Correspondence

PDZ domains: troubles in classification

Paola Vaccaro[a], Luciana Dente[b,*]

First published online 11 January 2002

Classification of protein interaction domains on the basis of the chemical characteristics of binding pocket residues is a difficult task, because multiple contact positions are usually involved in the recognition specificity mechanism. On the other hand, target peptides may be classified according to the (few) specific residues that constitute the binding motif, through analysis of molecular repertoires (libraries of synthetic peptides; phage display; two hybrid, etc.) that allow identifying collections of different ligands.

We have recently pointed out that, in order to characterize PDZ domains and to infer their binding specificity, it is necessary to exploit computational procedures, which simultaneously take into account all the contact positions of the domain binding pocket and the corresponding residues of the ligand [1]. PDZ domains are protein interaction modules that recognize and bind the C-terminal four residues of their target. The solution of X-ray crystallographic structure of PDZ domains complexed with their peptide ligands reveals at least 23 contact positions, whose interacting atoms are at a distance shorter than the sum of the van der Waals radii (r) +3 Å.

Some of these positions contain residues that are highly conserved in the PDZ domain family: for example the 'GLGF' loop and a positively charged residue that accommodate the terminal carboxylate group. The majority of the PDZ domains (identified in the proteome up to now) recognize ligands of class I (Table 1). These PDZ domains are provided with a hydrophilic pocket, where residues of the βB strand and of the αB helix (in particular a histidine, highly conserved at position αB1) are involved in contacting the peptide ligand.

Most of the remaining PDZ domains recognize a varied class of ligand peptides, characterized by aromatic or hydrophobic residues at position P^{-2}. Even residues mimicking part of a hydrophobic moiety at position P^{-2} (such as the arginine at P^{-2} of the peptide ligand in the crystallized structure of hCASK) can be accommodated in the large hydrophobic pocket that characterizes PDZ domains binding to class II peptides [2]. Main determinants of the binding, as derived from the contacts in the crystal structure, are residues at βB5, αB1 and αB3 positions.

Bezprozvanny and Maximov have recently proposed a classification of the PDZ domains listed in the SMART Website, based upon the type of residues present in only two contact positions – βB5 and αB1 – of the binding pocket [3]. By grouping the couples of residues on the basis of their polarity and/or bulkiness, they defined 25 groups and correlated them to experimentally determined ligands. Unfortunately, ligand sequences are available only for nine out of 25 groups and, while the first group (G,H) is enforced by the presence of 68 PDZ domains (those binding to class I motifs), the others are less clearly determined. Two of them (G,n) and (a,p) do not correspond to known PDZ domains; other 14 (G,h), (G,a), (n,H), (n,n), (n,p), (Sp,n), (Sp,h), (Sp,a), (Lh,H), (Lh,n), (Lh,p), (a,H), (a,n), (a,h) are not correlated to any ligand sequence; four other groups (p,H), (Sp,p), (Lh,h), (Lh,a) can be unified into canonical class II binding domains, because they all recognize ligands conforming to the consensus ([Φ/Ψ]XΦ*) (Table 1). One group (n,a) corresponds to the previously defined group of PDZ domains that bind to class III ligands (consensus [D/E]XV*). Another group (n,h) includes PDZ domains characterized by dual specificity. On this purpose we have to emphasize that, in some cases, different PDZ domain lists, available on the Web, may induce confusion in group assignment. In particular, in the course of the characterization of the PDZ domains present in the full-length clone hINADL (accession number AJ224747) we used the 'Pfam v5.5 version sequence alignment' [1]. Only seven PDZ domains were found by Pfam search tool, because the score of the region among residues 555 and 638 resulted less significant than the required threshold. By screening a combinatorial phage library, we classified the first four domains of hINADL as class II binding PDZ, and the last three domains as class I binding PDZ.

In contrast, Bezprozvanny and Maximov used a SMART classification that considers eight PDZ domains in hINADL. As a consequence of this different numbering, the binding preferences of hINADL-5 (hINADL-4 in our list) defined by Bezprozvanny and Maximov, using different techniques, results in agreement with the class II binding consensus, identified by phage library screening [1]. Therefore, the binding to neurexin Ia, whose terminus is 'EYYV*' (class II), is not surprising, while it is interesting to note that the same domain binds also to peptide 'DHWC' at the end of NC4 (N-type Ca^{2+} channel). Anyway, as previously mentioned, the large hydrophobic pocket of a class II binding domain can accommodate also residues that in part mimic a hydrophobic moiety [2]. Dual specificity is not a rare event in PDZ binding mode: we have recently described a new class of binding motifs, characterized by the presence of a negative residue at the C-terminus, which were selected by hINADL-3. This PDZ domain can bind with the same affinity also to class II peptide ligands [1].

In conclusion, we believe that a classification of PDZ domains, exclusively based on the chemical properties of only two positions, is not sufficient to describe the complexity of the domain family and to predict the specificity of binding. Several experimental data confirm this statement: we have shown that the substitution of the histidine at the crucial position αB1 of hINADL-7 (a class I binding PDZ) is not sufficient to change its binding specificity and to confer a clear ligand preference [1]. Furthermore, the first PDZ of MUPP1 is a member of the most abundant group (G,H); therefore it should bind to class I motifs. In contrast, our computer-aided analysis [1] and two hybrid (in vitro) and co-immunoprecipitation (in vivo) assays [4] indicate a preference of MUPP1-1 PDZ domain for class II ligands, such as NG2 proteoglycan. Another example is CIPP-3 PDZ domain that is classified as

PII: S0014-5793(02)02220-2

Table 1

Position	Consensus			
	[S/T]XΦ* class I	[Φ/Ψ]XΦ* class II	[D/E]XV* class III	XΨ[D/E]* class IV
P^0	hydrophobic or aromatic	hydrophobic or aromatic	hydrophobic or aromatic	negative or #
P^{-1}	any (Trp or Asp)	any (Trp or Asp)	any	aromatic
P^{-2}	Ser or Thr	hydrophobic or aromatic	negative	any
P^{-3}	any (Glu)	any (Glu)	Gly or Glu	any

PDZ domain ligands can be distinguished on the basis of residues at position P^{-2} (classes I, II, III) or position P^0 (class IV). Residues in brackets are the most abundant at that position. Φ: hydrophobic; Ψ: aromatic; *: COOH-terminus; #: any residue different from Φ/Ψ.

(G,p) and therefore it should prefer peptides related to the consensus (ΨDΦ*); in contrast, it interacts with Kir4.2 channel and NR2 (type A-B-C-D) receptors that have class I extremities [5].

At the moment PDZ domains can be classified only according to the classes of their ligands. Three classes are distinct by the kind of residue present at position P^{-2} and one class by the residue at position P^0 (Table 1). Ligands of class IV are peculiar, because they are in contrast with the general rule that PDZ domains bind to C-terminal hydrophobic residues. Actually, these ligands have been characterized only by in vitro binding assays; therefore their physiological binding relevance should also be proved in vivo. A further class definition, according to the residues present at positions P^{-3} and P^{-1} that often determine the unique specificity of each single PDZ domain, would be useful. At the moment, this is yet complicated by the variety of residues displayed in these positions (even if some residues, shown in brackets in Table 1, are more frequent).

The predictive power of computational methods aimed at inferring the recognition specificity of modular domains will continuously improve with the enrichment of domain/ligands structural and biochemical information and, hopefully, it will be possible to predict with sufficient confidence the putative targets of any PDZ domain.

References

[1] Vaccaro, P., Brannetti, B., Montecchi-Palazzi, L., Philipp, S., HelmerCitterich, M., Cesareni, G. and Dente, L. (2001) J. Biol. Chem. 276, 42122–42130.
[2] Daniels, D.L., Cohen, A.R., Anderson, J.M. and Brunger, A.T. (1998) Nat. Struct. Biol. 5, 317–325.
[3] Bezprozvanny, I. and Maximov, A. (2001) FEBS Lett. 509 (3), 457–462.
[4] Barritt, D.S., Pearn, M.T., Zisch, A.H., Lee, S.S., Javier, R.T., Pasquale, E.B. and Stallcup, W.B. (2000) J. Cell. Biochem. 79, 213–224.
[5] Kurschner, C., Mermelstein, P.G., Holden, W.T. and Surmeier, D.J. (1998) Mol. Cell Neurosci. 3, 161–172.

*Corresponding author.
E-mail addresses: dentel@dfb.unipi.it, dente@uniroma2.it

[a] *Kenton Laboratories, c/o Sigma-Tau, via Pontina Km 30.400, Pomezia, Rome, Italy*
[b] *Department of Physiology and Biochemistry, University of Pisa, Via Carducci, 56010 Pisa, Italy*

PII: S0014-5793(02)02220-2