An Expanded WW Domain Recognition Motif Revealed by the Interaction between Smad7 and the E3 Ubiquitin Ligase Smurf2

Received for publication, February 15, 2006, and in revised form, April 25, 2006 Published, JBC Papers in Press, April 26, 2006, DOI 10.1074/jbc.M601493200

P. Andrew Chong‡§1, Hong Lin‡, Jeffrey L. Wrana‡%, and Julie D. Forman-Kay†§

From the ‡Program in Structural Biology and Biochemistry, Hospital for Sick Children, Toronto M5G 1X8, %Department of Medical Genetics and Microbiology, University of Toronto, Toronto M5S 1A8, †Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto M5G 1X5, and §Department of Biochemistry, University of Toronto, Toronto M5S 1A8, Canada

Smurf2 is an E3 ubiquitin ligase that drives degradation of the transforming growth factor-β (TGF-β) and bone morphogenetic protein (BMP) pathways play key roles in many biological processes including tumor suppression and embryonic development (1). Signals from these extracellular ligands are transduced through transmembrane Ser/Thr kinases that phosphorylate intracellular transcription factors known as the receptor-regulated Smad (R-Smad) proteins (2). In addition to R-Smads (Smads 1–3, 5 and 8), the Smad family includes the common mediator Smad (Smad4) and inhibitory Smads (I-Smads: Smads 6 and 7). Control of these pathways is exerted to some degree by ubiquitin-mediated targeted degradation of Smad proteins and receptors, in part involving the Smad ubiquitination regulatory factor (Smurf) proteins (3, 4). Smurf1 and Smurf2 are HECT-type E3 ubiquitin ligases that ubiquitinate specific proteins, marking them for degradation by the proteasome (5). The Smad proteins can themselves be ubiquitinated by Smurf proteins. For example, Smurf1 efficiently targets BMP-responsive Smad1 and Smad5 for degradation (3). Smurf2 has been implicated in ubiquitination of Smad1 and Smad2 (6, 7). In addition, the Smad proteins can function as adapters to enhance Smurf-dependent ubiquitination of other target proteins (4, 8, 9). For example, Smad7, which binds to the TGF-β receptors and Smurfs, acts as an adapter allowing Smurf1 and Smurf2 to bind to and ubiquitinate the TGF-β receptors, targeting them for degradation (4, 9).

Recent reports suggest that the Smurf homologues Itch, Nedd4-2, and WWP1 also function in regulating TGF-β signaling pathways. Itch can both ubiquitinate Smad2 and enhance Smad2 phosphorylation in mouse embryonic fibroblasts (10). Nedd4-2 and WWP1 both bind to Smads 2, 3, 6, and 7 (11, 12) and, similar to Smurf1 and Smurf2, both interact with TGF-β receptor complexes via Smad7 to induce their ubiquitin-mediated degradation. In view of the combinatorial possibilities for interaction with these HECT-type E3 ubiquitin ligases, further characterization of the molecular basis of target recognition would help define specificity.

Most HECT ubiquitin ligases share a common architecture comprising an N-terminal C2 domain required for proper localization, multiple WW domains involved in target recognition, and a C-terminal HECT domain, which catalyzes the transfer of ubiquitin onto target proteins (5). As in other HECT ubiquitin ligases, Smurf recognition of target or adapter proteins is mediated in part by interactions between Smurf WW domains and PY motifs (PPXY) found in the linker regions of R- and I-Smads. In addition, the Smad7 N-terminal domain can contact the Smurf HECT domain, regulating E2 recruitment (13).

WW domains are small modular binding domains that recognize polyproline motifs. They are composed of a three-stranded anti-parallel β-sheet. WW domains are named for two conserved Trp residues (15). Muta-

*This research was supported by grants from the Canadian Institute for Health Research (to J. L. W.) and from the National Cancer Institute of Canada (to J. D. F.-K.). 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 online version of this article (available at http://www.jbc.org) contains supplemental data.

The atomic coordinates (code 2DJY) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

The chemical shift data will be deposited in the Biological Magnetic Resonance Data Bank, University of Wisconsin-Madison, Madison, WI (http://www.bmrb.wisc.edu).

1 Recipient of a scholarship from the Hospital for Sick Children Foundation.

2 A Howard Hughes international scholar.

3 To whom correspondence should be addressed: Dept. of Structural Biology and Biochemistry, Hospital for Sick Children, 555 University Ave., Toronto-M5G 1X8, Canada. Tel.: 416-813-5358; Fax: 416-813-5022; E-mail: forman@sickkids.ca.

4 The abbreviations used are: TGF-β, transforming growth factor-β; PY, polyproline-tyrosine; Nedd4, neuronal precursor cell expressed developmentally down-regulated 4; YAP1, Yes-associated protein 1; HECT, homologous to the E6-AP carboxyl terminus; r.m.s., root mean square; HPLC, high pressure liquid chromatography; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; SH, Src homology; TOCSY, total correlation spectroscopy.
nition of target ligands (15, 18, 19). WW domains can be divided into four groups based on their ligand binding specificity (20). Group I binds to PY motifs with the sequence PXpY. Group II binds to ligands with a PPLP motif, usually within the context of multiple Pro residues. Groups III and IV recognize polyproline motifs flanked by Arg or Lys and phospho-(Ser/Thr)-Pro sequences, respectively. Several structures of group I WW domain-PY motif complexes have been published (16, 17, 19, 21). These structures clearly delineate the molecular determinants involved in recognition of the PXpY sequence. However, the prevalence of group I WW domain-type interactions in cellular protein-protein interactions makes it difficult to understand how specificity is maintained.

To examine target binding by the Smurf proteins, we studied the recognition of Smad7 by the Smurf WW domains. Smad7 and Smurf2 functions are closely linked, because Smad7 acts as an adapter mediating Smurf2 interaction with targets and with the E2 enzyme UbcH7, which stimulates ubiquitin ligase activity (13). Our results demonstrate that the Smurf2 WW3 domain binds to a region in Smad7 that includes a PY motif. Determination of the solution structure of a complex between the Smad7 PY motif region (ELESPPPPYSRYPMD) and the Smurf2 WW3 domain by NMR revealed an extended binding motif that includes the PY motif and six residues C-terminal to the PY motif, denoted as the “PY-tail.” Binding experiments show that the Phe in the WW3 domain binding site decreases the affinity relative to the canonical Trp. However, this loss of binding affinity is balanced by additional interactions between the Smad7 PY-tail and the WW3 domain. We propose that this unusual WW domain-PY motif interaction allows the Smurf2 WW3 domain to recognize a subset of PY motif-containing proteins with specificity provided by an expanded interface.

**EXPERIMENTAL PROCEDURES**

**Protein and Peptide Preparation**—DNAs encoding the Smurf2 WW2 (amino acids 250–288) and WW3 (amino acids 297–333) domains and the Smad7 PY motif peptide (amino acids 203–217) were amplified using PCR and inserted into pGEX 6P1 (GE Healthcare) using BamHI and XhoI restriction sites. The resulting constructs were confirmed by sequencing. Smurf2 WW3 Phe-325 was mutated to Trp using the QuikChange mutagenesis kit (Stratagene). Glutathione S-transferase fusions were expressed in *Escherichia coli* BL21(DE3) cells at 25 °C with a 16-h isopropyl 1-thio-β-D-galactopyranoside induction period. WW3 and Smad7 PY peptide were labeled with 15N and 13C by expressing the proteins in M9 minimal medium containing 15N, 13C-labeled PY peptide and 1.9 mM unlabeled WW3 domain. To measure intermolecular NOEs, we recorded half-filtered NOESY spectra (27) with mixing times of 250 ms on two samples in 99% D2O, one containing ~1.3 mM 15N, 13C-labeled WW3 domain and 1.7 mM labeled PY peptide and one containing 1.5 mM 15N, 13C-labeled PY peptide and 1.9 mM unlabeled WW3 domain.

Structural restraints were restricted to NOEs and TALOS-based (28) dihedral angle restraints. Tolerances for dihedral angle restraints were set to the greater of two S.D. values or 20°. After initial assignment of 300 NOEs, which for symmetry related peaks were observed, ARIA1.2 (29) was used to facilitate the assignment of NOEs and calculate structures. Eight iterations were performed with the cutoff for ambiguous restraints decreasing from 1.0 to 0.8 and the violation tolerance decreasing from 1000 to 0.1 Å over the course of the calculations. The 30 lowest energy structures, of 200 calculated in the final iteration, were used for analysis. PROCHECK software (30) was used to analyze the Ramachandran plots.

**Binding Studies**—Fluorescence binding experiments were performed at 20 °C using an AVIV ratio spectrophotometer model ATF105 and a Microlab 500 automated titrator. Proteins were dialyzed into a filtered buffer containing 40 mM Na2HPO4, pH 7.2, and 20 mM NaCl. Fluorescence levels were monitored as peptide was added to the WW domain samples using an excitation wavelength of 295 nm and an emission wavelength of 330 nm. Data were fit as previously published (31). The temperature difference between the sample in the fluorometer cell and titrant in the automatic titrator had a negative effect on the precision of the results. It was necessary to correct for this temperature difference by subtracting a blank run in which WW protein without PY peptide was injected into a solution containing WW domain at the same concentration.

**RESULTS**

To address Smurf2 specificity, the sequences of Smurf2 WW2 and WW3 domains were analyzed, as previous data demonstrated their requirement for efficient binding to Smad7. Of note, WW1 is not required for Smurf2 to bind Smad7 (4). Sequence alignment of WW domains from a number of proteins indicated that both Smurf2 WW2 and WW3 domains are non-canonical (supplemental Fig. S1). The WW2 domain has a Tyr (Tyr-257) at the site normally occupied by the conserved core Trp. Indeed, analysis of the 201 WW domains identified by the Simple Modular Architecture Research Tool reveals that the Smurf2 WW2 and WW3 domains are the only ones with a residue other than Trp at position 17 (numbering according to YAP65) (38). This Tyr has the potential to destabilize the WW2 domain (15). In contrast, the WW3 domain of Smurf2 has a Trp at position 17, but in place of the canonical binding site Trp at position...
39 there is a Phe (supplemental Fig. S1, Phe-325 in Smurf2). Although residues other than Trp are observed in this position in a number of other WW domains, the Trp in this position is involved in binding to the Pro residues of the PY motif and is thought to be required for the interaction (15). Thus, it was not clear from sequence analysis which of the WW2 or WW3 domains would bind to the Smad7 PY motif.

The Smurf2 WW3 Domain Binds to the Smad7 PY Motif—To determine which WW domain binds the Smad7 PY motif, the affinity of a peptide containing the Smad7 PY motif for the isolated WW2 and WW3 domains was measured using intrinsic fluorescence (supplemental Table S1). There was no change in fluorescence when PY peptide was added to the WW2 domain, implying that the PY peptide does not bind to the isolated WW2 domain. To confirm this, we monitored the addition of PY peptide to the WW2 domain by NMR (not shown). Of note, WW2 domain amide proton chemical shift dispersion was consistent with a folded protein, but upon the addition of PY peptide, there were no chemical shift changes, corroborating the lack of interaction. In contrast, analysis of the WW3 domain by fluorescence revealed that it binds to the PY peptide with an affinity of $40.0 \pm 0.1 \mu M$. Because the Smurf2 WW3 domain has a Phe in place of the canonical binding site Trp, a mutant of this WW3 domain with the canonical binding site Trp (F325W) was examined and found to bind the PY peptide with an affinity of $0.17 \pm 0.01 \mu M$. The 200-fold higher affinity of the F325W WW3 mutant relative to wild type indicates that the naturally occurring Phe-325 significantly decreases the affinity of the Smurf2 WW3 domain for the Smad7 PY peptide relative to the canonical Trp. Nonetheless, the isolated Smurf2 WW3 domain binds to the PY peptide with an affinity that is similar to that observed for interactions between other PY peptides and canonical group I WW domains (15).

Solution Structure of a Smurf2 WW3 Domain-Smad7 PY Motif Complex—Our results suggested that the Smurf2 WW3 domain may bind the Smad7 PY peptide via a novel mechanism. To examine how an affinity similar to canonical WW domain-PY motif interactions is achieved, despite the absence of a binding site Trp, we determined the solution structure of the complex. The peptide and protein were docked using two half-filtered NOEY experiments that provided over 200 intermolecular distance restraints. The resultant structure is well defined for residues 299–329 of the WW3 domain and residues 207’–217’ (prime is used to distinguish PY peptide residues from WW domain residues) of the PY peptide with an overall backbone r.m.s. deviation of $0.5 \pm 0.1 \AA$ (Fig. 1a, Table 1). Although the structures are very well defined overall, two areas of the structure have poor structural statistics. In particular, Smurf2 Ala-309 and Thr-310, which are in a loop between two $\beta$-strands, have high backbone r.m.s. deviation values and $\phi, \psi$ angles that are in the generously allowed or disallowed area of the Ramachandran space in some structures. This is the result of an exchange process (discussed below) involving residues 309–310 and leads to NOEs that are inconsistent with a single structure. The ensemble of the 30 lowest energy structures must be interpreted with caution in this area and should not be construed as an accurate representation of the actual conformational pool but as the lowest energy fit to the NOE data. The backbone $\phi$ and $\psi$ angles for Smad7 Ser-212’ and Arg-213’ also lie in the generously allowed or disallowed regions of the Ramachandran plot in most of the 30 lowest energy structures, likely also because of the difficulty in fitting NOE data in regions undergoing exchange.

The structure of the Smurf2 WW3 domain is very similar to other known WW domain structures (Fig. 1b) (16, 17, 19). It consists of 30 well structured residues that form a three-stranded anti-parallel $\beta$-sheet, with a hydrophobic cluster on one face of the sheet and a binding pocket on the opposite face. As seen in other WW domain structures, the conserved core Trp-303 (Fig. 1b, shown in green) and the conserved Pro-328 (position 42 in YAP65), form part of the hydrophobic cluster together with Phe-315 and Gln-324. The $\beta$-sheet is twisted, forming a cradle on the binding face in which the Smad7 PY peptide binds.

Despite the presence of Phe-325 in place of the canonical binding site Trp, the topology of the PY motif (PPY) and Smurf2 WW3 domain is very similar to that observed for other group I WW domain-PY complexes, such as the dystrophin WW domain-$\beta$-dystroglycan complex (17) (Fig. 2). The first and second Pro residues in the PY motif, Pro-208’ and -209’ (Pro-889’ and -890’ of $\beta$-dystroglycan), which are in a PPII helical conformation, bond in a hydrophobic pocket, stacking against Phe-325 and Tyr-314, respectively. One difference is that the side chains of Pro-208’ and Phe-325 in Smurf2 are twisted almost perpendicular to the analogous Pro-889’ and Trp-3083 of the dystrophin complex. The positioning of the carbonyl group of Pro-209’ and the hydroxyl group of Thr-323 suggests the presence of a hydrogen bond as is seen in other
Smurf2 WW3-Smad7 PY Structure

**TABLE 1**
Structural statistics for the 30 final structures of the Smurf2 WW3 domain-Smad7 PY peptide complex

| r.m.s. deviations from distance restraints (Å) | 0.021 ± 0.002 |
|---------------------------------------------|----------------|
| All (1,723)                                 | 0.021 ± 0.002 |
| Unambiguous (1,451)                         | 0.020 ± 0.002 |
| Ambiguous (272)                             |                |

| r.m.s. deviations from dihedral angle restraints (°) | 0.10 ± 0.09 |
|-----------------------------------------------------|-------------|

| r.m.s. deviations from idealized geometry |
|-------------------------------------------|
| Bonds (Å)                                  | 0.003 ± 0.0002 |
| Angles (°)                                 | 0.425 ± 0.021 |
| Impropers (°)                              | 0.314 ± 0.037 |

**Ramachandran map: Smurf2-(299–329), Smad7-(207’–217’)**

Most favored regions: 69.3 ± 4.6%
Additionally allowed regions: 21.6 ± 5.3%
Generously allowed regions: 4.0 ± 2.7%
Disallowed regions: 5.0 ± 1.9%

**Atomic r.m.s. deviation (Å) from mean structure**

| Smurf2-(299–329), Smad7-(207’–217’) | 0.50 ± 0.13 |
|-------------------------------------|-------------|
| Backbone                             | 1.09 ± 0.16 |
| All heavy atoms                      | 1.11 ± 0.28 |

**FIGURE 2.** Superposition of the PY motifs and binding pockets from the Smad7-Smurf2 WW3 domain complex and the β-dystroglycan-dystrophin WW domain complex. Smad7 and Smurf2 are shown in red and blue and β-dystroglycan and dystrophin in black and gray, respectively. Labels in parentheses indicate residues from the β-dystroglycan-dystrophin complex.

**FIGURE 3.** Interaction between the PY-tail and WW3 domain. a, surface representation of the WW3 domain with Smad7 peptide as a ribbon. Blue and red represent positive and negative electrostatic potential, respectively. The backbone of the PY peptide is shown in green, with the side chains of the peptide in yellow. The PY-tail, comprising the β-dystroglycan-dystrophin complex is remarkably similar to other group I WW domain-PY motif complexes in the core PY motif binding region.

WW domain interactions (17, 19). The Tyr of the PY motif, Tyr-211’ (Tyr-892’, equivalent residues from the dystrophin-β-dystroglycan complex are in parentheses), lies in a hydrophobic pocket formed by Val-316 (Ile-3074), an uncharged His-318 (His-3076), and Arg-321 (Gln-3079). The uncharged status of His-318 was confirmed by analysis of the chemical shifts of the His side chain atoms (32). Furthermore, protonation of His-318 by decreasing the pH caused a decrease in binding affinity and dissociation of the WW3-PY complex observed by NMR (not shown). These data show that the topology of the Smurf2 WW3-Smad7 PY motif complex is remarkably similar to other group I WW domain-PY motif complexes in the core PY motif binding region.

Structure of the PY Peptide and Interaction Between the PY-tail and WW3 Domain—This topological similarity suggests that discrimination of PY-motif-containing peptides by WW domains might be associated with areas outside of the core PY motif and its cognate binding pocket. Indeed, analysis of the contact between the WW3 domain and the PY peptide revealed that it extended beyond the interface with the PY motif proper to include an interaction with the six Smad7 residues C-terminal to the PY motif (PY-tail)(Fig. 3a). The PY-tail, which has not been observed for other PY motif-containing peptides, lies parallel to the β-strands of the WW3 domain and binds residues in the first and second β-strands and the intervening β1–β2 loop (Figs. 1b and 3a). These are extensive backbone interactions between the PY-tail and the WW3 domain as well as specific side chain contacts involving Arg-213’ with His-318, Pro-215’ with Arg-306, and Asp-217’ with Thr-308 and Ala-309. To facilitate these interactions, the PY peptide adopts a loop structure with residues in the PY motif lying anti-parallel to the last four residues of the PY-tail. This loop structure is stabilized by extensive intrapeptide contacts involving Pro-209’, Pro-210’, Ser-212’, Tyr-214’, and Met-216’. NOEs are observed between side chain protons of Met-216’, which is centrally located, and side chain protons of Pro-209’, Ser-212’, and Tyr-214’. The full interaction has a buried surface area of 1011 ± 23 Å² and is more extensive than any previously observed WW domain-PY peptide interaction.
The smaller number of intermolecular NOEs between the WW3 domain and PY-tail, relative to the number with the PY motif, suggests this region interacts more weakly with the surface of the WW3 domain. To confirm that the PY-tail does contribute to the interaction with the WW3 domain, single Ala mutations at five positions in the tail were made, and the ability of these mutant peptides to bind to the WW3 domain was tested (Fig. 3b). Three of five of the mutations in the PY-tail caused a reduction in binding affinity, confirming the importance of the PY-tail for interaction with the WW3 domain. Specifically, mutation of Ser-212’, Tyr-214’, or Met-216’, which make important intramolecular interactions that stabilize the loop structure of the peptide, lead to more than a 3-fold reduction in affinity, despite the paucity of contacts between the side chains of these residues and the WW3 domain. Although Arg-213’ interacts with His-318, mutation to Ala had a negligible effect on the affinity of the WW3 domain for the PY peptide, possibly because the Arg is not optimal for binding. Mutation of Ser-206’ N-terminal to the PY motif had a negligible effect on affinity. Although Pro-215’ packs closely against Arg-306 in the WW3 domain, mutation of this residue does not decrease the binding affinity. The loss of contacts are likely balanced by an enhancement in affinity due to elimination of the cis form of the peptide bond, which is not compatible with the bound form of Pro-215’. That is, mutation of this Pro eliminates the cis form of this peptide bond, which is heavily populated by the uncomplexed peptide (see “Relevance of the Structure of the Free PY Peptide for the Interaction” in supplemental material), thus eliminating the energetic barrier to binding associated with cis to trans isomerization and enhancing the on-rate constant $k_{on}$. Together, these mutational data demonstrate that intramolecular interactions stabilizing the Smad7 PY loop structure are required for optimal binding with the WW3 domain.

**Dynamic Character of the PY-tail**—As described above, the PY-tail region seems to bind relatively weakly to the WW3 domain surface. In fact, the last few residues of the PY-tail exchange between three conformations when the peptide is complexed with the WW domain. The exchange occurs on the slow to intermediate NMR time regime (ms–μs), allowing the three conformations to be identified on the basis of differing chemical shift. Spectra of the complexed PY peptide show three sets of resonances for these three conformations, two of which overlap with the sets of resonances seen in the spectrum of free PY peptide (Fig. 4), suggesting that, although the core PY motif is fully bound, the C terminus of the PY-tail samples both bound and unbound states. Although overlap in the resonances makes it difficult to quantify populations precisely, analysis of the resonances for the last residue of the PY-tail, Asp-217’, revealed that ~60–70% of Asp-217’ is in the bound conformation. The 30% that is in the unbound conformation also exists in two different states, because of Pro cis/trans isomerization of Pro-215’, with ~70% in the major (trans) conformation. Met-216’ also appeared to be undergoing similar slow exchange processes. Due to degeneracy of the Pro resonances, it was difficult to identify distinct peaks for bound and unbound conformations of Pro-215’. Exchange between the bound and unbound peptide C terminus also caused broadening of the amide proton resonances for Thr-310 and Ala-309 in the β1-β2 loop of the WW3 domain, which is consistent with structural information showing that Asp-217’ interacts with this loop. The exchange suggests that this region makes a lower energetic contribution to the binding affinity than if the interactions were fully populated. It is possible that the interaction is stabilized in the context of full-length Smad7. Nonetheless, the PY-tail mutants provide clear evidence for the significant impact this region has on the affinity.

**DISCUSSION**

The WW2 and WW3 domains of Smurf2 are both required for Smurf2 function. Our data indicate that the isolated WW2 domain, which has a Tyr in place of the core Trp, is folded but does not bind a PY motif-containing peptide from Smad7. The role of the WW2 is thus unclear; it may bind to a region of Smad7 near to the PY motif and/or the WW3 domain. On the other hand, the Smurf2 WW3 domain does bind the PY motif region from Smad7. This WW3 domain is the first example of a WW domain that has a Phe in place of the canonical binding pocket Trp, yet still binds a PY motif. The Phe decreases the affinity of the WW3 domain for the PY motif, requiring additional interactions that are provided by the PY-tail, which adopts a loop conformation and binds the β1-strand and the β1–β2 loop of the WW3 domain. Intrapeptide interactions stabilize the loop structure and are essential for high affinity binding. These unique features likely contribute to the specific recognition of Smad7 by Smurf2.

Specificity in WW domain-PY motif interactions is central to understanding how this ubiquitous protein interaction domain functions in distinct pathways. Although the PY motif is critical for group I WW domain interactions, residues within the PY motif and the PY motif binding pockets are well conserved and are not a focal point for differentiating between ligands within the extensive group I. The requirement for the first two Pro residues and the Tyr of the PY motif is well documented, and mutation of these residues leads to a dramatic loss in binding affinity (19, 33). The key residues in the WW domain involved in binding to the PY motif (turquoise residues in Fig. 5, a and b) are also conserved or similar. An aromatic residue (Tyr/Phc-28, hYAP65 numbering) on the β2-strand and a Thr-37 and Trp-39 on the β3-strand form the XP groove, which binds the Pro residues of the PY motif. The Tyr is bound by a large hydrophobic residue (Val/Ile/Leu-30) on the β2-strand and a neutral His-32 and the aliphatic portion of an Arg or Lys (Arg/Lys/Gln-35) on the β2–β3 loop. In general, the XP grooves and Tyr binding pockets of group I WW domains (turquoise residues in Fig. 5, a and b) are very similar and are not sufficient for discrimination among group I ligands.

In a few cases, there are more substantive changes to the XP groove that contribute to specificity. For example, the dNedd4 WW3’ domain binds a LPSY motif-containing peptide with a 2-fold higher affinity than a PPSY motif-containing peptide (21). Preference for the Leu can be attributed to the larger size of the Leu side chain and to proper positioning of residues N-terminal to the LPSY motif, which bind to residues...
The similarity among PY motif binding sites implies that specificity determinants of group I WW domains lie primarily outside of the region that binds to the core PY motif. In fact, interactions with regions N- and C-terminal to the PY motif may be a general requirement for achieving optimal affinity. This is similar to what has been observed for SH3 domain-ligand complexes, with interactions outside of the core PXXP motif providing specific ligand affinity in a number of cases (34). In the case of Smurfl2, the need for binding regions outside the PY motif is exacerbated by the binding site Phe. This may explain why the interface between Smurfl2 WW3 and the Smad7 PY peptide is more extensive than any previously observed. Nonetheless, the YAP65 WW, rNedd4 WW4, and dNedd4 WW3* domains all bind to peptide residues outside the PY motif.

Regions N- and C-terminal to the PY motif can interact with surfaces on the β1-strand and the β1–β2 loop (pink residues in Fig. 5, a and b), regions which exhibit more sequence variability than the PY motif binding pocket and provide specificity determinants differentiating group I ligands, as observed in the known structures of group I complexes. The affinity of the YAP65 WW domain for peptides with PY motifs can be enhanced by up to 9-fold by modifying regions N- and C-terminal to the PY motif, although the molecular interactions responsible are unknown (19). In the dNedd4 WW4 domain, βENaC peptide complex, a Leu three residues C-terminal to the PY motif binds to a hydrophobic pocket formed by Phe-28 and Ile-30 on the β2-strand and Arg-20 on the β1-strand. This βENaC Leu residue and its cognate binding pocket enhance the specificity of the interaction, and mutation of this Leu leads to a significant decrease in binding affinity (33). This hydrophobic binding pocket is conserved in the Smurfl2 WW3 domain, but the additional Leu interaction may not be sufficient to compensate for the reduction in affinity for the PY motif caused by the binding site Phe. A peptide Ala binds a similar position in the dNedd4 WW3* domain-Commissureless PY motif complex (13). Specificity in the dNedd4 WW3* domain-Commissureless PY motif complex is further enhanced by Ala-22, Asn-24, Arg-26 in the β1–β2 loop, and Trp-39 of dNedd4 WW3*, which form a binding pocket for a Gly and Thr N-terminal to the PY motif. The conservation of positively charged residues on the surface of many HECT-type E3 ubiquitin ligase WW domains (Fig. 5b, Arg-20, -26, and -35) hints that electrostatic interactions might be one form of interaction with ligand residues outside the PY motif participating in specificity. Indeed, the R-Smad proteins have acidic residues N- and C-terminal to the PY motif (Fig. 5c). There is evidence, as well, that interaction with this class of WW domains is enhanced by acidic peptide residues or phosphorylation (33, 35, 36). Interactions with regions N- and C-terminal to the PY motif could also inhibit binding. Altogether these data support a model in which the PY motif functions as an anchor with regions N- and C-terminal to the PY motif increasing (or decreasing) the affinity and modulating specificity.

The β1–β2 loop appears to be a specificity hot spot. This region may explain subtle differences in the specificity of Smurfl1 and -2, which have 85% sequence conservation. Two of the five different residues between these domains are on the β1-strand pointing away from the binding surface. The remaining differences are clustered in the β1–β2 loop (Fig. 5, a and b). Two of these are conservative changes, a Thr to Ser and a Val to Ile; however, there is also a more dramatic change, Ala-309 of Smurfl2 is a Val in Smurfl1. Although these changes are unlikely to seriously impact binding of the PY motif itself, they likely affect interaction with the PY-tail.

Unfortunately, two facts make specificity predictions within group I WW domains difficult. First, the structure of the bound peptide is critical for determining which peptide residues will bind to the surface of the WW domain. In the case of the rNedd4-βENaC complex, proper placement of the peptide Leu residue is dependent on a turn formed by the PY motif Tyr and the three residues that follow. In the Smurfl2-Smad7 complex, the PY-tail interaction relies on the peptide loop structure and intrapeptide interactions that stabilize this loop. Of note, the loop structure is not observed in the free Smad7 peptide, underscoring the difficulty in predicting the structure of the bound peptide (see supplemental material). Second, interactions with regions N- and C-terminal to the PY motif can affect significant changes in affinity without making a dominant contribution to the overall binding energy. Relatively small energetic contributions can have a significant impact on the dissociation constant and, hence, specificity. In fact, given a base dissociation constant of 10 μM, a 20% increase in binding energy results in a 10-fold increase in Kd. Because regions N- and C-terminal to the PY motif only need to make a small contribution to binding energy to significantly enhance the affinity, these additional interactions may be satisfied by a larger, but still limited, subset of amino acid sequences.
In summary, the PY motif of Smad7 serves as an anchor for specificity determinants N- and C-terminal to the motif. Although the requirement for additional interactions with regions N- or C-terminal to the PY motif is mandated by the Phe in the Smurf2 WW3 binding pocket, these specificity determinants appear to be a common theme in WW domain-PY motif complexes. Continuing exploration of interactions between group I WW domains and regions N- and C-terminal to PY motifs will further our understanding of WW domain recognition and help in elucidating specificity in targeted ubiquitin-mediated degradation pathways that help control TGF-β and other signaling processes.

Acknowledgments—We thank Drs. R. Muhandiram for help in setting up NMR experiments, A. Davidson for use of his fluorimeter, and R. Hudson, V. Kanelis, A. Velyvius, and R. Sprangers for useful discussions.

REFERENCES
1. Massague, J., Blain, S. W., and Lo, R. S. (2000) Cell 103, 295–309
2. Massague, J. (1998) Annu. Rev. Biochem. 67, 753–791
3. Zhu, H., Kavask, P., Abdollah, S., Wrana, J. L., and Thomsen, G. H. (1999) Nature 400, 687–693
4. Kavask, P., Rasmussen, R. K., Causing, C. G., Bonni, S., Zhu, H., Thomsen, G. H., and Wrana, J. L. (2000) Mol. Cell 6, 1365–1375
5. Herskho, A., and Ciechanover, A. (1998) Annu. Rev. Biochem. 67, 425–479
6. Lin, X., Liang, M., and Feng, X. H. (2000)
7. Lin, X., Liang, M., and Feng, X. H. (2000)
8. Bonni, S., Wang, H. B., Causing, C. G., Kavask, P., Stroschein, S. L., Luo, K., and Wrana, J. L. (2001) Nat. Cell Biol. 3, 587–595
9. Ebiawa, T., Fukuchi, M., Muralakami, G., Chiba, T., Tanaka, K., Imamura, T., and Miyazono, K. (2001) J. Biol. Chem. 276, 12477–12480
10. Bai, Y., Yang, C., Hu, K., Elly, C., and Liu, Y. C. (2004) Mol. Cell 15, 825–831
11. Kuratomi, G., Komuro, A., Goto, K., Shinozaki, M., Miyazawa, K., Miyazono, K., and Imamura, T. (2000) Biochem. J. 366, 461–470
12. Komuro, A., Imamura, T., Saitoh, M., Yoshida, Y., Yamori, T., Miyazawa, K., and Miyazawa, K. (2004) Oncogene 23, 6914–6923
13. Ogunjimi, A. A., Briant, D. J., Pece-Barbara, N., Le Roy, C., Di Guglielmo, G. M., Kavask, P., Rasmussen, R. K., Seet, B. T., Slicheri, F., and Wrana, J. L. (2005) Mol. Cell 19, 297–308
14. Macias, M. I., Wiesner, S., and Sudol, M. (2002) FEBS Lett. 513, 30–37
15. Koorf, E. K., Petrassi, H. M., Ratnaswamy, G., Huff, M. E., Sudol, M., and Kelly, J. W. (1999) Biochemistry 38, 14338–14351
16. Kanelis, V., Rotin, D., and Forman-Kay, J. D. (2001) Nat. Struct. Biol. 8, 407–412
17. Huang, X., Poy, F., Zhang, R., Joachimiak, A., Sudol, M., and Eck, M. J. (2000) Nat. Struct. Biol. 7, 634–638
18. Chen, H. I., and Sudol, M. (1995) Proc. Natl. Acad. Sci. U. S. A 92, 7819–7823
19. Pires, J. R., Taha-Nejad, F., Toepert, F., Ast, T., Hoffmann, U., Schneider-Mergener, J., Kuhe, R., Macias, M. I., and Oschkinat, H. (2001) J. Mol. Biol. 314, 1147–1156
20. Sudol, M., Chen, H. I., Bougeret, C., Einbond, A., and Bork, P. (1995) FEBS Lett. 369, 67–71
21. Kanelis, V., Bruce, M. C., Skrynnikov, N. R., Rotin, D., and Forman-Kay, J. D. (2006) Structure 14, 543–553
22. Gill, S. C., and von Hippe1, P. H. (1989) Anal. Biochem. 182, 319–326
23. Kanelis, V., Forman-Kay, J. D., and Kay, L. E. (2001) IUBMB Life 52, 291–302
24. Kanelis, V., Donaldson, L., Muhandiram, D. R., Rotin, D., Forman-Kay, J. D., and Kay, L. E. (2000) J. Biomol. NMR 16, 253–259
25. Wang, H., and Zuidweg, E. R. (1995) J. Biomol. NMR 5, 207–211
26. Pascal, S., Muhandiram, D. R., Yamazaki, T., Forman-Kay, J. D., and Kay, L. E. (1994) J. Magn. Reson. B 101, 197–201
27. Zwaliden, C., Legault, O., Vincent, S., Greenblatt, J., Konrat, R., and Kay, L. E. (1997) J. Am. Chem. Soc. 119, 6711–6721
28. Cornilescu, G., Delaglio, E., and Bax, A. (1999) J. Biomol. NMR 13, 289–302
29. Nilges, M. (1998) Prog. Nucl. Magn. Reson. Spectrosc. 32, 107–139
30. Laskowski, R. A., Mullmann, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) J. Biomol. NMR 8, 477–486
31. Chong, P. A., Ozdamer, B., Wrana, J. L., and Forman-Kay, J. D. (2004) J. Biol. Chem. 279, 40707–40714
32. Pelton, J. G., Torchia, D. A., Meadow, N. D., and Roseman, S. (1993) Protein Sci. 2, 543–558
33. Henry, P. C., Kanelis, V., O’Brien, M. C., Kim, B., Gauthais, I., Forman-Kay, J., Schild, L., and Rotin, D. (2003) J. Biol. Chem. 278, 20019–20028
34. Bauer, T., Schweimer, K., Meiselbach, H., Hoffmann, S., Rosch, P., and Sticht, H. (2005) Protein Sci. 14, 2487–2498
35. Schleinkofer, K., Wiedemann, U., Otte, L., Wang, T., Krause, G., Oschkinat, H., and Wade, R. C. (2004) J. Mol. Biol. 344, 865–881
36. Otte, L., Wiedemann, U., Schlegel, B., Pires, J. R., Beyermann, M., Schmieder, P., Krause, G., Volkmer-Engert, R., Schneider-Mergener, J., and Oschkinat, H. (2003) Protein Sci. 12, 491–500
37. Koradi, R., Billeter, M., and Wuthrich, K. (1996) J. Mol. Graph. 14, 29–32
38. Schultz, J., Milpetz, F., Bork, P., and Ponting, C. P. (1998) Natl. Acad. Sci. U. S. A. 95, 5857–5864