Specificity Profiling of Dual Specificity Phosphatase Vaccinia VH1-related (VHR) Reveals Two Distinct Substrate Binding Modes*

Rinrada Luechapanichkul,1,2 Xinwen Chen,1 Hashem A. Taha,1 Shubham Vyas,3 Xiaoyan Guan,4 Michael A. Freitas,5 Christopher M. Hadad,1 and Dehua Pei1,3

From the 1Department of Chemistry and Biochemistry, The Ohio State University and 2Department of Molecular Virology, Immunology, and Medical Genetics, The Ohio State University Medical School, Columbus, Ohio 43210

Background: Factors that determine the in vivo substrate specificity of dual specificity phosphatases are currently unknown.

Results: Specificity profiling of VHR through peptide library screening identified two distinct classes of peptide substrates, which bind to VHR in opposite orientations.

Conclusion: VHR may act on a previously unrecognized class of protein substrates.

Significance: The results should help identify new VHR substrates and elucidate its biological function.

Vaccinia VH1-related (VHR) is a dual specificity phosphatase that consists of only a single catalytic domain. Although several protein substrates have been identified for VHR, the elements that control the in vivo substrate specificity of this enzyme remain unclear. In this work, the in vitro substrate specificity of VHR was systematically profiled by screening combinatorial peptide libraries. VHR exhibits more stringent substrate specificity than classical protein-tyrosine phosphatases and recognizes two distinct classes of Tyr(P) peptides. The class I substrates are similar to the Tyr(P) motifs derived from the VHR protein substrates, having sequences of (D/E/ϕ)(D/S/N/T/E)(P/I/M/S/A/V)pY(G/A/S/Q) or (D/E/ϕ)(T/S)(D/E)pY(G/A/S/Q) (where ϕ is a hydrophobic amino acid and pY is phosphotyrosine). The class II substrates have the consensus sequence of (V/A)p(1/M/V/F/X)1–4pY (where X is any amino acid) with V/A preferably at the N terminus of the peptide. Site-directed mutagenesis and molecular modeling studies suggest that the class II peptides bind to VHR in an opposite orientation relative to the canonical binding mode of the class I substrates. In this alternative binding mode, the Tyr(P) side chain binds to the active site pocket, but the N terminus of the peptide interacts with the carboxylate side chain of Asp164, which normally interacts with the Tyr(P) + 3 residue of a class I substrate. Proteins containing the class II motifs are efficient VHR substrates in vitro, suggesting that VHR may act on a novel class of yet unidentified Tyr(P) proteins in vivo.

Protein phosphorylation on tyrosine residues mediates regulates numerous physiological and pathological processes and is controlled by the opposing action of protein-tyrosine kinases and protein-tyrosine phosphatases (PTPs). Given the prevalence of this modification (~20,000 phosphotyrosine (Tyr(P)) sites in the human proteome have been documented in the PhosphoSite database) and the large number of kinases and PTPs in mammalian cells, identification of the physiological substrates of the kinases and phosphatases and understanding the molecular basis of their substrate specificity remain important and yet challenging goals of current research, especially for PTPs (1). It is now generally believed that PTPs, like their kinase counterparts, have exquisite substrate specificity in vivo that is combinatorially controlled by the sequence specificity of the PTP active site and protein-protein interaction via substrate-recruiting domains/docking sites outside the PTP active site (2). Recent specificity profiling of several classical PTPs by screening combinatorial peptide libraries as well as earlier kinetic studies with individual Tyr(P) peptides revealed that although no two PTPs share exactly the same substrate specificity profile they generally have broad and overlapping sequence specificity toward peptide substrates (3–6). Because most of these PTPs also contain other domains (e.g. SH2 domains and membrane localization motifs) in addition to the catalytic domain, it is likely that the in vivo substrate specificity of these enzymes is enhanced by their “recruiting” domains. In this regard, the substrate specificity of the subgroup of atypical dual specificity phosphatases raises intriguing questions (7). Because most members of this subgroup contain only the catalytic domain and do not have any recognizable substrate-recruiting domains or surfaces, do their PTP active sites have...
more stringent sequence selectivity compared with the classical PTPs? Or alternatively, do they simply act as relatively nonspecific phosphatases in vivo?

The 21-kDa vaccinia VH1-related (VHR) phosphatase is a member of the atypical dual specificity phosphatase subgroup and one of the first dual specificity phosphatases identified (8, 9). Other than the conserved PTP signature motif, HCXXGXXR, VHR has little sequence similarity to the classical PTPs. Knockdown of VHR by siRNA resulted in cell cycle arrest in G1/S and G2/M, indicating a critical role for VHR in cell cycle progression (10). A link between VHR and several types of cancers has recently been established (11–13). Despite the importance of VHR in cell cycle regulation and cell signaling pathways, only a few physiological VHR substrates have been identified. VHR has been shown to inactivate mitogen-activated protein kinases (MAPKs) Erk and Jnk in vivo by dephosphorylating both tyrosine and threonine phosphorylation sites within the Thr-X-Tyr motif of their activation loop (14–16). However, purified Erk proved to be a relatively poor in vitro substrate of recombinant VHR \((k_{cat}/K_m \sim 10^8 \text{M}^{-1} \text{s}^{-1})\) compared with its highly efficient dephosphorylation by other MAPK phosphatases \((k_{cat}/K_m \sim 10^6 \text{M}^{-1} \text{s}^{-1})\) (15, 17). This raises the question whether dephosphorylation of the MAPKs by VHR in vivo is aided by any recruiting strategy. Other than the MAPKs, signal transducers and activators of transcription 5 (STAT5) with a single Tyr(P) site, epidermal growth factor receptor, and ErbB2 have also been reported as VHR substrates (18, 19). Two recent studies showed that phosphorylation of VHR at Tyr\(^{138}\) is required for its efficient dephosphorylation of MAPKs and STAT5 in vivo (18, 20). In the case of STAT5, it was suggested that the Tyr(P)\(^{138}\) motif binds to one of the SH2 domains of the STAT5 homodimer, thereby recruiting VHR to the substrate protein (18, 21). In addition, this interaction displaces the STAT5 Tyr(P)\(^{694}\) peptide from the SH2 domain and renders it available for dephosphorylation by the VHR active site. How Tyr\(^{138}\) phosphorylation increases its activity toward Erk and Jnk is currently unknown. Kang and Kim (22) recently reported that the vaccinia-related kinase 3 binds to and activates VHR and that this direct interaction regulates Erk signaling, but again the molecular mechanism of activation remains a mystery. By using synthetic peptides derived from mitogen-activated protein kinases p38 and Jnk1, Dixon and co-workers (23) showed that VHR is highly active against Tyr(P) peptides but has only very weak activity toward Ser(P) and Thr(P) substrates. It was reported that VHR has a moderate preference (by 2–3-fold) for bisphosphorylated peptides (e.g. pTXXpY where pT is phosphothreonine and pY is phosphotyrosine) over monophosphorylated Tyr(P) peptides (e.g. TXXpY) (15, 24). To our knowledge, there has been no systematic study of the VHR substrate specificity. In this work, we profiled the sequence specificity of VHR by screening combinatorial peptide libraries. VHR recognizes two distinct classes of peptide sequences. Although the first class of substrates are similar to the Tyr(P) motifs derived from known VHR protein substrates, the class II substrates contain a (V/A)P motif and bind to the VHR active site in an opposite orientation relative to the class I substrates. Our results suggest that the sequence specificity of the VHR active site is a key determinant of its in vivo substrate specificity.

**EXPERIMENTAL PROCEDURES**

**Materials**—Reagents for peptide synthesis were from Advanced ChemTech (Louisville, KY), Peptides International (Louisville, KY), or NovaBiochem. N-(9-Fluorenylmethoxycarbonyloxy)succinimide was from Advanced ChemTech. α-Cyano-4-hydroxycinnamic acid, phenyl isothiocyanate, and 3-methyl-2-benzothiazolinonehydrazide were obtained from Sigma-Aldrich. N°-Fmoc-O-butyryl-3,5-difluoroacetorosine was synthesized as described previously (25). Peptide libraries were synthesized on polyethylene glycol acrylamide resin (0.4 mmol/g; 150–300-μm diameter in water) as described previously (6, 26). DNA primers for site-directed mutagenesis experiment were purchased from Integrated DNA Technologies (Coralville, IO). Pfu Turbo polymerase and DpnI enzymes were from Stratagene (Santa Clara, CA). Streptomyces antibioticus tyrosinase was purified as described previously (6, 27). GST-LASPI was obtained from Novus Biologicals (Littleton, CO), and roundabout homolog 1 (ROBO1)-Myc-DDK was from OriGene (Rockville, MD). Abl protein-tyrosine kinase was from New England Biolabs (Ipswich, MA). Anti-ROBO1 (clone 7E3.1) and anti-Tyr(P) (clone 4G10) antibodies were obtained from Millipore (Temecula, CA). Sequencing grade trypsin was from Promega (Madison, WI).

**VHR Purification**—Wild-type human VHR was expressed in Escherichia coli and purified to near homogeneity as described previously (23). VHR mutants were generated by the QuikChange mutagenesis method using the following primers: D18N, cgacctgctctgaaacggcagcggctgctac; D47N, gtctgtggctcagtcaatctgtcagcggctgctac; and roundabout homolog 1 (ROBO1) gene (Rockville, MD). Abl protein-tyrosine kinase was from New England Biolabs (Ipswich, MA). Anti-ROBO1 (clone 7E3.1) and anti-Tyr(P) (clone 4G10) antibodies were obtained from Millipore (Temecula, CA). Sequencing grade trypsin was from Promega (Madison, WI).

**Library Screening**—In a typical screening experiment, 200 mg of the proper peptide library (wet resin in N,N-dimethylformamide; ~40,000 beads) was placed in a plastic micro-BioSpin column (0.8 ml; Bio-Rad) and extensively washed with N,N-dimethylformamide and double distilled H2O. The resin was blocked for 1 h with a blocking buffer (30 mM Hepes, pH 7.4, 150 mM NaCl, 0.01% Tween 20, and 0.1% gelatin). The library was then incubated with VHR (final concentration, 200 nM) in the blocking buffer containing 5 mM tris(carboxyethyl)phosphine (total volume, 0.8 ml) at room temperature for 15 min with gentle mixing. The resin was drained, washed with 0.1 M KH2PO4, pH 6.8 buffer (3 × 1 ml), and resuspended in 1.6 ml of 0.1 M KH2PO4, pH 6.8 containing 1.2 μM tyrosinase and 6 mM 3-methyl-2-benzothiazolinonehydrazide. The resulting mixture was incubated at room temperature with gentle mixing and exposure to air. Intense pink/red color typically developed on positive beads within 30 min. The positive beads were removed from the library manually using a micropipette under a dissecting microscope and sequenced by the partial Edman degradation–mass spectrometry method (28). Control experiments without VHR produced no colored beads under identical conditions.

**Synthesis of Selected Peptides**—Individual peptides were synthesized on 100 mg of cross-linked ethoxylate acrylate resin.
VHR Substrate Specificity

(CLEAR)-amide resin using standard Fmoc/N,N,N',N'-tetramethyl-O-((H-benzotriazol-1-yl)uronium hexafluorophosphate/hydroxybenzotriazole chemistry. For the coupling reaction of Tyr(P), 2.0 eq of Fmoc-amino acid were used, whereas 4.0 eq were used for all other amino acids. The resin-bound peptides were cleaved from the resin and side chain-deprotected using modified reagent K (79:7.5:5:5:2.5:1 (v/v) trifluoroacetic acid/phenol/H2O/thioanisole/ethanedithiol/anisole) at room temperature for 2 h. The solvents were removed by evaporation under a N2 stream and trituration with cold diethyl ether. The precipitate was collected and dried under vacuum. The crude peptides were purified by reversed-phase HPLC on a semi-preparative C18 column. The identity of each peptide was confirmed by MALDI-TOF.

Enzymatic Assays—VHR activity assays were performed with synthetic Tyr(P) peptides as substrates in a quartz microcuvette. The buffer was a three-component system consisting of 0.1 M acetate, 0.05 M Tris, and 0.05 M Bis-Tris (pH 7.4). The buffer was carefully chosen to maintain constant ionic strength and avoid any inhibition of enzyme activity. The reaction was initiated by the addition of VHR (final concentration, 50–1200 nM) and monitored continuously at 282 nm (for Tyr, \( \Delta M \) = 1102 m \(^{-1} \) cm \(^{-1} \)) on a UV-visible spectrophotometer. The initial velocity was determined over the linear range and calculated by measuring the slope of the absorbance change. To determine the kinetic parameters \( k_{cat} \), \( K_m \), and \( k_{cat}/K_m \), the initial velocities were measured at various substrate concentrations, and the data were fitted to the Michaelis-Menten equation. For peptide substrates containing free N terminus, the reaction progress curves (as monitored at 282 nm) were directly fitted against the equation

\[
t = \frac{p}{k_{cat}E} + \left( \frac{K_m}{k_{cat}E} \right) \ln[p_\infty - p]
\]

(Eq. 1)

where \( t \) is time, \( p \) is the product concentration at time \( t \), \( E \) is the total enzyme concentration, and \( p_\infty \) is the product concentration at infinity to determine the kinetic parameters (29).

Molecular Modeling—We utilized the available crystal structure of the C124S mutant of VHR bound with a Tyr(P) peptide in the active site (Protein Data Bank code 1J4X) (24). The peptides studied herein were built using the backbone atoms of the native ligand in the crystal structure, keeping the Tyr(P) residue unmodified. Additionally, the C124S mutation was inverted using UCSF Chimera (30).\(^5\) The protonation states of titratable amino acids were identified using pdb2pqr utility (31) for pH 7. To prepare these systems for molecular dynamics (MD) simulations, the structures were solvated in an octahedron box of TIP3P water molecules with counter ions to make the system neutral in terms of charge (32, 33). The MD simulations involved a three-step minimization before the production MD simulations. In the first step, only water molecules were allowed to relax, whereas the protein and ligand were fixed. In the subsequent steps, the entire system was allowed to minimize followed by a gradual increase of the temperature of the system from 0 to 300 K. The resulting minimized system was used for production MD simulations using the isothermal–isobaric (NPT) ensemble. These MD simulations were performed using the AMBER ff03 force field as implemented in the AMBER 10.0 software package (34). The MD trajectories were then utilized to determine relative free energies of binding of peptides using the Poisson-Boltzmann (MM-PBSA) (35–37) and the generalized Born (MM-GBSA) methodologies (38) as implemented in AMBER.

In Vitro Protein Dephosphorylation by VHR—ROBO1 (1.0 \( \mu \)g) was phosphorylated with Abl (100 units) in the presence of ATP (10 mM) in NEBuffer for kinase (50 mM Tris-HCl, pH 7.5, 10 mM MgCl\(_2\), 1 mM EGTA, 2 mM dithiothreitol, and 0.01% Brij 35) overnight at 37 °C (total reaction volume, 20 \( \mu \)l). The kinase reaction was terminated by the addition of 50 mM EDTA, pH 7.0 (final concentration). Subsequent dephosphorylation of the protein was carried out by adding 0.25 ng of VHR to the above solution followed by incubation at 25 °C. At various time points (0–50 min), 8-\( \mu \)l aliquots were withdrawn and mixed with an equal volume of 2 \( \times \) SDS-PAGE loading buffer. The samples were boiled for 10 min, separated by 8% SDS-PAGE, transferred to PVDF membrane, and immunoblotted with anti-Tyr(P) antibody. The membrane was washed in 200 mM glycine, pH 2.2, 3.5 mM SDS, and 1% Tween 20 and reprobed with anti-ROBO1 antibody following the manufacturer’s recommendations. LASP1 was phosphorylated by Abl and evaluated as a potential VHR substrate in a similar manner.

In-solution Digestion—Abl-phosphorylated ROBO1 or LASP1 from above (1.0 \( \mu \)g) was reduced with 10 mM dithiothreitol at 25 °C for 30 min in 100 mM NH\(_4\)HCO\(_3\) and alkylated with 40 mM iodoacetamide in 100 mM NH\(_4\)HCO\(_3\) at 25 °C for 30 min. Excess iodoacetamide was quenched with 40 mM dithiothreitol at 25 °C for 30 min. LASP1 was digested at a 1:20 enzyme/substrate ratio in 100 mM NH\(_4\)HCO\(_3\), pH 7.6 with Glu-C (sequencing grade; Promega) at 37 °C for 16 h followed by trypsin digestion at 37 °C for 5 h. ROBO1 was digested with trypsin overnight at 37 °C. The resulting peptides were desalted with a C18 ZipTip and eluted with 50% acetonitrile containing 0.1% trifluoroacetic acid.

Mass Spectrometry—The peptides from the above proteolytic digests were separated by reversed-phase HPLC ( Dionex Ultimate 3000 capillary/nano-HPLC system, Dionex, Sunnyvale, CA) and analyzed on a Thermo Fisher LTQ Orbitrap XL (Thermo Finnigan, San Jose, CA). The peptides were separated on a 0.2 \( \times \) 150-mm C18 column (3 \( \mu \)m, 200 A; Michrom Bioreources Inc., Auburn, CA) at a flow rate of 2 \( \mu \)l/min with mobile phase A containing H\(_2\)O and 0.1% formic acid and mobile phase B containing acetonitrile and 0.1% formic acid. A blank run was carried out in between each sample injection. Protein and peptide sequences were determined by the MassMatrix database search engine (MassMatrix 2.4.2, February 22, 2012).

RESULTS

VHR Substrate Specificity (N-terminal to Tyr(P))—Because PTPs, which have been characterized previously, generally show stronger specificity on the N-terminal side of Tyr(P)

\(^5\) UCSF Chimera is developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco with support from the National Institutes of Health (National Center for Research Resources Grant 2P41RR01081 and NIGMS Grant 9P41GM103311).
(3–6), we first screened VHR against a peptide library containing five random residues on the N-terminal side of Tyr(P), NH<sub>2</sub>-ASXXXpYAABBRM-resin (library I; where B is β-alanine and X is 3,5-difluorotyrosine (F<sub>2</sub>Y), norleucine, or any of the 17 proteinogenic amino acids excluding Met, Cys, and Tyr) (Table 1). Library screening involved treatment of a portion of the one bead-one compound library (~40,000 beads/reaction) with a small amount of VHR (200 nM final concentration) for a short time (15 min) so that only the most preferred sequences would undergo significant dephosphorylation. The reaction product (i.e. exposed tyrosine side chain) was then selectively oxidized by tyrosinase into an orthoquinone, which was incubated with 3-methyl-2-benzothiazolinonehydrazone to form a reddish pigment that remained covalently bound to the positive beads (26, 39). The colored beads were removed from the library and individually sequenced by partial Edman degradation-mass spectrometry (28). The N-terminal dipeptide Ala-Ser was added to minimize any potential bias on the library screening by the positive charge associated with the free N terminus, whereas the BBRM motif facilitated peptide release (cleavage after the methionine by CNBr) and MS analysis (Arg provides a fixed positive charge).

Screening of a total of ~160,000 library beads (in four separate experiments) produced 98 positive sequences, which can be separated into two different classes on the basis of sequence similarity (Table 2). Inspection of the class I substrates (92 sequences) revealed that VHR has strong sequence selectivity immediately N-terminal to Tyr(P) (Fig. 1). At the Tyr(P) – 1 position, VHR strongly prefers a Pro residue (in 52% of all class I peptides) followed by Ile, Nle, and other amino acids of small side chains (Ser, Ala, Val, and Thr). At the Tyr(P) – 2 position, the most preferred residues are Asp, Ser, and Thr. Six of the selected peptides also contained the (S/T)(D/E)pY motif. Thus, the screening result is in agreement with the previously reported VHR recognition motif -pTXY- (where X is Glu, Gly, or Pro) (23, 24). Although Ser(P) and Thr(P) were not included in the library, Asp and Glu can often serve as effective Ser(P) and Thr(P) mimetics, respectively. Therefore, the strong selectivity for Asp/Glu at the Tyr(P) – 2 position suggests that Ser(P)/Thr(P) would be preferred at this position. Furthermore, the ability of VHR to tolerate both Ser/Thr and Ser(P)/Thr(P) at this position implies that VHR should be able to dephosphorylate the Tyr(P) residue of the mitogen-activated protein kinases even when the TXY motif is not stoichiometrically phosphorylated on both tyrosine and threonine residues. VHR shows a modest preference for acidic (Glu and Asp) and hydrophobic residues (Phe, Trp, and Ile) at Tyr(P) – 3 to Tyr(P) – 5 positions but tolerates most of the other amino acids (Fig. 1). The class II substrates, although present in only six of the selected sequences, have a tight consensus; they all contain an API (or APL) motif (Table 2).

To test whether VHR has any specificity for amino acid residues beyond the Tyr(P) – 5 position, we repeated the screening experiment against library II, NH<sub>2</sub>-AXXXXXppYAABB (80,000 beads) (Table 1), which contains eight random residues N-terminal to Tyr(P). Surprisingly, of the 16 positive sequences, seven contained an N-terminal API or APl motif (Table 3). The two groups of peptides clearly have different

| Library No. | Library design | Amino acid composition at random positions (X) |
|------------|---------------|---------------------------------------------|
| I          | NH<sub>2</sub>-ASXXXXXpYAABBRM | Ala, Arg, Asn, Asp, F<sub>2</sub>Y, Gly, Gln, Glu, His, Ile, Leu, Lys, Nle, Phe, Pro, Ser, Thr, Trp, Val |
| II         | NH<sub>2</sub>-AXXXXXXXpYAABBRM | Ala, Arg, Asn, Asp, Gly, Gln, Glu, His, Ile, Leu, Lys, Nle, Phe, Pro, Ser, Thr, Trp, Val |
| III        | Alloc-AXXXXXXXpYAABBRM | Ala, Arg, Asn, Asp, Gly, Gln, Glu, His, Ile, Leu, Lys, Nle, Phe, Pro, Ser, Thr, Trp, Val |
| IV         | NH<sub>2</sub>-XXYYYYpYAABBRM | Ala, Arg, Asn, Asp, F<sub>2</sub>Y, Gly, Gln, Glu, His, Ile, Leu, Lys, Nle, Phe, Pro, Ser, Thr, Trp, Val |
| V          | NH<sub>2</sub>-XXXXXXpYAABBRM | Ala, Arg, Asn, Asp, Gly, Gln, Glu, His, Ile, Leu, Lys, Nle, Phe, Pro, Ser, Thr, Trp, Val |
| VI         | NH<sub>2</sub>-AXXXXXXpYAABBRM | Ala, Arg, Asn, Asp, F<sub>2</sub>Y, Gly, Gln, Glu, His, Ile, Leu, Lys, Nle, Phe, Pro, Ser, Thr, Trp, Val |
| VII        | NH<sub>2</sub>-VXXXXXXpYAABBRM | Ala, Arg, Asn, Asp, F<sub>2</sub>Y, Gly, Gln, Glu, His, Ile, Leu, Lys, Nle, Phe, Pro, Ser, Thr, Trp, Val |
| VIII       | NH<sub>2</sub>-APXXXpYAABBRM | Ala, Arg, Asn, Asp, F<sub>2</sub>Y, Gly, Gln, Glu, His, Ile, Leu, Lys, Nle, Phe, Pro, Ser, Thr, Trp, Val |
| IX         | NH<sub>2</sub>-APIXXXXpYAABBRM | Ala, Arg, Asn, Asp, F<sub>2</sub>Y, Gly, Gln, Glu, His, Ile, Leu, Lys, Nle, Phe, Pro, Ser, Thr, Trp, Val |
| X          | NH<sub>2</sub>-IDFD(D/P)pYYYYNNBBRM | Ala, Arg, Asn, Asp, F<sub>2</sub>Y, Gly, Gln, Glu, His, Ile, Leu, Lys, Nle, Phe, Pro, Ser, Thr, Trp, Val |
| XI         | NH<sub>2</sub>-AXXXXXXXpYYYYNNBBRM | Ala, Arg, Asn, Asp, Gly, Gln, Glu, His, Ile, Leu, Lys, Nle, Phe, Pro, Ser, Thr, Trp, Val |
sequences. Although the class I peptides are rich in either acidic or basic residues throughout their sequences, the API-containing peptides (class II) typically have acidic residue(s) close to Tyr(P) and hydrophobic amino acids between the acidic residues and the N-terminal API motif. Because the class II peptides all have the API motif at the N terminus, we tested the importance of a free N-terminal amine for VHR recognition by screening it against library III, Alloc-A

\[ \text{pYAABBRM} \] (Table 1), which contained an N-terminal allyloxycarbonyl (Alloc) group but was otherwise identical to library II. Screening of library III (80,000 beads) produced 49 positive sequences (Table 3). Remarkably, none of the selected peptides contained an APImotif; they all have sequences similar to those of the class I peptides selected from libraries I and II. The drastically different screening results between libraries II and III suggest that VHR contains a specific binding site for the N-terminal API motif.

Identification of Other Class II Motifs and Optimal Distance from Tyr(P)—Because library II contained a fixed Ala at the N terminus that might have biased the library screening toward the N-terminal API motif, we asked whether VHR could recognize sequences other than API. To answer this question, we screened VHR against libraries IV (NH2-XXXXXpYAABBRM) and V (NH2-XXXXXXpYAABBRM) (~80,000 beads for each library) (Table 1). Both libraries had completely random sequences at the N terminus. Among the 17 sequences selected from library IV, 14 (82%) contained an N-terminal NH2-(V/A)P

\[ X \] motif where \( X \) is often a \( \beta \)-branched amino acid (e.g., Thr, Ile, and Val) (Table 3). The more frequent selection of an N-terminal Val than Ala suggests that Val is the preferred N-terminal residue for VHR. The remaining three peptides contained an N-terminal Thr, which is structurally similar to Val. Screening of library V did not result in an overwhelming selection of peptides with N-terminal (V/A)P

\[ X \] motifs (Table 3). Of the 27 selected sequences, only two peptides contained N-terminal VPT and APQ sequences. However, the selected sequences clearly showed an over-representation of Val at the N terminus (12 of the 27 peptides). A possible explanation for the above observations is that the NH2-(V/A)PX motif where \( X \) is often a \( \beta \)-branched amino acid (e.g., Thr, Ile, and Val) (Table 3). The more frequent selection of an N-terminal Val than Ala suggests that Val is the preferred N-terminal residue for VHR. The remaining three peptides contained an N-terminal Thr, which is structurally similar to Val. Screening of library V did not result in an overwhelming selection of peptides with N-terminal (V/A)PX motifs (Table 3). Of the 27 selected sequences, only two peptides contained N-terminal VPT and APQ sequences. However, the selected sequences clearly showed an over-representation of Val at the N terminus (12 of the 27 peptides). A possible explanation for the above observations is that the NH2-(V/A)PX motif needs to be situated at an appropriate distance from the Tyr(P) residue for VHR to engage in optimal interactions with both the NH2-(V/A)PX and Tyr(P) motifs; either increasing or decreasing the distance would diminish its binding affinity for VHR.

To determine the optimal distance between the (V/A)PX motif and Tyr(P), we modified library IV by adding Ala, Val,
| NH₂AX₃pY (library II) | AllocAX₃pY (library III) | NH₂X₃pY (library IV) | NH₂X₃pY (library V) | NH₂AX₃pY (library VI) | NH₂VX₃pY (library VII) | NH₂APX₃pY (library VIII) | NH₂APIX₃pY (library IX) |
|----------------------|--------------------------|----------------------|----------------------|----------------------|------------------------|------------------------|--------------------------|
| AFFIRRKDA            | ANDLEDA                  | APXX                  | WLIETEDE             | APAWDA               | VPIITT                 | APIMEDF                | *APIAWEDA                |
| ARDTTNEHA            | AAVEIVIDA                | APQRI                 | ANLKRWDI             | APFFNA               | VPIAWG                 | APMPNDI                | APIITMDD                 |
| AXXXSWSWP            | ARTWAGEA                 | APTIS                 | APQPWRFP             | APFKNI               | VPINHD                  | APIVPME                 | APIITMDD                 |
| ARRVHNWLL            | AEDEEFVNA                | APTHI                 | DDFMWWDNE            | APINLYP              | VPIELP                  | APIIDEEEM               | APIIPEE                  |
| AIISESWDQ            | AGVSPSEWAQ               | APFVT                 | DQQDDWEW             | APINLYP              | VPIINL                  | APIIVWE                 | APIIPEE                  |
| AEWEDELEEM           | AXXKXKDRS                | VPIDT                 | EAMWDFIDN            | APINVLI              | VPIISVT                 | APIIPINI                | APIIPEE                  |
| AVGWNNEDW            | AGWDMWDA                 | VPVNA                 | FPDFFDNF             | APINVF               | VPIFTN                  | APIIPINI                | APIIPEE                  |
| AMFFEDSW             | AFEWWEWLSD               | VPIVLD                | INRSMVED             | APIIPM                | VPIWQP                  | APIIPINI                | APIIPEE                  |
| APIVTDMA             | AEDWDWDDWE               | VPTAA                 | KEWEDEDF             | APIIQD               | VPVWHT                  | APILMDN                 | APIIRTX                  |
| APIIWPEDB            | ATWENNDF                 | VPSMF                 | MFEEOQWF           | APLNPF               | VPQSWG                  | APIREDX                 | APIIPEE                  |
| APIQVFQSE            | AMETEFPF                 | VPSW               | VSVDSVDWE            | APLFAG               | VPQHAF                  | APSWERI                 | APIIPEE                  |
| APIQVIEEI            | AGWDLWDD                  | VPGHD                 | VSWEWEVD            | AMMPDF               | VPSQV                  | APIIEMB                 | APIIPEE                  |
| APIARNLM             | ASEWGRADI                | VPFDF                 | QAVaweVD            | APIVRI               | VPTFQP                  | APIVDFD                 | APIIPEE                  |
| APIMWWDDV            | AIDLEWQDI                | VQQNV                 | VPTQNWMD             | APIVVT               | VALQES                  | APIVDFD                 | APIIPEE                  |
| APFGPSWGP            | ADSWELWEI                | TQMG                  | VEEDEQWS             | APIFNL               | VVREWQ                  | APIVDFD                 | APIIPEE                  |
| APRFKTLSA            | AWDQDMWDE               | TDAHG                 | VRRSDPA             | ARFQWQ               | VXWAX                   | APIVDFD                 | APIIPEE                  |
| AFRRKNFRI            | AFSFDFDEI               | TVGDSD                | VVQRGSFL             | AIIWXX               | VFRVNP                  | APIVDFD                 | APIIPEE                  |
| AENELAWI             | AFRKRNFRI                | AKRRFR                | VKMNAPTS             | VQHPN                 | APQNEF                  | APIVDFD                 | APIIPEE                  |
| AEKDWDWI             | AENELAWI                | VAFVRVNM             | VQRLANPS             | VTEHTG               | APIA                  | APIVDFD                 | APIIPEE                  |
| APANWDEL            | AETEFPF                 | VRDSRFS             | VSDFIG               | VTEFG                | APIA                  | APIVDFD                 | APIIPEE                  |
| AEWIDNMD            | APSVNDQDM                | VDRFFKT               | VTEFHT              | APIA                  | APIA                  | APIVDFD                 | APIIPEE                  |
| AFDGDWWQ             | AEWIDNMD                | WIDRXX                | VTIAGA              | APIA                  | APIA                  | APIVDFD                 | APIIPEE                  |
| AERDFEFDN            | AFDGDWWQ                | XEIFSFD               | VSDDEG               | APIA                  | APIA                  | APIVDFD                 | APIIPEE                  |
| ARRRKNNF             | AERDFEFDN                | XFPGEIEP             | XVNRIWS              | APIA                  | APIA                  | APIVDFD                 | APIIPEE                  |
| AKGRNFKNF            | AKGRNFKNF               | XNIRIF                 | XVPIATGF            | APIA                  | APIA                  | APIVDFD                 | APIIPEE                  |
| AKGRNFKNF            | AKGRNFKNF               | XNIRIF                 | XVPIATGF            | APIA                  | APIA                  | APIVDFD                 | APIIPEE                  |
| AWFGRKFP             | AWFGRKFP                | XNIRIF                 | XVPIATGF            | APIA                  | APIA                  | APIVDFD                 | APIIPEE                  |
| AFDGSEWQ            | AFDGSEWQ                | XNIRIF                 | XVPIATGF            | APIA                  | APIA                  | APIVDFD                 | APIIPEE                  |
| AFDGSEWQ            | AFDGSEWQ                | XNIRIF                 | XVPIATGF            | APIA                  | APIA                  | APIVDFD                 | APIIPEE                  |
| AFDGSEWQ            | AFDGSEWQ                | XNIRIF                 | XVPIATGF            | APIA                  | APIA                  | APIVDFD                 | APIIPEE                  |
| AFDGSEWQ            | AFDGSEWQ                | XNIRIF                 | XVPIATGF            | APIA                  | APIA                  | APIVDFD                 | APIIPEE                  |
| AFDGSEWQ            | AFDGSEWQ                | XNIRIF                 | XVPIATGF            | APIA                  | APIA                  | APIVDFD                 | APIIPEE                  |
| AFDGSEWQ            | AFDGSEWQ                | XNIRIF                 | XVPIATGF            | APIA                  | APIA                  | APIVDFD                 | APIIPEE                  |
| AFDGSEWQ            | AFDGSEWQ                | XNIRIF                 | XVPIATGF            | APIA                  | APIA                  | APIVDFD                 | APIIPEE                  |
| AFDGSEWQ            | AFDGSEWQ                | XNIRIF                 | XVPIATGF            | APIA                  | APIA                  | APIVDFD                 | APIIPEE                  |
| AFDGSEWQ            | AFDGSEWQ                | XNIRIF                 | XVPIATGF            | APIA                  | APIA                  | APIVDFD                 | APIIPEE                  |

* Peptides selected for further kinetic analysis.
Ala-Pro, or Ala-Pro-Ile to its N terminus to give libraries VI–IX, respectively (Table 1). The four libraries (~120,000 beads each) were screened against VHR under the same conditions. We noted that the positive beads from libraries VI and VII developed reddish color faster and of a greater intensity than those in libraries VIII and IX, although the former had a smaller number of colored beads due to their lower probability of having sequences containing the (V/A)PI combination. With either Ala or Val at the Tyr(P) – 6 position, the sequences selected from libraries VI and VII contained predominantly Pro at the Tyr(P) – 5 position and hydrophobic residues, most frequently Ile, at the Tyr(P) – 4 position (Table 3). Similarly, most of the peptides selected from library VIII had Ile or other hydrophobic residues (e.g. Nle, Leu, Val, and Phe) at the Tyr(P) – 5 position. On the other hand, the positive sequences selected from library IX did not show any obvious trend except for the frequent selection of Asp, Glu, Ser, and Thr at the Tyr(P) – 2 position. The selection of charged residues throughout the random region is reminiscent of the class I sequences derived from libraries I and II (Table 2). Taken together, the library screening results suggest that VHR contains a second binding site that specifically recognizes the peptide motif NH2-(V/A)P(I/L/M/V/F), and for optimal binding to VHR, this motif should be separated from the Tyr(P) residue by one to six amino acid residues.

**VHR Substrate Specificity (C-terminal to Tyr(P))**—To determine the substrate specificity of VHR on the C-terminal side of Tyr(P), we first screened it against library X, IDFD(D/P)pXY-XXXXNNBBRM (Table 1), which contains a preferred class I sequence of VHR on the N-terminal side of Tyr(P). Analysis of the selected sequences revealed several features. First, the results confirmed the preference of VHR for a Pro residue at the Tyr(P) – 1 position; 39 of the 43 peptides (91%) had a Pro at this position (supplemental Table S1 and Fig. 2), whereas Asp and Pro were equally populated in the original library X (50% each). Second, VHR prefers a small (Gly, Ala, and Ser) or to a lesser extent a hydrophilic residue (Gln, Glu, and Arg) at the Tyr(P) + 1 position. It does not show significant selectivity beyond the Tyr(P) + 1 position, although basic residues are dramatically under-represented (only one Lys in 43 sequences). To test whether the preferred N-terminal sequence biased the C-terminal sequences, we also screened VHR against a library containing five random residues on either side of Tyr(P), NH2-AXXXXPYXXXXNNBBRM (library XI; Table 1). The data confirmed the preference of VHR for an N-terminal AP(I/V/L) motif and broad specificity on the C-terminal side of Tyr(P) (supplemental Table S2).

**Kinetic Properties of Selected Peptides**—To confirm the library screening results, representative peptides selected from the peptide libraries were individually synthesized, and their kinetic properties against VHR were determined in solution and compared with those of previously reported peptide substrates. Peptide Ac-VSEDPpYAASAS (Table 4, peptide 1), a class I sequence selected from library I, has a $k_{cat}$ value of 4.5 s$^{-1}$, a $K_m$ value of 117 μm, and a $k_{cat}/K_m$ value of 3.8 × 10$^8$ M$^{-1}$ s$^{-1}$. Consistent with the screening data, substitution of Ser for Asp at the Tyr(P) – 2 position of peptide 1 had minimal effect on the kinetic activity, whereas phosphorylation of the Tyr(P) – 2 Ser reduced the activity by ~3-fold (Table 4, compare peptides 1–3). The latter observation is in contrast to previous reports that bisphosphorylated MAPK peptides (pTgP) were 2–3-fold more active than the monophosphorylated peptides (TXpY) (15, 24). Given the relatively small effect upon phosphorylation of the Tyr(P) – 2 residue, we believe that VHR readily accepts both (pS/pT)XPY (where pS is phosphoserine) and (S/T)XPY proteins as substrates, and the effect of phosphorylation is sequence-dependent. The importance of Pro at the Tyr(P) – 1 position is highlighted by the 2.5-fold decrease in activity when the Pro was replaced by Ile, the second most frequently selected amino acid at this position, and the 24-fold activity reduction upon replacement by Lys (which was not selected during screening) (Table 4, compares peptides 1, 4, and 5).

Replacement of the C-terminal AASAS sequence of peptide 1 with acidic motifs (YDFAS and DFEAS) resulted in a slight reduction in activity (Table 4, compare peptides 1, 9, and 10). Phosphorylated Erk and Jnk have a $k_{cat}/K_m$ value of 6.0 × 10$^5$ M$^{-1}$ s$^{-1}$. In comparison, control peptides Ac-SASpYSASAS and Ac-ARKIpYAA, which were not selected from the library, showed 100–500-fold lower activity (Table 4, compare peptides 1, 9, and 10). Phosphorylated Erk and Jnk have a $k_{cat}/K_m$ value of ~4 × 10$^5$ M$^{-1}$ s$^{-1}$ (14, 15), whereas the Tyr(P)694 peptide of STAT5 (peptide 11) has a second-order rate constant of 6.2 × 10$^3$ M$^{-1}$ s$^{-1}$. These results demonstrate that both class I and II peptides are efficient VHR substrates and that the active site of VHR displays a high level of sequence selectivity.

The structural determinants of class II substrates were further assessed by synthesizing and assaying peptides 12–24, which contained a free versus acetylated N terminus, Val versus Ala as the N-terminal residue, and different numbers of intervening residues between the (V/A)PI motif and Tyr(P) (Table 5). The results show that VHR strongly prefers a (V/A)PI motif with a free N-terminal amine; acetylation of the N terminus reduces the catalytic efficiency by 10–15-fold (compare peptides 12 and 13; peptides 14, 15, and 16; and peptides 17, 18, and 19). An N-terminal Val generally gives slightly higher activity than the corresponding APX peptides (1.5–2-fold). Finally, the kinetic activity increased as the number of intervening residues between the (V/A)PI motif and Tyr(P) decreased, reaching the
maximal activity at two residues (compare peptides 20–24). Further reduction in the distance decreased the activity. To our knowledge, peptide 22 is the most active VHR substrate reported to date ($k_{cat}/K_m$ value of $1.0 \times 10^6$ M$^{-1}$ s$^{-1}$).

**Class II Substrates Bind VHR in an Alternative Mode**—The fact that VHR is able to recognize class II substrates containing a varying number of intervening residues between the (V/A)PI and Tyr(P) motifs suggests that VHR has a separate binding site for the (V/A)PI motif. To locate the secondary binding site, we initially attempted to co-crystallize the catalytically inactive C124S mutant VHR with several class II substrates. Unfortunately, all of our efforts were unsuccessful. We therefore resorted to site-directed mutagenesis to map the (V/A)PI-binding site on VHR. Because a free N terminus increases the reactivity by 10–15-fold primarily through improved binding (lower $K_m$ value), we reasoned that the secondary binding site may contain at least one acidic residue that engages in charge-charge interactions with the positively charged N terminus of the substrate. Inspection of the co-crystal structure of VHR bound with peptide DDE(Nle)TGPYYVATR (24) identified four acidic residues within a 24-Å distance from the active site. Among these acidic residues, Asp$^{18}$, Asp$^{47}$, and Glu$^{159}$ residues are located near the binding surfaces for the N-terminal portion of the peptide substrate, whereas Asp$^{164}$ interacts with the C-terminal fragment by forming a hydrogen bond with the Thr residue at the Tyr(P) + 3 position. These four residues were individually mutated to the corresponding Asn or Gln, and the kinetic properties of the resulting VHR mutants were assessed against p-nitrophenyl phosphate and class II substrates containing both free and acetylated N termini.

The D47N, E159Q, and D164N mutants had essentially the same catalytic activity as WT VHR toward p-nitrophenyl phosphate, whereas the D18N mutant was ~2-fold less active due to a lower $k_{cat}$ value (Table 5). Thus, mutation of residue 47, 159, or 164 does not significantly affect the active site structure of VHR. When peptide NH$_2$-APINDIpYAA (peptide 18, which contains a free N terminus) was used as the substrate, WT VHR and the D18N, D47N, and E159Q mutants had similar activities with $K_m$ values of 15–30 μM and $k_{cat}/K_m$ values of 1.2–5.0 × 10$^5$ M$^{-1}$ s$^{-1}$. In contrast, the D164N mutant was 8.8-fold less active due to its increased $K_m$ value (170 μM instead of 23 μM for WT).

After peptide 18 was acetylated at the N terminus (Ac-APINDIpYAA), all five enzyme variants had similar $k_{cat}/K_m$ values ($0.47–1.9 \times 10^6$ M$^{-1}$ s$^{-1}$). Note that N-acetylation reduced the activity toward WT, D18N, D47N, and E159Q VHR by 15–55-fold but only by 3.8-fold for the D164N mutant. The same observation was made with another class II substrate, NH$_2$-APFPQDIpYAA (peptide 15). These data indicate that in the E5 complex, the N terminus of the class II substrates is near Asp$^{164}$; to make this possible, the peptide substrate must bind to VHR in an orientation opposite to what has been reported for class I substrates (24). To further illustrate the different binding modes of the two classes of substrates, we tested WT and D164N VHR against a class I peptide that bears a free N termi-
Thus, eight peptide comparison with the VHR co-crystal structure are shown in Fig. 5. The relaxed structures of peptide 22 bound to VHR in both the canonical and reverse binding modes and their comparison with the VHR co-crystal structure are shown in Fig. 3. In addition, atomic interactions involved in the active site of VHR are presented in Fig. 3D. These snapshots demonstrate that the key interactions between the peptides and protein exist primarily at the active site pocket with some conformations showing H-bonding interactions with the Asp164 residue (supplemental Fig. S3). This observation holds true only for ligands in the reverse orientation and for approximately 12–22% of the conformational ensembles (supplemental Table S3). In other words, binding is not significantly affected by other amino acids in the peptide. This is consistent with the experimental kinetic studies where the $K_m$ values of peptides 13, 15, 16, and 18–24 remained remarkably unaltered (Table 5).

Computed MM-PBSA and MM-GBSA energies for different binding modes are summarized in supplemental Table S4. The binding energies of peptides 21 and 22 are more favorable in the reverse mode (N-terminus pointed toward Asp164) than the canonical binding mode (when the Pro residue is in trans configuration), which is in agreement with the experimental observations. Furthermore, MD simulations show that both peptides prefer the trans configuration of the Pro residue over a cis-proline. Although both peptides prefer the reverse binding mode, modeling results do not favor one peptide over the other in agreement with the experimental observation that peptides 21 and 22 have rather similar $K_m$ values (Table 5).

In addition to the MD simulations with wild-type VHR, we also carried out MD simulations on the D164N mutant. The mutation was performed using the UCSF Chimera software suite and was subjected to the same simulation protocol as that used for the wild-type enzyme. There was no unusual structural change upon D164N mutation as shown in the root mean square deviation plot (supplemental Fig. S4). This suggests that there are no significant deviations in the binding modes of the ligands upon D164N mutation.

In Vitro Dephosphorylation of Class II Protein Substrates by VHR—To examine the physiological relevance of the class II substrates, we searched the human proteome for proteins with a free N-terminal (V/A)P(I/L/M/V/F) motif and a Tyr(P) residue near the N terminus. According to the PhosphoSite database, no human protein contains such an N-terminal sequence as most of the eukaryotic proteins are N-terminally acetylated (41). We next searched the database for proteins containing internal (V/A)P(I/L/M/V/F)$\alpha$-pY motifs as potential VHR

### VHR Substrate Specificity

**Table 6** Activity of WT and mutant VHR against Tyr(P) peptides

| Peptide | Sequence | VHR | $k_{cat}$ | $K_m$ | $k_{cat}/K_m$ |
|---------|----------|-----|-----------|-------|--------------|
| pNPP    |          | WT  | $1.7 \pm 0.1 \times 10^6$ | 5.200 ± 300 | 0.32 |
|         |          | D18N| $1.0 \pm 0.1 \times 10^6$ | 6.700 ± 400 | 0.15 |
|         |          | D47N| $2.3 \pm 0.1 \times 10^6$ | 5.600 ± 400 | 0.40 |
|         |          | E159Q | $2.1 \pm 0.1 \times 10^6$ | 7.300 ± 500 | 0.29 |
|         |          | D164N| $1.8 \pm 0.1 \times 10^6$ | 5.600 ± 400 | 0.31 |
|         |          | WT   | $6.3 \pm 0.4 \times 10^6$ | 23 ± 0.4 | 290 |
|         |          | D18N| $2.6 \pm 0.1 \times 10^6$ | 20 ± 0.1 | 120 |
|         |          | D47N| $7.3 \pm 1.0 \times 10^6$ | 15 ± 1.9 | 500 |
|         |          | E159Q | $4.9 \pm 0.4 \times 10^6$ | 30 ± 5.3 | 170 |
|         |          | D164N| $5.5 \pm 1.1 \times 10^6$ | 170 ± 34 | 33 |
| 18      | NH$_2$APINDIpYAA-NH$_2$ | WT  | $8.2 \pm 0.1 \times 10^6$ | 420 ± 8 | 19 |
|         |          | D18N| $100$ | $>100$ | 4.7 |
|         |          | D47N| $100$ | $>100$ | 9.0 |
|         |          | E159Q | $>100$ | $>100$ | 5.8 |
| 15      | NH$_2$APFPQDpYAA-NH$_2$ | WT   | $8.6 \pm 0.6 \times 10^6$ | 35 ± 0.1 | 270 |
|         |          | D164N| $3.0 \pm 0.2 \times 10^6$ | 220 ± 27 | 14 |
| 14      | Ac-ApFPQDpYAA-NH$_2$ | WT   | $7.6 \pm 0.4 \times 10^6$ | 360 ± 40 | 21 |
| 25      | NH$_2$VSEDpYAA-NH$_2$ | WT   | $6.0 \pm 0.2 \times 10^6$ | 630 ± 48 | 9.5 |
|         |          | D164N| $4.9 \pm 0.1 \times 10^6$ | 770 ± 50 | 6.5 |

**Note:** All values are in microMolar ($\mu$M).
substrates. Table 7 lists some of the potential substrates, which have been experimentally demonstrated to be phosphorylated at the putative class II VHR sites in vitro and/or in vivo. Two of the proteins, LASP1 (LIM and SH3 domain protein 1) and ROBO1, were selected for further study because their putative class II VHR sites (TSAPVpY171QQPQQ and VAPVQpY1114NIVEQ, respectively) are known to be phosphorylated by Abl kinase (42, 43), which is commercially available. Indeed, treatment of recombinant ROBO1 and LASP1 with Abl resulted in efficient tyrosyl phosphorylation of both proteins (Fig. 4, A and B). Protease digestion of the resulting phospho-ROBO1 followed by LC-MS/MS analysis confirmed the phosphorylation of Tyr1114 (Fig. 4C and supplemental Table S5). Due to lack of proper protease cleavage sites near LASP1 Tyr171, confirmation of Tyr171 phosphorylation by LC-MS/MS was unsuccessful. Western blot analysis with anti-Tyr(P) antibody showed that both ROBO1 and LASP1 were efficiently and completely dephosphorylated by VHR (Fig. 4, A and B). Thus, VHR is able to dephosphorylate class II motifs in intact proteins. Consistent with this observation, peptides corresponding to the Tyr(P)171 and Tyr(P)1114 sites were efficient VHR substrates with $k_{cat}/K_m$ values of $4.1 \times 10^3$ and $1.9 \times 10^3 \text{M}^{-1} \text{s}^{-1}$, respectively (Table 4).

**DISCUSSION**

The lack of any recognizable substrate-recruiting domain/ surface by the atypical dual specificity phosphatases implies that their PTP active site must possess substantial sequence selectivity to selectively dephosphorylate their cognate substrates. Screening of combinatorial peptide libraries against VHR, a representative member of the subgroup, reveals that VHR indeed has significantly narrower sequence specificity than the classical PTPs (3–6). The selected sequences, which should represent the most reactive substrates of VHR, exhibit remarkable sequence covariance. On the basis of sequence similarity, the peptide substrates can be divided into two major classes (classes I and II). Within the class I substrates, the peptides can be further separated into two subclasses with the class IA and IB substrates having sequences of (E/D/ϕ)(D/S/N/T/E)(P/I/M/S/A/V)pY(G/A/S/Q) and (E/D/ϕ)(T/S)(D/E)pY(G/A/S/Q) (where ϕ is a hydrophobic residue), respectively. To date, five protein substrates have been reported for VHR, and
the Tyr(P) sites dephosphorylated by VHR contain the following sequences: Erk, GFLpTEpYVAT (14); Jnk, FMMpTPpYVVT (15); STAT5, KAVDGpYVKP (18); epidermal growth factor receptor, VDADEpYLIP (19); and ErbB2, IDETEpYAAD (19). These in vivo VHR substrate motifs largely match the consensus sequences derived from our library screening. Two of the proteins, phosphorylated Erk and Jnk, have been purified to homogeneity and assayed against recombinant VHR in vitro (14, 15, 17). Both protein substrates had a $k_{cat}/K_m$ value of $\sim 4 \times 10^4$ M$^{-1}$ s$^{-1}$. Short peptides corresponding to the pTXpY
motifs of these two proteins gave similar $k_{cat}/K_m$ values (23, 24), suggesting that most of the interactions between VHR and these proteins are mediated through the VHR active site. On the other hand, the STAT5 Tyr(P)694 peptide is a substantially poorer substrate in vitro ($k_{cat}/K_m$ value of $6.2 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$). It has been proposed that the activity of VHR toward STAT5 in vivo is enhanced by additional interactions between Tyr(P) of VHR and the STAT5 SH2 domain (18, 21). Note that the class I peptides selected from the libraries generally have $k_{cat}/K_m$ values of $1–4 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$ (Table 4), which is in the same range as the protein substrates. These data suggest that the sequence specificity of the VHR catalytic site is a key determin-ant of its substrate specificity in vivo. The relatively low intrin-sic catalytic efficiency (defined as the $k_{cat}/K_m$ value of a PTP active site toward its optimal peptide substrate) of VHR toward class I substrates as compared with other PTPs (e.g. PTP1B, SHP-1, and SHP-2) have intrinsic catalytic efficiencies of $>10^7 \text{ M}^{-1} \text{ s}^{-1}$ (6)) suggests that the activity of VHR toward its phys-iological substrates is likely enhanced by substrate recruiting strategies. Nevertheless, the specificity profile of VHR should be useful for identifying additional protein substrates of VHR.

Our screening result (class I substrates) is consistent with a previous structural study of VHR bound with a class I substrate, DDE(Nle)pTGpYVATR (24). In the canonical binding mode, the Tyr(P) side chain occupies the relatively shallow active site pocket, making interactions with Arg130 and the amide nitro-gen of residues 125–130 (Fig. 3A). The Thr(P) side chain of the pTXpY motif points into a nearby basic pocket containing Arg128. It is expected that the frequently selected Asp at the Tyr(P) – 2 position would bind to this pocket and electrostat-ically interact with Arg128. The Tyr(P) + 1 Val stacks over VHR residue Tyr128, providing an explanation for the preference of the enzyme for a small residue (i.e. Gly, Ala, or Ser) at the Tyr(P) + 1 position as larger side chains would create steric clashes with Tyr128. The acidic residues at Tyr(P) – 3 to Tyr(P) – 5 positions make contacts with the positively charged protein surfaces, consistent with the selection of acidic and hydrophobic residues at these positions (Fig. 1). Finally, in the canonical mode, Asp164 of VHR forms a hydrogen bond with the side chain of Thr at the Tyr(P) + 3 position.

A surprising finding of this study is that VHR also recognizes a second class of Tyr(P) peptides (class II). The class II sub-strates have entirely different sequences, and all contain a (V/A)p(I/L/M/V/F) motif, preferably at the N terminus of the peptide. Mutagenesis and modeling studies suggest that the class II peptides bind to VHR in an orientation opposite to that of the class I substrates (Fig. 3C). In this alternative binding mode, the Tyr(P) side chain binds to the active site pocket, presumably making similar contacts with the active site resi-dues as in the canonical mode (Fig. 3A). The VPI motif binds to a shallow pocket near Asp164, engaging in charge-charge inter-actions between the peptide N-terminal ammonium ion and the carboxylate side chain of Asp164. The bipartite binding mode permits the intervening residues to form a loop, which may or may not directly contact the protein surface. This scenario explains how VHR can accommodate intervening sequences of varying amino acid composition and lengths (at least one to six amino acids). It is remarkable that the class II substrates are generally 1–2 orders of magnitude more reactive than the class I peptides. It should be noted that although our database search did not identify any intact protein containing the N-terminal class II motifs it is possible for some proteins to undergo proteolytic processing at internal positions to expose an N-terminal NH$_2$-p(Y/A)(I/L/M/V/F) motif and subsequently serve as VHR substrates (44–47). The ROBO1 and LASP1 results suggest that proteins containing internal (V/A)p(I/L/M/V/F)X$_n$-pY motifs may also act as in vivo VHR substrates as N-terminally acylated class II peptides are still efficient VHR substrates, often more reactive than the most efficient class I peptides (e.g. peptide 8 in Table 4).

REFERENCES

1. Tigani, T., and Bennett, A. M. (2007) Protein tyrosine phosphatase func-tion: the substrate perspective. Biochem. J. 402, 1–15
2. Tonks, N. K., and Neel, B. G. (2001) Combinatorial control of the specific-ity of protein tyrosine phosphatases. Curr. Opin. Cell Biol. 13, 182–195
3. Zhang, Z.-Y., Maclean, D., McNamara, D. J., Sawyer, T. K., and Dixon, J. E. (1994) Protein tyrosine phosphatase substrate specificity: size and phos-photyrosine positioning requirements in peptide substrates. Biochemistry 33, 2285–2290
4. Vetter, S. W., Keng, Y.-F., Lawrence, D. S., and Zhang, Z.-Y. (2000) As-sessment of protein-tyrosine phosphatase 1B substrate specificity using “inverse alanine scanning.” J. Biol. Chem. 275, 2265–2268
5. Wälchli, S., Espanel, X., Harrenga, A., Rossi, M., Cesareni, G., and Hooft van Huijstdijnen, R. (2004) Probing protein-tyrosine phosphatase sub-strate specificity using a phosphotyrosine-containing phage library. J. Biol. Chem. 279, 311–318
6. Ren, L., Chen, X., Luechapanichkul, R., Sehner, N. G., Meyer, T. M., Wavreille, A.-S., Chan, R., Iorio, C., Zhou, X., Neel, B. G., and Pei, D. (2011) Substrate specificity of protein tyrosine phosphatases 1B, RPTPα, SHP-1, and SHP-2. Biochemistry 50, 2339–2356
7. Bayon, Y., and Alonso, A. (2010) in Emerging Signaling Pathways in Tumor Biology (Lazo, P. A., ed) pp. 185–208, Transworld Research Network, Kerala, India
8. Alonso, A., Sasin, J., Bottini, N., Friedberg, I., Friedberg, I., Osterman, A., Godzik, A., Hunter, T., Dixon, J., and Mustelin, T. (2004) Protein tyrosine phosphatases in the human genome. Cell 117, 699–711
9. Ishibashi, T., Bottaro, D. P., Chan, A., Miksi, T., and Aaronson, S. A. (1992) Expression cloning of a human dual-specificity phosphatase. Proc. Natl. Acad. Sci. U.S.A. 89, 12170–12174
10. Rahmouni, S., Cerignoli, F., Alonso, A., Tsutji, T., Henkens, R., Zhu, C., Louis-Sully, C., Moutschen, M., Jiang, W., and Mustelin, T. (2006) Loss of the VHR dual-specific phosphatase causes cell-cycle arrest and senescence. Nat. Cell Biol. 8, 524–531
11. Arnoldussen, Y. J., Lorenzo, P. I., Pretorius, M. E., Waehre, H., Risberg, B., Maelandsmo, G. M., Danielson, H. E., and Saatcioglu, F. (2008) The mito-gen-activated protein kinase phosphatase vaccinia H1-related protein inhib-its apoptosis in prostate cancer cells and is overexpressed in prostate cancer. Cancer Res. 68, 9255–9264
12. Hao, L., and ElShamy, W. M. (2007) BRCAl-IRIS activates cyclin D1 ex-pression in breast cancer cells by downregulating the INK phosphatase DUSP3/VHR. Int. J. Cancer. 121, 39–46
13. Henkens, R., Delvenne, P., Araf, M., Moutschen, M., Zeddou, M., Tauz, L., Boniver, J., Mustelin, T., and Rahmouni, S. (2008) Cervix carcinoma is associated with an up-regulation and nuclear localization of the dual-specificity protein phosphatase VHR. BMC Cancer 8, 147–155
14. Todd, J. L., Tanner, K. G., and Denu, J. M. (1999) Extracellular regulated kinases (ERK) 1 and ERK2 are authentic substrates for the dual-specificity protein-tyrosine phosphatase VHR. A novel role in down-regulating the ERK pathway. J. Biol. Chem. 274, 13271–13280
15. Todd, J. L., Rigas, J. D., Rafy, L. A., and Denu, J. M. (2002) Dual-specificity protein tyrosine phosphatase VHR down-regulates c-Jun N-terminal ki-nase (INK). Oncogene 21, 2573–2583
16. Alonso, A., Saxena, M., Williams, S., and Mustelin, T. (2001) Inhibitory
role for dual specificity phosphatase VHR in T cell antigen receptor and CD28-induced Erk and Ink activation. J. Biol. Chem. 276, 4766–4771
17. Zhou, B., Wang, Z.-X., Zhao, Y., Brautigan, D. L., and Zhang, Z.-Y. (2002) The specificity of extracellular signal-regulated kinase 2 dephosphorylation by protein phosphatases. J. Biol. Chem. 277, 31818–31825
18. Hoyt, R., Zhu, W., Cerignoli, F., Alonso, A., Mustelin, T., and David, M. (2007) Cutting edge: selective tyrosine dephosphorylation of interferon-activated nuclear STAT5 by the VHR phosphatase. J. Immunol. 179, 3402–3406
19. Wang, J.-Y., Yeh, C.-L., Chou, H.-C., Fu, Y.-N., Chen, Y.-T., Cheng, H.-W., Huang, C.-Y., Liu, H.-P., Huang, S.-F., and Chen, Y.-R. (2011) Vaccinia H1-related phosphatase is a phosphatase of ErbB receptors and is down-regulated in non-small cell lung cancer. J. Biol. Chem. 286, 10177–10184
20. Alonso, A., Rahmouni, S., Williams, S., van Stipdonk, M., Jaroszewski, L., Godzik, A., Abraham, R. T., Schoenberger, S. P., and Mustelin, T. (2003) Tyrosine phosphorylation of VHR phosphatase by ZAP-70. Nat. Immunol. 4, 44–48
21. Jardin, C., and Sticht, H. (2012) Identification of the structural features that mediate binding specificity in the recognition of STAT proteins by dual-specificity phosphatases. J. Biomol. Struct. Dyn. 29, 777–792
22. Kang, T.-H., and Kim, K.-T. (2006) Negative regulation of ERK activity by VRK3-mediated activation of VHR phosphatase. Nat. Cell Biol. 8, 863–869
23. Denu, J. M., Zhou, G., Wu, L., Zhao, R., Yuwaniyama, J., Saper, M. A., and Dixon, J. E. (1995) The purification and characterization of a human dual-specific protein tyrosine phosphatase. J. Biol. Chem. 270, 3796–3803
24. Schumacher, M. A., Todd, J. L., Rice, A. E., Tanner, K. G., and Denu, J. M. (2002) Structural basis for the recognition of a bisphosphorylated MAP kinase peptide by human VHR protein phosphatase. Biochemistry 41, 3009–3017
25. Gopishetty, B., Ren, L., Waller, T. M., Wavreille, A. S., Lopez, M., Thakkar, A., Zhu, J., and Pei, D. (2008) Synthesis of 3,5-difluorotyrosine-containing peptides: application in substrate profiling of protein tyrosine phosphatases. Org. Lett. 10, 4605–4608
26. Chen, X., Ren, L., Kim, S., Carpino, N., Daniel, J. L., Kunapuli, S. P., Tsyganok, A. Y., and Pei, D. (2010) Determination of the substrate specificity of protein-tyrosine phosphatase TULA-2 and identification of Syk as a TULA-2 substrate. J. Biol. Chem. 285, 31268–31276
27. Li, S., and Zeng, D. (2007) Chemoenzymatic enrichment of phosphotyrosine-containing peptides. Angew. Chem. Int. Ed. Engl. 46, 4751–4753
28. Thakkar, A., Wavreille, A. S., and Pei, D. (2006) Traceless capping agent for peptide sequencing by partial Edman degradation and mass spectrometry. Anal. Chem. 78, 5935–5939
29. Zhang, Z.-Y., Maclean, D., Thieme-Selser, A. M., Roese, R. W., and Dixon, J. E. (1993) A continuous spectrophotometric and fluorometric assay for protein tyrosine phosphatase using phosphotyrosine-containing peptides. Anal. Biochem. 211, 7–15
30. Petersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612
31. Dolinsky, T. J., Nielsen, J. E., McCammon, J. A., and Baker, N. A. (2004) PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. Nucleic Acids Res. 32, W665–W667
32. Mahoney, M. W., and Jorgensen, W. L. (2000) A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. J. Chem. Phys. 112, 8910–8922
33. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. 79, 926–993
34. Duan, Y., Wu, C., Chowdhury, S., Lee, M. C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., Caldwell, J., Wang, J., and Kollman, P. (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. J. Comput. Chem. 24, 1999–2012
35. Sharp, S. A., and Honig, B. (1994) Accurate calculation of hydration free-energies using macroscopic solvent models. J. Phys. Chem. 98, 1978–1988
36. Luo, R., David, L., Gilson, M. K. (2002) Accelerated Poisson-Boltzmann calculations for static and dynamic systems. J. Comput. Chem. 23, 1244–1253
37. Fogolari, F., Brigo, A., and Molinari, H. (2003) Protocol for MM/PBSA molecular dynamics simulations of proteins. Biophys. J. 85, 159–166
38. Sigalov, G., Onufriev, P., and Scheffel, P. (2005) Incorporating variable dielcetric environments into the generalized Born model. J. Chem. Phys. 122, 94511–94951
39. Garaud, M., and Pei, D. (2007) Substrate profiling of protein tyrosine phosphatase PTP1B by screening a combinatorial peptide library. J. Am. Chem. Soc. 129, 5366–5367
40. Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., and Olson, A. J. (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. J. Comput. Chem. 19, 1639–1662
41. Bradshaw, R. A., Brickey, W. W., and Walker, K. W. (1998) N-terminal processing: the methionine aminopeptidase and Nα-acetyltransferase families. Trends Biochem. Sci. 23, 263–267
42. Bashaw, G. J., Kidd, T., Murray, D., Pawson, T., and Goodman, C. S. (2000) Repulsive axon guidance: Abelson and Enabled play opposing roles downstream of the roundabout receptor. Cell 101, 703–715
43. Lin, Y. H., Park, Z.-Y., Lin, D., Brahmbhatt, A. A., Rio, M.-C., Yates, J. R., 3rd, and Klemke, R. L. (2004) Regulation of cell migration and survival by focal adhesion targeting of Lasp-1. J. Cell Biol. 165, 421–432
44. Srinivasa, S. M., Hegde, R., Saleh, A., Datta, P., Shiozaki, E., Chai, J., Lee, R. A., Robbins, P. D., Fernandes-Alnemri, T., Shi, Y., and Alnemri, E. S. (2001) A conserved XIAP-interaction motif in caspase-9 and Smac/DIABLO regulates caspase activity and apoptosis. Nature 410, 112–116
45. Yu, C., Fang, M., Li, Y., Li, L., and Wang, X. (2000) Smac, a mitochondrial protein that promotes cytochrome c-dependent caspase activation by eliminating IAP inhibition. Cell 102, 33–42
46. Verhagen, A. M., Silke, J., Ekert, P. G., Pakusch, M., Kaufmann, H., Connolly, L. M., Day, C. L., Tikoo, A., Burke, R., Wrobel, C., Moritz, R. L., Simpson, R. J., and Vaux, D. L. (2002) HtrA2 promotes cell death through its serine protease activity and its ability to antagonize inhibitor of apoptosis proteins. J. Biol. Chem. 277, 445–454
47. Hegde, R., Srinivasa, S. M., Datta, P., Madesh, M., Wassell, R., Zhang, Z., Cheong, N., Nejmeh, J., Fernandes-Alnemri, T., Hoshino, S., and Alnemri, E. S. (2003) The polypeptide chain-releasing factor GSPT1/eRF3 is proteolytically processed into an IAP-binding protein. J. Biol. Chem. 278, 38699–38706