Identification of Novel Human WW Domain-containing Proteins by Cloning of Ligand Targets*

(Received for publication, January 31, 1997, and in revised form, April 2, 1997)

Gregorio Pirozzi‡§, Stephen J. McConnell‡¶, Albert J. Uveges‡, J. Mark Carter‡, Andrew B. Sparks∥**, Brian K. Kay∥, and Dana M. Fowlkes‡ ‡‡

From ‡Cytogen Corporation, Princeton, New Jersey 08540-5309 and the ||Curriculum in Genetics and Molecular Biology, Department of Biology, University of North Carolina, Chapel Hill, North Carolina 27599

A recently described protein module consisting of 35-40 semiconserved residues, termed the WW domain, has been identified in a number of diverse proteins including dystrophin and Yes-associated protein (YAP). Two putative ligands of YAP, termed WBP-1 and WBP-2, have been found previously to contain several short peptide regions consisting of PPPPY residues (PY motif) that mediate binding to the WW domain of YAP. Although the function(s) of the WW domain remain to be elucidated, these observations strongly support a role for the WW domain in protein-protein interactions. Here we report the isolation of three novel human cDNAs encoding a total of nine WW domains, using a newly developed approach termed COLT (cloning of ligand targets), in which the rapid cloning of modular protein domains is accomplished by screening cDNA expression libraries with specific peptide ligands. Two of the new genes identified appear to be members of a family of proteins, including Rsp5 and Nedd-4, which have ubiquitin-protein ligase activity. In addition, we demonstrate that peptides corresponding to PY and PY-like motifs present in several known signaling or regulatory proteins, including RasGAP, AP-2, p53BP-2 (p53-binding protein-2), interleukin-6 receptor- α , chloride channel CLCN5, and epithelial sodium channel ENaC, can selectively bind to certain of these novel WW domains.

The recognition and elucidation in recent years of various modular protein domains, along with their specific peptide ligands, have spawned a remarkable progress in our understanding of their role in signal transduction and other fundamental cellular processes (1, 2). Analysis of the SH (Src homology) domains, SH2 and SH3, present in a wide variety of proteins involved in cellular signaling and transformation has been particularly fruitful. The ligand specificity of many different SH2 and SH3 domains has been defined using combinatorial peptide libraries. SH2 domains bind with high affinity to phosphotyrosine residues within a specific sequence context (3). In contrast, SH3 domains bind to proline-rich peptides that share a conserved PXXP motif (4, 5).

A newly described protein module, termed the WW domain, has been reported (6-8). The WW domain consists of 35-40 amino acids and is characterized by four well conserved aromatic residues, two of which are tryptophan. The secondary structure of the WW domain has recently been determined and consists of a slightly bent three-stranded antiparallel β -sheet (9). This domain has been reported in a wide variety of proteins of yeast, nematode, and vertebrate origin, including Rsp5, Yesassociated protein (YAP),¹ human and murine Nedd-4, FE65, Pin1, and a human RasGAP-related protein (10-14). Although the precise physiological role of the WW domain remains undetermined, its presence in diverse proteins involved in signaling, regulatory, and cytoskeletal functions, as well as its rapidly emerging role in signaling mechanisms that underlie several human diseases, clearly underscores its importance (15, 16). Two ligand proteins for the YAP WW domain, WBP-1 and WBP-2 (WW domain-binding protein), have recently been cloned and found to contain a run of prolines followed by a tyrosine residue (PY motif) that mediate specific binding to the YAP WW domain (17). It is likely that WW domains have distinct ligand specificities, as the PY motifs of WBP-1 and WBP-2 did not bind to the dystrophin WW domain. Furthermore, the PY motif appears to be distinct from the PXXP ligand consensus sequence of SH3 domains. These observations suggest a direct role for the WW domain in mediating specific and distinct protein-protein interactions.

Recently, we described a method termed COLT (cloning of ligand targets), which enables the rapid cloning of ligandbinding modular protein domains using peptide ligand sequences as probes for screening cDNA expression libraries (18). While COLT has been used to clone several novel SH3 domaincontaining genes using proline-rich SH3 domain ligand peptides as probes, in this report, we describe its use in identifying three novel human WW domain-containing genes. In examining the ligand specificity of the individual WW domains, we demonstrate that they bind distinct PY motif peptide ligands with differential specificity and relative affinity. In addition, we demonstrate that peptides containing PY and PY-like motifs present in a variety of signaling or regulatory proteins can selectively bind to these novel WW domains.

EXPERIMENTAL PROCEDURES

Isolation and Characterization of WW Domain-encoding cDNA Clones by COLT—All peptides used were synthesized with an N-terminal biotin-SGSG linker and purified by high pressure liquid chromatography, and their structures were confirmed by mass spectroscopy and amino acid analysis. Multivalent peptide-streptavidin/alkaline phosphatase complexes were assembled as described (18) with the exception of the phosphotyrosine-containing peptide (pWBP-1), for

^{*} This work was supported by a grant from Cytogen Corp. (Princeton, NJ). 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.

[¶] Present address: Chugai Biopharmaceuticals Inc., 6275 Nancy Ridge Rd., San Diego, CA 92121.

^{**} Present address: Johns Hopkins Oncology Center, 424 North Bond St., Baltimore, MD 21231.

^{‡‡} Present address: Novalon Pharmaceutical Corp., 214 West Cameron Ave., Chapel Hill, NC 27599-7525.

[§] To whom correspondence should be addressed: Cytogen Corp., 201 College Rd. East CN 5309, Princeton, NJ 08540-5309. Tel.: 609-520-3060; Fax: 609-951-9779; E-mail: gpirozzi@mru3.cytogen.com.

 $^{^{1}}$ The abbreviations used are: YAP, Yes-associated protein; GST, glutathione S-transferase.





which a streptavidin-horseradish peroxidase conjugate was used (Sigma). Human bone marrow and brain λ cDNA libraries (CLONTECH, Palo Alto, CA) were plated at a density of $\sim 1 \times 10^5$ plaque-forming units/10-cm plate, and positive plaques were detected and purified as described previously (15). Bound streptavidin-horseradish peroxidase conjugate was detected with the IBI EnzygraphicTM Web (Kodak Scientific Imaging Systems) as described by the manufacturer. DNA was sequenced on both strands using ABI PRISMTM dye terminator cycle chemistry (Perkin-Elmer) on an ABI 373A automated DNA sequencer.

Preparation of WW Domain-GST Fusion Proteins and Peptide Crossaffinity Mapping-Polymerase chain reaction fragments encoding individual WW domains were subcloned into the SalI and NotI sites of pGEX-4T-2 (Pharmacia Biotech Inc.), and fusion proteins were expressed and purified as described by the manufacturer. Enzyme-linked immunosorbent assay-based cross-affinity experiments were performed essentially as described (5), with the following modifications. Briefly, microtiter wells were coated with 1–5 μ g of fusion protein in 100 mM NaHCO₃, blocked with SuperBlock Tris-buffered saline (Pierce), and washed four times with phosphate-buffered saline and 0.05% Tween 20. Specific peptide-streptavidin/alkaline phosphatase complexes were added as described above, and unbound complexes were washed five times with phosphate-buffered saline and 0.05% Tween 20. Following addition of p-nitrophenyl phosphate substrate (Kirkegaard & Perry Laboratories), peptide binding was quantitated after 30 min at $A_{405 \text{ nm}}$. Relative binding measurements from three independent determinations were assigned to a scale as follows: A units of 0-0.5 = (-), 0.5-1.0 = (+), 1.0-2.0 = (++), 2.0-3.0 = (+++), and > 3.0 = (++++).Peptide binding to full-length human Fyn and Lyn was assessed by a filter binding assay (5). Peptide sequences used in cross-affinity experiments correspond to segments of the following proteins: RasGAP (GenBank accession number P20936); the AP-2 transcription factor (P05549); p53BP-2 (p53-binding protein-2) (U09582); interleukin-2 receptor- γ (D11086); interleukin-6 receptor- α (P22272); voltage-gated chloride channel CLCN5 (X91906); Rous sarcoma virus Gag (D10652); human T-cell lymphotropic virus type 1 Gag (D13784); β -dystroglycan (L19711); Formin (X53599); amiloride-sensitive epithelial Na⁺ channel ENaC α (P37089), ENaC β (X87159), and ENaC γ (X87160); muscarinic acetylcholine receptor M4 (P08173); and c-Abl (P00522). Src and Crk SH3-binding peptide sequences were derived from a phage display random peptide library screen (5). Protein sequence homology searches were performed using BLAST (19) and PROFILES (20) programs.

RESULTS AND DISCUSSION

Isolation of Novel WW Domain-containing Proteins—In an effort to identify novel WW domain-containing proteins, four putative WW domain peptide ligand sequences including WBP-1 (PGT<u>PPPPY</u>TVGPGY), WBP-2A (YVQP<u>PPPPY</u>PGPM), WBP-2B (PGTP<u>YPPPPEFY</u>), and a PY peptide segment of Ras-GAP (GGGFPPL<u>PPPPY</u>LPPLG) were used as a mixed probe to screen human brain and bone marrow cDNA expression libraries by COLT. Thirteen positive clones were identified, including sibling and overlapping clones originating from the same mRNA. These clones were sequenced and found to encode three distinct proteins, WWP1, WWP2, and WWP3 (<u>WW</u> domaincontaining protein), containing a total of nine novel WW domains (Fig. 1). Data base homology searches revealed that both WWP1 and WWP2 contain four tandem WW domains and share a similar modular domain architecture with Nedd-4 and Rsp5 (10). The smaller partial clone, WWP3, was isolated from the brain cDNA library and contains a single WW domain.

In addition to the high overall amino acid homology (55%), alignment of the nine novel WW domain sequences with several previously identified domains reveals two significant blocks of homology flanking the core of the domain (Fig. 2). These blocks include N-terminal tryptophan and C-terminal proline residues that are absolutely conserved in all WW domains identified to date. The WW domains of WWP1, WWP2, WWP3, Nedd-4, Rsp5, and YAP are more similar to each other than to WW domains found in other proteins (21). Furthermore, contrary to what one would expect from a recent evolutionary duplication event of the WW domain within a gene, individual WW domains from WWP1 and WWP2 show a greater similarity to the corresponding WW domains in either Nedd-4 or Rsp5 than to each other. The above observations suggest that these WW domains are perhaps functionally similar and that multiple WW domains within the same protein are not redundant but may have evolved to perform divergent specialized roles. Interestingly, we did not isolate a human YAP cDNA from these screens. However, murine YAP was detected in a screen of a mouse embryonic cDNA library with the WBP-2A peptide as a COLT probe.²

In addition to the WW domains, primary structure analysis of the clones revealed several other interesting features. Complete and partial C-terminal HECT (homologous to the E6associated protein carboxyl terminus) domains of ~300 amino acids (Fig. 3) are contained within clones WWP2 and WWP1, respectively. This domain has been shown in vitro to have E3 ubiquitin-protein ligase activity in several proteins including rat p100, yeast Rsp5, and human papilloma virus E6-AP (22, 23). Encoded within the last 40 amino acids of the HECT domain is a conserved cysteine residue that is the likely site for ubiquitin thioester formation. The presence of a HECT domain is noteworthy since structurally and functionally related E3 ubiquitin-protein ligases are thought to serve a major role in defining the substrate specificity of the ubiquitin degradation system (24). In fact, Rsp5 was recently shown to be involved in the induced degradation of several nitrogen permeases in yeast (25). WWP2 also encodes an N-terminal C2-like domain characteristic of a large family of proteins including protein kinase C (26) and synaptotagamins (27). The C2 domain has been shown to bind membrane phospholipids in a calcium-dependent manner and is thought to function in the intracellular compartmentalization of proteins (28). Although the different modular domains present within WWP1 and WWP2 are highly homologous to those found in Nedd-4 and Rsp5, there is no significant homology among these proteins in regions flanking these domains. Also of interest is the presence in clone WWP3

² R. Gao and B. K. Kay, unpublished data.

Gene/Species	pos.	WW Domain Sequences	Acc. No.
WWP1-1		LPSGWGWEQRKDPHGRTYYVDHNTRTTTWERPQPLPPG	U96113
WWP1-2		QP LP P GWE RRVDDRRRV YY VD HNT R TT T W QR P TMESVR	
WWP1-3		GPLPPGWEKRVDSTDRVYFVNHNTKTTQWEDPRTQGLQ	
WWP1-4		EPLPEGWEIRYTREGVRYFVDHNTRTTFKDPRNGKSS	
WWP2-1		DALPAGWEQRELPNGRVYYVDHNTKTTTWERPLPPGWE	U96114
WWP2-2		RPLPPGWEKRTDPRGRFYYVDHNTRTTTWQRPTAEYVR	
WWP2-3		GPLPPGWEKRQD.NGRVYYVNHNTRTTQWEDPRTQGMI	
WWP2-4		PALPPGWEMKYTSEGVRYFVDHNTRTTTFKDPRPGFES	
WWP3		GPLPENWEMAYTENGEVYFIDHNTKTTSWLDPRCLNKQ	U96115
YAP/Human	171	VPLPAGWEMAKTSSGQRYFLNHIDQTTTWQDPRKAMLS	X80507
YAP/Mouse-1	151	VPLPAGWEMAKTSSGQRYFLNHNDQTTTWQDPRKAMLS	X80508
YAP/Chick-1	169	VPLPPGWEMAKTPSGQRYFLNHIDQTTTWQDPRKAMLS	X76483
YAP/Mouse-2	218	GPLPDGWEQAMTQDGEVYYINHKNKTTSWLDPRLDPRF	X80508
YAP/Chick-2	229	GPLPDGWEQAMTQDGEIYYINHKNKTTSWLDPRLDPRF	X76483
Nedd4/Human-	1 218	SPLPPGWEERQDILGRTYYVNHESRRTQWKRPTPQDNL	D42055
Nedd4/Mouse-	1?	SPLPPGWEERQDVLGRTYYVNHESRRTQWKRPSPDDDL	D10714
Rsp5/Yeast-1	228	GRLPPGWERRTDNFGRTYYVDHNTRTTTWKRPTLDQTE	L11119
Ned4/Human-2	375	SGLPPGWEEKQDERGRSYYVDHNSRTTTWTKPTVQATV	D42055
Ned4/Mouse-2	?	SGLPPGWEEKQDDRGRSYYVDHNSKTTTWSKPTMQDDP	D10714
Rsp5/Yeast-2	331	GELPSGWEQRFTPEGRAYFVDHNTRTTTWVDPRRQQYI	L11119
Ned4/Human-3	448	GFLPKGWEVRHAPNGRPFFIDHNTKTTTWEDPRLKIPA	D42055
Ned4/Mouse-3	?	GP LP P GWE ERTHTD G RVFFIN HN IKK T Q W ED PR LQNVA	D10714
Rsp5/Yeast-3	387	GPLPSGWEMRLTNTARVYFVDHNTKTTTWDDPRLPSSL	L11119
Ned4/Human-4	500	GPLPPGWEERTHTDGRIFYINHNIKRTQWEDPRLENVA	D42055
Dmd/Human	3052	TSVQGPWERAISPNKVPYYINHETQTTCWDHPKMTELY	P11532
Dmd/Torca	253	TSVQGPWERAISPNKVPYYMNHQTQTTCWDHPKMTELY	M37645
Utro/Human	2813	TSVQLPWQRSISHNKVP YY INHQTQTTCWDHPKMTELF	X69086

LP GWE tttGt YYh HNTtTTtW tPt

Consensus

FIG. 2. Deduced amino acid sequence and alignment of novel **WW domains**. The name, species, amino acid position (if known), WW domain sequence, and accession number of representative WW domain sequences are shown. Amino acids conserved in >75% of the sequences are *shaded*. The consensus shows conserved amino acids including hydrophobic (*h*) and polar (*t*) residues.

of a partial N-terminal guanylate kinase-like domain that is involved in GMP binding in several membrane-associated proteins including the human erythrocyte membrane protein p55 (29) and rat postsynaptic density protein PSD-95 (30).

WW Domain-Peptide Ligand Selectivity-To examine the peptide ligand binding preferences of all nine individual novel WW domains, an enzyme-linked immunosorbent assay-based cross-affinity map experiment was performed with each domain expressed as a GST fusion protein and WBP-derived peptides (Fig. 4). Peptides WBP-1, WBP-2A, and WBP-2C bound to several individual WW domains to varying degrees. The WBP-2B peptide with an N-terminal tyrosine residue relative to the run of prolines had no binding activity, indicating the necessity for a C-terminal tyrosine in the PY motif. The relative importance of individual proline residues within the PY motif for binding to various WW domains was assessed by alanine substitution for both the WBP-1 and WBP-2A peptides. All of the variant WBP-1 peptides, with the exception of the third proline substitution (WBP-1-Pro3), retained binding activity to WW domains present in clones WWP1 and WWP2, suggesting a critical role for the third proline residue. Interestingly, substitution of the second proline residue (WBP-1-Pro2) did not abolish binding to WW domains WWP1.1 and WWP2.3. This was unanticipated in light of the results observed for binding of the WBP-1 protein PY motif to the YAP WW domain, in which both the second and third proline residues are crucial for binding (9, 17). This difference suggests that WW domains WWP1.1 and WWP2.3 possess a more promiscuous binding specificity than does the YAP WW domain. Similarly, proline substitution of the WBP-2A peptide indicates that the third proline residue (WBP-2A-Pro3) is absolutely essential for WW domain binding, whereas substitution of the second proline (WBP-2A-Pro2) is not.

The specificity of individual WW domains for PY motif sequences was demonstrated by the ability to discriminate between peptides containing SH3 domain PXXP ligand consensus sequences (Src and Crk) as well as generally proline-rich control peptides derived from several proteins including acetylcholine receptor M4 and c-Abl. In addition, none of the PY motif peptides bound to either full-length Fyn or Lyn (which contain both SH3 and SH2 domains) in filter binding assays. Taken together, these results suggest that the PY motif represents a distinct binding sequence for WW domains.

The presence of a critical tyrosine residue in the PY motif raised the question of whether tyrosine phosphorylation could modulate WW domain binding. Although it is not known whether PY motifs are phosphorylated *in vivo*, the presence of a phosphotyrosine residue in the pWBP-1 peptide abolishes WW domain binding. Moreover, binding of the pWBP-1 peptide could be restored by removal of the phosphate moiety with prior treatment of either the free peptide or peptide bound to a streptavidin-horseradish peroxidase conjugate with alkaline phosphatase (data not shown). These results suggest a potential regulatory role for tyrosine phosphorylation in modulating WW domain-ligand interactions.

Potential WW Domain-PY Motif Interactions-Data base searches revealed that PY and PY-like motif sequences are found in a wide variety of regulatory proteins. Included among these proteins are the GTPase-activating protein RasGAP, the AP-2 transcription factor, p53BP-2, renal chloride channel CLCN5, the dystrophin-interacting molecule β -dystroglycan, the interleukin-2 receptor and interleukin-6 receptor- α , and the retroviral Gag proteins from human T-cell lymphotropic virus type 1 and Rous sarcoma virus type 1 (30-36). We tested the ability of peptides containing these motifs to bind to the novel WW domains (Fig. 5). Interestingly, although all of these peptides displayed an ability to bind WW domains in general, differences in specificity and relative binding were evident. For example, of all the peptides tested, only the CLCN5 peptide showed appreciable binding to the WWP1.4 and WWP2.4 domains. The observation that PY motif-containing peptides from several other proteins did not bind to any WW domain tested indicates that these interactions are specific and potentially biologically relevant (data not shown).

Of particular note is the demonstration that the human T-cell lymphotropic virus type 1 and Rous sarcoma virus type 1 peptides derived from the Gag protein proline-rich "L domain" bound to several WW domains. L domain regions are highly conserved in retroviruses and have been shown to function in a positionally independent manner essential for retroviral budding (37). Our results, coupled with a recent report demonstrating the interaction of the YAP WW domain with the L domain of Rous sarcoma virus (38), suggest a direct role for a WW domain(s)-Gag protein interaction in this process. The interaction of a β -dystroglycan peptide with several WW domains is also of interest. β -Dystroglycan, which contains a C-terminal PY motif, was previously shown to interact with the single WW domain present in dystrophin (15). Our results suggest that perhaps several different WW domain-containing proteins can interact with the β -dystroglycan C-terminal PY motif. Recently, a 12-amino acid proline-rich region of Formin, a protein encoded by the mouse limb deformity locus (39), was shown to bind to both SH3 domain- and several novel WW domaincontaining proteins (40). Significantly, a peptide encompassing the same proline-rich region of Formin did not bind to any of our novel WW domains. Since this peptide does not contain a PY motif, this suggests that our WW domains, unlike those present in the Formin-binding proteins, require a PY or PY-like motif for binding.

Taken together, the above observations suggest that interactions between these proteins and WW domain-containing proteins may play a role in the former's regulation *in vivo*. For example, given the likelihood that WWP1 and WWP2 function as E3 ubiquitin-protein ligases, one could invoke a simple model whereby initial substrate-specific recognition occurs via WW domain-substrate protein interaction followed by ubiquitin transfer and subsequent proteolysis. On a general level,

Gene	C-Terminal HECT Domain Sequences							
	-100 -50							
WWP1	YR-HYTRNSKQIIWFWQFVKETDNEVRMRLLQFVTGTCRLPLGGFAELMGSN *							
WWP2	YR-HYTKNSKQIQWFWQVVKEMDNEKRIRLLQFVTGTCRLFVGGFAELIGSNGPQKFCID-KVGKETWLFRSHTCFNRLDLPPYKSYEQLREKLLYAI-EETE	GEGQE						
RSP5	YR-GYQESDEVIQWFWKCVSEWDNEQRARLLQFTTGTSRIFVNGFKDLQGSDGPRRFTIE-KAGEVQQLPKSHTCFNRVDLFQYVDYDSMRQRLTLAV-EETI	GFGQE						
NEDD-4	FNDESGENAEKLL-IHWFWKAVWMDSEKRIRLLQFVTGTSRVPMNGFAELYGSNGPQSFTVE-QWGTPDRLPRAHTCFNRLDLPPYESFDELWDRLOMAI-ENTQ	GFDHVD	L11119					
E6-AP	YDGGYTRDSVLIREFWEIVHSFTDEORRLFLOFTTGTDRAPVGGLGRLKMIIAKNGPDTERLPTSHTCFNVLLLPEYSSKEKLRERLLKAIT-YAR	GFGML	D10714					
Rat p100	YKGDYSATHPT-QFKRWFWSIVERMSMTERQDLVYFWTSSPSLPASEEGFQPMPSITIRPPDDQHLPTANTCISRLYVPLYSSRQILRQRLLLAIKTR	NFGFV	L07557					
D13635	YSGGYSADHPVIRVFWRVVEGFTDEEKRKLLKFVTSCSRPPLLGFKELYPAFCIHNGGSDLERLPTASTCMNLLKLPEFYDETLLRSRLLYAI-ECAA	GFELS	S22659					
D25215	PDHGYTHDSRAVK-VRLFWETFHEFPLEKKRKFLLFLTGSDRIPIYGMASLPSITIQSTASGEEYLPVAHTCYNLLDLPKYSSREILSARLTQAL-DNYE	GESLA /	D13635					
D28476	AEHGYTMDSSIFLF-EILSSFDNEQQRLFLQFVTGSPRLPVGGFRSLNPPLQIVIESTENPDDFLPSVMTCVNYLKLPDYSSIEIMRERLLIAAREGQQ	SFHLS	D25215					
Dros.HYD	FNDESSEGPDKLKFKRWFWSIVEKMNIMERQHLVYFWTGSPALPASEEGFQPLPSVTIRPADDSHLPTANTCISRLYIPLYSSRSILRSKNIMAIKSR	NFGFV I	D28476					
YKL162	YKNGYSMNHQVIHDFISIISAFGKHERRLFLQFLTGSPRLPIGGFKSLNPKFTIVRKTFEDGLTADEYLPSVMTCANYLKLPKYTSRDIMRSRLCQAIEEGAG	AFLLS	P33202					
ScORF	YVGGFSDDSRAVCWFWEIIESWDYPLQRKLLQFVTASDRIPATGISTIPFKISLLGSHDSDDLPLAHTCFNEICLWNYSSRKRLELRLLWAI-NESE	GYGFR	L26910					
UreBl	Y-HKYOSNSIOIOWEWRALRSFDOADRAKFLOFVTGTSRVPLOGFAALEGMNGIOKFOIHRDDRSTDRLPSAHTCFNOLDLPAYESFEKLRHMLLLAIOECSE	GFGLANK	008214					

FIG. 3. Alignment of WWP1 and WWP2 with representative C-terminal HECT domains. The C-terminal 100 amino acids of HECT domain-containing genes are shown aligned with the partial and complete C termini of WWP1 and WWP2, respectively. Completely and partially (>60%) conserved residues are *shaded* and in *boldface*, respectively. The *asterisk* denotes the conserved cysteine residue that serves as a site for ubiquitin thioester formation.

		WW Domain/GST Fusion Proteins										
Peptide	Sequence	WWP1.1	WWP1.2	WWP1.3	WWP1.4	WWP2.1	WWP2.2	WWP2.3	WWP2.4	WWP3	Fyn	Lyn
WBP-1	PGT <u>PPPPY</u> TVGPGY	++++	-	+++	-	+++	-	++++	++++	++++	-	-
WBP-2A	YVQP <u>PPPPY</u> PGPM	++++	+	+++	-	+++	-	++++	++++	++++	_	-
WBP-2B	PGYP <u>YPPPP</u> EFY	-		-	-	-	-	-	_	-	-	-
WBP-2C	TSQP <u>PPPPY</u> YPP	+++	-	+++	-	+++	-	++++			-	-
WBP-1-Pro1	PGT <u>APPPY</u> TVGPGY	·+··	-	+++	-	+++	-	++++	++++	++++	-	-
WBP-1-Pro2	PGT <u>PAPPY</u> TVGPGY	+++	-	-	-	-	-	+	-	-	-	-
WBP-1-Pro3	PGT <u>PPAPY</u> TVGPGY		-			-		-		-	-	-
WBP-1-Pro4	PGT <u>PPPAY</u> TVGPGY	+++	-	+++	-	+++	-	++++	++	++++	-	-
WBP-2A-Pro1	YVQ A<u>PPPPY</u>PGP M	++++	+	++++	-	++++	-	++++	++	++++		-
WBP-2A-Pro2	YVQP <u>APPPY</u> PGPM	++++		++++	-	+++	-	++++	+++	++++	-	-
WBP-2A-Pro3	YVQP <u>PAPPY</u> PGPM	+	-	+	-	+	-	++	-	++	-	-
WBP-2A-Pro4	YVQP <u>PPAPY</u> PGPM		~~	-		-	-				-	-
WBP-2A-Pro5	YVQP <u>PPPAY</u> PGPM	+++	-	+++	-	+++	-	+++	-	+++	-	-
pWBP-1	PGTPPPPpYTVGPGY	-				-	-	-	_			
M4 AChR	PPPALPPPPRPVADK	-	-	-	-	-	-	-	-	-		
c-Abl	SRLKPAPPPPPAASAG	-	-	_	-	-	-	-	-	_	-	_
Src	GILAPPVPPRNTR	-	-	-	-	-	-	-	-	_	++	++
Crk	SVPAPPPLPPKSGG	-	-	-	-	_	_	_	_	_	+	-

FIG. 4. **WW domain-WBP peptide interactions.** Biotinylated peptides were tested for their relative binding to individual WW domains expressed as GST fusion proteins and to full-length Fyn and Lyn in filter binding assays (see "Experimental Procedures"). PY and PY-like motifs within the peptide sequences are *underlined*, and specific alanine substitution variants of the WBP-1 and WBP-2A peptides are indicated in *boldface*. pWBP-1 denotes phosphotyrosine-containing peptide. Relative binding was assessed from three independent determinations. All PY motif peptides displayed no detectable binding to GST control protein or to bovine serum albumin. *AChR*, acetylcholine receptor.

			WW Domain/GST Fusion Proteins									
Peptide	Sequence	WWP1.1	WWP1.2	WWP1.3	WWP1.4	WWP2.1	WWP2.2	WWP2.3	WWP2.4	WWP3	Fyn	Lyn
RasGap	GGGFPPL <u>PPPPY</u> LPPLG	++++	-	+++	-	+++	-	+++	++++	++++		-
AP-2	ADFQPPY <u>FPPPY</u> QPIYPQS	+++	+	-	-	-	-	-	++++	++	-	
p53BP-2	EYPPY <u>PPPPY</u> PSGE	+++++	+++	+++		++++	+++	+++	+++++	++	_	-
CLCN5	HSP <u>PLPPY</u> TPPTL	++	-	+	++	++	-	-	++++	++++	-	-
IL-2Rγ	QHSPYW <u>APPCY</u> TLKPET	+	-	-	-	-	-	++	-	-	—	—
IL-6Ra	SKTTS <u>PPPPY</u> SLGPLK	+++	+	++++	-	++++	++	+++	++++	++++	-	-
β-Dystroglycan	MTPYR <u>SPPPY</u> VPP	+	-	+	-	-	-	+++	++++	++++	-	-
HTLV-1 Gag	PDSDPQIPPPYVEPTA	+	-	-	-	++	-	-	++++	-	-	-
RSV-1 Gag	TATASA <u>PPPPY</u> VGSGL	++++	-	+++	-	+++	+	++	++++	++++	-	-
Formin	APPTPPPLPP	-	-		-	-	-	-		-	-	-

FIG. 5. Naturally occurring WW domain-PY motif interactions. PY and PY-like motif peptides derived from representative proteins (see "Experimental Procedures") were tested for their relative binding to WW domain-GST fusion proteins. $IL-2R\gamma$ and $IL-6R\alpha$, interleukin-2 receptor- γ and interleukin-6 receptor- α , respectively; HTLV-1, human T-cell lymphotropic virus type 1; RSV-1, Rous sarcoma virus type 1.

the identification of WW domain PY motif ligand sequences in various candidate proteins can lead to testable predictions of specific WW domain-mediated interactions *in vivo*.

WW Domain-Epithelial Na⁺ Channel Interactions—The demonstration that peptides containing PY-like motifs derived from the cytoplasmic domains of both the wild-type β - and γ -subunits of the epithelial Na⁺ channel (ENaC β -WT and ENaC γ -WT) bind to several WW domains is of particular interest (Fig. 6). Recently, a number of mutations in both ENaC β and ENaC γ have been demonstrated in patients with an autosomal dominant form of hypertension characterized by elevated renal Na⁺ reabsorption, termed Liddle syndrome (41). Specifically, several nonsense mutations leading to the truncation of the cytoplasmic domain of both subunits, in addition to two missense mutations (P616L and Y618H) contained within a conserved proline-rich segment of the cytoplasmic domain of the β -subunit, have been identified in patients (42–44). Moreover, expressed proteins containing these mutations resulted in a 3–8-fold increase in ENaC activity that was directly related to an increase in the total number of active channels. These results suggest the hypothesis that specific cytoplasmic regions of the β - and γ -subunits are involved in the normal negative regulation of channel activity via interactions with a modulatory protein(s). In fact, Nedd-4 was recently identified as a binding partner with the C terminus of rat ENaC β using the yeast two-hybrid system (45, 46). In addition, we have recently

14010

		WW Domain/GST Fusion Proteins										
Peptide	Sequence	WWP1.1	WWP1.2	WWP1.3	WWP1.4	WWP2.1	WWP2.2	WWP2.3	WWP2.4	WWP3	Fyn	Lyn
ENaCα-WT	LTA <u>PPPAY</u> ATLGP	+++	-	++	-	+++		+++	-		-	-
ENaCα	LTA <u>PPPAA</u> ATLGP	-	-	-	-	-	-	-	-	-		1
ENaCβ-WT	PGT <u>PPPNY</u> DSLRL	++	-	-	-	-	-	++	-	++	-	-
ENaCβ-P616L	PGT <u>PPLNY</u> DSLRL	-	-	-	-		-	-	-	-	-	
ENaCβ-Y618H	PGT <u>PPPNH</u> DSLRL	-	-	-	-	-	-	-	-	-	-	-
ΕΝαCβ	PGT <u>APPNY</u> DSLRL	++		-	-		—	+	-	++	_	-
ΕΝαCβ	PGT <u>PAPNY</u> DSLRL	-	-	-	-	-	-	-	-	-	-	1
ΕΝαCβ	PGT <u>PPPNA</u> DSLRL	-	-	-	-	-	-			-	-	-
ΕΝαCβ	PGT <u>PPPNY</u> D A LRL	+	-	-	-	-	-	+	-	-	-	-
ΕΝαCβ	PGT <u>PPPNY</u> DSARL	+	-	-	-	-		-	-	+	-	-
ENaCβ	PGT <u>PPPNY</u> DS E RL	-	-	-	-	-	-	-	-	-	-	-
ENaCγ-WT	PGT <u>PPPKY</u> NTLRL	++	-	+		++	-	+		+++	-	-
ENaCγ	PGTPPPKANTLRL	-	-	-	-	-	-	-	_	-		-

FIG. 6. WW domain-human epithelial sodium channel interactions. PY motif-containing peptides derived from the wild-type α -, β -, and γ -subunits of human ENaC and several variants were tested for their relative binding to WW domain-GST fusion proteins. ENaC β -P616L and ENaC β -P618H denote peptides containing specific missense substitutions found in Liddle syndrome patients. Amino acid substitutions are indicated in *boldface*.

isolated WWP1 in COLT screens using $ENaC\beta$ and $ENaC\gamma$ peptides (data not shown).

Our observation that mutant peptides (ENaC\beta-P616L and ENaC_B-Y618H) containing missense substitutions found in Liddle syndrome patients do not bind to the WW domains in clones WWP1 and WWP2 is consistent with the above hypothesis. This result also confirms the observation that the third proline residue and the tyrosine within the PY motif are critical for binding to the WW domain. Other substitutions of the β -subunit PY motif and flanking sequences were also shown to diminish binding to specific WW domains. Thus, substitution of the second proline residue of the core PY motif completely abrogated WW domain binding. In addition, mutation of specific residues flanking the C terminus of the PY motif also led to diminished WW domain binding. These results directly correlate with the activity of various $ENaC\beta$ mutants measured by a functional assay in Xenopus oocytes (47). A PY motifcontaining peptide from the cytoplasmic domain of the wildtype α -subunit of ENaC (ENaC α -WT) was also shown to bind to several WW domains, suggesting that this subunit may also be regulated by a WW domain-mediated interaction(s). Taken together, the above observations suggest a mechanism whereby a WW domain-mediated interaction(s) of a Nedd-4 family member(s) leads to the eventual ubiquitin-mediated degradation and negative regulation of the Na⁺ channel.

We have used the COLT approach to identify two new members (WWP1 and WWP2) of a family of human Nedd-4-like proteins associated with ubiquitin-protein ligase activity. Moreover, we demonstrate that individual WW domains of these proteins can clearly bind distinct PY motif peptide ligands with differential specificity and relative affinity. In addition to demonstrating the expanded utility of COLT methodology, it appears likely that additional proteins containing the modular WW domain remain to be found. In fact, we have recently isolated a third novel Nedd-4-like family member, containing three WW domains, from a human prostate cDNA library.3 Our results, coupled with knowledge of the NMR structure of the WW domain-ligand complex, provide a framework with which to examine the peptide ligand specificity and structure/function activity of individual WW domains. In this regard, optimal ligand preferences of the WW domains are currently being deduced by screening combinatorial phage display peptide libraries.⁴ Given the small size and high degree of sequence conservation of the WW domain, it is extraordinary

that exquisite ligand selectivity is observed. The NMR structure of the human YAP WW domain and its peptide ligand reveals that the hydrophobic residues Leu-190, His-192, and Trp-199 (see Fig. 2) form a binding site in contact with the ligand (9). In light of these data, it is interesting to note that domains WWP1.4 and WWP2.4, which contain a C-terminal phenylalanine residue instead of a tryptophan, display a more restrictive ligand binding preference. In addition, the presence of a valine or isoleucine residue instead of Leu-190 may also play a role in determining the distinct ligand specificity of the novel WW domains. The presence of multiple WW domains with distinct ligand specificities in WWP1 and WWP2 suggests that these proteins may bind to a broad range of cellular targets. Alternatively, multiple WW domains may confer additive binding affinity to target molecules that contain multiple PY motif ligands.

Finally, the interaction of peptides containing PY and PYlike motifs from several proteins with the WW domains in WWP1 and WWP2 directly suggests a role for the ubiquitinmediated degradation of these proteins. In this respect, it is noteworthy that several cell membrane proteins, including the platelet-derived growth factor receptor (48) and yeast α -factor receptor Ste2p (49), are subject to ubiquitination and eventual degradation upon ligand binding. This is particularly relevant in light of the observed interaction of PY motif-containing peptides from ENaC subunits with specific WW domains and may lead to an understanding of the molecular pathology of Liddle syndrome.

Acknowledgments—We thank Jeremy Kasanov and Rong Gao for sharing unpublished results and Greg Miller for expert assistance in peptide synthesis.

REFERENCES

- 1. Cohen, G. B., Ren, R., and Baltimore, D. (1995) Cell 80, 237-248
- 2. Pawson, T. (1995) Nature 373, 573-560
- Songyang, Z., Shoelson, S. E., Chaudhuri, M., Gish, G., Pawson, T., Haser, W. G., King, F., Roberts, T., Ratnofski, S., Lechleider, R. J., Neel, B. G., Birge, R. B., Fajardo, J. E., Chao, M. M., Hanafusa, H., Schaffhausen, B., and Cantley, L. C. (1993) Cell 72, 767–778
- Chen, J. K., Lane, W. S., Brauer, A. W., Tanaka, A., and Schreiber, S. L. (1993) J. Am. Chem. Soc. 115, 12591–12952
- Sparks, A. B., Rider, J. E., Hoffman, N. G., Fowlkes, D. M., Quilliam, L. A., and Kay, B. K. (1996) Proc. Natl. Acad. Sci. U. S. A.93, 1540–1544
- Bork, P., and Sudol, M. (1994) Trends Biochem. Sci. 19, 531–533
 André, B., and Springael, J.-Y. (1994) Biochem. Biophys. Res. Commun. 205,
- 1201–1205 8. Hofmann, K., and Bucher, P. (1995) *FEBS Lett.* **358**, 153–157
- Hormann, K., and Bucher, P. (1995) FEBS Lett. 398, 153–157
 Macias, M. J., Hyvönen, M., Baraldi, E., Schultz, J., Sudol, M., Saraste, M.,
- and Oschkinat, H. (1996) Nature **382**, 646–649 10. Sudol, M., Bork, P., Einbond, A., Kastury, K., Druck, T., Negrini, M., Huebner,
- Sudol, M., Bork, P., Einbond, A., Kastury, K., Druck, T., Negrini, M., Huebner, K., and Lehman, D. (1995) J. Biol. Chem. 270, 14733–14741
- Kumar, S., Tomooka, Y., and Noda, M. (1992) Biochem. Biophys. Res. Commun. 185, 1155–1161

³ A. J. Uveges and G. Pirozzi, unpublished data.

⁴ J. Kasanov, G. Pirozzi, and B. K. Kay, unpublished data.

- Lu, K. P., Hanes, S. D., and Hunter, T. (1996) Nature 380, 544–547
 Duilio, A., Zambrano, N., Mogavero, A. R., Amendola, R., Cimino, F., and Russo, T. (1991) Nucleic Acids Res. 19, 5269–5274
- Weissbach, L., Settleman, J., Kalady, M. F., Snijders, A. J., Murthy, A. E., Yan, Y.-X., and Bernards, A. (1994) J. Biol. Chem. 269, 20517-20521
- 15. Einbond, A., and Sudol, M. (1996) FEBS Lett. 384, 1-8
- 16. Sudol, M. (1996) Trends Biochem. Sci. 21, 161–163
- Chen, H. I., and Sudol, M. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 7819–7823
 Sparks, A. B., Hoffman, N. G., McConnell, S. J., Fowlkes, D. M., and Kay, B.
- K. (1996) Nat. Biotechnol. 14, 741–744
- 19. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) J. Mol. Biol. 215, 403-410
- 20. Gribskov, M., McLachlan, A. D., and Eisenberg, D. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 4355-4358
- 21. Sudol, M., Chen, H. I., Bougeret, C., Einbold, A., and Bork, P. (1995) FEBS Lett. 369, 67–71
- Huibregtse, J. M., Scheffner, M., Beaudenon, S., and Howley, P. M. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 2563–2567
- 23. Scheffner, M., Nuber, U., and Huibregtse, J. M. (1995) Nature 373, 81-83
- 24. Ciechanover, A. (1994) Cell 79, 13-21
- 25. Hein, C., Springael, J.-Y., Volland, C., Haguenaur-Tsapis, R., and André, B.
- Hell, C., Springaei, S. I., Vohand, C., Haguehau, Psapis, K., and Linde, D. (1995) Mol. Microbiol. 18, 77–87
 Kaibuchi, K., Fukamoto, Y., Oku, N., Takai, Y., Arai, K., and Muramatsu, M. (1989) J. Biol. Chem. 264, 13489–13496
 Sutton, R. B., Davletov, B. A., Berghuis, A. M., Südhof, T. C., and Sprang, S.
- R. (1995) Cell 80, 929-938
- 28. Davletov, B. A., and Südhof, T. C. (1993) J. Biol. Chem. 268, 26386-26390
- Ruff, P., Speicher, D. W., and Husain-Chishti, A. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 6595–6599
- 30. Cho, K.-O., Hunt, C. A., and Kennedy, M. B (1992) Neuron 9, 929-942
- Trahey, M., Wong, G., Halenbeck, R., Rubenfeld, B., Martin, G. A., Ladner, M., Long, C. M., Crosier, W. J., Watt, K., Koths, K., and McCormick, F. (1988) Science 242, 1697-1700
- 32. Williams, T., Admon, A., Luscher, B., and Tjian, R. (1988) Genes Dev. 2, 1557 - 1569
- 33. Iwabuchi, K., Bartel, P. L., Li, B., Marraccino, R., and Fields, S. (1994) Proc.

- Natl. Acad. Sci. U. S. A. 91, 6098-6102
- Helps, N. R., Barker, H. M., Elledge, S. J., and Cohen, P. T. W. (1995) FEBS Lett. 377, 295–300
- Sugita, T., Totsuka, T., Saito, M., Yamasaki, K., Taga, T., and Hirano, T. (1990) J. Exp. Med. 171, 2001–2009
- 36. Lloyd, S. E., Pearce, S. H. S., Fisher, S. E., Steinmeyer, K., Schwappch, B., Scheinman, S. J., Harding, B., Bolini, A., Deveto, M., Goodyer, P., Rigden, S. P. A., Wrong, O., Jentsch, T. J., Craig, I. W., and Thakker, R. V. (1996) Nature 379, 445-449
- Parent, L. J., Bennett, R. P., Craven, R. C., Nelle, T. D., Krishna, N. K., Bowzard, J. B., Wilson, C. B., Puffer, B. A., Montelaro, R. C., and Wills, J. W. (1995) J. Virol. 69, 5455-5460
- 38. Garnier, L., Willis, J. W., Verderame, M. F., and Sudol, M. (1996) Nature 381, 744 - 745
- 39. Woychik, R. P., Stewart, T. A., Davis, L. G., D'Eustachio, P., and Leder, P. (1985) Nature 318, 36-40
- Chan, D. C., Bedford, M. T., and Leder, P. (1996) *EMBO J.* 15, 1045–1054
 Shimkets, R. A., Warnock, D. G., Bositis, C. M., Nelson-Williams, C., Hansson,
- J. H., Schambelan, M., Gill, J. R., Ulick. S., Milora, R. V., Findling, J. W.,
- Canessa, C. M., Rossier, B. C., and Lifton, R. P. (1994) Cell 79, 407-414 42. Schild, L., Canessa, C. M., Schimkets, R. A., Gautschi, I., Lifton, R. P., and
- Hanson, J., Schröder, S. C. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 5699–5703
 Hansson, J. H., Schild, L., Lu, Y., Wilson, T. A., Gautschi, I., Schimkets, R. A.,
- 44. Tamura, H., Schild, L., Enomoto, N., Matsui, N., Marumo, F., and Rossier, B. C. (1996) J. Clin. Invest. 97, 1780-1784
- 45. Staub, O., Dho, S., Henry, P. C., Correa, J., Ishikawa, T., McGlade, J., and Rotin, D. (1996) *EMBO J.* **15**, 2371–2380
- 46. Schild, L., Lu, Y., Gautschi, I., Schneeberger, E., Lifton, R. P., and Rossier, B. C. (1996) *EMBO J.* **15**, 2381–2387
 Snyder, P. M., Price, M. P., McDonald, F. J., Adams, C. M., Volk, K. A., Zeiher,
- B. G., Stokes, J. B., and Welsh, M. J. (1995) Cell 83, 969-978
- 48. Mori, S., Heldin, C. H., and Claesson-Walsh, L. (1992) J. Biol. Chem. 267, 6429-6434
- 49. Hicke, L., and Riezman, H. (1996) Cell 84, 277-288