EXPERIMENTAL and MOLECULAR MEDICINE, Vol. 28, No 2, 65-69, June 1996

The WW domain binds polyprolines and is involved in human diseases

Marius Sudol

Department of Biochemistry, Mount Sinai School of Medicine, One Gustave Levy Place, New York, NY 10029, U.S.A.

Accepted 20 June 1996

Abbreviation: ENaC, epithelial sodium channel

Abstract

Our study of transforming genes led us to the identification of a new protein module, the WW domain, that is found in a wide range of structural, regulatory and signaling proteins. The WW domain is functionally similar to the SH3 module, in that it binds polyproline containing ligands. However, its structure is distinct. Unexpectedly, the WW domain-polyproline ligand connection was recently implicated in several human genetic disorders including the Mendelian form of hypertension known as Liddle's syndrome, and indirectly in muscular dystrophy, Alzheimer's disease and also in limb and kidney formation. Since both the WW domain and the core motif of its ligand are relatively short (40 and 5 amino acid residues), one could suggest that human diseases that involve mutations of these sequences could be treated successfully, not only by gene therapy approaches, but also by low molecular weight compounds. The rigid molecular shapes represented by the polyproline cores of the WW domain ligands could serve as guides for rational drug design.

Identification of the WW domain

The WW module is a small globular domain which is composed of 38-40 semi-conserved amino acids and is involved in mediating protein-protein interaction (Bork and Sudol, 1994; Chen and Sudol, 1995). Functionally, it is similar to the SH3 domain, in that it binds polyproline ligands (Sudol, 1996a,b). However, the structure of the WW domain is distinct from that of the SH3 domain (Macias *et al.*, 1996).

The WW domain was identified based on computer-aided analysis of imperfectly repeated sequences in the mouse isoform of YAP, and also in a yeast factor RSP5 that encodes a protein with ubiquitin-ligase activity (reviewed by Sudol *et al.*, 1995b). The name refers to one of the distinguishing features of the domain: the presence of two highly-conserved tryptophan residues (W) which are spaced 20-22 amino acids apart (Bork and Sudol, 1994). Similar to other intracellular modules, one of the important features of the WW domain is its widespread occurrence (Sudol *et al.*, 1995a). The WW domain is shared by proteins of diverse functions including structural, regulatory and signaling proteins in yeast, nematode and mammals (see discussion in Sudol *et al.*, 1995a,b). By now almost 30 proteins, excluding numerous orthologs, have been identified which contain from 1 to 4 copies of the WW domain. In anticipation of the rapidly growing number of WW domain-containing sequences being uncovered, we have been providing updated information on the WW domain via the World Wide Web (WWW) with the address: http://www.embl-heidelberg.de/~bork/ww1.html/ or http://swan.embl-heidelberg.de:8080/Modules/ww-gif.html since December 1994. Both the alignment and a diagram with the modular structure of the proteins containing the WW domain(s) are available via the WWW network and are updated by us (Peer Bork and Marius Sudol) frequently.

Identification of the ligand to the WW domain

Since we proposed that the WW domain of YAP might mediate protein-protein interactions (Sudol *et al.*, 1995 a), we performed a functional screen of a cDNA expression library to identify its ligand. Two putative ligands that bind to the WW domain were identified. We named them WBP-1 and WBP-2 (Chen and Sudol, 1995). Sequence comparison of the ligands revealed that they share a proline-rich region that binds strongly and specifically to the WW domain of human YAP. This region consists of a five amino acid sequence, PPPPY, which is perfectly conserved between the two ligands. By sequentially replacing each of the these five positions with alanine for *in vitro* binding assays, a preliminary consensus of XPPXY was established for binding (Chen and Sudol, 1995). This consensus is distinct from the SH3-binding motif PXXP. With these data, we have implicated the WW domain in mediating protein-protein interactions, as a variant of the paradigm set by SH3 domains and their proline-rich ligands (Chen and Sudol, 1995; Musacchio *et al.*, 1994c).

NMR structure of the WW domain in complex with its ligand

Although the WW and SH3 domains are functionally similar, their structures are distinct. The NMR structure of the WW domain of human YAP in complex with its cognate peptide was recently solved (Macias *et al.*, 1996). Interestingly, the WW domain seems to represent one of the most compact globular structures known to date. As with the SH3, the WW domain appears to be modular in the sense that the amino- (N) and carboxy-termini (C) are juxtaposed on one side of the domain, opposite from the ligand-binding surface (Macias *et al.*, 1996; Musacchio *et al.*, 1994c). The hallmarks of the WW domain of human YAP are as follows: a three-stranded antiparallel β sheet which is somewhat bent, and a hydrophobic pocket composed of leucine, tyrosine, and the second (W) tryptophan. These amino acids are part of the ligand binding interface, as deduced from structural and mutational studies (Macias *et al.*, 1996; Chen, H.I. and Sudol, M., unpublished). Two central prolines of the ligand XPPX form Van-der-Waals contacts with the second tryptophan, whereas the terminal tyrosine of the ligand fits into a hydrophobic pocket of the WW domain formed by the conserved leucine and histidine.

Biology of the WW domain-polyproline ligand link

Unexpectedly, in the last year the WW domain-ligand link was implicated, directly or indirectly, in several human genetic disorders (Sudol, 1996). The best example comes from the study of Liddle's syndrome - a form of hypertension with Mendelian inheritance, which results from dominant genetic lesions of the renal tubule's amiloride-sensitive epithelial sodium channel (ENaC) (Botero-Velez *et al.*, 1994; Lifton, 1995; Rossier, 1996).

Liddle's syndrome

In 1963, Liddle and colleagues described a rare syndrome of a salt-sensitive form of hypertension with clinical symptoms corresponding to a primary hyperaldosteronism (Liddle *et al.*, 1963). The affected patients suffered from severe hypertension and metabolic alkalosis, both of which are normally associated with hyperproduction of aldosterone by the adrenal glands. However, Liddle was able to demonstrate that the hypertension was not due to the production of mineralocorticoids but rather was connected with the function of a sodium channel. Using an amiloride analog that selectively blocks the sodium channel, Liddle was able to significantly decrease blood pressure in patients ingesting a low-salt diet. At that point he proposed the term "pseudoaldosteronism" - today this name is synonymous with Liddle's syndrome (Lifton, 1995).

The mutations that give rise to Liddle's syndrome are usually deletions or stop codon-created truncations and less frequently point mutations, that affect the carboxy terminal region of either the β or γ subunit of the ENaC and lead to constitutively increased channel activity (Shimkets *et al.*, 1994; Schild *et al.*, 1995; Hansson *et al.*, 1995). Work from three independent laboratories headed by Rossier, Lifton and Welsh pointed clearly to the highly conserved proline-rich regions in all three subunits of the channel (including the α subunit) as potential target sites for mutations (Canessa *et al.*, 1994; Snyder *et al.*, 1995; Schild *et al.*, 1996). Experimentally introduced missense mutations that alter the conserved PPPXY consensus sequence within the α, β or γ subunits, when assayed in the *Xenopus* oocyte system, reproduced the increase in channel activity that had been originally found in mutants in which the entire cytoplasmic carboxytermini were deleted (Snyder *et al.*, 1995; Schild *et al.*, 1996). Moreover, two Liddle's syndrome patients were recently characterized with single amino acid substitutions in either the third proline (P616L; giving rise to PPLNY) or tyrosine (Y618H; giving rise to PPPNH) in the PPPNY motif of the beta subunit of ENaC (Hansson *et al.*, 1995; Tamura *et al.*, 1996). The "alanine scan" of the proline-rich region of the ENaC followed by a functional assay in frog oocytes definitively confirmed that the PPPNY behaves in the same way as our XPPXY consensus, which was established for the WBP-1 binder to the WW domain of YAP (Schild *et al.*, 1996). When the proline-rich region of the β subunit of ENaC was used as bait in the yeast two hybrid screen, a known protein, NEDD4, composed of three WW domains and a ubiquitin ligase domain was isolated as an interactive partner (Kumar *et al.*, 1992; Staub *et al.*, 1996). The very modular structure and the overall topology of the NEDD4 protein is suggestive of its possible molecular function as a regulator and suppressor of the ENaC (Staub *et al.*, 1996; Sudol *et al.*, 1995a) (Figure 1A). In the most simple scenario, the three WW domains bind three ENaC subunits and this binding is required for the ubiquitin ligase to modify the channel and to target it for regulated degradation. Lack of binding of any of the three subunits to the corresponding WW domain would affect the entire complex and prevent proper ubiquitination of the channel.

Muscular dystrophy

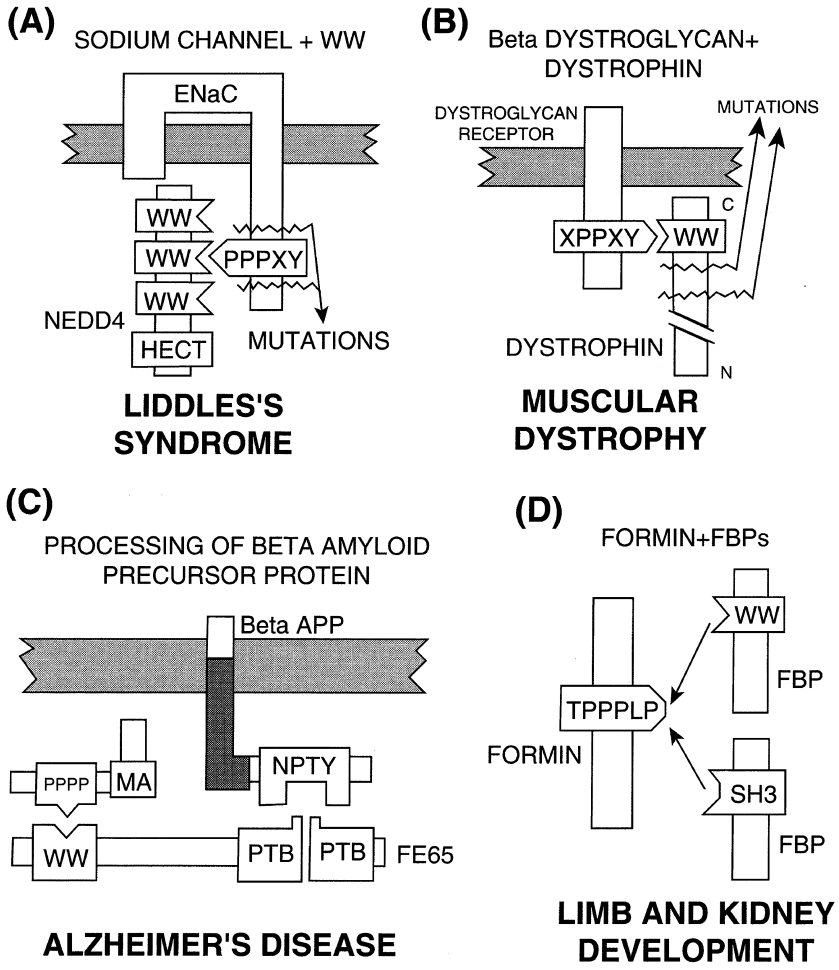

Figure 1. Schematic diagram depicting examples of biological functions for proteins containing WW domains. As with other known signaling domains, the WW domain represents a molecular adhesive that is specific for ligands containing polyproline sequences, frequently having the XPPXY consenus. (**A**) Each of the three WW domains of NEDD4 interacts with the XPPXY motifs of at least one of the epithelial sodium channel subunits. Deletion of the XPPXY motif or point mutation within the motif results in Liddle's syndrome. In all models, the cell membrane is represented by a cross-hatched end-broken rectangle. (**B**) One of the proline-rich motifs of the β dystroglycan receptor interacts with the WW domain of dystrophin, localizing dystrophin to the membrane of a normal cell. Mutations resulting in deletion of the C-terminal region of dystrophin prevent this interaction, causing Duchenne's muscular dystrophy. (**C**) The FE65 factor contains one WW domain and two PTB domains. The PTB domains bind to the β-amyloid precursor protein. A ligand to the WW domain of FE65 may contain modifying (processing) activity (MA) toward the beta amyloid precursor. (**D**) A formin interacts with two formin binding-proteins (FBPs). The WW domain of one FBP and the SH3 domain of the other compete to bind to the proline rich sequence TPPPLP (which contains versions of the consensus cores for WW and SH3 domains: XPPX and PXXP; "binary switch") resulting in the regulation of limb and kidney development.

The muscular dystrophies are a group of diseases that primarily affect skeletal muscle and are characterized by progressive muscle wasting (Campbell, 1995). Duchenne and Becker muscular dystrophies (DMD and BMD) are X-linked recessive diseases that are caused by mutations in the dystrophin gene (Ahn and Kunkel, 1993).

We have identified a WW domain in the carboxy terminal region of dystrophin just between the fourth proline-rich "hinge" region and the beginning of the cysteine-rich stretch (Bork and Sudol, 1994). Knowing that the carboxy terminal region of dystrophin is of particular importance because it interacts with a large, oligomeric, membrane-spanning protein complex (Tinsley *et al.*, 1994; Matsumura and Campbell, 1994), we focused our attention on two components of this complex: β-dystroglycan (43-kDa protein) and syntropin (59-kDa protein). Both β-dystroglycan and syntropin were shown to interact directly with the carboxy terminal region of dystrophin, but the precise sites of their interaction have not been delineated to date (Tinsley *et al.*, 1994). Using various approaches, including computer searches for conserved proline-rich sequences (potential binding sites for WW domains), *in vitro* precipitations of cell lysates with either the bacterially expressed and purified WW domain or the WW domain plus the cysteine-rich polypeptide, and phage display libraries to predict ligand preference of the WW domain of dystrophin, we have shown that one of the three proline rich motifs in the β-dystroglycan molecule forms a stable complex with the WW domain of dystrophin (Bougeret, C. and Sudol, M., unpublished; Einbond and Sudol, 1996) (Figure 1B). At present, we are investigating the biological ramifications of this link. Following the example of molecular lesions causing the Liddle's syndrome, it is tempting to speculate that mutations in the proline-rich motif of the β-dystroglycan receptor might result in dystrophic phenotypes (Sudol, 1996).

Alzheimer's disease

Alzheimer's disease (AD) is a neurodegenerative disorder characterized by the progressive deterioration of memory (Schellenberg, 1995). The onset of Alzheimer's disease correlates well with the formation of amyloid plaques and vascular deposits consisting of the 4-kDa amyloid β peptide (Aβ). Aβ peptide is proteolytically derived from a large membrane protein, the β amyloid precursor protein (βAPP).

Our interest in the molecular aspect of the βAPP processing came from the identification of the WW domain in the FE65 adapter protein (Bork and Sudol, 1994). The FE65 factor was cloned and characterized by the Russo laboratory at the University of Naples as a brain-specific gene (Duilio *et al.*, 1991; Faraonio *et al.*, 1994). The modular structure of the FE65 protein is reminiscent of known signaling adapters. At its amino terminal portion it contains the WW domain while at its carboxy-terminal end it contains two PTB domains (Bork and Margolis, 1995; Kavanaugh *et al.*, 1995; Van Der Geer and Pawson, 1995). Recently, the PTB domains of FE65 were shown to bind specifically to the βAPP molecule (Fiore *et al.*, 1995). Prompted by this result, we have cloned a cDNA for a potential ligand to the WW domain of FE65, which may participate in the processing of βAPP (Ermekova, K., Russo, T. and Sudol, M., unpublished). The scenario of a docking protein that forms a tri-partite complex, including the βAPP is attractive and amenable for direct experimental attack (Figure 1C).

In addition to providing insights into the mechanism of the βAPP processing in normal and Alzheimer's disease-derived tissues, the identification of specific domains in FE65 and in its WW ligand suggests that mutations in the relevant genomic, coding sequences could contribute to the phenotype of Alzheimer's disease (Sudol, 1996).

Limb and kidney development

A recent report from the Leder laboratory proposed a functional interplay between the WW and SH3 domains (Chan *et al.*, 1996). This hypothesis emerged from the study of formins, a set of protein isoforms encoded by the mouse limb deformity locus (ld) (Woychik *et al.*, 1985). The ld locus was identified as result of transgene insertion that disrupted embryonic pattern formation, resulting in a reduction and fusion of the distal bones and digits of fore- and hindlimbs as well as in variable incidence of renal aplasia. An exhaustive screen of expression libraries from the mouse limb bud with the polyproline region of formin was performed (Chan *et al.*, 1996). This resulted in the identification of two sets of formin-binding proteins (FBPs); one set, as expected, had SH3 domains and another had WW modules. In a series of well controlled experiments, the WW domains of FBPs were shown to compete with the SH3 domain of ABL in binding the polyproline peptide present in formin. The authors suggested that in general, WW domains might regulate the function of SH3 domains by modulating their interaction with ligands through direct competition for the same proline-rich sequences on target proteins (Chan *et al.*, 1996) (Figure 1D). This is an attractive scenario for a molecular mechanism regulating developmental processes. If indeed these interactions are confirmed *in vivo*, they will provide one of the first examples of degeneracy in the "protein-protein interaction code" (Sudol, 1996).

Acknowledgements

Supported by grants from National Institutes of Health, Human Frontier Science Program Organization, Council for Tobacco Research, USA, Inc., and The Muscular Dystrophy Association. Special thanks are to Professor Jeong-Sun Seo for stimulating discussions and friendship.

References

Ahn, A. H. and Kunkel, L. M. (1993) The structural and functional diversity of dystrophin. *Nature Genet.* 3: 283-291

Bork, P. and Margolis, B. (1995) A phosphotyrosine interaction domain. *Cell* 80: 693-694

Bork, P. and Sudol, M. (1994) The WW domain: a signaling site in dystrophin? *TIBS* 19: 531-533

Botero-Velez, M., Curtis, J. J., and Warnock, D. G. (1994) Liddle's syndrome revisited: a disorder of sodium reabsorpotion in the distal tubule. *N. Engl. J. Med.* 330: 178-181

Campbell, K. P. (1995) Three muscular dystrophies: loss of cytoskeleton-extracellular matrix linkage. *Cell* 80: 675-679

Canessa, C. M., Schild, L. Buell, G., Thorens, B., Gautschi, I., Horisberger, J.-D. and Rossier, B. C. (1994) Amiloride-sensitive epithelial Na^+ channel is made of three homologous subunits. *Nature* 367: 463-467.

Chan, D. C., Bedford, M. T. and Leder, P. (1996) Formin binding proteins bear WWP/WW domains that bind proline-rich peptides and functionally resemble SH3 domains. *EMBO J.* 15: 1045-1054

Chen, H. I. and Sudol, M. (1995) The WW domain of Yes-associated protein binds a proline-rich ligand that differs from the consensus established for Src homology 3-binding modules. *Proc. Natl. Acad. Sci. U.S.A.* 92: 7819-7823

Duilio, A., Zambrano, N., Mogavero, A. R., Ammendola, R., Cimino, F. and Russo, T. (1991) A rat brain mRNA encoding a transcriptional activator homologous to the DNA binding domain of retroviral integrases. *Nucl. Acid Res.* 19: 5269-5274

Faraonio, R., Minopoli, G., Porcellini, A., Costanzo, F., Cimino, F. and

Russo, T. (1994) The DNA sequence encompassing the transcription start site of a TATA-less promoter contains enough information to drive neuron-specific transcription. *Nucl. Acid Res.* 22: 4876-4883

Fiore, F., Zambrano, N., Minopoli, G., Donini, V., Duilio, A. and Russo, T. (1995) The regions of the Fe65 protein homologous to the phosphotyrosine interaction/phosphotyrosine binding domain of Shc bind the intracellular domain of the Alzheimer's amyloid precursor protein. *J. Biol. Chem.* 270: 30853-30856

Hansson, J. H., Schild, L., Lu, Y., Wilson, T. A., Gautschi, I., Shimkets, R., Nelson-Williams, C., Rossier, B. C. and Lifton, R. P. (1995) A *de novo* missense mutation of the β subunit of the epithelial sodium channel causes hypertension and Liddle syndrome, identifying a proline-rich segment critical for regulation of channel activity. *Proc. Natl. Acad. Sci. U.S.A .* 92: 11495-11499

Kavanaugh, W. M., Turck, C. W. and Williams, L. T. (1995) PTB domain binding to signaling proteins through a sequence motif containing phosphotyrosine. *Science* 268: 1177-1179

Kumar, S., Tomooka, Y. and Noda, M. (1992) Identification of a set of genes with developmentally down-regulated expression in the mouse brain. *Biochem. Biophys. Res. Comm.* 185: 1155-1161

Liddle, G. W., Bledsoe, T. and Coppage, W. S. J. (1963) A familial renal disorder simulating primary aldosteronism but with negligible aldosterone secretion. *Trans. Assoc. Am. Physicians* 76: 199-213

Lifton, R. P. (1995) Genetic determinants of human hypertension. *Proc. Natl. Acad. Sci. U.S.A.* 92: 8545-8551

Macias, M. J., Hyvonen, M., Schultz, J., Chen, H. I., Saraste, M., Sudol, M. and Oschkinat, H. (1996) The structure of the WW domain in complex with a proline-rich peptide. *Nature*, submitted.

Matsumura, K. and Campbell, K. P. (1994) Dystrophin-glycoprotein complex: Its role in the molecular pathogenesis of muscular dystrophies. *Muscle Nerve* 17: 2-15

Musacchio, A., Wilmanns, M. and Saraste, M. (1994) Structure and function of the SH3 domain. *Prog. Biophys. Mol. Biol.* 61: 283-297

Rossier, B. C. (1996) The renal epithelial sodium channel: new insights in understanding hypertension. *Adv. Nephrol.* 25: 275-286

Schild, L., Canessa, C. M., Shimkets, R. A., Gautschi, I., Lifton, R. P. and Rossier, B. C. (1995) A mutation in the epithelial sodium channel causing Liddle disease increases channel activity in the *Xenopus laevis* oocyte expression system. *Proc. Natl. Acad. Sci. U.S.A .* 92: 5699-5703

Schild, L., Lu, Y., Gautschi, I., Schneeberger, E., Lifton. R. P. and Rossier, B. C. (1996) Identification of a PY motif in the epithelial Na^+ channel subunits as a target sequence for mutations causing channel activation found in Liddle's Syndrome. *EMBO J.*, in press.

Shimkets, R. A., Warnock, D. G., Bostis, C. M., Nelson-Williams, C., Hansson, J. H., Schambelan, M., Gill, JR., J. R., Ulick, S., Milora, R. V., Finding, J. W., Canessa ,C. M., Rossier, B. C. and LiftonI, R. P. (1994) Liddle's syndrome: heritable human hypertension caused by mutations in the β subunit of the epithelial sodium channel. *Cell* 79: 407-414

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) Mechanism by which Liddle's syndrome mutations increase activity of a human epithelial Na^+ channel. *Cell* 83: 969-978

Staub, O., Dho, S., Henry, P. C., Correa, J., Ishikawa, T., McGlade, J. and Rotin, D. (1996) WW domains of Nedd4 bind to the proline-rich PY motifs in the epithelial Na^+ channel deleted in Liddle's syndrome. *EMBO J.*, in press.

Sudol, M. (1994) Yes-associated protein (YAP65) is a proline-rich phosphoprotein that binds to the SH3 domain of the Yes proto-oncogene product. *Oncogene* 9: 2145-2152

Sudol, M. (1996a) The WW module competes with the SH3 domain? *TIBS* 21: 161-163

Sudol, M. (1996b) Structure and function of the WW domain. *Prog. Biophys. Mol. Biol.*, in press.

Sudol, M., Alvarez-Buylla, A. and Hanafusa, H. (1988) Differential developmental expression of c-*Yes* and c-*Src* protein in cerebellum. *Oncogene Res.* 2: 345-355

Sudol, M., Bork, P., Einbond, A., Kastury, K., Druck, T., Negrini, M., Huebner, K. and Lehman, D. (1995a) Characterization of the mammalian YAP (Yes-associated protein) gene and its role in defining a novel protein module, the WW domain. *J. Biol. Chem.* 270: 14733-14741

Sudol, M., Chen, H. I., Bougeret, C., Einbond, A. and Bork, P. (1995b) Characterization of a novel protein-binding module - the WW domain. *FEBS Lett.* 369: 67-71

Tamura, H., Shild, L.., Enomoto, N., Matsui, N., Marumo, F., Rossier, B. C. and Sasaki, S. (1996) Liddle disease caused by a missense mutation of the β subunit of the epithelial sodium channel gene. *J. Clin. Invest.* 97: 1780-1784

Tinsley, J. M., Blake, D. J., Zuellig, R. A. and Davies, K. E. (1994) Increasing complexity of the dystrophin-associated protein complex. *Proc. Natl. Acad. Sci. U.S.A.* 91: 8307-8313

Van Der Geer, P. and Pawson, T. (1995) The PTB domain: a new protein module implicated in signal transduction. *TIBS*, 20: 277-280

Woychik, R. P., Stewart, T. A., Davis, L. G., D'eustachio, P. and Leder, P. (1985) An inherited limb deformity created by insertional mutagenesis in a transgenic mouse. *Nature* 318: 36-40