THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 279, No. 6, Issue of February 6, pp. 4962–4969, 2004

Printed in U.S.A.

Structural Basis for the Specific Recognition of RET by the Dok1 Phosphotyrosine Binding Domain*

Received for publication, October 7, 2003, and in revised form, November 5, 2003
Published, JBC Papers in Press, November 7, 2003, DOI 10.1074/jbc.M311030200

Ning Shi‡§, Sheng Ye‡, Mark Bartlam‡, Maojun Yang§, Jing Wu§, Yiwei Liu‡, Fei Sun‡, Xueqing Han‡, Xiaozhong Peng§, Boqing Qiang§, Jiangang Yuan§¶, and Zihe Rao‡‖

From the ‡Laboratory of Structural Biology, Tsinghua University and National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Science, Beijing 100084 and the §National Laboratory of Medical Molecular Biology, Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences, Peking Union Medical College, National Human Genome Center, Beijing 100005, China

Dok1 is a common substrate of activated protein-tyrosine kinases. It is rapidly tyrosine-phosphorylated in response to receptor tyrosine activation and interacts with ras GTPase-activating protein and Nck, leading to inhibition of ras signaling pathway activation and the c-Jun N-terminal kinase (JNK) and c-Jun activation, respectively. In chronic myelogenous leukemia cells, it has shown constitutive phosphorylation. The N-terminal phosphotyrosine binding (PTB) domain of Dok1 can recognize and bind specifically to phosphotyrosine-containing motifs of receptors. Here we report the crystal structure of the Dok1 PTB domain alone and in complex with a phosphopeptide derived from RET receptor tyrosine kinase. The structure consists of a β-sandwich composed of two nearly orthogonal, 7-stranded, antiparallel β-sheets, and it is capped at one side by a C-terminal α-helix. The RET phosphopeptide binds to Dok1 via a surface groove formed between strand β5 and the C-terminal α-helix of the PTB domain. The structures reveal the molecular basis for the specific recognition of RET by the Dok1 PTB domain. We also show that Dok1 does not recognize peptide sequences from TrkA and IL-4, which are recognized by Shc and IRS1, respectively.

Protein-protein interactions play key roles in signal transduction. These interactions are often mediated by adapter proteins, which simultaneously associate with several kinases of a signaling pathway, forming an ordered module that permits sequential activation of each enzyme and by anchoring proteins, which are tethered to subcellular structures and localize their complement of enzymes close to their site of action. The docking protein Dok1 appears to function as an adapter. It has been identified as the highly phosphorylated 62-kDa protein that interacts with ras GTPase-activating protein in chronic myelogenous leukemia progenitor cells and v-Abl-transformed preB cells (1, 2). The expression of v-Abl or the chimeric protein p210bcr-Abl in chronic myelogenous leukemia cells has been shown to lead to constitutive Dok1 phosphorylation (1, 2). Recent studies have shown that Dok1 is a common substrate of activated protein-tyrosine kinases such as v-Abl (2), v-Src (3), BCR (4), EphRs (5), RET (6), and integrin β (7). It is rapidly tyrosine-phosphorylated in response to receptor tyrosine activation in various cell systems.

Dok1 contains an N-terminal pleckstrin homology (PH)[1] domain followed by a central phosphotyrosine binding (PTB) domain and a proline- and tyrosine-rich C-terminal tail. The PH domain is known to bind to acidic phospholids and localize proteins to the plasma membrane, whereas the PTB domain is known to mediate protein-protein interactions by binding to phosphotyrosine-containing motifs (8). The C-terminal part of Dok1 contains multiple tyrosine phosphorylation sites. When phosphorylated, they become potential docking sites for Src homology 2-containing proteins such as ras GTPase-activating protein and Nck, leading to inhibition of ras signaling pathway activation and the c-Jun N-terminal kinase (JNK) and c-Jun activation, respectively (6).

The many proteins that have been identified to contain PTB domains fall into two major groups. The first group contains PTB domains that have primary sequence similarity to the Shc PTB domain. The second group contains insulin receptor substrate (IRS)-like proteins such as IRS, Dok, and SNT/FRS2, which contain PTB domains with limited sequence similarity to the Shc PTB domain but similar binding characteristics (9). The Dok1 PTB domain belongs to the second group. It is 17% identical in sequence to the IRS PTB domain and was supposed to recognize sequences containing the NKLpY motif (10). To better understand the PTB domain specificity of Dok and the interaction between Dok1 and RET, we have determined the x-ray crystal structure of the murine Dok1 PTB domain alone and in complex with a phosphopeptide derived from RET.

MATERIALS AND METHODS

Peptide Synthesis and Binding Studies—The following peptides were synthesized by Sigma: the Shc-specific TrkA Tyr(P)-490 (Ac-HIIENPQpYFSDAGGK-NH_2), the IRS1-specific IL-4R Tyr(P)-497 (Ac-LVIAGNPApYRSGGK-NH_2), RET Tyr(P)-1062 (Ac-STWIEN-

* This work was supported by National Sciences Foundation of China (Grants 30200045 and 30221003); the National Program for Key Basic Research Project ("973" Grants G1999075602 and G1999011902); and the National High Technology Research and Development Program ("863" Grant 2002BA711A12). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 1P5T) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

¶ To whom correspondence may be addressed. Tel.: 86-10-65296411; Fax: 86-10-65240529; E-mail: yuanjiangang@pumc.edu.cn.

‖ To whom correspondence may be addressed: Laboratory of Structural Biology, School of Life Science and Engineering, Tsinghua University, Beijing 100084, P. R. China. Tel.: 86-10-6277-1493; Fax: 86-10-6277-3145; E-mail: raozh@xtal.tsinghua.edu.cn.

[1] The abbreviations used are: PH, pleckstrin homology; PTB, phosphotyrosine binding; JNK, c-Jun N-terminal kinase; IRS, insulin receptor substrate; MES, 4-morpholineethanesulfonic acid; PIPES, 1,4-piperazinediethanesulfonic acid; pY, phosphotyrosine; MEN, multiple endocrine neoplasia.

TABLE I
X-ray data collection, phasing, and refinement statistics
Numbers in parentheses correspond to the highest resolution shell (2.59–2.50 Å).

Data set	MAD data			RET peptide complex
	Peak	Edge	Remote	
Wavelength (Å)	0.9798	0.9800	0.9000	0.9000
Resolution (Å)	50–2.5	50–2.5	50–2.5	50–2.5
Space group		$P2_12_12_1$		$P2_12_12_1$
Unit Cell				
a/*b*/*c* (Å)		41.1/56.2/99.6		45.5/55.7/99.1
Reflections				
Total	52658	49486	55735	44613
Unique	8330 (780)[a]	8423 (740)[a]	8238 (670)[a]	8410 (764)[a]
Redundancy	7.2 (4.8)[a]	7.1 (3.9)[a]	7.3 (2.5)[a]	5.3 (4.2)[a]
Completeness (%)	99.0 (94.5)[a]	97.6 (89.0)[a]	96.6 (80.1)[a]	91.0 (85.7)[a]
R_{merge}	0.106 (0.348)[a]	0.102 (0.361)[a]	0.108 (0.460)[a]	0.104 (0.349)[a]
Mean I/σ(I)	6.5 (1.7)[a]	5.6 (1.6)[a]	5.4 (1.3)[a]	13.4 (2.7)[a]
Refinement statistics				
Resolution range (Å)		50.0–2.5		50.0–2.5
R_{work}/R_{free} (%)[b]		21.8/26.5		21.3/27.7
r.m.s.d.[c] from ideal values				
Bonds (Å)		0.014		0.016
Angles (°)		1.78		2.07
Number of atoms				
Protein		1689		1811
Water		16		17
Ramachandran plot				
Most favored (%)		86.7		82.8
Additionally allowed (%)		12.7		16.1

[a] $R_{merge} = \Sigma_h\Sigma_I| I_{Ih} - \langle I_h\rangle |/\Sigma_h\Sigma_I \langle I_h\rangle$, where $\langle I_h\rangle$ is the mean of the observations I_{ih} of reflection h.
[b] $R_{work} = \Sigma(\|F_p(obs)| - |F_p(calc)\|)/\Sigma|F_p(obs)|$; R_{free} = *R* factor for a selected subset (5%) of the reflections that was not included in prior refinement calculations.
[c] r.m.s.d., root mean square deviation.

KLpYGMSDGGK-NH_2) and RET Tyr-1062 non-phosphopeptide (Ac-STWIENKLYGMSDGGK-NH_2). A C-terminal GGK extension was added to each of the peptides for coupling to the CM5 chip via the lysine side chain amino group. Binding analyses of the Dok1 PTB domain and the peptides were performed using a Biosensor BIAcore instrument (BIACORE 1000) (BIAcore) according to the manufacturer's instructions. CM5 research grade sensor chips (BIAcore) were used. All buffers were filtered before use. The peptide concentration of 200 μg/ml and a contact time of 13.3 min at a flow rate of 3 μl/min gave ~200 resonance units. Three phosphopeptides were coupled to different flow cells of the CM5 chip, respectively. A reference surface was generated simultaneously under the same conditions but without peptide injection and used as a blank to correct for instrument and buffer artifacts. All measurements were conducted in HEPES-buffered saline buffer (10 mM HEPES, pH 7.4, containing 0.15 M NaCl, 3 mM EDTA, and 0.005% Tween 20) at a flow rate of 20 μl/min at 25 °C. After each measurement, the chip surface was regenerated with 5 μl of 6 M guanidine-HCl (pH 7.0) buffer at a flow rate of 10 μl/min at 25 °C. The Dok1 PTB domain was injected at variable concentrations at 20 μl/min flow rate, and binding to the peptides immobilized on the chip was monitored in real time. Response curves were prepared by subtracting the signal generated from the control flow cell. Kinetic parameters were determined using the software BIA evaluation 3.0.

Protein Expression, Purification, and Crystallization—The His-tagged murine Dok1 PTB domain (residues 154–266) was expressed and purified by Ni^{2+}-chelation chromatography. The N-terminal His tag was removed by thrombin digestion, and the protein was purified as described previously.[2] Se-Met-substituted Dok1 PTB domain was produced in the methionine auxotrophic *Escherichia coli* strain B834 (DE3) (Novagen). Crystals of Se-Met-derived Dok1 PTB domain were grown in a hanging drop by mixing 1 μl of protein solution (7 mg/ml, 10 mM MES (pH 6.5), 50 mM NaCl, 10 mM DTT) and 1 μl of reservoir solution containing 28% (v/v) polyethylene glycol 6000, 0.1 M MES (pH 6.0), 10 mM dithiothreitol (DTT). Crystals of Dok1 PTB domain in a complex with RET peptide were grown by the same method using 1 μl of protein solution (10 mg/ml, 1 mM RET peptide, 10 mM MES (pH 6.5), 50 mM NaCl, 10 mM DTT) with 1 μl of reservoir solution (30% (v/v) polyethylene glycol 6000, 0.1 M PIPES (pH 6.0), 10 mM dithiothreitol (DTT)). The resulting crystals grew after 1 week at 16 °C.

[2] N. Shi, submitted for publication.

Data Collection and Structure Determination—Data were collected from a flash-frozen crystal after soaking the crystal in a reservoir solution containing 20% (v/v) glycerol. The MAD data were collected at the BL41XU beamline at SPring-8. Three different wavelengths were used to obtain the multiwavelength anomalous diffraction data: 0.9798 Å (peak), 0.9800 Å (edge), and 0.9000 Å (remote). Data were integrated, scaled, and merged using the HKL programs DENZO and SCALEPACK (12). Crystals of the Dok1 PTB domain belong to the space group $P2_12_12_1$, with unit cell parameters a = 41.1 Å, b = 56.2 Å, c = 99.6 Å, $\alpha = \beta = \gamma = 90°$, containing two molecules per asymmetric unit. Three selenium sites were located and refined at 2.5-Å resolution using SOLVE (13), which produced a mean figure of merit of 0.32. After auto-modeling with RESOLVE (14), about 50% of all the residues were easily modeled into the experimental map. The remaining residues were traced manually with O (15). CNS (16) was used for refinement and the addition of solvent molecules.

Data from the RET peptide complex crystals were collected at a wavelength of 0.9000 Å at the BL41XU beamline at SPring-8. The crystal also belonged to the space group $P2_12_12_1$, containing two molecules per asymmetric unit, but with different unit cell parameters, a = 45.5 Å, b = 55.7 Å, c = 99.1 Å, $\alpha = \beta = \gamma = 90°$. The structure of the complex was phased by molecular replacement using CNS with the model of the free Dok1 PTB domain as the starting model. The RET phosphopeptides bound to the Dok1 PTB domain were located using an $F_o - F_c$ difference electron density map. Model building and fitting were carried out using O, and refinement and addition of water molecules were performed using CNS. Data collection, processing, and refinement statistics are given in Table I. The complex model consists of residues 154–256 of mouse Dok1, the 10 residues of the RET phosphopeptide, and 17 water molecules. Model quality was checked with PROCHECK (17).

Coordinates—Coordinates and structure factors for the Dok1 PTB domain have been deposited in the Protein Data Bank (accession number 1P5T). Coordinates and structure factors for the Dok1 PTB domain and RET peptide complex have been deposited in the Protein Data Bank (accession number 1EUF).

RESULTS AND DISCUSSION

Specificity of Phosphopeptide Binding—Affinity analysis was performed by means of surface plasmon resonance. The synthetic peptides derived from TrkA (residues 483–494), IL-4R

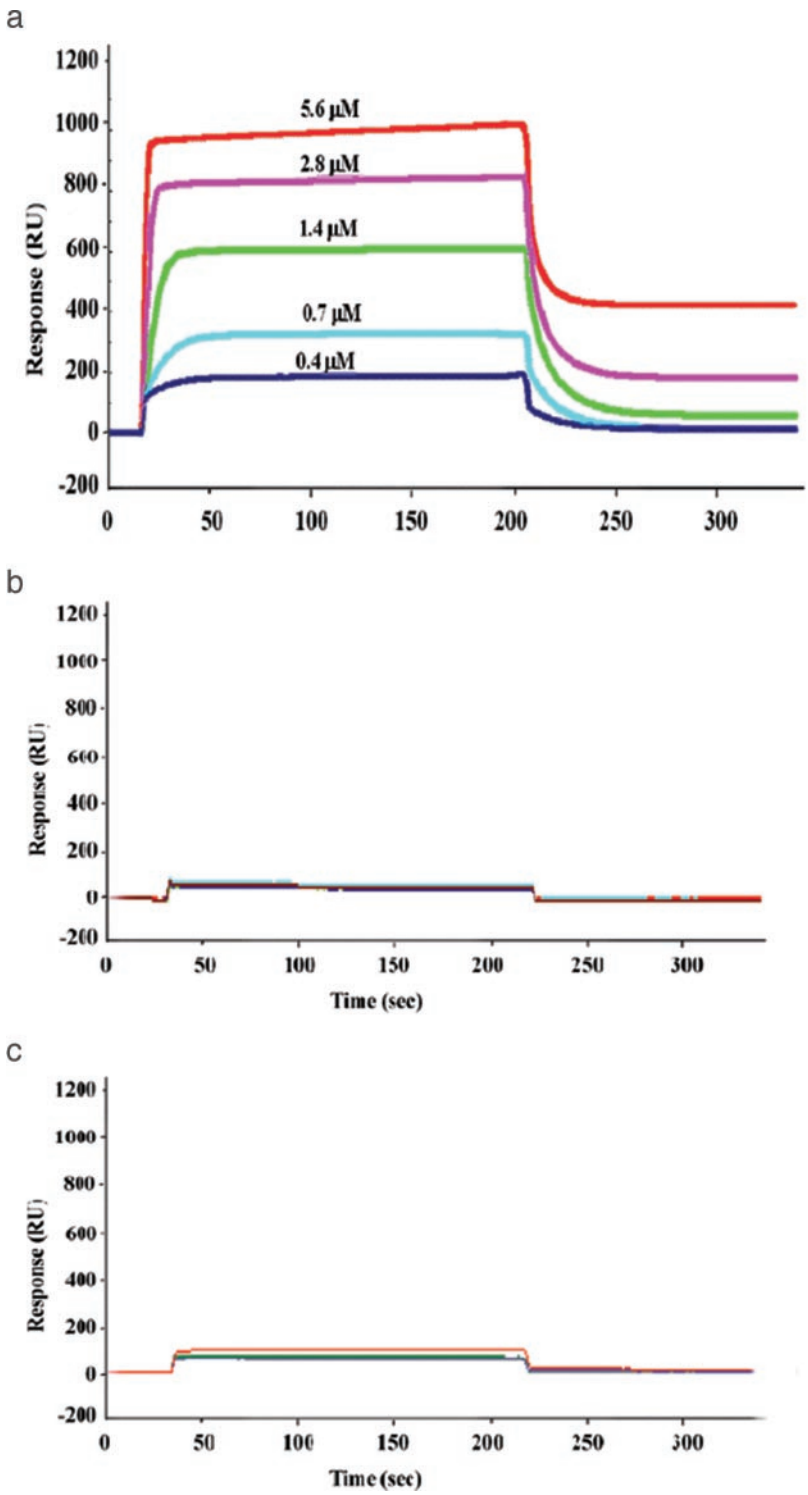

FIG. 1. **Biosensor analysis of the Dok1 PTB domain with immobilized phosphopeptides.** Five different concentrations of Dok1 PTB domain were injected over three flow cells with different phosphopeptides and the reference flow cell. The sensorgram shows the relative response in resonance units (*RU*) after background subtraction *versus* time in seconds are recorded for the following peptide: RET (*a*), TrkA (*b*), and IL-4R (*c*). The concentrations of PTB domain are indicated by *numbers* in the *corresponding graphs*.

(residues 489–499), and RET (residues 1054–1064) were coupled to the sensor chip, CM5 of BIAcore, and various concentrations of Dok1 PTB domain solutions were run over the chip. The dissociation constant (K_d) of binding of RET phosphopeptide to the Dok1 PTB domain was determined to be 3.2 μM from the data in Fig. 1*a*. Measurements made in the presence of 100–500 μM of its non-phosphorylated counterpart were unchanged. However, no binding could be detected for immobilized Trka and IL-4 peptides (Fig. 1, *b* and *c*), indicating binding specificity of the Dok1 PTB domain to the receptor.

Structural Overview—The native structure of the Dok1 PTB domain was determined by MAD phasing to 2.5-Å resolution with Se-Met derivative data. Final statistics for the structure are given in Table I. The electron density was of good quality and well defined for most of the structure. The final model consists of residues 154–256 of mouse Dok1 in chain A, residues 154–254 in chain B, and 16 water molecules. The PTB domain of Dok1 adopts a "PH domain-like" fold, with seven strands forming a β-sandwich composed of two nearly orthogonal antiparallel β-sheets (Fig. 2*a*). The β-sandwich is capped at one end by a C-terminal α-helix.

Structure of the Dok1 PTB Domain-RET Peptide Complex—To gain further insight into the molecular basis for the binding properties of the Dok1 PTB domain, we determined the

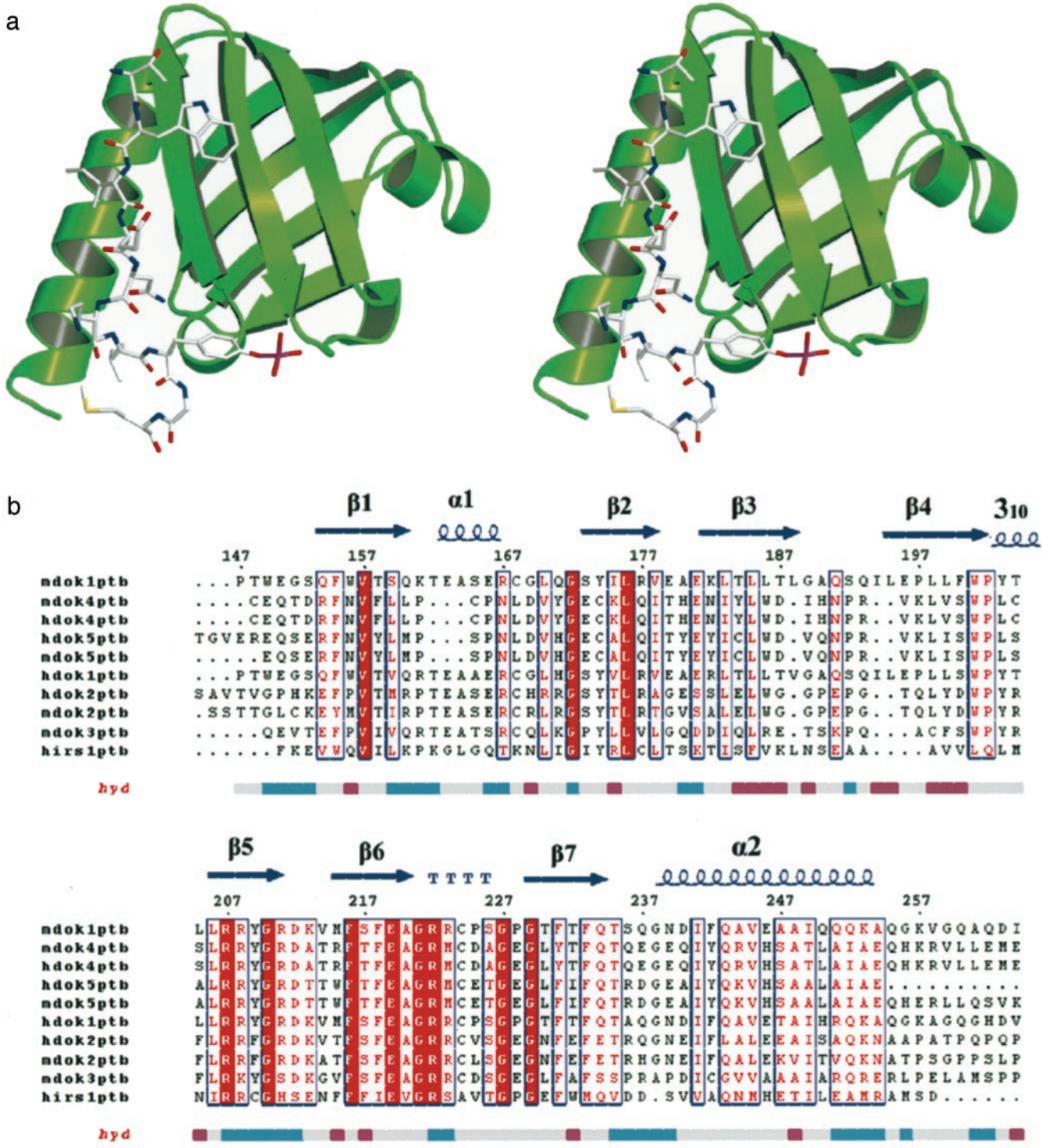

FIG. 2. **Overall structure of dok1 PTB domain.** *a*, ribbon stereo diagram showing the fold of the Dok1 PTB domain (*green*) and the orientation of the bound RET phosphopeptide (*white*). The ribbon diagram was generated with the program BOBSCRIPT (11). *b*, structure-based sequence alignments of the nine Doks and hIRS1 PTB domains. Sequences of mouse Dok1-(147–264), human dok1-(147–264), mouse Dok2-(144–259), human Dok2-(141–257), mouse Dok3-(156–266), mouse Dok4-(133–242), human Dok4-(133–242), mouse Dok5-(134–242), human Dok5-(129–232), and human IRS1-(160–262) were aligned. *Numbers* refer to mouse Dok1. The conserving residues were *boxed* in *red* and *blue*. Critical arginines for phosphotyrosine recognition are indicated by *green dots*. Alignment was generated using CLUSTAL X (1.8).

structure of a 1:1 complex of the Dok1 PTB domain (residues 154–256) with an 11-residue peptide derived from the C-terminal of RET (residues 1054–1064). The structure of the complex was determined by molecular replacement using the native structure as a search model. The structure of the complex is displayed in Fig. 2*a*, and statistics for the structure determination are given in Table I. Clear density was observed for all residues of the RET peptide, with the exception of Ser in the −8 position relative to the phosphotyrosine (pY-8). The peptide-binding site on the Dok1 PTB domain is characterized by an L-shaped surface groove formed by residues from strand β5 and the C-terminal α-helix, α2. The peptide forms a β-turn to occupy the L-shaped binding site (Figs. 2*a* and 3).

Phosphopeptide Recognition—Although it is known that PTB

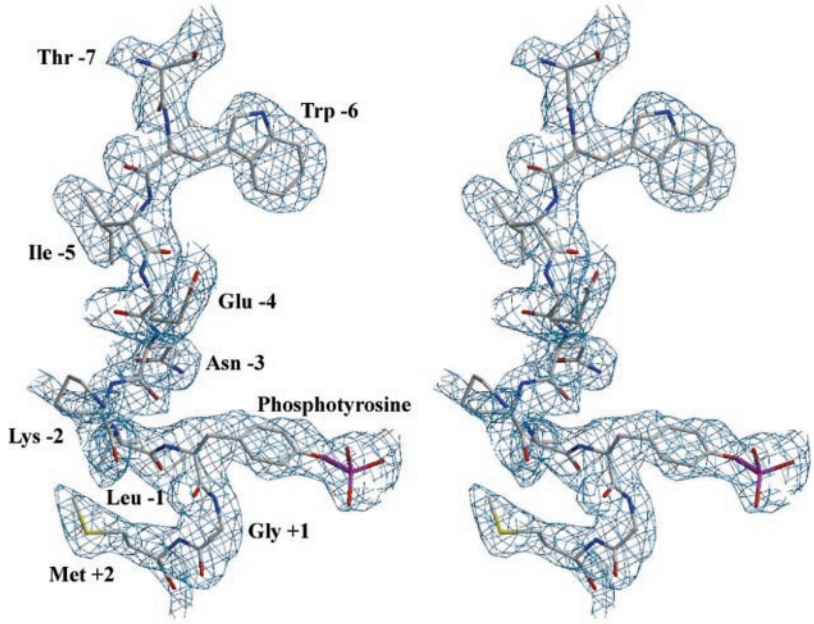

FIG. 3. **Stereo view of the electron density map covering the RET peptide.** A $2|F_o - F_c|$ map is shown at 2.5-Å resolution using phases calculated from the final, refined model and contoured at 1.0σ

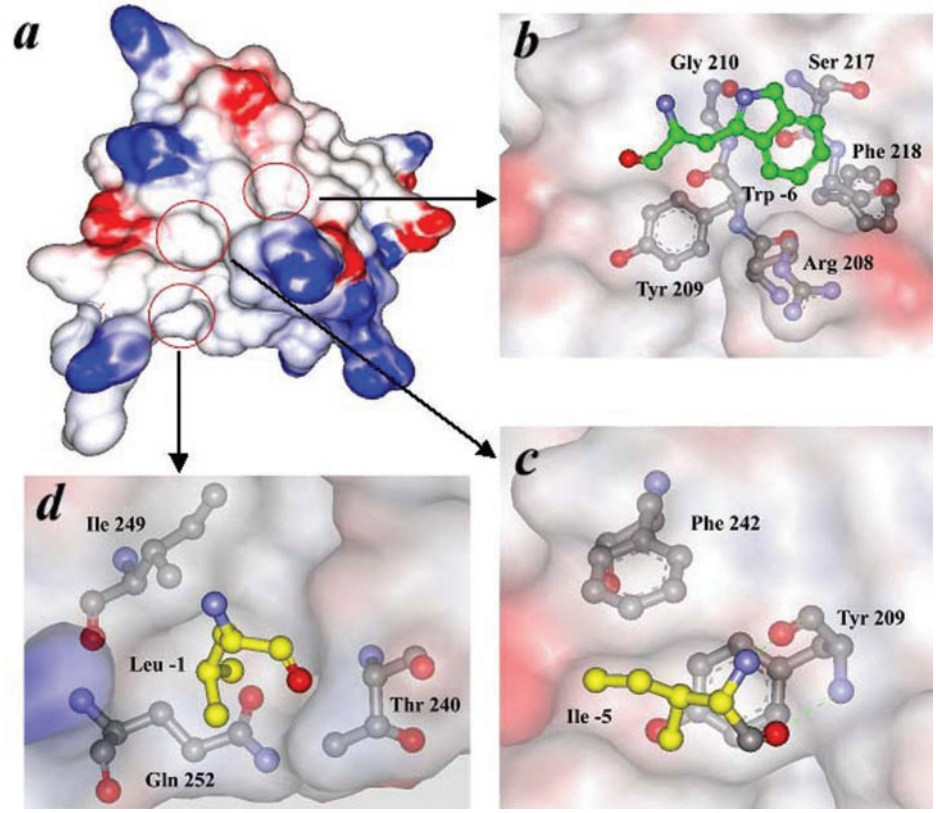

FIG. 4. **The contacts between Dok1 PTB domain and RET peptide side chains that contribute specificity to the interaction.** *a*, molecular surface representation of the Dok1 PTB domain structure calculated and shaded according to electrostatic potential using the program ViewerPro (Accelrys). As shown in *b*, Arg-208, Tyr-209, Gly-210, Ser-217, and Phe-218 of Dok1 PTB domain form a hydrophobic pocket, which may show a preference for large side chain hydrophobic residues such as Trp, Tyr, Phe, and Met in position pY-6 of the peptide. As shown in *c*, large hydrophobic side chains are present at pY-5 in the Dok1 PTB domain recognition motifs similar to Shc. As shown in *d*, Gln-252, Ile-249, and Thr-204 of Dok1 PTB domain form a hydrophobic pocket, which may prefer Leu or Ile in position pY-1 of the peptide. The key residues are shown in *ball-and-stick* representation.

domains mainly recognize NPXpY motifs, careful analysis of binding indicates that these domains have slightly different binding specificities (9). Asparagine in position −3 relative to phosphotyrosine (pY-3) and the phosphotyrosine group are necessary for binding to most PTB domains. A hydrophobic residue at position −5 and a proline at −2 are crucial for the Shc PTB domain, but the amino acids from −6 to −8 residues N-terminal to the phosphotyrosine are important for IRS1 binding to the NP*X*pY motifs. The proline in the NP*X*pY motifs also appears to be more important for IRS1 PTB binding than for Shc PTB binding. In addition, IRS1 PTB favors a small hydrophobic amino acid such as alanine at the −1 position. Substituting this alanine by a glutamate (such as insulin receptor) leads to a 30-fold loss of affinity for the IRS1 PTB domain. Studies with a combinatorial phosphopeptide library have indicated that the Dok1 PTB domain recognizes distinct sequences as compared with the IRS1 and Shc PTB domains. Leu at position −1 and hydrophobic amino acids Tyr, Met, and Phe at −6 were strongly selected for binding by the Dok1 PTB domain. Similar preferences for hydrophobic residues at position −5 to −8 have also been reported for other PTB domains.

Our binding studies show that the Dok1 PTB domain can bind only with the RET peptide and not with the IL-4 receptor and TrkA peptides (Fig. 1). Previous experiments indicated

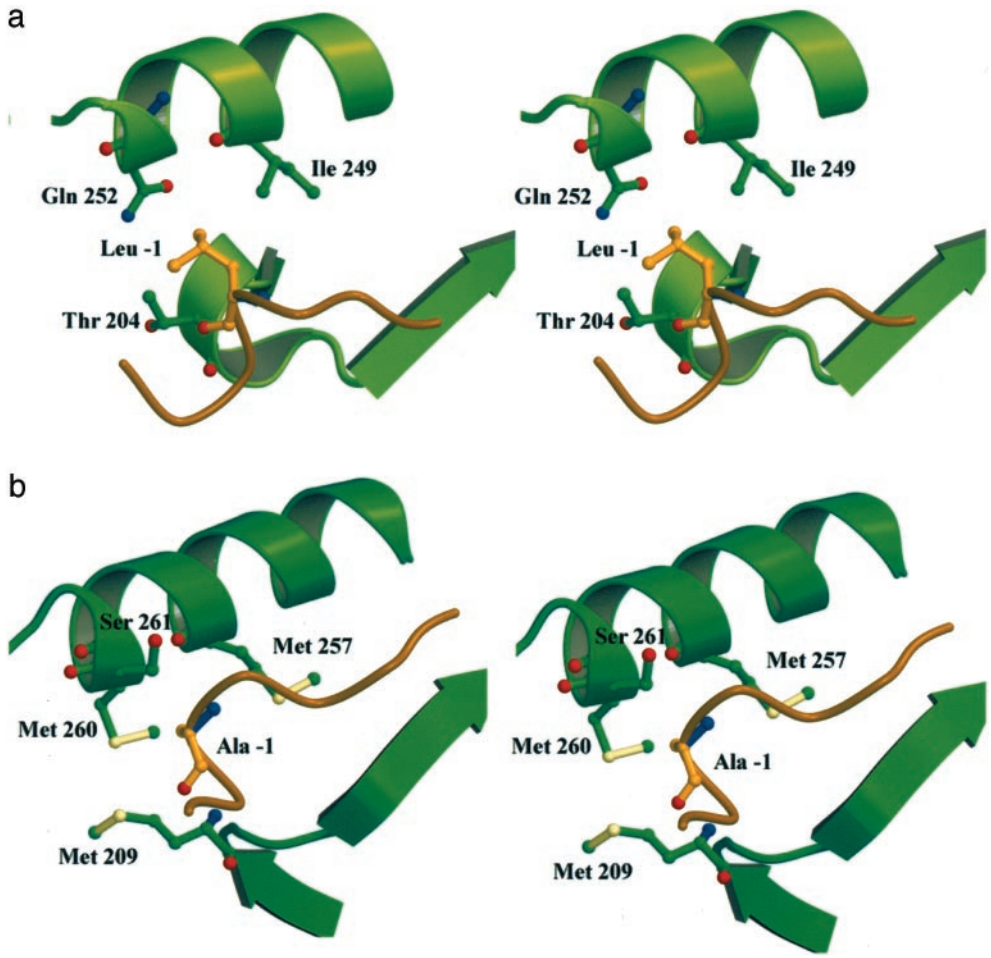

FIG. 5. **Stereo view of the interactions between residues at pY-1 of the phosphopeptide, shown in *brown*, and Dok1 (*a*) or IRS1 (*b*) PTB domain.** Residues involved in important interactions are shown in *ball-and-stick* representation. The residues interacting with pY-1 are represented as *green*; the sulfur atom is represented in *yellow*.

that IRS1 can bind with IL-4 and insulin receptor peptides and also with the RET peptide but not with the middle T, TrkA, Erb4, or epidermal growth factor receptor peptides that have hydrophobic residues at position −5 relative to Tyr(P) (9, 18). The Shc PTB domain can bind with mT, TrkA, Erb4, or epidermal growth factor receptor peptides and also with IL-4 and RET peptides (9, 19). The distinct specificities of these PTB domains correlate with and may account for some biological differences between these cytoplasmic substrates of tyrosine kinase-linked receptors.

Interactions between RET Peptide and the Dok1 PTB Domain—The RET peptide forms a β-turn and fills an L-shaped groove on the surface of the PTB domain that is formed by residues from the β5 strand and the C-terminal α helix. The estimated surface area of Dok1 PTB buried by the bound peptide is 761 Å^2. The recognition groove is composed of residues from the β5 strand, the C-terminal α-helix, and the 3_{10} turn connecting strands β4 and β5, including Tyr-203, Thr-204, Leu-205, Leu-206, Arg-207, Arg-208, Tyr-209, Arg-211, Ser-217, Phe-218, Gly-221, Arg-222, Phe-242, Ile-249, Gln-252, Lys-253. These residues make extensive contacts with all 10 residues of the RET peptide, through both hydrogen bonds and hydrophobic interactions. The phosphotyrosine is coordinated by Arg-207 and Arg-222, which extend from the β5 and β6 strands, respectively, and which are conserved in all Dok family proteins (Fig. 2*b*). The pY side chain lies in an open pocket created by the 3_{10} turn and residues at the end of strands β5 and β6 (Fig. 2*a*). An extensive network of hydrogen bonds and ionic interactions coordinate the phosphate oxygens, consistent with the observation that phosphorylation of the tyrosine is necessary for peptide binding. Replacing Arg-207 with alanine eliminates the ability of the Dok1 PTB domain to bind phosphopeptides (6). In addition, integrin β_3 and β_7 can bind to the Dok1 PTB domain with their tails containing the Dok1 PTB domain recognition motifs (7). The replacement with alanine of the Tyr-747 at pY position of integrin β_3 tails or the Tyr-778 at pY position of integrin β_7 tails also disrupted binding to the Dok1 PTB domain (7).

The backbone of N-terminal residues of the RET peptide, including residues pY-7 Thr, pY-6 Trp, pY-5 Ile, pY-4 Glu, pY-3 Asn, form a strand that hydrogen-bonds with strand β5 in an antiparallel orientation. In addition to backbone interactions, there are numerous contacts between the domain and peptide side chains that contribute specifically to the interaction. The indole ring of Trp-6 is bound in a pocket between β5 and β3 that is composed of Arg-208, Tyr-209, Gly-210, Ser-217, and Phe-218. This large pocket suggests that hydrophobic residues with large side chains might be selected here (Fig. 4*b*). Using a combinatorial peptide library approach, Songyang *et al.* (10) found that Tyr, Met, and Phe were strongly selected at this site. The side chain of Ile-5 shows numerous contacts with Phe-242 in the C-terminal α-helix (Fig. 4*c*). Large hydrophobic side chains are present at pY-5 in the Dok1 PTB domain recognition motifs. Integrin β_3 and β_7 can bind to the Dok1 PTB domain via their tails, which contain Dok1 PTB domain recognition motifs (7). Replacement of Asp-773 at the −5 position of integrin β_7 tails with more hydrophobic Ala or Phe residues dramatically increased Dok1 PTB domain binding to β7 tails, and con-

versely, a substitution of Ala-742 to Asp at the −5 position in integrin β_3 resulted in reduced binding to Dok1 PTB domain. The Asn at pY-3 is similar to that in other PTB domain recognition motifs NPXpY and appears to play an important structural role in stabilizing the β-turn of the peptides formed. The side chains of pY-4 Glu and pY-2 Lys extend away from the surface of the domain. In Dok1 PTB domain, Leu at the −1 position extends into a hydrophobic pocket composed of Gln-252, Ile-249, and Thr-204 and was exclusively selected (Fig. 4*d*). In addition, pY+1 Gly forms a hydrogen bond with Thr-204, and the side chain of pY+2 Met interacts with Lys-253 (Fig. 5*a*).

Comparison with Other PTB Domains—There is 17% sequence identity between the PTB domains of Dok1 and IRS1, whereas there is no significant sequence homology between the PTB domains of Dok1 and Shc. Despite the low sequence homology, the overall structure of the Dok1 PTB domain is similar to its IRS1 (20) and Shc (21) counterparts. Dok1 shares the PH domain-like fold of the PTB domain family (22) (Fig. 6) and a common mode of peptide binding, with the same β-turn conformation and orientation of phosphopeptide observed in each of the PTB domains. There are further similarities between IRS1 and the Dok1 PTB domain. Arg-212 and Arg-227, which recognize the phosphotyrosine in IRS1, are equivalent to Arg-207 and Arg-222 in the Dok1 PTB domain. These two residues are also conserved throughout the IRS protein family (Fig. 2*b*).

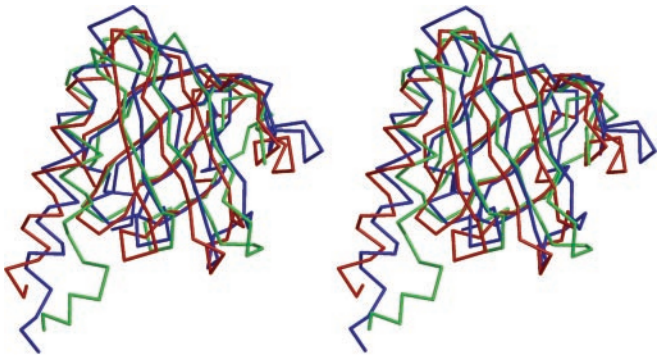

FIG. 6. **Stereo view of the superposition of Dok1 (*red*), IRS1 (*blue*), and Shc (*green*) PTB domains.** Dok1, IRS1, and Shc share a common PH domain-like fold. For clarity, selected residues from the Shc PTB domain have been omitted. Cα atoms of core residues of the structures superimpose with an root mean square deviation of < 1.0Å.

Interestingly, the Dok1 PTB domain has a different set of residues for recognizing the peptide. In IRS1, pY-1 of the peptide interacts with a hydrophobic patch composed of Met-209, Met-260, Ser-261, and Met-257 (Fig. 5*b*) (20). Ala was selected in this position, and although pY-1 can be substituted for Glu or Leu, they would result in an unfavorable interaction with this patch. When the pY-1 Ala in IL-4R is substituted by a Glu, as in the case of the insulin receptor, the result is a 30-fold loss in binding to IRS1 (23). In Dok1, pY-1 of the peptide interacts with a hydrophobic pocket composed of Gln-252, Ile-249, and Thr-204, and Leu was exclusively selected in this position (Fig. 5*a*). The different binding of TrkA and RET phosphopeptides to Dok1 may be due to the replacement of Gln by Leu at the pY-1 position. It is demonstrated that pY-1 Leu is very important to the Dok1 PTB domain binding motif.

The proline in position pY-2 is known to be crucial for high affinity binding for Shc and IRS1. Substitution of this residue reduces but does not abolish binding for Shc PTB domain. Meanwhile, substitution of the pY-2 proline with alanine abolishes binding for IRS1 (9). The side chain of pY-2 Lys extends away from the surface of the domain in Dok1 PTB, where it seems that a proline at position pY-2 is not essential.

Large hydrophobic side chains are present at pY-5 in the Dok1 PTB domain recognition motifs, similar to Shc. However, there is insufficient space in the IRS1 domain complex to accommodate large, hydrophobic side chains at peptide position pY-5. As with

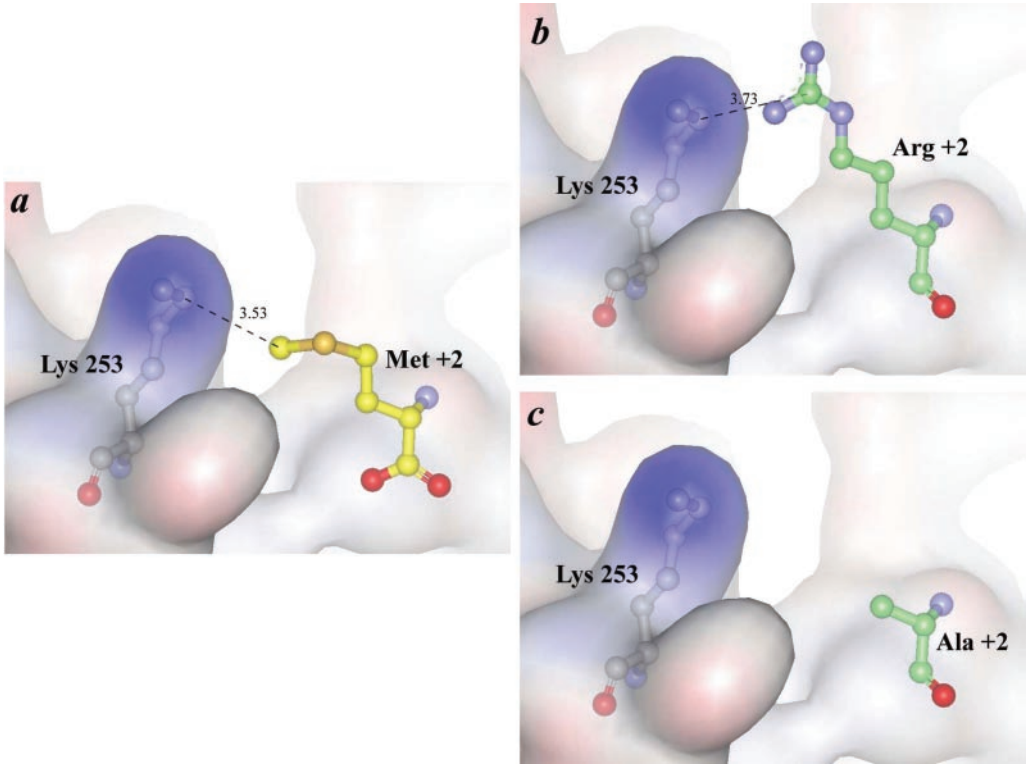

FIG. 7. **The interaction between Dok1 PTB domain and three isoforms of RET.** Arg-1064 in RET9 and Ala-1064 in RET 43 were modeled from our structure of the Dok1 PTB domain complexed with the RET51 phosphopeptide. Met-1064 in RET51 (*a*) and Arg-1064 in RET9 (*b*) both form an interaction with Lys-253 in the Dok1 PTB domain (3.53 and 3.73 Å, respectively), but Ala-1064 in RET 43 (*c*) does not.

the Shc PTB domain, the Dok1 PTB domain can also recognize the motifs of growth factor receptors and transforming proteins that possess large hydrophobic side chains at pY-5, whereas IRS1 does not bind to growth factor receptors. These differences indicate that Dok1 PTB recognizes distinct sequences (N*X*LpY) as compared with the Shc and IRS1 PTB domain (NP*X*pY).

Dok1 PTB Domain Binding to RET Isoforms—The RET proto-oncogene encodes a tyrosine kinase receptor that is essential for the development of the enteric nervous system and the kidney. Germline mutations of the RET proto-oncogene cause multiple endocrine neoplasia (MEN) 2A or 2B (24). RET has three isoforms, RET51, RET9, and RET43, formed by alternative splicing at a site just downstream of tyrosine 1062 (pY) (25). These isoforms play different roles in tumor development. RET51-MEN2A and RET51-MEN2B mutant proteins have stronger transforming activity than RET9-MEN2A and RET9-MEN2B mutant proteins, respectively (26). The activity of RET43 is very low (27, 28). The Tyr-1604 (pY+2) residue is different in each of these RET isoforms. The RET9 isoform has arginine in the pY+2 position, whereas RET43 has alanine and RET51 has methionine in the equivalent position. In our model of the Dok1 PTB domain-RET peptide complex, the RET peptide is derived from RET51, and the Tyr-1604 (pY+2) residue is a methionine that forms a hydrophobic interaction with residue Lys-253 that extends from the C-terminal α-helix (the distance between C_ϵ of Met +2 to C_ϵ of Lys-253 is 3.53Å) (Fig. 7*a*). When Met in the pY+2 position is replaced by Arg, there is still an interaction between Arg+2 and Lys-253, but it is weakened (the distance between C_ζ of Arg+2 to C_ϵ of Lys-253 is 3.73Å) (Fig. 7*b*). However, the substitution of Ala for Met at the pY+2 position abolishes the hydrophobic interaction altogether (Fig. 7*c*). These findings are consistent with the relative transforming activities of the RET isoforms.

Conclusions—A detailed analysis of the structure of the Dok1 PTB domain and its complex with a RET phosphopeptide has revealed the basis for ligand recognition by the Dok1 PTB domain. We also show that the recognition of peptides by the Dok1 PTB domain is specific since Dok1 cannot bind IL-4 receptor and TrkA peptides that are recognized by Shc and IRS1 PTB domains, respectively. A structural comparison of the Dok1 PTB domain with other PTB domain structures explains their different peptide binding specificities. Furthermore, the distinct specificities of the PTB domains correlate with and should account for key biological differences between these cytoplasmic substrates of tyrosine kinase-linked receptors.

Acknowledgments—We thank Dr. Min Yao for assistance during data collection at beamline 41 XU at SPring-8, Hyogo, Japan. We also thank Dr. George F. Gao of Oxford University, Oxford, UK, for help with synthetic peptides.

REFERENCES

1. Carpino, N., Wisniewski, D., Strife, A., Marshak, D., Kobayashi, R., Stillman, B., and Clarkson, B. (1997) *Cell* **88,** 197–204
2. Yamanashi, Y., and Baltimore, D. (1997) *Cell* **88,** 205–211
3. Shah, K., and Shokat, K. M. (2002) *Chem. Biol.* **9,** 35–47
4. Kato, I., Takai, T., and Kudo, A. (2002) *J. Immunol.* **168,** 629–634
5. Becker, E., Huynh-Do, U., Holland, S., Pawson, T., Daniel, T. O., and Skolnik, E. Y. (2000) *Mol. Cell Biol.* **20,** 1537–1545
6. Murakami, H., Yamamura, Y., Shimono, Y., Kawai, K., Kurokawa, K., and Takahashi, M. (2002) *J. Biol. Chem.* **277,** 32781–32790
7. Calderwood, D. A., Fujioka, Y., de Pereda, J. M., Garcia-Alvarez, B., Nakamoto, T., Margolis, B., McGlade, C. J., Liddington, R. C., and Ginsberg, M. H. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100,** 2272–2277
8. Dhe-Paganon, S., Ottinger, E. A., Nolte, R. T., Eck, M. J., and Shoelson, S. E. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 8378–8383
9. Wolf, G., Trub, T., Ottinger, E., Groninga, L., Lynch, A., White, M. F., Miyazaki, M., Lee, J., and Shoelson, S. E. (1995) *J. Biol. Chem.* **270,** 27407–27410
10. Songyang, Z., Yamanashi, Y., Liu, D., and Baltimore, D. (2001) *J. Biol. Chem.* **276,** 2459–2465
11. Esnouf, R. M. (1997) *J. Mol. Graphics* **15,** 132–134
12. Otwinowski, Z., and Minor, W. (1997) in *Macromolecular Crystallography, Part A* (Carter, C. W., Jr., and Sweet, R. M., eds) Vol. 276, pp. 307–326, Academic Press, Orlando, FL
13. Terwilliger, T. C., and Berendzen, J. (1999) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **55,** 849–861
14. Terwilliger, T. C. (2000) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **56,** 965–972
15. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. Sect. A* **47,** 110–119
16. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **54,** 905–921
17. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Cryst.* **26,** 283–291
18. Melillo, R. M., Carlomagno, F., De Vita, G., Formisano, P., Vecchio, G., Fusco, A., Billaud, M., and Santoro, M. (2000) *Oncogene* **20,** 209–218
19. Asai, N., Murakami, H., Iwashita, T., and Takahasi, M. (1996) *J. Biol. Chem.* **271,** 17644–17649
20. Zhou, M. M., Huang, B., Olejniczak, E. T., Meadows, R. P., Shuker, S. B., Miyazaki, M., Trub, T., Shoelson, S. E., and Fesik, S. W. (1996) *Nat. Struct. Biol.* **3,** 388–393
21. Zhou, M. M., Ravichandran, K. S., Olejniczak, E. F., Petros, A. M., Meadows, R. P., Sattler, M., Harlan, J. E., Wade, W. S., Burakoff, S. J., and Fesik, S. W. (1995) *Nature* **378,** 584–592
22. Murzin, A. G., Brenner, S. E., Hubbard, T., and Chothia, C. (1995) *J. Mol. Biol.* **247,** 536–540
23. He, W., O'Neill, T. J., and Gustafson, T. A. (1995) *J. Biol. Chem.* **270,** 23258–23262
24. Watanabe, T., Ichihara, M., Hashimoto, M., Shimono, K., Shimoyama, Y., Nagasaka, T., Murakumo, Y., Murakami, H., Sugiura, H., Iwata, H., Ishiguro, N., and Takahasi, M. (2002) *Am. J. Pathol.* **2002,** 249–256
25. Ivanchuk, S. M., Myers, S. M., and Mulligan, L. M. (1998) *Oncogene* **16,** 991–996
26. Pasini, A., Geneste, O., Legrand, P., Schlumberger, M., Rossel, M., Fournier, L., Rudkin, B. B., Schuffenecker, I., Lenoir, G. M., and Billaud, M. (1997) *Oncogene* **15,** 393–402
27. Carter, M. T., Yome, J. L., Marcil, M. N., Martin, C. A., Vanhorne, J. B., and Mulligan, L. M. (2001) *Cytogenet. Cell Genet.* **95,** 169–176
28. Lee, D. C., Chan, K. W., and Chan, S. Y. (2002) *Oncogene* **21,** 5582–5592