see Supplementary Figs. S1 and S2[1]). The selected compounds shown in Supplementary Fig. S2 represented the lowest pseudo-energy solutions from our screening process. The general structure of the compounds and the general location of the peripheral site again suggest that it might constitute a protein–protein interaction site.

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: GX5162).

References

Backer, M. de, McSweeney, S., Rasmussen, H. B., Riise, B. W., Lindley, P. & Hough, E. (2002). *J. Mol. Biol.* **318**, 1265–1274.

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

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.

Hoylaerts, M. F., Manes, T. & Millán, J. L. (1992). *Biochem. J.* **286**, 23–30.

Hoylaerts, M. F., Manes, T. & Millán, J. L. (1997). *J. Biol. Chem.* **272**, 22781–22787.

Kozlenkov, A., Hoylaerts, M. F., Ny, T., Le Du, M.-H. & Millán, J. L. (2004). *J. Bone Miner. Res.* **19**, 1862–1872.

Kozlenkov, A., Manes, T., Hoylaerts, M. F. & Millán, J. L. (2002). *J. Biol. Chem.* **27**, 22992–22999.

Le Du, M.-H. & Millán, J. L. (2002). *J. Biol. Chem.* **277**, 49808–49814.

Le Du, M.-H., Stigbrand, T., Taussig, M. J., Menez, A. & Stura, E. A. (2001). *J. Biol. Chem.* **276**, 9158–9165.

Llinas, P., Stura, E. A., Ménez, A., Kiss, Z., Stigbrand, T., Millán, J. L. & Le Du, M.-H. (2005). *J. Mol. Biol.* **350**, 441–451.

Millán, J. L. (2006*a*). *Purinergic Signal.* **2l**, 335–341.

Millán, J. L. (2006*b*). *Mammalian Alkaline Phosphatases: From Biology to Applications in Medicine and Biotechnology*, pp. 1–322. Weinheim: Wiley–VCH.

Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S. & Dodson, E. J. (1999). *Acta Cryst.* D**55**, 247–255.

Stec, B., Holtz, K. M. & Kantrowitz, E. R. (2000). *J. Mol. Biol.* **299**, 1303–1311.

Whyte, M. P. (1995). *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., edited by C. R. Scriver, A. L. Beaudet, W. S. Sly & D. Valle, pp. 4095–4112. New York: McGraw–Hill.