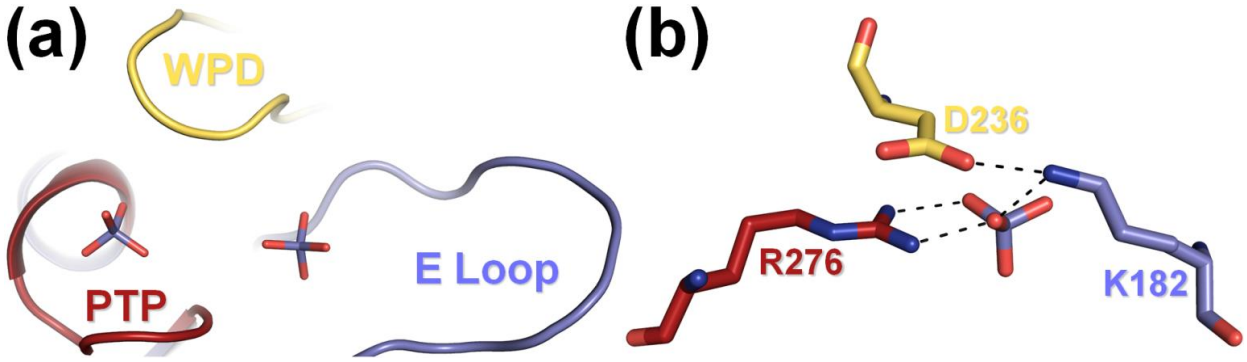
## Supplementary Data

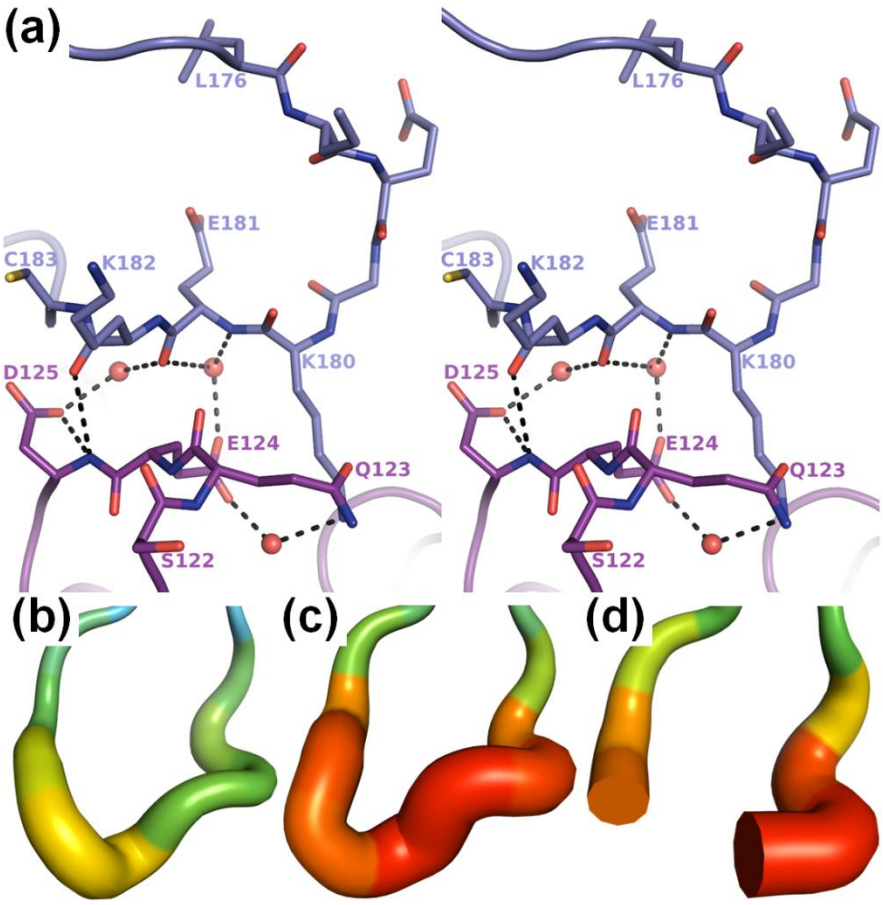
# Visualizing Active Site Dynamics in Single Crystals of HePTP: Opening of the WPD Loop Involves Coordinated Movement of the E Loop

**David A. Critton[a], Lutz Tautz[b] and Rebecca Page[a]***

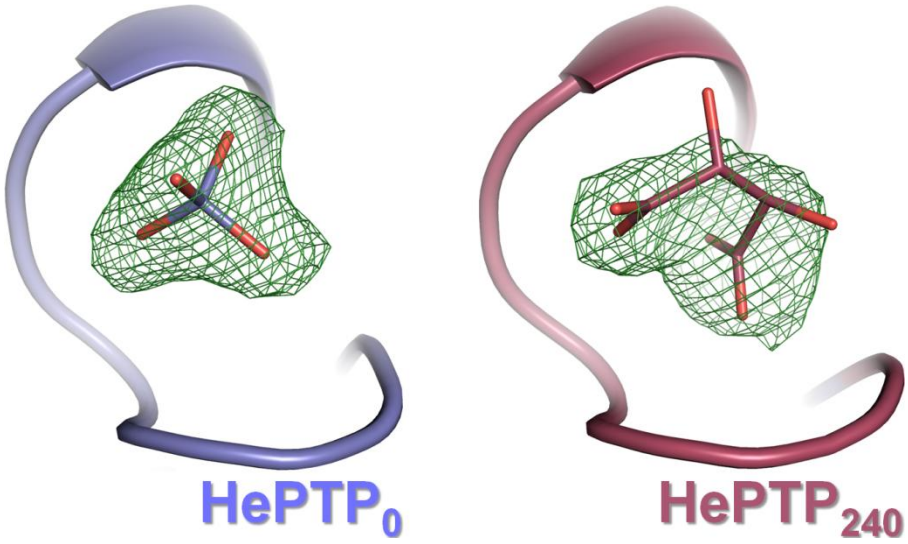
[a]Department of Molecular Biology, Cell Biology and Biochemistry**,** Brown University, Providence, RI 02912, USA, [b]Infectious and Inflammatory Disease Center, Sanford-Burnham Medical Research Institute, La Jolla, CA 92037, USA



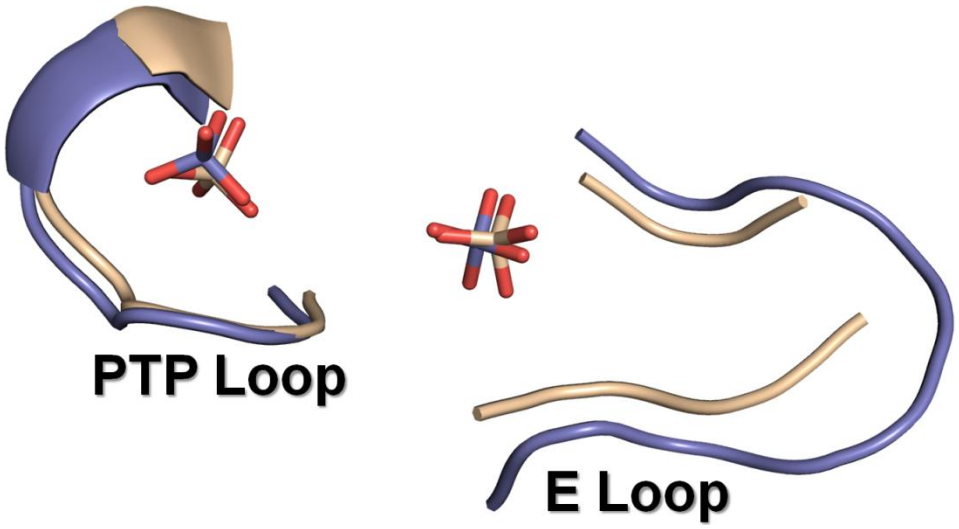

**Supplementary Figure S1.** HePTP binds tetrahedral oxyanions at the active site and at the E loop phosphate-binding site. (a) Position of the bound sulfate molecules (shown in stick representation) relative to the PTP loop (red), WPD loop (yellow) and E loop (blue) in the structure of $HePTP_0$. (b) The E loop sulfate binds to HePTP via multiple highly conserved residues. In the structure of $HePTP_0$, the E loop sulfate forms a bipartite hydrogen bonding interaction with the 100% conserved PTP loop residue Arg276 (red), and is also coordinated by the 80% conserved E loop residue Lys182 (blue). The catalytic Asp236 of the WPD loop (yellow) also hydrogen bonds with Lys182. Thus, the PTP, WPD and E loops form a network of interactions in HePTP. Protein models prepared using PyMOL.[1]



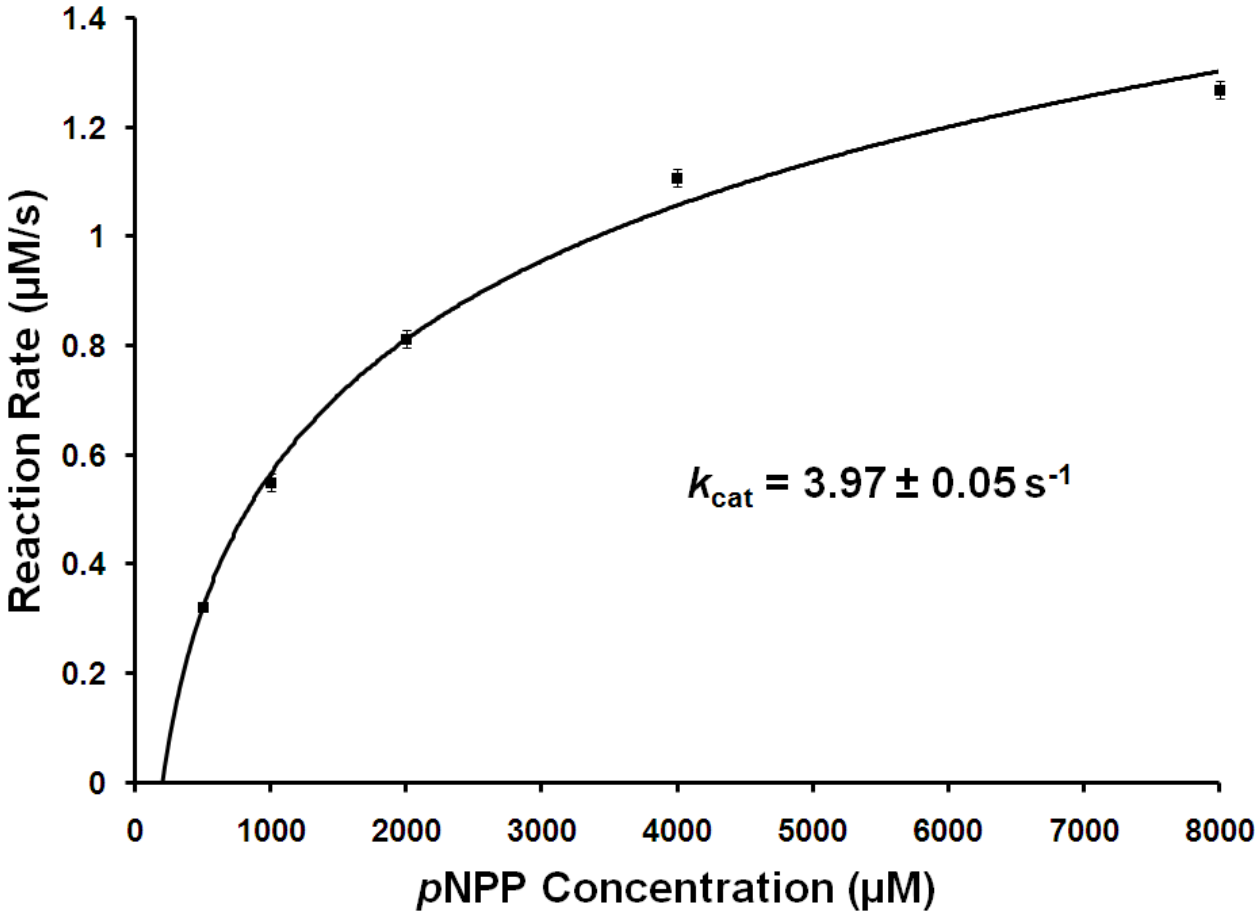

**Supplementary Figure S2.** The E loop interacts at crystal contacts in undepleted crystals, and disordering of the E loop in sulfate-depleted crystals destabilizes symmetry mate residues. (a) The E loop interacts at crystal contacts in undepleted crystals (i.e. $HePTP_0$). Stereo image of the interaction of E loop residues (blue, stick representation) with symmetry mate residues 122–125 (purple, stick representation) and water molecules (red, sphere representation). Hydrogen bonds and polar contacts between E loop residues and symmetry mate residues are illustrated by dashed lines. (b-d) Ordered-to-disordered transition of residues 122–125 in sulfate-depleted HePTP crystals. B-factor putty representation of residues 122–125 in $HePTP_0$ (b), $HePTP_{24}$ (c) and $HePTP_{240}$ (d). B-factors (from lowest-to-highest) correspond to color scale (from blue-to-red) and thickness (from thin-to-thick). Protein models prepared using PyMOL.[1]




**Supplementary Figure S3.** Exchange of active site-bound sulfate with tartrate in soaked HePTP crystals. m$F_O$-D$F_C$ electron density maps (green meshes) for bound sulfate in $HePTP_0$ (left; contoured at 3.0 σ to 1.90 Å) and bound tartrate in $HePTP_{240}$ (right; contoured at 3.0 σ to 2.25 Å). Protein models prepared using PyMOL.[1]




**Supplementary Figure S4.** Both HePTP and PTP1B contain an E loop oxyanion-binding pocket. Superposition of the structures of $HePTP_0$ (blue) and PTP1B (beige; PDB ID: **2B4S**). Tetrahedral oxyanion (e.g. sulfate) is bound at both the active site (e.g. PTP loop) and at the E loop binding site.




**Supplementary Figure S5.** The K182A mutation inhibits HePTP catalytic activity. Michaelis-Menten plot of reaction rate versus substrate concentration for the HePTP K182A-catalyzed dephosphorylation of the general PTP substrate *para*-nitrophenyl phosphate (*p*NPP), which was assayed in triplicate. The turnover number ($k_{cat}$) of HePTP K182A (i.e. 3.97 $s^{-1}$) is less than one third that of WT HePTP (i.e. 12.55 $s^{-1}$).[3]

**Supplementary Table S1**
E Loop Conformation in Different Human PTP Structures.

| Protein | PDB ID(s) | E Loop Conformation |
|---|---|---|
| HePTP | 2A3K, 1ZC0, 2GP0, 2HVL, 2QDM, 2QDC, 3D42, 3D44 | Disordered |
| STEP | 2BIJ, 2CJZ | Disordered |
| STEP | 2BV5 | Helical |
| PCPTP1 | 1JLN, 2A8B | β-hairpin |
| CD45 | 1YGR, 1YGU | β-hairpin |
| RPTPκ | 2C7S | β-hairpin |
| RPTPμ | 1RPM | β-hairpin |
| RPTPρ | 2OOQ | β-hairpin |
| LAR | 1LAR | β-hairpin |
| RPTPσ | 2FH7 | β-hairpin |
| RPTPγ | 2NLK | β-hairpin |
| GLEPP | 2G59, 2GJT | β-hairpin |
| RPTPβ | 2HC1, 2HC2, 2I3R, 2I4E, 2I4G, 2I4H, 2I5X | β-hairpin |
| RPTPβ | 2I3U | Disordered |
| RPTPη | 2CFV, 2NZ6 | Disordered |
| PTP1B | 1BZH, 1G1G, 1G7F, 1NO6, 1NWL, 1NZ7, 1OEM, 1OEO, 1OES, 1OET, 1OEU, 1OEV, 1ONZ, 1PH0, 1PTT, 1PTU, 1PYN, 1T48, 1T49, 1T4J, 2F6F, 2HNP, 2HNQ | β-hairpin |
| PTP1B | 2B4S, 2FJM, 2FJN | Disordered |
| TCPTP | 1L8K | Extended |
| PTPH1 | 2B49 | β-hairpin |
| PTP-MEG1 | 2I75 | β-hairpin |
| PTP-BAS | 1WCH | β-hairpin |
| Shp1 | 1GWZ, 2B3O, 2OC3 | β-hairpin |
| Shp1 | 1FPR | Extended |
| Shp2 | 2SHP, 3B7O | β-hairpin |
| PTP-MEG2 | 2PA5 | β-hairpin |
| Lyp | 2QCJ, 2QCT, 3BRH, 3H2X | β-hairpin |

**Supplementary Table S2**
Data Collection and Refinement Statistics.

| | HePTP C270S* |
|---|---|
| PDB ID | **2QDP** |
| Data Collection | |
| Space group | $P6_1$ |
| Unit cell | |
| a, b, c (Å) | 127.1, 127.1, 60.5 |
| α, β, γ (°) | 90.0, 90.0, 120.0 |
| Wavelength (Å) | 0.9795 |
| Resolution (Å) | 50.0–2.72 (2.82–2.72)[a] |
| No. protein molecules/ASU | 1 |
| Total/unique reflections | 63604/15103 |
| Redundancy | 4.2 (4.1)[a] |
| Completeness (%) | 99.2 (98.4)[a] |
| $R_{merge}$ (%)[b] | 12.2 (60.1)[a] |
| Mean $I/\sigma(I)$ | 9.2 (2.2)[a] |
| Refinement | |
| Resolution range | 20.0–2.72 |
| No. reflections (total) | 14314 |
| No. reflections (test) | 752 |
| $R_{work}$ (%)[c] | 15.7 |
| $R_{free}$ (%)[d] | 22.7 |
| RMS deviations from ideal geometry | |
| Bonds (Å) | 0.020 |
| Angles (°) | 2.00 |
| Ramachandran plot | |
| Residues in allowed regions (%) | 99.3 |
| Residues in disallowed regions (%) | 0.7 |
| Mean B Value | |
| Protein | |
| Total | 29.0 |
| Active Site[e] | 19.8 |
| Water | |
| Active Site Phosphate | 19.0 |
| No. Atoms | |
| Protein | 2322 |
| Water | 132 |
| Phosphate | 1 |

[a]Values in parentheses are for the highest resolution shell.
[b]$R_{merge} = \Sigma|I_i-\langle I_i\rangle|/\Sigma|I_i|$ where $I_i$ is the scaled intensity of the $i^{th}$ measurement, and $\langle I_i\rangle$ is the mean intensity for that reflection.
[c]$R_{work} = \Sigma||F_{obs}|-|F_{calc}||/\Sigma|F_{obs}|$ where $F_{calc}$ and $F_{obs}$ are the calculated and observed structure factor amplitudes, respectively.
[d]$R_{free}$ = as for $R_{work}$, but for 5.0% of the total reflections chosen at random and omitted from refinement.
[e]Calculated for residues 270–276 of the HePTP PTP loop.
[f]Calculated for ligand bound at the HePTP PTP loop.

## Supplementary Materials and Methods

### Cloning, Expression and Purification

HePTP (residues 44–339) containing either the S72D mutation, the C270S mutation or the K182A mutation was subcloned into a derivative of the pET28a bacterial expression vector (Novagen) containing an N-terminal expression and hexahistidine purification tag (MGSDKIHHHHHH).[2] The constructs were verified by sequencing (SeqWright). The expression plasmids was transformed into BL21-CodonPlus (DE3)-RIL (Stratagene) cells and expression carried out in LB medium containing kanamycin. Cell cultures were grown at 37°C with vigorous shaking to an $OD_{600}$ of 0.8, at which point the cells were incubated on ice for 1 hour. Expression was induced by the addition of 1 mM IPTG (final concentration) and the cultures grown for an additional 18 hours at 18°C with vigorous shaking. The cells were harvested by centrifugation and resuspended in extraction buffer (50 mM Tris pH 8.0, 500 mM NaCl, 5 mM imidazole, 0.1% Triton X-100, and EDTA-free protease inhibitor tablets, Roche) and lysed by high pressure cell homogenization (Avestin C3 Emulsiflex). The cell debris was removed by centrifugation (35,000 g/30 minutes/4°C). The filtered supernatant was loaded onto a HisTrap HP column (GE Healthcare) equilibrated with 50 mM Tris pH 8.0, 5 mM imidazole and 500 mM NaCl and eluted with a 5–300 mM imidazole gradient. Fractions corresponding to purified HePTP were pooled, concentrated and further purified using size exclusion chromatography (Superdex75 26/60, GE Healthcare) equilibrated in protein buffer (10 mM Tris pH 7.8, 100 mM NaCl, 0.5 mM TCEP). Fractions corresponding to monomeric HePTP were pooled, concentrated, and frozen in liquid nitrogen and stored at −80°C until needed. The presence of the S72D mutation was verified by electrospray ionization (ESI)-mass spectrometry (MS).

### Crystallization, Data Collection and Structure Determination for $HePTP_{44–339}$ C270S

$HePTP_{44–339}$ C270S crystallized 0.07 M ammonium acetate, 12% (w/v) polyethylene glycol (PEG) 10,000, 0.07 M Bis-Tris pH 5.5 using the sitting drop vapor diffusion method at 4°C. A peptide corresponding to Erk2 residues 182–189 ($^{182}$LTEYVATR$^{189}$), phosphorylated at residues T183 and Y185, was also present during crystallization. These crystals did not contain bound peptide, and instead contained a phosphate molecule bound at the active site. We determined that free phosphate appeared in the Erk2 peptide solution over the course of time required for crystal formation (2–14 days) using the malachite green assay (data not shown). Because we previously determined that $HePTP_{44–339}$ C270S is catalytically inactive,[3] the phosphate molecule bound at the active site in the current structure most likely originated from water hydrolysis of the Erk2 peptide phosphothreonine residue. Crystallographic data for $HePTP_{44–339}$ C270S was collected at BNL-NSLS Beamline X6A at 100K using an ADSC QUANTUM 4 CCD detector. All crystallographic data were indexed, scaled and merged using HKL2000 0.98.692i.[4] The structure of $HePTP_{44–339}$ C270S (HePTP C270S*) was solved by molecular replacement using the program Phaser 1.3.2[5] and the structure of WT $HePTP_{44–339}$ (PDB ID: **1ZC0**) as an input model, after omitting solvent molecules, resulting in rotation- and translation-function Z-scores >20. The structure was completed by cycles of manual building using the program Coot 6.0.2[6] coupled with structure refinement using RefMac 5.2.0019[7] against the native datasets. The structure of HePTP C270S* was determined to 2.72 Å resolution and refined to $R_{work}$ = 15.7% and $R_{free}$ = 22.7%, and contains 1 molecule of HePTP, 132 water molecules, and 1 phosphate molecule per asymmetric unit (HePTP residues 336–339 were not observed in the electron density map and so were not modeled). The stereochemical quality of the model was analyzed using MolProbity,[8] which performs Ramachandran plot, Cβ deviation, and rotamer analyses. The agreement of the model to the diffraction data was analyzed using SFCheck 7.2.02.[9] Atomic coordinates of the final models and experimental structure factors for $HePTP_{44–339}$ C270S have been deposited with the Protein Data Bank (PDB) as entry **2QDP**.

**Michaelis-Menten Kinetic Assay.** The general PTP substrate *para*-nitrophenyl phosphate (*p*NPP) was purchased from Sigma. All other chemicals and reagents were of the highest grade commercially available. The HePTP K182A-catalyzed hydrolysis of *p*NPP was assayed in triplicate at 30°C in 0.15 M Bis-Tris pH 6.0, with ionic strength adjusted to 150 mM with NaCl. The reaction was initiated by addition of various concentrations of *p*NPP (ranging from 0.1 to 5 $K_m$) to the reaction mixture to a final volume of 100 µl. The reaction was quenched by addition of 100 µl of 1 M NaOH. The non-enzymatic hydrolysis of the substrate was corrected by measuring the control without addition of enzyme. The amount of *para*-nitrophenolate product was determined from the absorbance at 405 nm detected by a microplate reader (SpectraMax M5, Molecular Devices or PowerWaveX340, Bio-Tek Instruments), using a molar extinction coefficient of 18,000 $M^{-1}cm^{-1}$. The turnover number ($k_{cat}$) was evaluated by fitting the data to the Michaelis-Menten equation, using nonlinear regression and the program SigmaPlot (version 8.0).

**REFERENCES**


1. DeLano, W. L. (2002). The PyMOL Molecular Graphics System. DeLano Scientific, Palo Alto, CA, USA.
2. Peti, W. & Page, R. (2007). Strategies to maximize heterologous protein expression in Escherichia coli with minimal cost. *Protein Expr Purif* **51**, 1-10.
3. Critton, D. A., Tortajada, A., Stetson, G., Peti, W. & Page, R. (2008). Structural basis of substrate recognition by hematopoietic tyrosine phosphatase. *Biochemistry* **47**, 13336-45.
4. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol. (Part A)* **276**, 307-326.
5. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. (2005). Likelihood-enhanced fast translation functions. *Acta Crystallogr D Biol Crystallogr* **61**, 458-64.
6. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-32.
7. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240-55.
8. Lovell, S. C., Davis, I. W., Arendall, W. B., 3rd, de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S. & Richardson, D. C. (2003). Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins* **50**, 437-50.
9. Vaguine, A. A., Richelle, J. & Wodak, S. J. (1999). SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta Crystallogr D Biol Crystallogr* **55**, 191-205.