Structure-assisted discovery of *Variola major* H1 phosphatase inhibitors

Phan *et al.*

Synopsis

The structure of the dual-specificity H1 phosphatase encoded by the smallpox virus has been determined and used to identify small-molecule inhibitors by *in silico* screening methods.

Queries and comments
Please supply or correct as appropriate all **bold underlined** text.

Author index
Authors' names will normally be arranged alphabetically under their family name and this is commonly their last name. Prefixes (van, de *etc.*) will only be taken into account in the alphabetization if they begin with a capital letter. Authors wishing their names to be alphabetized differently should indicate this below. **Author names may appear more than once in this list; it is not necessary to mark this correction on your proofs.**

Phan, J.
Tropea, J.E.
Waugh, D.S.

$\sigma\Sigma$

Structure-assisted discovery of *Variola major* H1 phosphatase inhibitors

Jason Phan,‡ Joseph E. Tropea and David S. Waugh*

Macromolecular Crystallography Laboratory, Center for Cancer Research, National Cancer Institute at Frederick, PO Box B, Frederick, MD, USA

‡ Present address: Structural Biology Laboratory, Department of Drug Discovery, H. Lee Moffitt Cancer Center and Research Institute, Tampa, Florida, USA.

Correspondence e-mail: waughd@ncifcrf.gov

Variola major virus, the causative agent of smallpox, encodes a dual-specificity H1 phosphatase. Because this enzyme is essential for the production of mature virus particles, it is an attractive molecular target for the development of therapeutic countermeasures for this potential agent of bioterrorism. As a first step in this direction, the crystal structure of H1 phosphatase has been determined at a resolution of 1.8 Å. *In silico* screening methods have led to the identification of several small molecules that inhibit *Variola* H1 phosphatase with IC_{50} values in the low micromolar range. These molecules provide novel leads for future drug development.

1. Introduction

Reversible protein phosphorylation, mediated by kinases and phosphatases, is a central mechanism of signal transduction in higher eukaryotes that regulates cell development, differentiation and death (Hunter, 1995). As such, the modulation of phosphorylation offers a potential means of therapy for a wide variety of maladies. Most protein kinases phosphorylate either tyrosine or serine and threonine residues (Taylor *et al.*, 1995). Protein phosphatases come in three common varieties: the protein tyrosine phosphatases (PTPases), the serine/threonine phosphatases and the dual-specificity phosphatases (DUSPs), which are capable of dephosphorylating tyrosine, serine and threonine residues (Martell *et al.*, 1998). Although the relatively large number of protein kinases (428) and phosphatases (99) encoded by the human genome (Ducruet *et al.*, 2005) naturally raises some concerns about the specificity of inhibitors, potent and specific inhibitors of both classes of enzymes have been developed in recent years, providing proof of principle for this method of therapeutic intervention (*e.g.* Sohn *et al.*, 2003; Sawyer *et al.*, 2005).

Certain viruses and bacterial pathogens of plants and animals also encode kinases and phosphatases that contribute to virulence by modulating host signal transduction pathways (Leader, 1993; Cozzone, 2005). Accordingly, these enzymes are considered to be attractive molecular targets for the development of therapeutic agents to combat the infectious diseases that they cause. For instance, the protein tyrosine phosphatase YopH has been the focus of efforts to develop countermeasures for *Yersinia pestis*, the causative agent of plague and a potential agent of bioterrorism (Lee *et al.*, 2003, 2005; Liang *et al.*, 2003; Phan *et al.*, 2003; Sun *et al.*, 2003; McCain *et al.*, 2004; Hu & Stebbins, 2005; Tautz *et al.*, 2005). Similarly, although officially eradicated (Stewart & Devlin,

2005), the smallpox virus (*Variola major*) is still considered to represent a potential bioterrorist threat (Harrison *et al.*, 2004). *Variola* virus encodes a dual-specificity protein phosphatase (H1) that is essential for viral replication (Liu *et al.*, 1995) and may therefore present an opportunity for therapeutic intervention. In addition to providing potential therapy for infected people, the availability of antiviral drugs could decrease the risks associated with the smallpox vaccine by providing an alternative for vaccine-sensitive individuals.

Common starting points for the development of enzyme inhibitors include substrate analogs, random screening and *in silico* screening (Bleicher *et al.*, 2003). Irrespective of which approach is taken, detailed information about the three-dimensional structure of the target enzyme and its complexes with inhibitors is highly desirable. Although the structures of several dual-specificity protein phosphatases have been determined (*e.g.* Yuvaniyama *et al.*, 1996; Fauman *et al.*, 1998; Farooq *et al.*, 2001, 2003; Kozlov *et al.*, 2004; Jeong *et al.*, 2005; Sun *et al.*, 2005; Yoon *et al.*, 2005; Jeong, Cho *et al.*, 2006; Jeong, Yoon *et al.*, 2006), none of them are similar enough to H1 phosphatase to enable an accurate homology model to be constructed. In order to overcome this difficulty, we successfully determined the structure of this crucial viral protein and employed an *in silico* screening strategy to identify several lead compounds with IC_{50} values in the low micromolar range for further development.

2. Materials and methods

2.1. Crystallization, derivatization and data collection

Recombinant H1 phosphatase was overproduced and purified as described by Tropea *et al.* (2006), while the incorporation of selenomethionine followed the procedure of Doublié (1997). The peak fractions from a gel-filtration column corresponding to monomeric enzyme were pooled and concentrated to 8.5 mg ml^{-1} in 50 mM MES buffer pH 6.0. Sparse-matrix screens from Hampton Research were set up using a Hydra II Plus One (Matrix Technologies) crystallization robot and multiple hits were observed after a few days. At a 1:4 ratio of protein to precipitant, Hampton Research (HR) Index Screen condition No. 47 (28% polyethylene glycol monomethyl ether 2000, 0.1 M bis-tris pH 6.5) and HR PEG/Ion Screen condition No. 2 (20% polyethylene glycol 3350, 0.2 M potassium fluoride) yielded thin trapezoid and rhombic plates that scattered X-rays, but not well enough for data collection. Therefore, HR Index Screen condition No. 47 was used in a PEG MME 2000 grid to scan for the optimal potassium fluoride concentration, leading to the identification of two crystal forms: $P4$ (24% PEG MME 2000, 1.7 M KF, 0.1 M bis-tris pH 6.5) and $C222_1$ (20% PEG MME 2000, 0.6 M KF, 0.1 M bis-tris pH 6.5). Additionally, attempts to cocrystallize H1 phosphatase with sodium orthovanadate, a competitive inhibitor of PTPases, yielded yet another type of tetragonal crystal ($I4_122$) even though the ion did not bind. These conditions consisted of 26% PEG MME 2000, 1.5 M KF,

Table 1
Crystallographic data, phasing and refinement statistics.

Values in parentheses are for the last shell.

Space group	$I4_122$	
Unit-cell parameters (Å)	$a = b = 101.4, c = 95.1$	
Data set	Se peak	Se inflection
Wavelength (Å)	0.9792	0.9794
Resolution (Å)	25–1.8 (1.86–1.8)	25–2.0 (2.07–2.0)
Unique reflections	22420	16386
Completeness (%)	96.0 (97.8)	95.2 (98.0)
Redundancy	7.5 (7.2)	7.6 (7.3)
$I/\sigma(I)$	12.9 (5.0)	16.1 (4.4)
R_{merge}† (%)	10.0 (39.0)	8.6 (39.0)
No. of sites	9	
Phasing power		
Anomalous	1.30	
Isomorphous (centric/acentric)	0.42/0.49	
Correlation coefficient (25–1.8 Å)		
After density modification	0.92	
R factor for F_c *versus* F_o	0.14	
Refinement resolution (Å)	25–1.8	
R_{cryst} (%)	17.8	
R_{free} (%)	19.5	
Ramachandran plot		
Most favorable (%)	88.6	
Allowed (%)	11.4	
Average B factor (Å^2)	31.8	
B factor from Wilson plot (Å^2)	28.9	
R.m.s.d. bonds (Å)	0.021	
R.m.s.d. angles (°)	1.62	
No. of molecules in ASU		
Polypeptide	1	
Solvent	190	

† Friedel-related reflections were treated as equivalent.

3 mM Na$_3$VO$_4$ and 0.1 M bis-tris pH 6.5 with a protein:precipitant ratio of 3:1.

Native data were collected for all three space groups to 2.0 Å with high completeness and the coordinates of human vaccinia H1-related phosphatase (PDB code 1vhr) were used as a search model in an attempt to solve the structure by molecular replacement. Since no credible solution was found, conventional heavy-atom compounds were screened using all crystal forms, but no useful derivative was obtained. Finally, the structure was solved by the multiwavelength anomalous dispersion (MAD) technique using selenomethionine-substituted protein. Fortuitously, the selenomethionyl protein crystallized reproducibly in the space group with the highest symmetry ($I4_122$) under the same conditions as the native protein. The mother liquor was supplemented with 5% glycerol as a cryoprotectant before flash-freezing the crystals for data collection at SER-CAT insertion-device beamline 22-ID (Advanced Photon Source, Argonne National Laboratory). A fluorescence scan of the selenomethionine-derivative crystal confirmed the incorporation of Se atoms. Data were collected at the absorption edge and the peak wavelengths and were reduced using the *HKL*-2000 suite of programs (Otwinowski & Minor, 1997).

2.2. Structure determination and refinement

Highly redundant 1.8 Å MAD from the $I4_122$ crystal were used in the *SHARP* program (de La Fortelle & Bricogne,
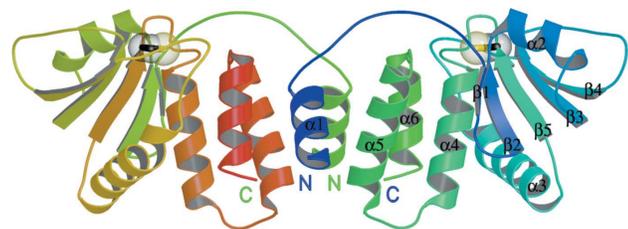
1997) to solve the structure. High-occupancy selenium positions were located for all but the N-terminal selenomethionine. After density modification, the correlation coefficient was 0.92 and the R factor for F_c versus F_o was 0.14 for data in the resolution range 25–1.8 Å (Table 1). The structure was traced automatically using the program *ARP/wARP* (Morris *et al.*, 2004) as implemented in the *SHARP* interface. Manual correction of the model was carried out using the graphics software *O* (Jones *et al.*, 1991) and the SeMet-substituted structure was refined against the peak data using *REFMAC*5 (Murshudov *et al.*, 1997). Structure validation was accomplished using *PROCHECK* (Laskowski *et al.*, 1993).

2.3. Molecular modeling

The well refined model of *Variola* H1 phosphatase was loaded into *SYBYL* 7.0 (Tripos Inc.), where solvent molecules were removed, H atoms added and charges assigned. The structure was then used as the receptor for a virtual screen of the Diversity Set of 1990 compounds from the Developmental Therapeutics Program of the National Cancer Institute using the program *AutoDock* 3.05 (Morris *et al.*, 1998). The Olson group at the Scripps Research Institute have composed a virtual library of these compounds by adding H atoms and partial atomic charges and have made it available for use with *AutoDock*. The docking grid was centered on the S atom of the catalytic cysteine with a radius of 7.4 Å. The Lamarckian genetic algorithm was used at default settings for ligand conformational search. Top hits with inhibition constants in the micromolar range were re-scored with the program *X-score* for additional filtering of the active compounds (Wang *et al.*, 2002). Although the scoring function in *AutoDock* is known to be reasonably accurate, *X-score* uses a large training set based on real experimental data for macromolecular ligand-binding interactions.

2.4. Enzyme assays

The *Variola* H1 phosphatase-catalyzed reaction was conducted at room temperature using *p*-nitrophenol phosphate (pNPP) as the substrate in 1 ml of assay buffer containing 100 m*M* bis-tris pH 6.5, 1 m*M* dithiothreitol, 1 m*M* ethylenediaminetetraacetic acid and 10% dimethylsulfoxide.
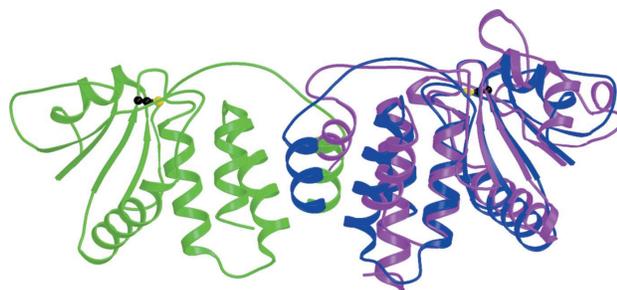
For inhibition assays, approximately 5 µg enzyme was pre-incubated with various concentrations of the inhibitor for 10 min before adding substrate to a final concentration of 200 µ*M* (approximately equal to the K_M) to initiate the reaction. The developing product was monitored continuously for 3 min at 405 nm on a Shimadzu UV–Vis spectrophotometer. The average absorbance per unit time was calculated using the installed kinetics program at various inhibitor concentrations and converted to the amount of product formed using a molar extinction coefficient of 18 000 M^{-1} cm^{-1} for pNPP. The relative activity was plotted against the inhibitor concentration using the program *SigmaPlot* to determine IC$_{50}$ values.

3. Results and discussion

3.1. Crystallization and structure solution

The *Variola* H1 PTPase, which is strongly inhibited by monovalent salts *in vitro* (Tropea *et al.*, 2006), responded to changes in ionic strength by crystallizing in different space groups. Increasing the potassium fluoride concentration caused the enzyme to shift from the orthorhombic space group $C222_1$ to the primitive tetragonal group $P4$ or, in the presence of sodium orthovanadate, to $I4_122$. Surprisingly, the latter crystal form did not have vanadate bound in the active site even though the concentration of the ligand in the mother liquor exceeded 1 m*M*. This may have been a consequence of the relatively low affinity of H1 phosphatase for vanadate (data not shown) and/or a chemical reaction with some component of the crystallization cocktail (a yellow tint was observed upon mixing).

The structure of H1 phosphatase was solved in space group $I4_122$ by the multiwavelength anomalous diffraction (MAD) technique using selenomethionine-substituted protein (Table 1). The MAD-phased experimental map calculated at a resolution of 1.8 Å was of excellent quality, allowing automatic tracing of all residues except the N-terminal serine (a non-native residue introduced during cloning) and the four C-terminal amino acids. The structure in space group $C222_1$ was subsequently solved by molecular replacement (data not shown). Apart from differences in a few disordered side chains

Figure 1
Ribbon model of the *V. major* H1 phosphatase shown with a rainbow coloring scheme. The enzyme crystallized as a domain-swapped dimer. The N- and C-termini of the two subunits are labeled in blue and green. The catalytic residue Cys110 is highlighted in bond and CPK format. Figs. 1 and 2 were prepared using the programs *MOLSCRIPT* (Kraulis, 1991) and *RASTER3D* (Merritt & Murphy, 1994).

Figure 2
Comparison of *Variola* H1 phosphatase with human VHR phosphatase. The ribbon model of the VHR monomer (magenta) is superimposed on a subunit of the H1 dimer (blue and green). The N-terminal helix of VHR has a similar conformation to that of the swapped molecule (green). The catalytic Cys110 of H1 is shown in ball-and-stick representation.

and loop regions, the $I4_1 22$ and $C222_1$ crystal forms yielded almost identical structures. The $C222_1$ structure was not refined because the virtually identical $I4_1 22$ structure was clearly of superior quality. The data obtained from the $P4$ crystals were not good enough to enable structure solution.

3.2. Description of the structure

In both crystal forms *Variola* H1 phosphatase crystallized as a domain-swapped dimer (Fig. 1), with the intermolecular axis lying along the crystallographic twofold axis. The swapped element is the N-terminal α-helix, which rotates around a single bond in order to interact with the corresponding part of the neighboring molecule. As observed in other domain-swapped structures, the intermolecular interactions are mainly hydrophobic in nature (Nioche *et al.*, 2002). In the swapped position, the N-terminal helix forms a four-helix bundle with the three C-terminal helices from the adjacent molecule. It is oriented perpendicularly to the helix adjacent to the C-terminal helix and interacts with the corresponding element from the neighboring subunit at its base, while the two N-termini point away from each other at a right angle (Fig. 1). The N-terminal α-helix could presumably adopt a very similar conformation to that of its swapped counterpart in an intramolecular configuration. The swapped helices are connected to the remainder of the polypeptides from which they originate by an 11-residue-long loop, giving *Variola* H1 phosphatase a distinctive appearance. It is unclear whether the observed dimer is biologically relevant or is just a crystallographic artifact (Tropea *et al.*, 2006). In any case, the environment of the enzyme active site is not perturbed by the helix swapping.

3.3. Comparison with human vaccinia H1-related phosphatase (VHR) and other DUSPs

As expected, the overall fold of H1 phosphatase is similar to that of human vaccinia H1-related phosphatase (VHR; Fig. 2). A structural comparison with coordinates available from the Protein Data Bank using the program *DALI* (Holm & Sander, 1995) yielded five phosphatases with high alignment Z scores: 18.0, 17.4, 13.9, 13.8 and 13.0 for human DUSP22 or JSP-1 (PDB code 1wrm), VHR (PDB code 1vhr), phosphoinositide phosphatase PTEN (PDB code 1d5r), cyclin-dependent kinase inhibitor 3 (PDB code 1fpz) and Cdc14b2 phosphatase (PDB code 1ohc), respectively. In such comparisons, Z scores greater than 3.0 are usually considered to indicate significant structural similarity. However, all of these proteins have only limited sequence identity to H1 phosphatase, ranging from 13% to 28% for the aligned segments, which constitute 75–85% of their total lengths. Despite the high degree of structural similarity between H1 phosphatase, DUSP22 and VHR, the distribution of electrostatic potential and topographical details of the surfaces are quite different owing to the variation in their amino-acid sequences (Fig. 3). The negatively charged residues are concentrated at one end of the molecule (bottom) in VHR and, to a lesser extent, in DUSP22, whereas they are localized toward the middle in *Variola* H1. In the *Variola* H1 molecule, instead of negatively charged residues there is a long hydrophobic groove (bottom) left exposed by
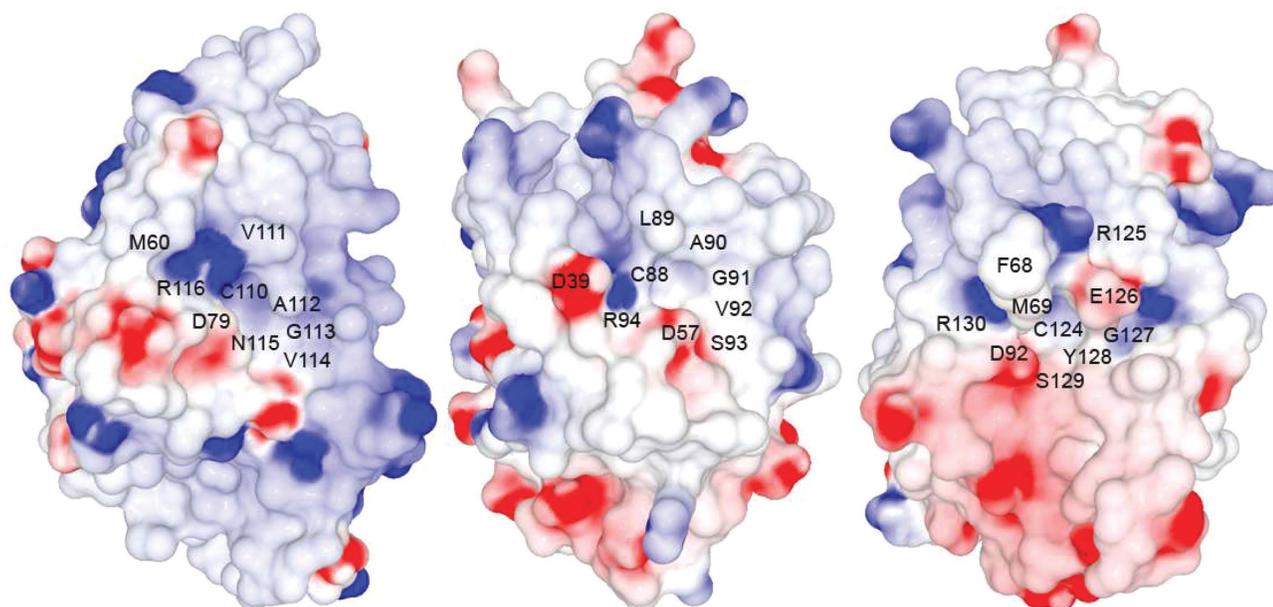
Figure 3
Space-filling models of the *Variola* H1 (left), human DUSP22 (center) and VHR (right) phosphatases with electrostatic potentials mapped onto their molecular surfaces. The potentials are colored from red ($-18kT/e$) to blue ($+18kT/e$), where k is the Boltzmann constant, T is the temperature and e is the charge of the electron. White areas represent neutral and hydrophobic amino acids. The active-site residues, including the catalytic cysteine located at the bottom of the phosphate-binding loop, are labeled. This figure was prepared using *MOLSCRIPT* (Kraulis, 1991), *GRASP* (Nicholls, 1992) and *RASTER3D* (Merritt & Murphy, 1994).

the N-terminal helix that is involved in the intermolecular swapping interaction.

An alignment of multiple DUSP structures (Fig. 4) reveals a conserved core around the phosphate-binding loop comprised of the central β-sheet and the four C-terminal α-helices and several variable regions at both termini and the $\beta3/\beta4$, $\beta4/\alpha3$ and $\alpha4/\alpha5$ loops (see Fig. 1 for numerical assignments of α-helices and β-strands). Additionally, there are two inserts that take the form of twisted loops between $\beta2/\alpha2$ and $\beta3/\beta4$ that vary in length relative to the H1 phosphatase. The marked differences in the length and orientation of the N- and C-termini are a consequence of the fact that many of these DUSPs are multi-domain enzymes. Accordingly, rather than playing a role in substrate binding or catalysis, the sequences adjacent to the ends of these catalytic domains serve as linkers to other domains.

3.4. Analysis of the *Variola* H1 phosphatase active site

At the heart of all phosphatases is the highly conserved phosphate-binding loop comprised of the consensus sequence $HC(X)_5R(S/T)$, with the catalytic cysteine located in the center of the loop (Aqvist & Kolmodin, 2001; Denu & Dixon, 1998; Zhang, 1998). Superposition of the C^α traces of H1 and related phosphatases reveals that while the backbone amide groups of the loop residues (X_5) interact with the phosphate ion, their side chains determine the topology of the binding surface and therefore contribute to substrate specificity by lining the mouth of the active-site pocket. Furthermore, residues from the two insert loops $\beta2/\alpha2$ and $\beta3/\beta4$ and the equivalent of the catalytic WPD loop in PTPases (loop $\beta4/\alpha3$) shape the periphery of the substrate-binding pocket. More significantly, the aspartic acid Asp79 in the WPD-equivalent loop of H1 phosphatase (Asp57 and Asp92 in DUSP22 and VHR, respectively) occupies a similar position to the aspartic acid residues in the classical PTPs (Asp356 in YopH; Asp181

in PTP1B; Asp425 in SHP-2; Asp1490 in LAR), which are known to act as the general acid during the enzyme-catalyzed hydrolysis of phosphotyrosine by these PTPases. The carboxylate groups of the WPD aspartates point toward the phenolic oxygen of the substrate phosphotyrosines. In the unbound H1 structure, the corresponding loop seems to be in the closed conformation, in which Asp79 instead interacts with the amide groups of Asn115 and Asn155.

The X_5 residues are very similar in H1 phosphatase and human DUSP22, but differ significantly from their equivalents in VHR and other enzymes (Fig. 3). They are Val, Ala, Gly, Val and Asn for H1, Leu, Ala, Gly, Val and Ser for DUSP22 and Arg, Glu, Gly, Tyr and Ser for VHR. The glycine in the X_3 position is invariant because it accommodates a necessary steric constraint on the phosphate-binding loops. The large and charged/polar side chains at positions X_1, X_2 and X_4 in VHR create a deeper and electrostatically different binding pocket compared with those of H1 and DUSP22. In addition, the side chain of Met69 in VHR is in hydrophobic contact with a methylene group of the bound HEPES molecule [4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid], while the rest of the piperazineethane moiety stacks against the aliphatic part of Glu126 and the aromatic ring of Tyr128. In the PTEN structure, the side chain of Val45 lines the mouth of the active site, while the phenyl ring of Tyr46 packs against the aliphatic part of Lys125, a phosphate-binding loop residue. None of these interactions are possible in the H1 binding site because it lacks the insert loops and possesses Ala112 and Val114 instead of the longer side chains found in VHR and PTEN. Nevertheless, the side chains of Ala112, Val114 and Val111 effectively mimic the aliphatic portions of Glu126, Tyr128 and Arg125 in VHR or Ala126, Lys128 and Lys125 in PTEN, respectively, because the charged ends of these side chains in the latter two enzymes point away from the active site in a manner that maximally exposes their hydrophobic trunks. In effect, the active site of H1 phosphatase is broader and shallower than the active sites of the other DUSPs, with the exception of human DUSP22, which has almost identical amino acids in these positions. The only two X_5-loop residues that differ between H1 and DUSP22 are Val111 and Asn115, and Leu89 and Ser93 (X_1 and X_5), respectively. The serine in DUSP22 forms a hydrogen bond to the carboxylate group of the catalytic aspartate Asp57 (Asp79 and Asp92 in H1 and VHR, respectively) as opposed to the asparagine amide group, which makes a stronger hydrogen-bond donor. Furthermore, this amide group can also form a hydrogen bond with the phenolic oxygen of pTyr more readily than can the serine side chain. Otherwise, the two active sites are very similar in shape and size compared with those of other DUSPs, including human VHR. Therefore, DUSP22 appears to be the
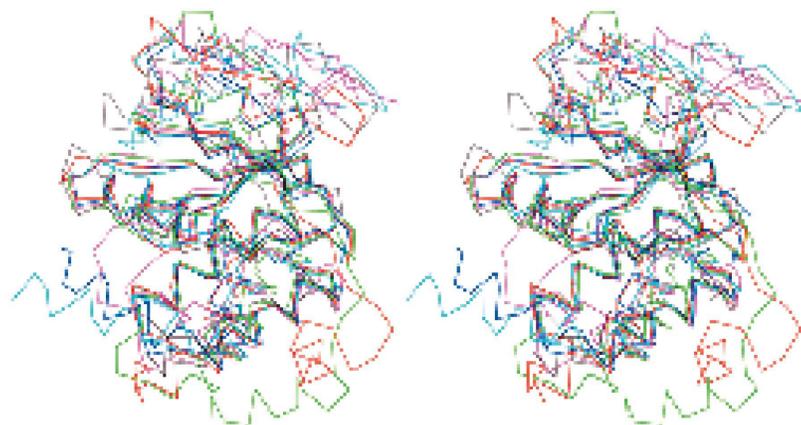
Figure 4
Please provide an improved version of this figure – Superposition of the C^α traces of the six structurally similar DUSPs *Variola* H1 (green), human DUSP22 (blue), human VHR (red), phosphoinositide phosphatase PTEN (grey), cyclin-dependent kinase inhibitor 3 (magenta) and Cdc14b2 phosphatase (cyan) shown in stereoview. This figure was prepared using the program *O* (Jones *et al.*, 1991). See Fig. 1 for the numbering of α-helices and β-strands.

best enzyme with which to gauge the specificity of H1 phosphatase inhibitors.
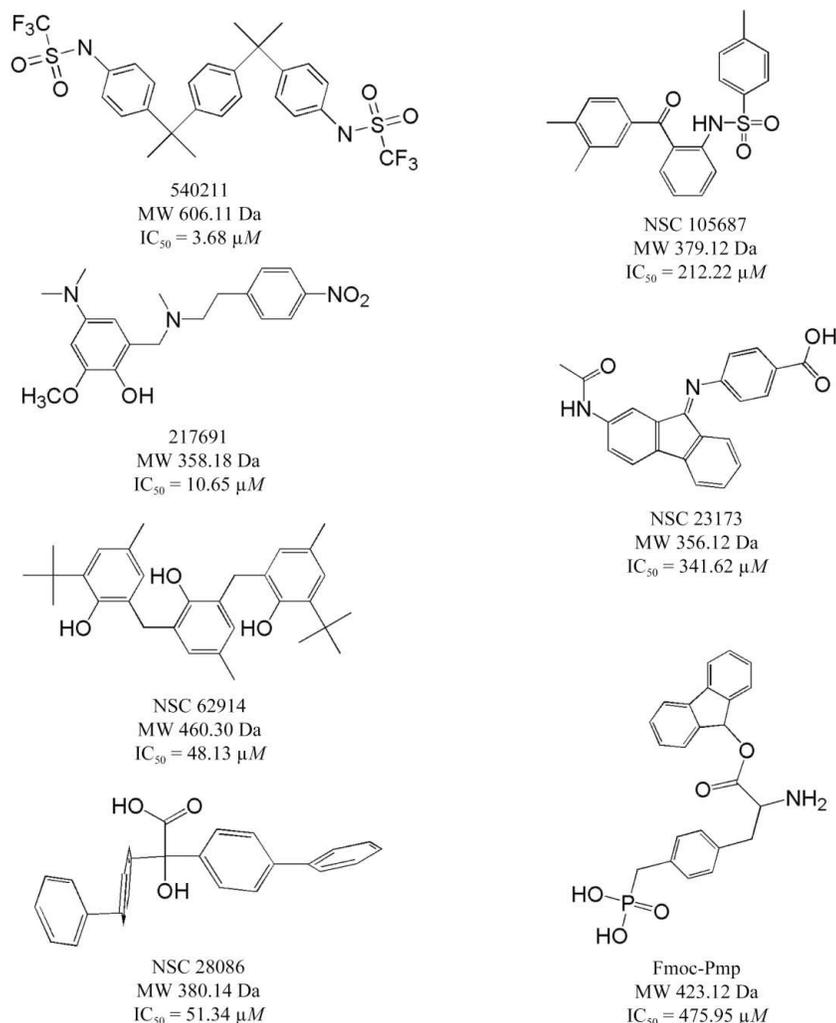
3.5. Lead discovery

In an attempt to identify potential inhibitors, we used the high-resolution structure of H1 phosphatase as a starting point for *in silico* screening of the Structural Diversity Set of Compounds from the Developmental Therapeutics Program of the National Cancer Institute (Fig. 5). Of approximately two dozen top-scoring compounds that were tested, one third were found to be insoluble in aqueous solution. Four of the soluble compounds exhibited measurable inhibitory activity. These were NSC-62914, NSC-28086, NSC-105687 and NSC-23173, with IC_{50} values of 48, 51, 212 and 342 μM, respectively. *In silico* screening of the CalBiochem phosphatase-inhibitor library identified two additional compounds (catalog Nos.

540211 and 217691) as possible leads and these were subsequently shown to have IC_{50} values against H1 phosphatase of 4 and 11 μM, respectively.

All of the inhibitors identified above possess an aromatic ring adjacent to a polar head group, which is reminiscent of phosphotyrosine. Accordingly, this may indicate that the enzyme preferentially dephosphorylates phosphotyrosine residues rather than serine or threonine. The polar head groups of the inhibitors may interact with the phosphate-binding loop of the enzyme while the aromatic moiety stacks against the active-site arginine. Ligands with additional hydrophobic groups could potentially be accommodated by the shallow groove formed by alanine and valine residues at the mouth of the active site in H1 phosphatase. Developing inhibitors of H1 phosphatase and other DUSPs with high potency may present a significant challenge because of their comparatively shallow substrate-binding grooves. However, the prospects for developing inhibitors with high selectivity for H1 phosphatase seem promising because the substrate-binding sites of the dual specificity phosphatases exhibit a much greater degree of variability than do those of their PTPase counterparts.

A potential advantage of the H1 phosphatase inhibitors identified in this study is that they have been pre-filtered for favorable drug-like properties. Efforts to cocrystallize these inhibitors with the *Variola* enzyme are currently under way. It is hoped that the resulting structural information will reveal a common pharmocophore and suggest ways in which the potency and specificity of the inhibitors can be further improved. It will also be important to obtain at an early stage information about the specificity, toxicity and bioavailability of these compounds and their derivatives and appropriate resources are currently being assembled to address these issues. Although we hope that one will never be needed, the availability of an effective therapeutic countermeasure for the smallpox virus would be very reassuring indeed.

540211
MW 606.11 Da
IC_{50} = 3.68 μM

217691
MW 358.18 Da
IC_{50} = 10.65 μM

NSC 62914
MW 460.30 Da
IC_{50} = 48.13 μM

NSC 28086
MW 380.14 Da
IC_{50} = 51.34 μM

NSC 105687
MW 379.12 Da
IC_{50} = 212.22 μM

NSC 23173
MW 356.12 Da
IC_{50} = 341.62 μM

Fmoc-Pmp
MW 423.12 Da
IC_{50} = 475.95 μM

Figure 5
Structures, molecular weights and experimentally determined IC_{50} values of the most potent inhibitors identified by virtual screening of *Variola* H1 phosphatase. This figure was prepared using *ChemDraw* (CambridgeSoft).

References

Aqvist, J. & Kolmodin, K. (2001). *FEBS Lett.* **498**, 208–213.

Bleicher, K. H., Bohm, H. J., Muller, K. & Alanine, A. I. (2003). *Nature Rev. Drug Discov.* **2**, 369–378.

Cozzone, A. J. (2005). *J. Mol. Microbiol. Biotechnol.* **9**, 198–213.

Denu, J. M. & Dixon, J. E. (1998). *Curr. Opin. Chem. Biol.* **2**, 633–641.

Doublié, S. (1997). *Methods Enzymol.* **276**, 523–530.

Ducruet, A. P., Vogt, A., Wipf, P. & Lazo, J. S. (2005). *Annu. Rev. Pharmacol. Toxicol.* **45**, 725–750.

Farooq, A., Chaturvedi, G., Mujtaba, S., Plotnikova, O., Zeng, L., Dhalluin, C., Ashton, R. & Zhou, M.-M. (2001). *Mol. Cell*, **7**, 387–399.

Farooq, A., Plotnikova, O., Chaturvedi, G., Yan, S., Zeng, L., Zhang, Q. & Zhou, M.-M. (2003). *Structure*, **11**, 155–164.

Fauman, E. B., Cogswell, J. P., Lovejoy, B., Rocque, W. J., Holmes, W., Montana, V. G., Piwnica-Worms, H., Rink, M. J. & Saper, M. A. (1998). *Cell*, **93**, 617–625.

Harrison, S. C., Alberts, B., Ehrenfeld, E., Enquist, L., Fineberg, H., McKnight, S. L., Moss, B., O'Donnell, M., Ploegh, H., Schmid, S. L., Walter, K. P. & Theriot, J. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 11178–11192.

Holm, L. & Sander, C. (1995). *Trends Biochem. Sci.* **20**, 478–480.

Hu, X. & Stebbins, C. E. (2005). *Bioorg. Med. Chem.* **13**, 1101–1109.

Hunter, T. (1995). *Cell*, **80**, 225–236.

Jeong, D. G., Cho, Y. H., Yoon, T. S., Kim, J. H., Son, J. H., Ryu, S. E. & Kim, S. J. (2006). *Acta Cryst.* D**62**, 582–588.

Jeong, D. G., Kim, S. J., Kim, J. H., Son, J. H., Park, M. R., Lim, S. M., Yoon, T. S. & Ryu, S. E. (2005). *J. Mol. Biol.* **345**, 401–413.

Jeong, D. G., Yoon, T. S., Kim, J. H., Shim, M. Y., Jeong, S. K., Son, J. H., Ryu, S. E. & Kim, S. J. (2006). *J. Mol. Biol.* **360**, 946–955.

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

Kozlov, G., Cheng, J., Ziomek, E., Banville, D., Ghering, K. & Ekiel, I. (2004). *J. Biol. Chem.* **279**, 11882–11889.

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

La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Leader, D. P. (1993). *Pharmacol. Ther.* **59**, 343–389.

Lee, K., Boovanahalli, S. K., Nam, K. Y., Kang, S. U., Lee, M., Phan, J., Wu, I., Waugh, D. S., Zhang, Z.-Y., No, K. T., Lee, J. J. & Burke, T. R. Jr (2005). *Bioorg. Med. Chem. Lett.* **15**, 4037–4042.

Lee, K., Gao, Y., Yao, Z. J., Phan, J., Wu, I., Liang, J., Waugh, D. S., Zhang, Z.-Y. & Burke, T. R. Jr (2003). *Bioorg. Med. Chem.* **13**, 2577–2581.

Liang, F., Huang, Z., Lee, S. Y., Liang, J., Ivanov, M. I., Alonso, A., Bliska, J. B., Lawrence, D. S., Mustelin, T. & Zhang, Z.-Y. (2003). *J. Biol. Chem.* **278**, 41734–41741.

Liu, K., Lemon, B. & Traktman, P. (1995). *J. Virol.* **69**, 7823–7834.

McCain, D. F., Wu, L., Nickel, P., Kassack, M. U., Kreimeyer, A., Gagliardi, A., Collins, D. C. & Zhang, Z.-Y. (2004). *J. Biol. Chem.* **279**, 14713–14725.

Martell, K. J., Angelotti, T. & Ullrich, A. (1998). *Mol. Cell*, **8**, 2–11.

Merritt, E. A. & Murphy, M. E. P. (1994). *Acta Cryst.* D**50**, 869–873.

Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K. & Olson, A. J. (1998). *J. Comput. Chem.* **19**, 1639–1662.

Morris, R. J., Zwart, P. H., Cohen, S., Fernandez, F. J., Kakaris, M., Kirillova, O., Vonrhein, C., Perrakis, A. & Lamzin, V. S. (2004). *J. Synchrotron Rad.* **11**, 56–59.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.

Nicholls, A. (1992). *GRASP. Graphical Representation and Analysis of Surface Properties.* Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY, USA.

Nioche, P., Liu, W. Q., Broutin, I., Charbonnier, F., Latreille, M. T., Vidal, M., Roques, B., Garbay, C. & Ducruix, A. (2002). *J. Mol. Biol.* **315**, 1167–77.

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

Phan, J., Lee K., Cherry, S., Tropea, J. E., Burke, T. R. Jr & Waugh, D. S. (2003). *Biochemistry*, **42**, 13113–13121.

Sawyer, T. K., Shakespeare, W. C., Wang, Y., Sundaramoorthi, R., Huang, W. S., Metcalf, C. A. III, Thomas, M., Lawrence, B. M., Rozamus, L., Noehre, J., Zhu, X., Narula, S., Bohacek, R. S., Weigele, M. & Dalgarno, D. C. (2005). *Med. Chem.* **1**, 293–319.

Sohn, J., Kiburz, B., Li, Z., Deng, L., Safi, A., Pirrung, M. C. & Rudolph, J. (2003). *J. Med. Chem.* **46**, 2580–2588.

Stewart, A. J. & Devlin, P. M. (2005). *J. Infect.* **52**, 329–334.

Sun, J. P., Wang, W. O., Yang, H., Liu, S., Liang, F., Fedorov, A. A., Almo, S. C. & Zhang, Z.-Y. (2005). *Biochemistry*, **44**, 12009–12021.

Sun, J. P., Wu, L., Federov, A. A., Almo, S. C. & Zhang, Z.-Y. (2003). *J. Biol. Chem.* **278**, 33392–33399.

Tautz, L., Bruckner, S., Sareth, S., Alonso, A., Bogetz, J., Bottini, N., Pellecchia, M. & Mustelin, T. (2005). *J. Biol. Chem.* **280**, 9400–9408.

Taylor, S. S., Radzio-Andzelm, E. & Hunter, T. (1995). *FASEB J.* **9**, 1255–1266.

Tropea, J. E., Phan, J. & Waugh, D. S. (2006). *Protein Expr. Purif.* **50**, 31–36.

Wang, R., Lai, L. & Wang, S. (2002). *J. Comput.-Aided Mol. Des.* **16**, 11–26.

Yoon, T. S., Jeong, D. G., Kim, J. H., Cho, Y. H., Son, J. H., Lee, J. W., Ryu, S. E. & Kim, S. J. (2005). *Proteins*, **61**, 694–697.

Yuvaniyama, J., Denu, J. M., Dixon, J. E. & Saper, M. A. (1996). *Science*, **272**, 1328–1331.

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

<div align="center">

INTERNATIONAL UNION OF CRYSTALLOGRAPHY

Electronic Proof Instructions

</div>

These proofs should be returned **within 14 days of April 19 2007**. After this period, the Editors reserve the right to publish articles with only the Managing Editor's corrections.

Please

(1) Read these proofs and assess if any corrections are necessary.

(2) Check that any technical editing queries have been answered.

(3) Return any corrections immediately

(a) by e-mail to

lj@iucr.org

giving a full description of the corrections in plain text and indicating the page, column and line numbers where appropriate.

(b) by fax to

+44 1244 314888

If the corrections are too complicated to send by e-mail please mark your corrections on the proofs as indicated overleaf. It is recommended that authors check that the fax transmission has been successful.

(4) Authors are not required to return proofs by mail. However, if you wish to do this please send them to

Mr P. Strickland
Managing Editor,
International Union of Crystallography,
5 Abbey Square,
Chester CH1 2HU,
England

Telephone: +44 1244 342878

Substantial alterations, apart from occasioning delay in publication, are much more expensive than many authors would suppose. Authors may therefore be required to pay for any major alterations from their original copy, and it may sometimes be necessary to disallow such changes. Where alterations are unavoidable every effort should be made to substitute words or phrases equal in length to those deleted.

Please note that in order to save postal expenses and clerical work, the typescript, drawings and photographs of articles are normally destroyed after publication. Drawings and photographs will only be returned to authors if their return is specifically requested.

Authors will be informed by e-mail when their paper is published and may then download an electronic offprint of their paper from the author services page of **Crystallography Journals** **Online** (**http://journals.iucr.org**)

Printed offprints may be purchased using the attached form which should be returned as soon as possible. It will not be possible to supply offprints for orders received after the journal is printed.

Proof Corrections

If returning your corrections by fax (or mail) please use the marking system below. Please ensure that you use black ink (not pencil) and that you do not make corrections too near the edge of the page as these may be lost in transmission.

Please do not make corrections to the pdf file electronically and please do not return the pdf file.

The recommended system of marking corrections is as follows.

(1) Each place in the text where a correction is needed should be indicated either by crossing out the characters to be corrected or by an insert mark (⅄) if an insertion is needed.
(2) The correct character, word, instruction or insertion should be shown in the margin at one end of the line and should be terminated with a long slash mark (/).
(3) The corrections in the margins (both left and right margins may be used) should be arranged in the same sequence as they are required in the line of text.
(4) Characters to be printed in the superior position, such as superscripts, apostrophes and quotation marks, are to be so identified by the symbol ∨ under them, e.g. $\overset{2}{\vee}$ or $\overset{\prime\prime}{\vee}$. Characters to be printed in the inferior position, such as subscripts, are to be identified by the symbol ∧ over them, e.g. $\overset{\wedge}{2}$.
(5) Instructions written in the margins should be encircled. This indicates that this text is not to be printed. This does not apply to the full stop (period), the colon and the solidus (slash mark), which are circled for clarity.
(6) Some instructions are needed so frequently that it is convenient to indicate them simply with special symbols or abbreviations. The symbols are recognized as instructions without being circled. Some frequently used instructions are given in the following list.

Alteration	Marginal mark	Mark in text
Insert or substitute:		
Space	#/	
Full stop (period)	⊙/	
Colon	⊙/	
Solidus (slash mark)	⊘/	
Hyphen	� - ⅃	⅄ (for an insertion) or delete material to be replaced
Rule	⚫/	
Chemical bond	⟨bond⟩/	
Superior (e.g. superscript 2 or apostrophe	⅂³/ or ∨/	
Inferior (e.g. subscript 2)	⅄/	
Change to:		
Capitals	⟨Caps⟩	≡
Small capitals	⟨s c⟩	＝ ⎫ under characters
Italic type	⟨Ital⟩	— ⎬
Bold type	⟨Bold⟩	⁓ ⎭
Lower case letters	⟨l c⟩	⎫ Circle characters
Roman type	⟨Rom⟩	⎬
Delete	ℐ	Cross out unwanted material
Delete and close up	ℐ	Cross out unwanted material and surround with ⊃
Close up	⌣	⊃ around space to be closed up
Invert type	⊚⟩	Circle inverted characters
Transpose	⟨trs⟩	⎍⎑ between letters or words
Faulty setting (e.g. broken type)	✕	Circle defective characters
Leave as printed	⟨stet⟩	.... under material to be left
New paragraph	⟨n p⟩	⌐ before first word of new paragraph
No new paragraph or line	⟨run on⟩	⌒ between lines

Please note
(a) No correction should be made in the text without an accompanying mark in the margin, or the correction may be easily missed.
(b) Complicated corrections can be explained in a covering letter.

ORDER FORM
Return within 14 days of April 19 2007

YOU WILL AUTOMATICALLY BE SENT DETAILS OF HOW TO DOWNLOAD
AN ELECTRONIC OFFPRINT OF YOUR PAPER, FREE OF CHARGE.
PRINTED OFFPRINTS MAY BE PURCHASED USING THIS FORM.

Please keep a copy of this order and send the original to:

INTERNATIONAL UNION OF CRYSTALLOGRAPHY
5 Abbey Square
Chester CH1 2HU, England. Fax: +44 1244 314888

VAT No. GB 161 9034 76

Article No.: D071490-DZ5103

Title of article	Structure-assisted discovery of *Variola major* H1 phosphatase inhibitors
Name	Professor David S. Waugh
Address	Macmolecular Crystallography Laboratory, Center for Cancer Research, National Institute at Frederick, PO Box B, Frederick, MD- 21702-1201, USA
E-mail address (for electronic offprints)	waughd@ncifcrf.gov

OPEN ACCESS

IUCr journals now offer authors the chance to make their articles open access on **Crystallography Journals Online**. For full details of our open access policy, see http://journals.iucr.org/services/openaccess.html.

The charge for making an article open access is **900 United States dollars**.

☐ I wish to make my article open access.

OFFPRINTS*

I wish to order . . . . paid offprints

These offprints will be sent to the address given above. If the above address or e-mail address is not correct, please indicate an alternative:

PAYMENT

Charge for open access . . . . USD   Charge for offprints . . . . USD   Total charge . . . . USD

☐ A cheque for . . . . USD payable to **INTERNATIONAL UNION OF CRYSTALLOGRAPHY** is enclosed

☐ An official purchase order made out to **INTERNATIONAL UNION OF CRYSTALLOGRAPHY** ☐ is enclosed ☐ will follow

Purchase order No.

☐ Please invoice me

Date	Signature

OPEN ACCESS

The charge for making an article open access is **900 United States dollars**.

A paper may be made open access at any time after the proof stage on receipt of the appropriate payment. This includes all back articles on **Crystallography Journals Online**. For further details, please contact support@iucr.org. Likewise, organizations wishing to sponsor open-access publication of a series of articles or complete journal issues should contact support@iucr.org.

To check whether your organization has open-access membership, see http://scripts.iucr.org/cgi-bin/oa_inst_list

OFFPRINTS*

An electronic offprint is supplied free of charge.

Printed offprints without limit of number may be purchased at the prices given in the table below. The requirements of all joint authors, if any, and of their laboratories should be included in a single order, specifically ordered on the form overleaf. All orders for offprints must be submitted promptly; it will not be possible to supply offprints for orders received after the journal is printed.

Please note that normally offprints are sent about one month after publication of the article, and normally by surface mail. If offprints are wanted by air mail additional costs will be invoiced.

Prices for offprints are given below in **United States dollars** and include postage.

Number of offprints required	Size of paper (in printed pages)				
	1–2	3–4	5–8	9–16	Additional 8's
50	90	130	180	272	120
100	132	184	254	388	158
150	172	238	332	504	198
200	214	294	410	624	246
Additional 50's	42	54	78	118	40

PAYMENT AND ORDERING

Cheques should be in **United States dollars** payable to **INTERNATIONAL UNION OF CRYSTALLOGRAPHY**. Official purchase orders should be made out to **INTERNATIONAL UNION OF CRYSTALLOGRAPHY**.

Orders should be returned by fax to:

INTERNATIONAL UNION OF CRYSTALLOGRAPHY
5 Abbey Square
Chester CH1 2HU, England. Fax: +44 1244 314888

ENQUIRIES

Enquiries concerning offprints should be sent to support@iucr.org.