Minireview

PDZ Domains: Structural Modules for Protein Complex Assembly*

Published, JBC Papers in Press, December 10, 2001, DOI 10.1074/jbc.R100065200

Albert Y. Hung§ and Morgan Sheng†

From the §Center for Learning and Memory, RIKEN-MIT Neuroscience Research Center, Departments of Brain and Cognitive Sciences and Biology, Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 and the †Department of Neurology, Massachusetts General Hospital, Boston, Massachusetts 02114

PDZ domains are modular protein interaction domains that play a role in protein targeting and protein complex assembly. Once termed Discs-large homology regions (DHRs) or GLGF repeats (after a conserved Gly-Leu-Gly-Phe sequence found within the domain), these domains of ~90 amino acids are now primarily known by an acronym of the first three PDZ-containing proteins identified: the postsynaptic protein PSD-95/SAP90, the Drosophila septate junction protein Discs-large, and the tight junction protein ZO-1. Since their initial identification, PDZ and PDZ-like domains have been recognized in numerous proteins from organisms as diverse as bacteria, plants, yeast, metazoans, and Drosophila. In fact, they are among the commonest protein domains represented in sequenced genomes. Analysis of the human, Drosophila, and Caenorhabditis elegans genomes estimates the presence of 440 PDZ domains in 259 different proteins, 133 PDZ domains in 86 proteins, and 138 PDZ domains in 96 proteins, respectively (2). The structural features of PDZ domains allow them to mediate specific protein-protein interactions that underlie the assembly of large protein complexes involved in signaling or subcellular transport. Not surprisingly, disrupting these interactions can play a role in human diseases. Mutations in a gene encoding harmonin, a PDZ-containing protein, cause Usher syndrome type 1C, an autosomal recessive disorder characterized by congenital sensorineural deafness, vestibular dysfunction, and blindness (3–5). This was the first mutation in a PDZ-encoding gene linked to a human disease. Subsequently, mutations in the periaxin gene, which also encodes a PDZ-containing protein, have been identified as a cause of the rare autoimmune peripheral neuropathy, a severe demyelinating form of peripheral neuropathy (6, 7).

The Structural Basis of PDZ Binding and Specificity

The notion that PDZ domains serve as protein interaction modules emerged from the finding that the first and second PDZ (PDZ1 and -2) domains of PSD-95 can bind the extreme C-terminal peptide sequence of Shaker-type K+ channels (8) and NMDA receptor NR2 subunits (9, 10). PDZ3 of the protein tyrosine phosphatase FAP-1/PTP1E similarly was identified as a binding site for the C-terminal peptide of the cell surface receptor Fas (11). These studies further demonstrated that PDZ domains maintained their activity and selectivity when expressed in heterologous proteins, establishing these motifs as modular domains that bind the C termini of target proteins in a sequence-specific manner. The structural basis for PDZ specificity became apparent with the solution of the x-ray crystallographic structure of PDZ domains complexed with their cognate peptide ligands. First solved for PDZ3 of PSD-95 (12, 13), numerous additional PDZ crystal structures have been determined in recent years, including PDZ2 of PSD-95 (14), the single PDZ domain of CASK (15), syntrophin (16) and neuronal nitric-oxide synthase (nNOS) (17), PDZ2 from human phosphatase hPTP1E (18), PDZ1 of the Na+/H + exchanger regulatory factor (NHERF) (19), and PDZ1 of InaD (20). The common structure of PDZ domains comprises six ß strands (ßA–ßF) and two ÿ helices (ßA and ÿB), which fold in an overall six-stranded ß sandwich (Fig. 1A). C-terminal peptides bind as an antiparallel ß-strand in a groove between the ßB strand and the ÿB helix, in essence extending one of the ß-sheets. The conserved Gly-Leu-Gly-Phe (GLGF) sequence of the PDZ domain is found within the ßA–ßB connecting loop and is important for hydrogen bond coordination of the C-terminal carboxylate (COO–) group. The N and C termini of the PDZ domain are located near each other on the opposite side of the PDZ domain from the peptide-binding site, a feature shared with other protein interaction modules such as SH2 domains.

PDZ-like domains identified in plants and bacteria have a similar overall secondary and tertiary structure but show a different topology (Fig. 1C). Specifically, the ßA strand, derived from the N-terminal sequences in conventional PDZ domains, is formed by the C terminus in the PDZ-like domain of the photosystem II D1 C-terminal protease (21). This circularly permuted fold, also found in the Tsp protease from Escherichia coli, retains the ability to bind C-terminal sequences (22).

The crystallographic data indicate that the C-terminal four residues of PDZ ligands interact directly with the peptide-binding groove. The main chain atoms of the ßA strand form hydrogen bonds with the extended peptide ligand and stabilize the interaction although they do not account for sequence specificity. The specific interaction of C-terminal sequences suggests that recognition of the carboxyterminal group is critical for PDZ binding. Indeed, a highly conserved positively charged residue (e.g. arginine 318 of PDZ3 of PSD-95) and the main chain amides of the Gly-Leu-Gly-Phe motif form hydrogen bonds with the terminal carboxylate group (12). The side chain of the C-terminal residue (position 0) projects into a hydrophobic pocket accounting for PDZ domains binding preferentially to sequences ending with a hydrophobic residue (such as valine, isoleucine, or leucine (23)).

In the crystal structure of PSD-95 PDZ3 with its peptide ligand, the side chain of the residue at position −1 points away from the interaction surface (12). This correlates with the relative lack of PDZ specificity for recognition at the −1 position. Nevertheless, substitutions at this site can affect binding preference for individual PDZ domains, albeit to a lesser degree than the 0 and −2 positions (23–25). In contrast, the guanido group of arginine at position −1 of the cystic fibrosis transmembrane regulator (CFTR) C terminus (C-terminal sequence -DTRL) forms two salt bridges and two hydrogen bonds with residues in PDZ1 of NHERF, indicating that in some cases the −1 residue contributes directly to the specificity and affinity of the interaction (19). The crystal structure of the InaD PDZ1 with a C-terminal peptide from NorpA (C-terminal sequence -EPCAl further demonstrates a critical role for a cysteine residue at position −1, as intermolecular disulfide bond formation of this residue with a cysteine residue in the PDZ domain is required for high affinity interaction (20).

The binding specificity of PDZ domains is critically determined by the interaction of the first residue of helix ÿB (position ÿB1) and the side chain of the −2 residue of the C-terminal ligand; this forms the basis for PDZ classification (Table I) (23). In Class I PDZ
FIG. 1. Structure of PDZ and PDZ-like domains. A, ribbon diagram of PDZ3 of PSD-95 complexed with a C-terminal peptide from CRIP1 (12). The structure demonstrates the six β-strands (turquoise) and two α-helices (red) with the peptide (yellow) binding as a β-strand between the αβ helix and βB strand. The N and C termini are labeled. B, structure of α1-syntrophin PDZ domain complexed to nNOS (green) (17). Note the overall similarity in PDZ structure with replacement of the C-terminal peptide ligand by a β-finger. C, ribbon diagram of the photosystem II D1 protease PDZ-like domain from Scenedesmus obliquus (21). Although the overall topology is similar to conventional PDZ domains, the βA strand is derived from the C terminus of the domain.

Table I
Examples of PDZ domains classified by specificity for C-terminal peptide ligands

| Class   | C-terminal sequencea | Ligand protein       | PDZ protein | Ref. |
|---------|----------------------|----------------------|-------------|------|
| Class I | X-S/T-X-Φ             | ETIDV Shaker K⁺ channel | PSD-95 (PDZ2) | 8    |
|         |                      | ESDV NMDA receptor subunits | PSD-95 (PDZ2) | 9    |
|         |                      | TTIV Neurologin       | PSD-95 (PDZ3) | 68   |
|         |                      | ESLV Voltage-gated Na⁺ channel | Synthrophin | 69   |
|         |                      | QSAV Protein kinase C-α | PICK1 | 27   |
|         |                      | SAVH Syntrophin β-adrenergic receptor | NHERF (PDZ1) | 13   |
|         |                      | DTRL CFTR             | NHERF (PDZ1) | 70   |
|         |                      | QTRL GRAP             | Shank/ProSAP | 71   |
|         |                      | SSSL Metabotropic glutamate receptor subunit | Shank/ProSAP | 72   |

| Class II | X-Φ-X-Φ               | SVKI AMPA receptor subunit GluR2 | PICK1 | 28   |
|          |                      | EFPI Glycinorin C        | Erythrocyte p55 | 73   |
|          |                      | EYYV Neurexin           | CASK | 74   |
|          |                      | EYFA Syndecan-2         | CASK | 75   |
|          |                      | YYKV Ephrin B1          | PICK1 | 76   |
|          |                      | GRIP (PDZ6)            | Syntentin (PDZ2) |     |

| Class III | X-D/E-X-Φ               | VDSV Melatonin receptor | nNOS | 24   |
|           |                      | GEFL KIF17             | mLin10/Mint1/1 | 64   |

a Last four amino acids of the ligand proteins are shown although specificity can involve more proximal residues as well (see text). Φ, hydrophobic amino acid; X, unspecified amino acid.

interactions, such as those of PSD-95, a serine or threonine residue occupies the −2 position. The side chain hydroxyl group forms a hydrogen bond with the N-3 nitrogen of a histidine residue at position αB1 (12), which is highly conserved among Class I PDZ domains. In contrast, class II PDZ interactions are characterized by hydrophobic residues at both the −2 position of the peptide ligand and the αB1 position of the PDZ domain (23). A third class of PDZ domains, such as nNOS, prefers negatively charged amino acids at the −2 position (24). This specificity is determined by the coordination of the hydroxyl group of a tyrosine residue at position αB1 with the side chain carboxylate of the −2 residue (24, 26). Other classes of PDZ domain specificities are likely to be distinguished with further research.

PDZ domains vary in their range and stringency of specificity. For example, the PDZ domain of PICK1, which has a lysine residue at the αB1 position, can bind the C termini of both protein kinase C (ending in -QSAV (27)) and the AMPA receptor subunit GluR2 (ending in -SVKI (28)). Thus, a single PDZ domain can show both Class I and Class II specificity; the three-dimensional structural basis of this promiscuity remains to be determined.

Whereas the residue at the −2 position is a key determinant of PDZ-ligand interactions, more N-terminal residues in PDZ-binding peptides also contribute to specificity. Crystallographic data indicate that the −3 side chain also directly contacts the peptide-binding groove (12, 19, 26), and the residue at position −3 is important in determining the binding of ligands selected from a peptide library (23). Moreover, several studies have demonstrated that residues beyond the last four amino acids are also important, up to position −8 (18, 23, 25).

Although many examples of PDZ-peptide ligand interactions have now been identified, uncertainty remains regarding the binding affinity of the PDZ domain for its ligand. Using solid phase methods such as surface plasmon resonance (Biacore) or modified enzyme-linked immunosorbent assays, affinities of PDZ binding to their cognate peptides have been measured in the 10−100 nM range. On the other hand, solution methods such as fluorescence polarization suggest that the binding affinity may be weaker in the low micromolar range (25, 29). It is likely that a wide range of affinities is applied to the diverse peptides in vivo. Moreover, especially because many PDZ ligands are membrane-associated and clustered, it is difficult to know whether solid phase or solution methods better approach the in vivo situation.

**Regulation of PDZ Binding to C-terminal Sequences**

The regulation of protein-protein interactions between modular elements such as SH3 domains and their binding partners is often critical for cell signaling (30). What regulates the binding of PDZ domains to their respective C-terminal ligands? One mode of control is the phosphorylation of residues within the C-terminal sequences that bind the PDZ domain. For example, serine phosphorylation at position −2 in the inward rectifier K⁺ channel Kir2.3 by protein kinase A disrupts binding to the PDZ domains of PSD-95 (31). The association of β-adrenergic receptor with NHERF is similarly abolished by phosphorylation at position −2 by G-protein-coupled receptor kinase GRK5 (32). Intriguingly, binding of a single C-terminal peptide ligand to multiple PDZ domains can be differentially regulated. For instance, phosphorylation at serine 880 of the AMPA receptor subunit GluR2 (C-terminal sequence -SVKI) by protein kinase C inhibits binding to the PDZ domain of GRIP1 but not to PICK1 (33, 34). PDZ interactions may also be regulated by extracellular signals, as the binding of the β-adrenergic receptor C terminus to NHERF is stimulated by β-adrenergic agonists (35). Thus, it is likely that both intracellular and extracellular signals regulate the temporal and spatial organization of PDZ-based interactions.

**Binding of PDZ Domains to Internal Sequences**

Although binding to C-terminal peptides appears to be the typical mode of interaction, PDZ domains can also interact with internal peptide sequences. The best example of this is the interaction of nNOS with the PDZ domain of PSD-95 or syntrophin (36, 37). In the crystal structure of the nNOS-syntrophin PDZ complex, amino acid residues adjacent to the canonical PDZ domain of nNOS form a two-stranded β-hairpin “finger,” which docks in the peptide-binding groove of the syntrophin PDZ domain (Fig. 1B) (14, 26). The sharp β turn of the β-finger binds to the same site as the terminal carboxylate group of peptide ligands (36). Introduction of point mutations that destabilize the nNOS β-finger conformation results in decreased binding to the syntrophin PDZ domain, supporting the model that the β-finger is required for proper recognition (29). Screening of combinatorial phage libraries has also identified cyclic peptides as potential ligands for PDZ domains; PDZ binding of these peptides depends on intramolecular disulfide bond formation (38). These examples suggest that PDZ domains may
PDZ Domains as Organizers of Protein Complexes

The multidomain structure of PDZ-containing proteins enables them to interact with multiple binding partners simultaneously, thereby assembling larger protein complexes (recently reviewed in Refs. 52 and 53). PDZ-based complexes are often localized to specific subcellular compartments. PDZ-based scaffolds have been shown to organize signal transduction pathways such as phototransduction in Drosophila, where ion channels and signaling molecules are co-assembled by the multi-PDZ protein InaD (54–57). MAGUKs appear to play a similar role in the postsynaptic density, a specialized structure at excitary synapses enriched in glutamate receptors and associated signaling proteins (58–60). In both the examples of InaD and MAGUKs, the ability of PDZ-containing proteins to interact with multiple binding partners creates a protein complex specialized for local signaling functions.

PDZ proteins have also been implicated in the establishment of cell polarity. In Drosophila, two PDZ-containing proteins, Bazooka and PAR-6, form a ternary complex with atypical protein kinase C that is required for proper establishment and maintenance of apical-basal polarity in epithelial tissues (reviewed in Ref. 61). The orthologues of these binding partners in C. elegans are likewise necessary for polarization of the one-cell embryo. Importantly, depletion of any one of the proteins results in mislocalization of the other two, underscoring the functional importance of the complex. The LIN-2-LIN-7-LIN-10 tripartite complex also demonstrates a role for PDZ proteins in protein sorting (62, 63). This complex specifies the basolateral targeting of the epidermal growth factor receptor homolog LET-23 in C. elegans, and the mammalian homologs have been proposed to play a role in NMDA receptor trafficking through an interaction with the kinesin superfamily motor protein KIF17 (64).

These examples illustrate the role of PDZ-containing proteins in determining the subcellular location of their binding partners. Two recent studies also indicate that PDZ domains can modulate the function of their associated proteins as well as the localization, as interaction of the CFTR with either the multi-PDZ adaptor protein CAP70 or NHERF increases chloride channel activity of the CFTR protein, perhaps by promoting multimerization of the channel protein (65, 66). Thus PDZ domains may have a functional role in modulating the activity of their target ion channels and membrane receptors (67).

Conclusions

It is now clear that PDZ domain proteins play an important role in the targeting of proteins to specific membrane compartments and their assembly into supramolecular complexes. There is also evidence that they can regulate the function of their ligands in addition to serving as scaffolds. Their ability to bind short extreme C-terminal sequences offers a facile way for PDZ proteins to interact with target proteins without disrupting the overall structure and function of their ligands. Key questions remain in understanding how PDZ domains serve their function. For example, most PDZ domains seem to bind multiple ligands. What determines the interaction with specific binding partners and how is this regulated spatially and temporally in the cell? How is ligand binding and protein stoichiometry affected by PDZ-based homo- and heteromultimerization, and what is the structural basis for multimer formation? What role does the tandem organization of PDZ domains play in determining scaffolding function? The answers to these questions and others await investigation in this exciting and evolving field.

REFERENCES

1. Ponting, C. P. (1997) Protein Sci. 6, 464–468
2. Schultz J., Copley, R. R., Doerks, T., Ponting, C. P., and Bork, P. (2000) Nucleic Acids Res. 28, 231–234
3. Verpy, E., Leibovici, M., Zwaenepoel, I., Liu, X. Z., Gal, A., Salem, N., Mansour, A., Blanchard, S., Kobayashi, I., Krets, B. J., Slim, R., and Petit, C. (2000) Nat. Genet. 26, 51–55
4. Biener-Glinszicz, M., Lindner, K. J., Rutland, P., Blaydon, D., Smith, V. V., Milla, J. P., Hussain, K., Forth-Lavi, J., Coggrove, K. E., Shepherd, R. M., Barnes, P. D., O’Brien, R. E., Farndon, P. A., Sowden, J., Liu, X. Z., Scanlan, M. J., Malcolm, S., Dunne, M. J., Aynsley-Green, A., and Glaser, B. (2000) Nat. Genet. 26, 56–60
5. Montell, C. J. (2000) Nat. Genet. 26, 6–7
6. Boerkoel, C. F., Takashima, H., Stankiewicz, P., Garcia, C. A., Leber, S. M., Rhee-Morris, L., and Lupski, J. R. (2001) Am. J. Hum. Genet. 68, 325–333
PDZ Domains: Structural Modules for Protein Complex Assembly
Albert Y. Hung and Morgan Sheng

J. Biol. Chem. 2002, 277:5699-5702.  
doi: 10.1074/jbc.R100065200 originally published online December 10, 2001

Access the most updated version of this article at doi: 10.1074/jbc.R100065200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 76 references, 31 of which can be accessed free at http://www.jbc.org/content/277/8/5699.full.html#ref-list-1