FEBS Letters 586 (2012) 2648–2661

journal homepage: www.FEBSLetters.org

Review

The emerging contribution of sequence context to the specificity of protein interactions mediated by PDZ domains

Katja Luck *, Sebastian Charbonnier, Gilles Travé *

UMR 7242, Institut de Recherche de l'Ecole de Biotechnologie de Strasbourg, Bd Sébastien Brant, BP 10413, 67412 Illkirch, Cedex, France

ARTICLE INFO

Article history:
Received 29 February 2012
Revised 26 March 2012
Accepted 27 March 2012
Available online 4 April 2012

Edited by Marius Sudol, Gianni Cesareni, Giulio Superti-Furga and Wilhelm Just

Keywords:
PDZ
Specificity
Sequence context
Extension
β2–β3 Loop
Linear motifs

ABSTRACT

The canonical binding mode of PDZ domains to target motifs involves a small interface, unlikely to fully account for PDZ-target interaction specificities. Here, we review recent work on sequence context, defined as the regions surrounding not only the PDZ domains but also their target motifs. We also address the theoretical problem of defining the core of PDZ domains and the practical issue of designing PDZ constructs. Sequence context is found to introduce structural diversity, to impact the stability and solubility of constructs, and to deeply influence binding affinity and specificity, thereby increasing the difficulty of predicting PDZ-motif interactions. We expect that sequence context will have similar importance for other protein interactions mediated by globular domains binding to short linear motifs.

1. Introduction

Interactions between proteins are essential for most of the processes that happen in living cells. Many protein interactions in cell signalling are mediated by interactions between globular domains and short linear motifs (SLiMs) [1]. SLiMs are short disordered protein sequence segments that often become folded in their bound state [1,2]. Important insights on domain–SLiM interactions have been gained from studies on PDZ (PSD95-DLG1-ZO1) domains. PDZ domains constitute a large family of globular domains found in prokaryotes and eukaryotes [3] with about 270 occurrences in the human proteome (see next section). PDZ-domain containing proteins are implicated in diverse cellular functions such as establishment and maintenance of cell polarity [4], signal transmission in neurons [5] or in visual and auditive processes in the eye and ear [6,7], cell migration [8], and regulation of cell junctions [9] (for reviews see [10,11]).

The core PDZ fold adopts an antiparallel β barrel structure [12] comprising 5–6 β strands and 1–2 α helices (Fig. 1A). PDZ domains mainly recognize PDZ-binding motifs (PBMs) that are situated at the very C-terminus of proteins. Some PDZ domains may also bind internal (i.e. non-C-terminal) PBMs [13,14] or lipids [15]. PBMs bind via β augmentation to PDZ domains, e.g. PBMs adopt a β strand that pairs in an antiparallel manner with the β2 strand of the PDZ domain (Fig. 1A). The carboxylate group of the last residue of the SLiM (here, the term peptide will be equally used) is hydrogen-bonded to backbone amides of residues from the carboxylate binding loop (β1–β2 loop), thereby determining the C-terminal peptide selectivity of PDZ domains (Fig. 1A). Based on the recognition of C-terminal SLiMs, peptide positions are numbered starting from the last residue (position 0, p0) going backwards (p−1, p−2, and so forth). The last residue is almost always a hydrophobic residue, mainly Val, Leu or Ile. The third last peptide residue (p−2) can be either Thr or Ser (class I), hydrophobic (class II), or Glu or Asp (class III), thereby defining three main categories of PDZ-binding motifs [16,17]. Thus, recognition of SLiMs by PDZ domains is based on residues of two key peptide positions, p0 and p−2.

Indeed, it has been generally observed that SLiMs have on average less than four defined positions [2]. Given this small binding interface, numerous studies addressed the question about how SLiMs can fulfill the need for specific protein interactions in cell signalling [18–22]. An increasing number of studies now suggests that protein interactions in cell signalling are not only determined by their minimal interacting fragments (e.g. core globular domain

* Corresponding authors.
E-mail addresses: katja.luck@unistra.fr (K. Luck), gilles.trave@unistra.fr (G. Travé).

0014-5793 © 2012 Federation of European Biochemical Societies. Published by Elsevier B.V. Open access under CC BY-NC-ND license.
http://dx.doi.org/10.1016/j.febslet.2012.03.056

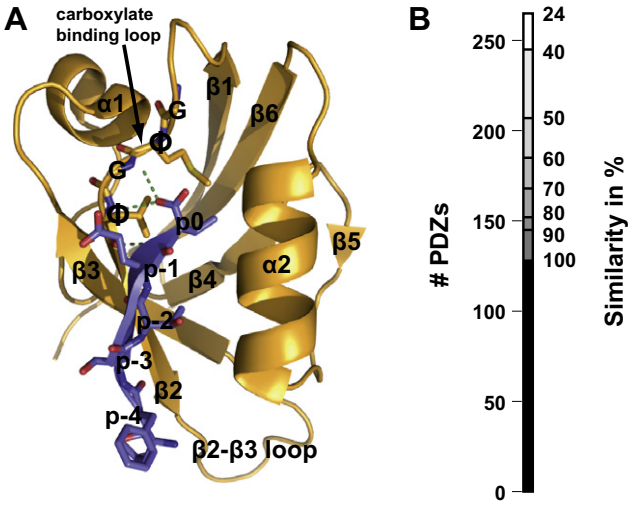

Fig. 1. Available structural information on PDZ domains. (A) Structure of the PDZ domain of AF6 (Afadin) bound to a C-terminal peptide (LFSTEV) derived from Bcr (PDB ID: 2AIN [112]). The secondary structure elements, peptide positions, and the common signature of the carboxylate binding loop (GϕGϕ where ϕ represents a hydrophobic residue) are indicated. Green dashed lines represent hydrogen bonds that are established between Val at p0 and the carboxylate binding loop. Figure was created with Pymol [113]. (B) Diagram that illustrates the number of human PDZ domains for which there is at least a structure of a PDZ domain in the PDB with X % sequence similarity based on local sequence alignments from BLAST searches (e.g. for 152 human PDZs there is a structure of a PDZ in the PDB with at least 80% sequence similarity.)

and SLiM) but also by their *context* [23–26]. The context can be either encoded within the *sequences* of the two interacting proteins; or defined by the *cellular environment*. In our definition, cellular context comprises factors that influence the temporal and spacial distribution of proteins, thereby determining when and under which conditions (e.g. local concentration) two proteins will meet in order to bind each other [25]. Sequence context comprises the regions in proteins that surround SLiMs or globular domains and that were shown to have an impact on the domain–SLiM interactions. Here, we refer to sequence context as *extensions*, if they occur directly upstream or downstream of the SLiM or domain and if they are not part of other domains. Sequence context that is not considered as extensions consists of neighbouring domains and regions that are not in the neighbourhood of the SLiM or domain.

The repertoire of structures on PDZ domains deposited in the Protein Data Bank (PDB) [27] is immense (Fig. 1B) providing us with a unique perspective on the structural diversity that exists within one protein domain family. Given the accumulated knowledge about PDZ–peptide interactions, they can serve as a model system to understand how domain–SLiM interactions are influenced by their sequence context. In the following, we will review studies that provide insights on extensions of PBMs and PDZs as well as studies that investigated the interplay between PDZ domains and their neighbouring domains.

2. Precision of the total number of PDZ domains in the human proteome

There has been considerable confusion about the total number of PDZ domains in the human proteome (hereafter called human PDZome). Numbers that are frequently referred to in the "PDZ literature" range from about 250 up to 900. Based on the articles that justify the number of human PDZs claimed in their text, we identified three different sources. Articles that claim a total number of 450 human PDZ domains (see suppl. data for references), refer to the database SMART [28]. Indeed, SMART provides a list of the human PDZome. However, upon closer inspection, this list turned out to be highly redundant, probably because it has been based on a set of protein sequences containing several isoforms for the same gene that have identical PDZ sequences. By now, SMART decreased the original number from 450 down to 364 (as of February 2012). Spaller [29] represents the second source, which claimed a total number of 918 human PDZ domains. In an erratum published four years later, Spaller corrected this number down to 234 and explained that he had been misled by erroneous numbers that were present in the preprint (but not the final version) of Bhattacharyya et al. [30]. Nonetheless, articles are still being published that seem to ignore the erratum [31]. Finally, several independent studies [32–34] that aimed at finding all human PDZ domains for bioinformatic analyses converge on a total number of about 270 PDZs in the human proteome without counting alternatively spliced forms of PDZs. This latter number, which we confirmed in our own investigations, is most probably the best estimate of the size of the human PDZome.

3. Extensions of PBMs

In numerous low-scale and large-scale studies as well as biocomputational analyses researchers sought to decipher the specificity rules of PDZ–peptide interactions [19,35–38]. Apart from peptide residues at p0 and p−2, which are the hallmarks for recognition by PDZs, the role of other residues of the PBM is more variable. Some studies suggested that the contribution of residues at p−1 and p−3 to the overall binding event is minor whereas in other studies PDZ domains exhibited clear preferences for certain residues at these sites over others [39–43]. More and more studies now converge on the idea that PBMs should at least be extended to p−4 as residues at this position have also been observed to significantly contribute to the binding to PDZs [19,35]. By extending the PBM further and further, a few interesting studies indicate that peptide residues up to p−10 are also implicated in PDZ binding. In the following, we define the core PBM as consisting of the last four residues (these residues clearly bind in the binding pocket of PDZ domains between the β2 strand and the α2 helix) and any longer PBMs will be considered as being extended.

3.1. Residues upstream of the core PBM modulate the binding affinity to PDZ domains

Interesting observations about PDZ–peptide recognition involving an extended PBM were obtained when comparing the binding of peptides derived from Wnt-signalling protein βCatenin and inward rectifier K(+) channel protein Kir2.3 to the PDZ domain of TIP1 (tax-interacting protein 1). A long βCatenin peptide bound stronger than a short peptide to the TIP1 PDZ [44,45]. The main contributor to this difference in affinity is most probably Trp at peptide position p−5 (in the following, we will write W_{-5}). Mutation of W_{-5} to Ala significantly weakened binding, as did mutation of Pro to Ala or Ser in the β2–β3 loop of the PDZ domain facing W_{-5} [44]. Interestingly, mutation of R_{-5} to Trp in Kir2.3 peptide led to an astonishing increase in binding affinity from 6.4 μM to 8.5 nM [46] (Table 1). As for the TIP1-βCatenin interaction, W_{-5} in the βPix C-terminus (Rho guanine nucleotide exchange factor 7) contributes to the binding to PDZ1 of SHANK1 (SH3 and multiple ankyrin repeat domains protein 1) as its mutation to Ala reduced binding affinity, too [47] (Table 1).

3.2. Electrostatic interactions between residues of extended peptides and of PDZ domains

The charge of the residue at position p−5 of the C-terminus of two inward rectifier K(+) channel proteins GIRK3 and IRK1

Table 1
Mutagenesis performed to study extended PBMs. The table summarizes the mutational data obtained from studies that analysed interactions between residues from extended PBMs and PDZ domain residues. If available, measured binding affinities are indicated. wt = wild type, Cter = C-terminal extension of PDZ, alt.spl. = alternatively spliced, Nter-PDZ = construct comprising the PDZ extended at its N-terminus, AA = amino acids, CC = coiled-coil.

PDZ	peptide	peptide sequence	PDZ modification	peptide modification	affinity in μM	ref
PAR3-3/3	PTEN	DEDQHSQITKV	wt	11AA, wt	19	[58]
			wt	8AA, wt	/	
			wt	D-10A/E-9A/D-8A	weaker	
			β2-β3:K609/R611A	11AA, wt	240	
			β2:K606/β2-β3:K609/R611A	11AA, wt	550	
			β2:K606A	11AA, wt	200	
	Cadherin	YGSDPQEELII	wt	12AA, wt	6	[57]
			wt	6AA, wt	28	
			wt	D-7A	42	
			wt	Y-10A	15	
			β2-β3:R609A	12AA, wt	25	
			wt	phosphoS-8	2	
Erbin-1/1	ErbB2	EYLGLDVPV	wt	9AA, wt	50	[61]
			wt	phosphoY-7	128	
TIP1-1/1	Catenin	QLAWFDTDL	wt	9AA, wt	0.19	[44]
			β2-β3:P45A/S	wt	3	
			wt	W-5A	20	
			wt	5AA, wt	21	[45]
	kir2.3	ISYRRESAI	wt	9AA, wt	6.4	[46]
			wt	10AA, wt	45	[114]
			wt	R-5W	0.0085	[46]
			wt	S-2D	/	
			wt	R-5W/S-2D	binding	
			α2:D91A	9AA, wt	29	
SHANK1-1/1	βPix	AWDETNL	wt	CC domain	2.6	[47]
			wt	W-5A	weaker	
DLG1-2/3	E6	RRETQV	wt	6AA, wt	1	[54]
DLG1-2/3	APC	RHSGSYLVTSV	wt	11AA, wt	1	[53]
			β2-β3:Q339P	11AA, wt	10	
DLG1-1/3	APC	RHSGSYLVTSV	wt	11AA, wt	18	[53]
DLG1-2/3	NR2B	KLSSIESDV	wt	9(14?)AA, wt	1	[55]
DLG1-1/3	NR2B	KLSSIESDV	wt	9(14?)AA, wt	10	[55]
			β2-β3:P245Q	9(14?)AA, wt	5	
			β2-β3:P245A	9(14?)AA, wt	10	
MAGI1-2/6	E6	SSRTRRETQL	wt	10AA, wt	3	[51]
			wt	R-7A/R-5A/R-4A	7	[51]
			wt	L/V at p0	0.25	[52]
			β2-β3:DEPDE/NQPNQ	L/V at p0	35	[51]
			β2-β3:DEPDE/AAPAA	L/V at p0	60	[51]
			Cter:S113R/L114K/V115R	L/V at p0	0.75	[52]
			Cter:S113G/L114G/V115G	L/V at p0	1.25	[52]
hPTP1E-2/5	APC	HSGSYLVTSV	wt	10AA, wt	200	[64]
			alt.spl.	10AA, wt	/	
	RIL	VAVYPNAKVELV	wt	12AA, wt	1100	[64]
			alt.spl.	12AA, wt	19000	
Harmonin-1/3	Sans	PALEDTEL	Nter-PDZ	SAM-PBM	0.001	[66]
			Nter-PDZ	PBM alone	1	
			Nter-PDZ	SAM alone	40	
ZO1-2/3	Connexin43	ASSRPRPDDLEI	dimer	9AA, wt	17	[69]
			dimer	12AA, wt	7	
			β2-β3: GGGAinsertion	9AA, wt	/	
			dimer	phosphoS-9	>100	

determined the binding of these two proteins to either the PDZ domain of SNX27 (sorting nexin 27) or the first two PDZ domains of PSD95 (postsynaptic density protein 95, DLG4). Exchanging Glu at position p−5 of GIRK3 with Arg, as observed in IRK1, was sufficient to induce the binding of GIRK3 to PSD95 and to disrupt its interaction with SNX27 [48].

In a recent study [49], molecular dynamics simulations suggested ionic contacts between the β2–β3 loop of PDZ3 of DLG4 (disks large homolog 4) and of a C-terminal peptide derived from CRIPT (cysteine-rich interactor of PDZ3) that seemed to be important for peptide binding.

Structural and mutagenesis studies provided evidence for electrostatic interactions between PDZ2 (also referred to as PDZ1) of MAGI1 (membrane-associated guanylate kinase inverted 1) and an extended C-terminal peptide derived from Human Papillomavirus (HPV) 16 E6 [50–52]. Mutation of the negative charges in the β2–β3 loop of PDZ2 to Gln and Asn or Ala significantly reduced the binding affinity as did mutation of the positive charges at position -4, -5 and -7 to Ala in the E6 peptide [51] (Table 1). In another study, the contribution of upstream peptide residues to the binding affinity to PDZ2 of MAGI1 has been more generally assessed. By measuring binding affinity of several peptides of different length, the preference of PDZ2 of MAGI1 for peptides with positive charges upstream the core PBM has been confirmed [26]. This is one clear case where an increase in specificity is driven by interactions between residues of extended PBMs and residues of the β2–β3 loop. In contrast, the β2–β3 loop of PDZ3 of MAGI1 did not seem to contribute at all to peptide binding [26]. The same peptides were also assayed for binding to PDZ3 of the cell polarity protein hScrib revealing that extended peptides generally bound with higher affinity to PDZ3. This study demonstrates that depending on the PDZ in question, interactions between upstream peptide residues and the β2–β3 loop can have different implications on peptide binding [26].

3.3. The sequence of the β2–β3 loop influences peptide binding

C-terminal peptides of APC (adenomatous polyposis coli protein), the glutamate receptor subunit NR2B, and HPV18 E6 bind stronger to PDZ2 of DLG1 (disks large homolog 1) as compared to PDZ1 of the same protein [53–55] (Table 1). Several studies suggest that amino acid differences in the β2–β3 loop of these two PDZ domains mainly account for the different binding preferences via interaction with residues of extended PBMs. Mutation of a Gln in the β2–β3 loop of PDZ2 to Pro (the corresponding residue in PDZ1) decreased binding affinity for APC [53]. The converse mutation (Pro to Gln) directed at the equivalent position of the β2–β3 loop of PDZ1 increased binding affinity of PDZ1 for NR2B whereas mutation of the same Pro of PDZ1 to Ala did not alter the binding affinity [55]. The Pro in the β2–β3 loop of PDZ1 might also be responsible for weaker binding to E6 in comparison to PDZ2. Liu et al. [54] noticed that the neighbouring residues of the Gln in PDZ2 adopt a particular conformation that contributes to peptide binding, and suggested that the Pro in PDZ1 did not allow its neighbouring residues to adopt a similar favorable conformation. One reason for the weaker binding displayed by PDZ3 to these peptides might be its shorter β2–β3 loop, which does not provide such a platform for extended peptide binding as does PDZ2 [53–55].

3.4. Extended PBMs confer dual binding specificity to PDZ domains

A few PDZ domains were shown to have dual specificity, e.g. binding to PBMs of class I and class II. The dual specificity of PICK1 (protein interacting with C kinase 1) has been mainly attributed to specific residues of the α2 helix [56]. In contrast, the dual specificity of PDZ3 of the cell polarity protein Par3 can be attributed to an extended binding pocket [57]. PDZ3 of Par3 binds to long C-terminal peptides derived from both the vascular endothelial Cadherin (class II PBM, 12 residues) and phosphatase PTEN (class I PBM, 11 residues). However, it significantly binds weaker to a shorter (6-residue long) Cadherin peptide and does not detectably bind anymore to a shorter (8-residue long) PTEN peptide [58,57] (Table 1).

Analysis of the two available NMR structures of these two complexes revealed several contacts between negatively charged residues in the extended peptides and positively charged residues of the β2 and β3 strand as well as the β2–β3 loop of PDZ3 [58,57] (Fig. 2A and B). Whereas the Cadherin peptide mediated more favourable interactions by means of its last four peptide residues, the PTEN peptide established more favourable contacts to PDZ3 with its upstream residues. At first sight this might be interpreted as an example where a longer binding pocket leads to more promiscuous binding behaviour. Yet, class I peptides such as PTEN have to fit very well to the extended binding site with their upstream residues in order to be bound by PDZ3, a constraint that might only be fulfilled by a few peptide sequences.

3.5. Conformational changes of the β2–β3 loop upon peptide binding

As indicated by the previous examples, residues from extended PBMs mainly modulate binding affinity to PDZ domains via interaction with residues of the β2–β3 loop. Comparison of available apo (unbound) and holo (bound) NMR and crystal PDZ structures revealed that the β2–β3 loop either changes conformation upon peptide binding or remains unchanged. Two examples of the latter case are represented by PDZ2 of DLG1 and Erbin (Erbb2-interacting protein) PDZ of which the β2–β3 loops exist in a stable conformation when no peptide is bound and this conformation remains unchanged upon peptide binding, also when complexed to different peptides. In Erbin PDZ, there is a chain of aliphatic contacts from the β2–β3 loop to the β3 strand and the β4 strand (Fig. 2C). Additionally, N1345 of the β2–β3 loop seems to establish a hydrogen bond to the backbone of the loop. These interactions between residues of the Erbin PDZ are probably the driving forces that keep that loop very rigid providing a stable platform for peptide binding. The β2–β3 loop of PDZ1 of ZO1 (Zonula occludens protein 1) is equally long as that of Erbin PDZ, and it adopts a similar conformation upon peptide binding, but it displays a different conformation in its unbound form. Here, the loop seems to restructure upon peptide binding allowing for accommodation of upstream peptide residues (Fig. 2D). Similar observations were obtained for the β2–β3 loop of the TIP1 PDZ domain (Fig. 2E2) and the PDZ of SHANK1 [47].

3.6. The β2–β3 loop can form an additional peptide binding pocket that accommodates upstream peptide residues

In many of the examples mentioned in the previous subsections, the β2–β3 loop contributed to the formation of an additional peptide binding pocket together with residues from strands β2, β3, and sometimes β4. In particular, an aromatic residue (mainly F or Y), located right at the beginning of the β3 strand, is often involved in the formation of this additional pocket. This residue contributed to the binding of peptide residues upstream position -3 and seemed to serve as an anchoring point for the structuring of the β2–β3 loop. The conserved aromatic character of this position in the family of PDZ domains suggests that residues at this position might be of more general importance for the structure and function of PDZs [49].

In general, this additional pocket was frequently observed to have hydrophobic character being occupied by upstream peptide residues with large aliphatic side chains such as Trp, Tyr, or Arg (Erbin PDZ, PDZ1 of ZO1, TIP1 PDZ, Fig. 2C–E1, respectively). However, in the case of the PDZ of SNX27, this additional pocket is rather of hydrophilic character being formed by three arginines

	PDZ	peptide	peptide sequence	pep. pos.	domain position
A	PAR3-3/3	PTEN	DEDQHSQITKV	-8	β2-β3:R609
				-9	β2:K606, β3:K611
				-10	β2-β3:K611
B	PAR3-3/3	Cadherin	YGSDPQEELII	-4	β2-β3:R609
				-7	β2-β3:R609, β2-β3:K611
				-9	β2-β3:K611
				-10	β2-β3:K611
C	Erbin-1/1	ErbB2	EYLGLDVPV	-7	β2-β3:P1305
	Erbin-1/1	artificial peptide	TGWETWV	-4	β2:S28, β3:T48, β2-β3:P37, β2-β3:F38
D	ZO1-1/3	artificial peptides	WRRTT[WY]L	-4	β2-β3:N44, α2:H88
				-5	β2-β3:N44, β2-β3:P45, β2-β3:F47
				-6	β3:V55, β2:S39, β2-β3:H46, β3:S57, β2-β3:E51
E	TIP1-1/1	Catenin	QLAWFDTDL	-4	α2:H90, β2-β3:Q39
				-5	β2:G34, β2-β3:N44, β2-β3:P45, β2-β3:F46, β3:Y56, β3:T58
	TIP1-1/1	Kir2.3	ISYRRESAI	-4	α2:H90, α2:D91
				-5	β2-β3:Q43, β2-β3:N44, β2-β3:P45
F	SHANK1-1/1	βPix	AWDETNL	-4	β2-β3:K682, α2:R736
				-5	β2:R679, β2-β3:F696, β3:Y701
				-8	β2-β3:K682
G	SHANK1-1/1	GKAP1a	EAQTRL	-5	β2:R679
	SNX27-1/1	GIRK	ESESKV	-5	β2:R56, β2-β3:R66, β3:H78, β4:R98
H	DLG1-2/3	E6	RRETQV	-4	β2-β3:G334, β2-β3:N338, α2:E384
				-5	β2-β3:H340
	MAGI1-2/6	E6	RSSRTRRETQL	-4	β2-β3:D35
				-5	β2-β3:E39, C-ter:S113, C-ter:L114, C-ter:V115
				-6	C-ter:S113, C-ter:L114, C-ter:V115
				-7	β2-β3:E36, β2-β3:D38
	hPTP1E-2/5	RA-GEF2	EADENEQVSAV	-5	β2-β3:V26, β2-β3:N27
I	Harmonin-1/3	Sans	PALEDTEL	-5	β2:R103, β2-β3:F108, β3:C110, C-ter:V172, C-ter:K173, C-ter:L180
				-6	C-ter:K173
				-7	β2-β3:F108
J	Harmonin-2/3	Cadherin23	TPLEITEL	-4	α2:K73
				-5	β2-β3:P31, β3:F38, C-ter:L97, C-ter:F98
K	ZO1-2/3	Connexin43	RPRPDDLEI	-4	β5-α2:N239
				-5	allows for kink to avoid clashes with β5-α2 loop
				-6	β5-α2:E238

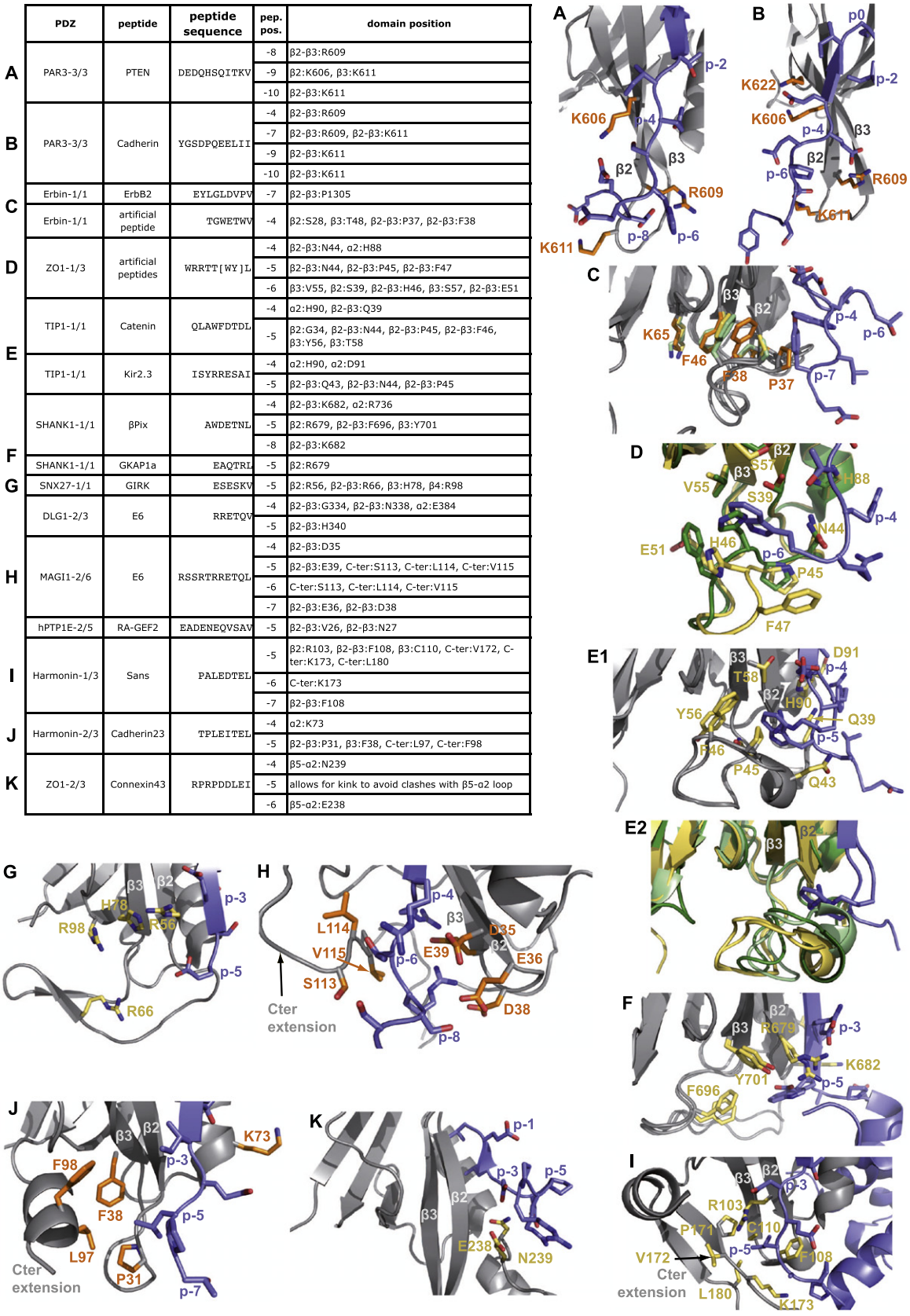

Fig. 2. Contacts between residues of extended PBMs and PDZ residues. The table summarizes the contacts that were observed between residues in extensions of PBMs and residues of PDZ domains or their extensions. A structural representation is provided for most of these PDZ–peptide complexes. Colour code: peptide residues in blue, PDZ residues from holo NMR structures in orange, from apo NMR structures in dark-green, from holo crystal structures in yellow, and from apo crystal structures in light-green. Structural information was used from the following PDB entries: (A) 2K20 [58]; (B) 2KOH [57]; (C) 1MFG [61], 1N7T [59], 2H3L [60]; (D) 2H2B, 2H3M [60]; E1: 3DIW [44], 3GJ9 [46]; E2: 3DJ1, 3DIW [44], 3GJ9 [46], 2KG2 [114]; (F) 3L4F [47], 1Q3P [115]; (G) 3QGL [48]; (H) 2KPL [52], 2I04 [50]; (I) 3K1R [66]; (J) 2KBS [67]; (K) 3CYY [69]. Figures were created with Pymol [113] (see suppl. data for the pymol session files).

from the β2 strand, β2–β3 loop, and β4 strand, and accommodates E_{-5} of the C-terminal peptide of GIRK3 (Fig. 2G). Another interesting example is that of the trimer of βPix bound to the SHANK1 PDZ. The homotrimer is formed by the coiled-coil domain of βPix leading to three closely located C-termini all carrying the PBM of βPix [47]. However, due to steric hindrance, only one of the three C-termini is accessible for binding to SHANK1 PDZ [47]. Here, the additional pocket formed by the β2–β3 loop, β2, and β3 strand adopts a hemispheric shape and interacts with one side of the ring of W_{-5} from βPix (Fig. 2F). The other side seems to be covered from solvent by residues from the coiled-coil domains. In that way, W_{-5} may contribute to the stabilisation of the whole complex.

In most cases, this additional pocket is occupied by peptide residues from position -5. However, it can sometimes be occupied by residues at different positions, such as W_{-4} for a phage display-derived peptide bound to Erbin PDZ (Fig. 2C) [59], W_{-6} for a phage display-derived peptide bound to PDZ1 of ZO1 (Fig. 2D) [60], and Y_{-7} for ErbB2 peptide bound to Erbin PDZ (Fig. 2C) [61]. Together with observations that peptide residues at position -4 can either contact domain residues from the β2 or β3 strand, from α2 helix or from the β2–β3 loop (see table in Fig. 2), this demonstrates to which extent peptides can adapt to PDZ domains. A rigid definition of pairs of domain and peptide residues as often considered in PDZ–peptide interaction predictors, would not reflect this adaptability and is therefore likely to be an inappropriate model.

3.7. Alternative splicing of the β2–β3 loop can modulate the peptide binding properties of PDZ domains

A very different example of influence of the β2–3 loop on peptide binding is provided by PDZ2 of the tyrosine protein phosphatase hPTP1E (or its mouse homolog PTP-BL). This PDZ domain represents one of the few PDZs subject to alternative splicing, and to our knowledge is the only one that was studied in detail (with in total 8 structures deposited in the PDB). The alternatively spliced form (PDZ2as) exhibits an insertion of 5 residues (VLFDK) at the start of the β2–β3 loop that abrogates binding to the C-terminal peptide of APC and RIL (reversion-induced LIM protein) [62–64] (Table 1). Two NMR structures of PDZ2as exist but the fine structural interpretation of how the insertion negatively influences peptide binding remains a matter of debate as the two structures considerably diverge from each other (Fig. S1) [63,64]. The alternative splicing event leads to replacement of the GG hinge region at the beginning of the β2–β3 loop (a conserved feature in the PDZ fold) by the more conformationally restricted VL sequence [64]. This appears to induce a displacement of the loop, which, together with a global destabilisation of the domain, seems to be the main cause for the observed inability of PDZ2 as to bind to C-terminal peptides [64]. It will be interesting to see how alternative splicing might modulate the binding behaviour of other PDZ domains and whether these are used in vivo to regulate PDZ–peptide interactions [65].

3.8. Upstream residues of a PDZ-bound peptide can interact with residues that do not belong to the core of that PDZ domain

In most of the cases upstream peptide residues interacted with residues of the core PDZ domain but there are also a few very interesting instances where they interact with residues from PDZ extensions or from other proteins. Atomic contacts could be observed by NMR between residues (Ser, Leu, and Val) of the C-terminal extension of PDZ2 (PDZ1) of MAGI1 and R_{-5} and T_{-6} of a bound C-terminal peptide derived from HPV16 E6 [52] (Fig. 2H). This significantly restricted the disordered conformation of the C-terminal extension of PDZ2 in the presence of bound peptide [52]. Mutation of the three residues SLV to either RKR or GGG resulted in threefold and fivefold reduced binding affinity in comparison to the wild type, respectively [52] (Table 1).

Harmonin and Sans are two proteins that are implicated in the Usher syndrome. The N-domain, PDZ1, and a C-terminal extension of PDZ1 of Harmonin form an integral domain that binds to a PBM located at the C-terminus of Sans [66]. This interaction is further stabilised by interaction between the SAM (sterile alpha motif) domain of Sans and PDZ1 of Harmonin leading to an extremely tight complex (reviewed in detail by Wang et al. [33]) (Table 1). L_{-5} inserts into a hydrophobic pocket formed by residues from the β2 and β3 strand, from the β2–β3 loop, and from the C-terminal extension (Fig. 2I). Additionally, hydrogen bonds were observed between Lys from the C-terminal extension and the backbone of A_{-6} [66]. The PDZ2 domain of Harmonin, complexed to the C-terminal peptide of the cell adhesion protein Cadherin23, exhibits an additional C-terminal α helix [67]. Interestingly, the C-terminal tails of Cadherin23 and of Sans are very similar with identical residues at p−5, −4, −2, −1, 0. L-5 of Cadherin23 interacts with residues of the extension and the core PDZ domain in a similar way as L-5 of Sans, although in the case of PDZ2 it is a hydrophobic patch that is formed instead of the pocket reported for PDZ1 (Fig. 2J). Comparing the structure of PDZ2 of Harmonin with the one from PDZ3 of ZO1 revealed a very similar hydrophobic patch formed from the additional C-terminal α helix of PDZ3 and residues from the β2–β3 loop and the β2–β3 sheet. No difference in binding affinity could be observed to the C-terminal peptide of the gap junction protein Connexin45 when the helical extension of PDZ3 of ZO1 was removed [68]. Yet we speculate that an increase in affinity might have been observed for peptides possessing residues at upstream peptide positions that would have been capable to interact with this hydrophobic surface.

ZO1 and ZO2 PDZ2 were shown to exist as homodimers [69,70]. The complex of ZO1 PDZ2 bound to the PBM of Connexin43, revealed that due to dimerization the extended peptide binding pocket is altered. This has been reviewed more in detail by Wang et al. [33]. The β5–α2 loop of the other PDZ is placed in front of the end of the binding pocket forcing the peptide to make a bend at position -5 that is occupied by a Pro in Cx43. D_{-4} and R_{-6} are involved in several interactions with residues from the β5–α2 loop (Fig. 2K, Table 1) [69].

Unfortunately, the Harmonin-Sans/Cadherin23 and ZO1-Cx43 complexes have neither been mutated nor compared to complexes of Harmonin and ZO1 with PBMs of other known binding partners. Such analyses would have helped to better understand the contribution of observed interactions between residues from PBM and PDZ extensions to complex formation. However, given the numerous precise residue contacts observed, it is tempting to speculate that these additional interactions between PBM and PDZ might increase the specificity of the PDZ domains for their peptide targets.

3.9. General remarks

It clearly emerges that residues of PBM extensions influence the affinity of peptides to PDZ domains and some studies provide clear evidence that this increases the specificity of PDZ–peptide interactions. The β2–β3 loop contributes to an extended binding pocket. Owing to the huge variability of this loop in length (up to 36 residues, PDZ of ARHGAP21, PDB ID: 2YUY [71]) and sequence composition, observed interactions between residues of the loop and of the peptide were of impressive diversity. It seems rather impossible to derive any rule about how the β2–β3 loop and the peptide extensions contribute to PDZ–peptide binding. Most likely, this complex system of interactions is at the moment unpredictable.

Interestingly, the contribution of upstream peptide residues for the binding to PDZ domains seems to be subject to regulation as

they were observed to overlap with phosphorylation sites. Reported phosphorylation of upstream PBM residues led either to an increase in binding affinity (S_{-8} of Cadherin bound to PDZ3 of Par3 [57]) or to a decrease (Y-7 of ErbB2 bound to Erbin PDZ [61], S-9 of Cx43 bound to PDZ2 of ZO1 [69]) [72] (Table 1). It appears that experimental studies on PDZ–peptide interaction specificities should systematically include peptides of different length (e.g. core PBMs and extended PBMs) to account for the role that upstream peptide residues play for PDZ binding.

In principle, extending the binding interface between PDZs and peptides is likely to increase affinity between both. However, none of the cellular peptides that we have described above possessed an optimal sequence for the entire binding site. This observation is reflected by the fact that three artificial peptides (a phage display peptide, mutated Kir2.3 peptide, mutated HPV16 E6 peptide) displayed much higher affinity to PDZ domains than the identified natural binding partners (Table 1). Noteworthy enough, a high affinity interaction is not necessarily specific, as one or both of the interaction partners might bind many other proteins with similarly high affinity. Conversely, a low affinity interaction can be specific, if most other interactions engaged by the two partners are of even lower affinity. As mentioned in the introduction, SLiMs seem too short to fully account for specific protein interactions. It is tempting to speculate that the extended binding interfaces reviewed herein serve to provide a balanced mix of advantageous and disadvantageous interface contacts, which guarantee the weak yet specific interactions typically observed in cell signalling.

4. Sequence context of PDZ domains

4.1. Defining the core and boundaries of PDZ domains: theory and practise

The canonical core PDZ fold is defined by the arrangement of its secondary structure elements, namely six β strands and two α helices (Fig. 1A). In theory, the boundaries of a PDZ domain are thus defined by the first residue of the first β strand and the last residue of the last β strand of the canonical PDZ fold. The web servers SMART [28] and Pfam [73], both based on Hidden Markov Models, efficiently detect PDZ domains, including quite divergent instances presenting structural rearrangements (such as the PDZ-like domains from HtrA and EpsC (see below) [74,75]). However, these programs have not been developed for accurately predicting the boundaries of the core domain structures.

In practise, the boundaries of most experimental PDZ constructs that have been used for structural studies or interaction assays, generally extend beyond the strict boundaries of the core PDZ structure. Such extensions vary from a few residues to longer stretches, which may be structured. Whereas Bhattacharya et al. [76] and Wang et al. [33] proposed approaches for predicting some secondary structure elements within PDZ extensions, the conformation of extensions and their impact on the solubility, stability and peptide binding properties of PDZ constructs (see following paragraphs) remain largely unpredictable.

4.2. A PDZome-wide database of suggested PDZ construct boundaries derived from structural information

The lack of a general rule for designing boundaries of PDZ domain constructs raises a practical issue concerning past and future experimental studies aimed at producing and comparing the functional properties of large numbers of PDZ domains. How can hundreds of PDZ constructs be properly designed? One solution to this problem may be to take advantage of the impressive structural data already available on PDZ domains. Indeed, successful structure determination probably represents the most relevant *a posteriori* quality proof for construct boundaries. Based on this principle, we propose here a list of manually curated domain boundaries sequences for the 266 PDZs that constitute to our knowledge the human PDZome (Table S1). For each known PDZ domain, we identified the three most similar PDZ domains for which structures were available and applied a set of hierarchized criteria (see suppl. material for protocol) to finally define the construct boundaries of the considered PDZ domain.

Noteworthy enough, the structures of about 120 human PDZs have been solved, and for most of the other human PDZs, the structure of one or several very closely related orthologous PDZ domains is available in the Protein Data Bank (PDB) (Fig. 1B). Therefore, the majority of constructs proposed in our curated list are either identical or highly similar to constructs for which a structure has been solved. On average, the constructs proposed in our list, including those of known structure, are 16 and 5 residues longer at the N- and C-terminus, respectively, than the start and end of the sequence predicted by SMART. In particular, SMART-predicted boundaries consistently excluded the first β strand of PDZ domains. Therefore, the raw SMART output is not sufficient for domain boundary prediction. This database provides a first suggestion of boundaries for the cloning of PDZ constructs from human or related proteomes. Of course, semi-empirical optimization of construct solubility and/or stability may still remain necessary in some cases [77].

4.3. PDZ extensions

The emerging roles of extensions of PDZ domains was addressed in a review by Wang et al. [33]. Here, we aim at complementing this review by focussing on recently published studies and by putting emphasis on the implications that extensions have on peptide binding.

4.3.1. Extensions that influence the dynamics, stability and solubility of PDZ domains

PDZ2 of NHERF1 (Na(+)/H(+) exchange regulatory cofactor 1) and PDZ3 of DLG4 (disks large homolog 4) have helical C-terminal extensions that were shown to be important for the stability and dynamics of the PDZ domain, respectively [76,78]. In both cases, removal of the extension did not alter the fold of the core PDZ domain but led to significantly decreased affinity to C-terminal peptides [76,78] (Table 2). Petit et al. [78] showed that a construct of PDZ3 of DLG4 lacking the extension, displays more side chain flexibility leading to increased entropy that makes peptide binding energetically less favourable. Both studies were reviewed more in detail by Wang et al. [33]. A recent molecular dynamics study [49] suggested that the helical extension of PDZ3 of DLG4 establishes ionic contacts with core PDZ residues. These contacts seemed to restrain the backbone flexibility of the β2–β3 loop and the carboxylate binding loop, thereby facilitating peptide binding.

The PDZ2 (PDZ1) domain of MAGI1 was shown to have a structured N-terminal and an unstructured C-terminal extension [52]. The C-terminal extension seemed to be essential for obtaining soluble and stable constructs. The N-terminal extension was shown to shield from solvent a hydrophobic patch that comprised a cysteine residue, thereby preventing the formation of soluble aggregates via non-native intermolecular disulfide bonds [77,52]. Mutations in the C-terminal (see previous section dealing with PBM extensions) and N-terminal (unpublished data) extensions weakened the binding affinity of PDZ2 to the HPV16 E6L/V peptide.

Table 2
Mutagenesis performed to study sequence context of PDZ domains. The table summarizes the mutational data obtained from studies that analysed the influence of extended PDZ domains or multido-main constructs on peptide binding. If available, measured binding affinities are indicated. SS-bridge = disulfide bridge.

PDZ	peptide	peptide sequence	PDZ modification	affinity in μM	ref
DLG4-3/3	CRIPT	TKNYKQTSV	extended	1	[78]
			core	26	
		NYKQTSV	extended	3.6	[80]
			core	81	
			α3:phosphY397	14	
NHERF1-2/2	CFTR	...TEEEVQDTRL	core	5	[76]
			extended	0.26	
			β1:R153Q	0.93	
INADL-5/10	Kon/Perd	LLRRNQYWV	reduced	1.2	[7]
			oxidized	20.4	
INADL-45/10			reduced	2.5	
INADL-45/10	TRP	TGRMISGWL	reduced	140	
INADL-345/10	TRP	...TGRMISGWL	reduced	0.1	
INADL-5/10	PLCβ	KTQGKTEFYA	reduced	/	
INADL-45/10			reduced	30	
			oxidized	/	
Par6-1/1	Rhodamine	VKESLV	core	72	[88]
			CRIB-PDZ	54	
			CRIB-PDZ SS-bridge	13	
			CRIB-PDZ+Cdc42	6	
			CRIB-PDZ SS-bridge+Cdc42	6	
ZO1-PSG	JAM-A	EGEFKQTSSFLV	extended	35	[107]
ZO1-3/3				/	
ZO1-PSG			SH3:β2-β3:L549R	/	[68]
ZO1-PSG	Connexin45	SGDGKTSVWI	extended	10	
ZO1-3/3				75	
ZO1-PSG			SH3:β2-β3:L549R	63	
DLG-3/3	CRIPT	DTKNYKQTSV	-	14	[108]
DLG-PSG			-	0.8	
DLG-3/3	neuroligin	KRVHIQEISV	-	900	
DLG-PSG			-	71	

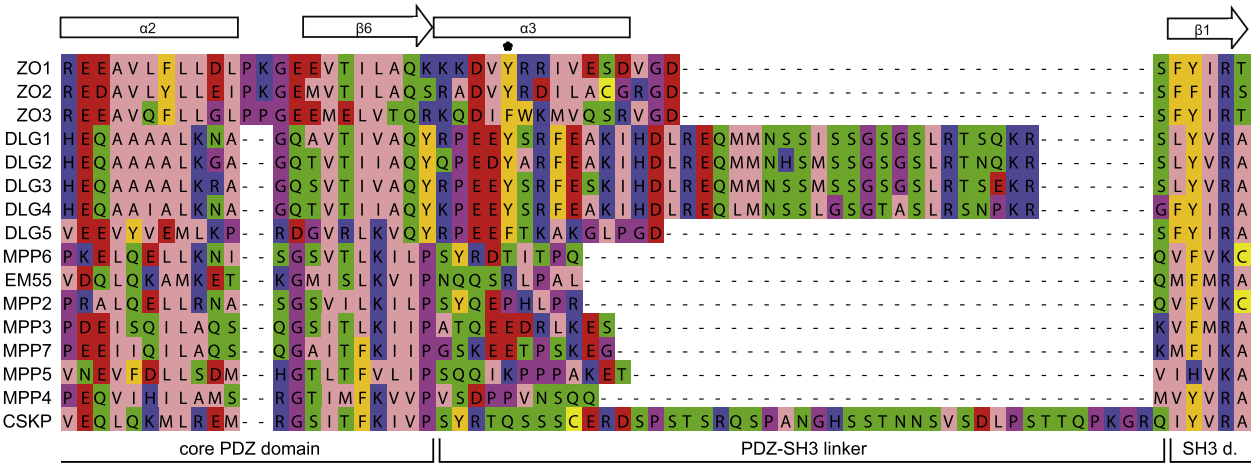

Fig. 3. Sequence alignment of the PDZ-SH3 linker of the human MAGUK family of proteins. Secondary structure elements on top of the alignment are indicated based on the structure of the PDZ3-SH3-GK module from ZO1 (PDB ID: 3SHW [68]). An initial alignment was built with Mafft [116] and corrected by hand using Jalview [117]. The asterisk indicates the phosphorylation site described for ZO1, which seems to be conserved for some members of the MAGUK family.

4.3.2. PDZ extensions as regulatory elements

DLG4 has a phosphorylation site (Y397) located in the helical C-terminal extension (hereafter called α3 helix) of PDZ3 [79]. Zhang et al. [80] studied the effect of this phosphorylation site on the structure of the extended PDZ3 domain and on ligand binding. Phosphorylated Y397 led to an equilibrium between a locally

unfolded and folded state of the α3 helix reflected by a fourfold decrease in affinity of phosphorylated PDZ3 in comparison to the unphosphorylated extended PDZ3 domain (Table 2).

DLG4 is a member of the membrane associated guanylate kinase (MAGUK) family of proteins together with four other DLG proteins, 3 ZO proteins, CASK (calcium/calmodulin-dependent serine protein kinase), and 7 MPP (membrane protein, palmitoylated) proteins [81]. All of these proteins share a common domain arrangement consisting of a PDZ domain followed by an SH3 (Src homology 3) and GK (guanylate kinase) domain (hereafter called PSG module). In addition to DLG4, the PDZ domain of the PSG module of ZO1 and DLG1 was also shown to possess an α3 helix. The α3 helix as well as the phosphorylation site seem to be conserved in the ZO and DLG subfamily of proteins (except for DLG5) (Fig. 3). Remarkably, the linker sequence between PDZ3 and SH3 of DLG1-4 is much longer than for the other members of the MAGUK-family. Three phosphorylated serines are reported in the linker region of DLG1-4 [82,83]. Furthermore, by using the ELM resource [84], we predicted in this linker a very likely actin binding site (LIG_Actin_WH2_2). To our knowledge, DLG proteins have not yet been demonstrated to bind to actin, yet they are already known to be involved in the regulation of actin filaments [85]. Therefore, it might be interesting to investigate the functional role of this actin binding site and its potential influence on peptide binding by the neighbouring PDZ3 and SH3 domains. The linker sequence of the MPP family is shorter than for the ZO and DLG proteins and does not display any conserved features. Furthermore, the region in MPP proteins that corresponds to the α3 helix observed in ZO and DLG proteins, tends to be proline-rich (Fig. 3) and is therefore unlikely to adopt a helical conformation.

In summary, this data suggests that a helical extension of the PDZ of the PSG module is a property shared by DLG1-4 and all three ZO proteins and that it might be used as a regulatory element to control peptide binding to PDZ3 via phosphorylation and actin binding.

4.3.3. PDZ extensions that modulate the conformation of the binding pocket

The PDZ domain of Par6 (partitioning defective 6 homolog) has in its unstructured N-terminal extension a CRIB (Cdc42/Rac-interactive binding) domain that adopts a β strand when bound by Cdc42 (Cell division control protein 42 homolog) [86]. The association of Cdc42 with the extended PDZ of Par6 leads to a 10-fold increase in affinity for C-terminal peptides bound by the PDZ [87] (Table 2). Several studies have been published investigating the mechanism and implications of the Cdc42-Par6 interaction that are in detail reviewed in Wang et al. [33]. Very recently, Whitney et al. [88] published structural details about how the signal resulting from Cdc42 binding to the CRIB domain is propagated to the binding pocket of the PDZ, thereby altering PDZ–peptide binding. By introducing a disulfide bridge between the otherwise very flexible CRIB domain and the core PDZ, they obtained a construct that appeared to mimick the structure of the CRIB-PDZ module when bound to Cdc42 and that was amenable for NMR studies. By comparing the structure of the disulfide-bridged mutant with a structure of the single PDZ and a structure from the Cdc42–Par6 complex, they revealed a switch in conformation of two residues (Lys and Leu) in the β1–β2 loop that reshapes the binding pocket upon Cdc42 binding and thereby facilitates peptide binding [88].

PDZ5 of INADL, a core component of the phototransduction pathway, has a pair of cysteines located in the peptide binding pocket, that is in reduced form in absence of light and forms a disulfide bridge upon light exposure [89]. C-terminal peptides were shown to bind significantly weaker to the oxidized form of PDZ5 in comparison to its reduced form [7] (Table 2). Interestingly, whereas the isolated PDZ5 was stable in its oxidized form, the reduced form of PDZ5 was prevailing within a PDZ4-PDZ5 tandem construct [7]. The crystal structure of PDZ4 and PDZ5 shows that they form a tight module that is stabilised by a C-terminal extension of PDZ5 that binds to the surface of PDZ4, and a N-terminal extension of PDZ4 that folds back onto PDZ4 [7]. Liu et al. [7] further provide interesting data that suggests that the C-terminal extension of PDZ5 is involved in a regulatory switch that changes the redox state of PDZ5 in the cell. Rapid light-induced acidification happens during signal transduction in the microvilli of eyes. This can lead to a local significant decrease of the pH that would be enough to protonate a histidine residue of PDZ4 leading to the disruption of its hydrogen bond with a threonine residue of the C-terminal extension. Consequently, the interaction between the C-terminal extension and PDZ4 is disturbed leading to a destabilisation of the whole module. This may result

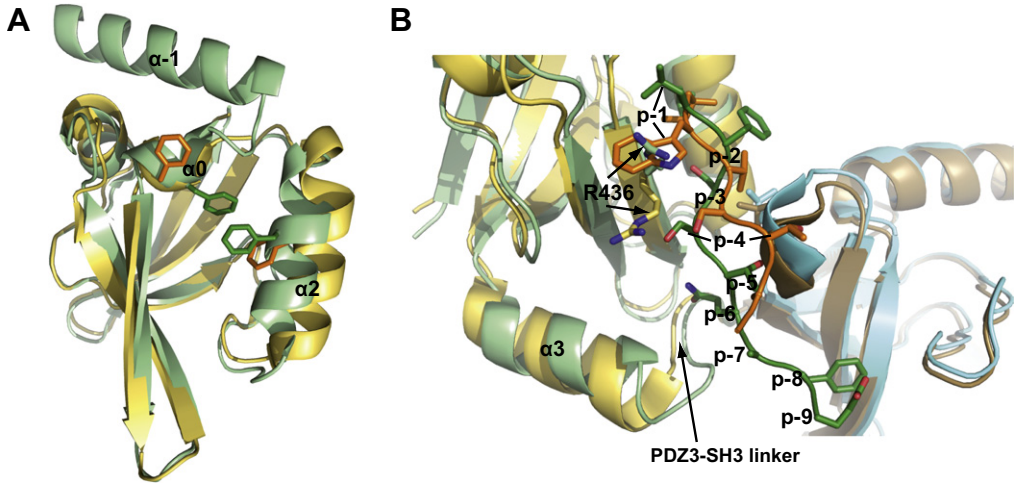

Fig. 4. Sequence context of PDZ domains. (A) Short (yellow, sPDZ, PDB ID: 2I6V) and long (green, lPDZ, PDB ID: 2I4S) version of the PDZ domain of the bacterial EpsC protein [75]. lPDZ has an additional N-terminal α helix (α−1) that leads to different conformations of α0 and α2. Two Phe are differently positioned (indicated in sticks) that lead in lPDZ to a closed peptide binding pocket. (B) Comparison of two complexes between the PDZ3-SH3-GK module of ZO1 and C-terminal peptides derived from JAM-A (forest-green) [107] and Cx45 (orange) [68]. The PDZ3 domains bound to JAM-A and Cx45 are shown in light-green and yellow, respectively, the corresponding SH3 domains in cyan and brown. The additional C-terminal α helix of PDZ3 is labelled α3. R436 possesses different conformations in the two complexes due to different residues at peptide position p−1 (W in Cx45 and L in JAM-A). The figures were created with Pymol [113] (see suppl. data for the pymol session files).

in a change of the local environment of the cysteine pair allowing for the formation of the disulfide bond. Possibly, the redox switch is used to initiate the dissociation, from INADL, of proteins such as the Ca^2+-permeable channel protein TRP and the phospholipase PLCβ. These dissociation events would serve to mediate signalling in photoreceptors upon light exposure.

Several bacterial proteins such as EpsC, a component of the type 2 secretion system of *V. cholerae,* contain PDZ-like domains that display a circularly permuted topology as compared to canonical PDZ folds [10,75]. In addition to this structural particularity, the carboxylate binding loop of the PDZ-like domain of EpsC is replaced by a small helical structure (hereafter called α0) [75] (see Fig. 4A). Both, the isolated PDZ-like core domain (sPDZ) and a longer construct (lPDZ) were crystallized. The lPDZ construct revealed an additional N-terminal α helix, which establishes many hydrophobic interactions with residues of the PDZ-like core structure [75]. The structures of sPDZ and lPDZ vary in the positioning of the α 2 helix (average difference of Cα atoms is about 3), which is further apart from β2 in sPDZ, thereby opening a deep and narrow hydrophobic groove formed between β2, α0, and α2 [75] (the numbering of secondary structure elements corresponds to the canonical topology of PDZs). The two structures also display significant differences in the positioning of α0. These two restructuring events cause a change in conformation of two Phe residues (one from α0, one from α2) that close the hydrophobic groove in lPDZ (Fig. 4A). It will be very interesting to investigate the binding properties of this very special PDZ-like domain that might have evolved to bind different molecules as has been suggested by the authors [75]. These three latter examples illustrate very different mechanisms by which peptide binding to PDZ or PDZ-like domains can be directly modulated via extensions that influence the conformation of the peptide binding pocket. Other mechanisms of peptide binding pocket modulation via extensions may exist for other PDZ domains, which would be interesting to find and investigate.

4.3.4. PDZ extensions that influence the folding of the PDZ domain

The assembly of PDZ4 and PDZ5 of INADL into a tight module, which was described in the previous section, also represents an example where the N- and C-terminal extensions were essential for obtaining soluble and folded constructs [7]. Very similar observations were obtained for PDZ4 and PDZ5 of GRIP1 (Glutamate receptor-interacting protein 1). Removal of the N-terminal extension of PDZ4 of GRIP1 led to spontaneous unfolding of the PDZ45 tandem [90] and prevented peptide binding [91] (more in detail reviewed elsewhere [5,33]). Together with the N-terminal extension of PDZ1 of Harmonin [66], these are also three examples where extensions of PDZ domains contribute to the construction of multidomain arrangements that ultimately alter the peptide binding behaviours of the PDZ domains being involved [33].

4.3.5. Non-exhaustive collection of structured and unstructured extensions that would deserve further investigation

- PDZ1 of INADL (PDB ID: 2DB5, NMR, [92]) has an additional N-terminal α helix that folds into a hole formed between the β2–β3 sheet and the β1–β5 sheet, and that is parallel to the β4 strand.
- PDZ1 of MPDZ (PDB ID: 2O2T, Xray, [93]) has an additional N-terminal and C-terminal α helix that both, contact the PDZ. The N-terminal helix is very similar to the one mentioned for PDZ1 of INADL.
- PDZ3 of Harmonin (PDB ID: 1V6B, NMR, [94]) has an additional N- terminal α helix that does not fold back on the PDZ core but establishes hydrophobic contacts with Val at the beginning of the β1 strand and Phe of the partially structured N-terminal extension that precedes the N-terminal α helix.
- The PDZ of MPP5 (PDB ID: 1VA8, NMR, [95]) has a small β hairpin that is formed at the N-terminus of the PDZ.
- PDZ7 of INADL (PDB ID: 2DAZ, NMR, [96]) has an N-terminal and C-terminal α helix that both, fold back on the PDZ core. The N-terminal helix is in slightly different orientation in comparison to PDZ1 of INAD and MPDZ. Here, the helix rather aligns to the β1- β5 sheet.
- PDZ2 of Harmonin (see section PBM extensions, PDB ID: 2KBS, NMR, [67]).
- PDZ2 (PDZ1) of MAGI2 (PDB ID: 1UEQ, NMR, [97]) has structured N-and C-terminal extensions that fold back onto the PDZ core and that are similar to those observed for PDZ2 of MAGI1 (see sections PBM extensions and PDZ extensions).
- The PDZ of PLCO (PDB ID: 1UJD, NMR, [98]) has a structured C-terminal extension that folds back onto the PDZ core.

4.3.6. General remarks

We have seen that PDZ extensions can influence the binding *affinity* of peptides to PDZ domains via a wide range of possibilities. Yet, do extensions also affect the binding *specificity* of peptides to PDZs? PDZ extensions that either directly interact with peptide residues (see section dealing with PBM extensions), as well as extensions that directly or indirectly alter the conformation of the peptide binding pocket, are very likely to increase the binding specificities of the corresponding PDZs. By contrast, PDZ extensions affecting the general fold, stability, dynamics of the PDZ, or participating in its general regulation, may more often impact indifferently the general binding behaviour of the PDZ to any of its targets, thereby not increasing binding specificity. Overall, these uncertainties strongly call for studies that will further focus on the impact of PDZ extensions on peptide binding specificities.

4.4. Influence of neighbouring domains on peptide binding to a PDZ

4.4.1. PDZ tandems

It is a well known property of PDZ domains that they often occur in multiple copies within one protein sequence. More and more evidence accumulates that neighbouring PDZ domains influence each other's structure and binding behaviour, especially those that are connected by very short linkers (reviewed in [5]). As structural data of such tandem PDZ domains indicate, they can be tightly packed and form one unit. The term supramodule was introduced to account for this property [5]. Two examples of such supramodules (PDZ45 of INADL and of GRIP1) were already mentioned in the previous section.

The PDZ12 tandem of DLG4 (PSD95) has been the subject of several publications. Based on the observation that PDZ1 and PDZ2 of DLG4 can bind to the same set of C-terminal peptides, it has been suggested that the two PDZs can bind simultaneously to the C-termini of homo- or heteromeric channel proteins of the postsynaptic density [99]. It has been shown that such synergistic binding leads to an increase in affinity and specificity as those ligands occurring in dimeric form are favoured as compared to monomeric ligands [99] (see review of Feng et al. [5] for more details). This property was successfully used to develop biomimetic divalent ligands that were much more efficient than monovalent ligands in disrupting the binding of DLG4 to its interaction partners [100,101]. Such a binding model would be favoured by an arrangement of PDZ1 and PDZ2 where the peptide binding pockets point to the same direction. Five very recent studies concentrated on the domain orientation of PDZ12 of DLG4 and revealed interesting findings. Using various techniques (NMR, crystallography, molecular dynamics simulations, single molecule FRET), very different conformations of the PDZ12 tandem were suggested ranging from a parallel alignment of the binding pockets to an antiparallel arrangement [99,100,102]. In addition, Wang et al. [103] provided data that

suggests that the tandem PDZ being restricted in its interdomain orientations in the ligand-free state, encounters a dramatic increase in flexibility when bound to a C-terminal peptide. Equivalent observations have been obtained for PDZ12 of DLG4 when bound to a divalent ligand [101]. Remarkably, McCann et al. [102] could further demonstrate that domain orientation and flexibility of the single PDZ tandem was comparable to those obtained in full length DLG4. It seems that, depending on the cellular context, parallel alignment of the binding pockets of PDZ12, which would favour multivalent ligand binding, is as likely as antiparallel alignment, which would instead enable recruitment of cytosolic proteins to membrane receptors [102]. More studies are needed to investigate this model. The fact that very different conclusions can be drawn from studies using different biophysical methods may indicate that methodologies remain of limited accuracy for the study of large and dynamic systems such as tandem PDZs.

4.4.2. Autoinhibition of PDZ domains

NHERF1 and X11 (amyloid beta A4 precursor protein-binding family A member 1, reviewed in [5]) have PDZ domains whose peptide binding is regulated via an autoinhibitory mechanism that involves PBMs at their own C-termini [104,105]. In an interesting NMR study, Bhattacharya et al. [76] investigated the molecular details of this autoinhibition for PDZ2 of NHERF1. NHERF1 has two PDZ domains and a C-terminal EB (Ezrin binding) domain that binds to Ezrin (involved in linking cytoskeletal structures to the plasma membrane) and that overlaps with a PBM at the C-terminus of NHERF1. Binding of Ezrin to the EB domain increases affinity of PDZ2 to the C-terminal peptide of CFTR (Cystic fibrosis transmembrane conductance regulator) by 24-fold [106,76]. NMR data indicates that the PBM at the C-terminus of NHERF1 can very weakly and transiently bind into the peptide binding pocket of PDZ2. When Ezrin binds to the EB domain, the whole EB domain including its PBM adopts a helical conformation and dissociates from PDZ2, making the peptide binding pocket fully accessible for binding to PBMs of other proteins [76].

4.4.3. Influence on PDZ–peptide binding from neighbouring domains of different type

In 2011, two articles were published within less than a month, reporting the crystal structure of the PDZ3-SH3-GK (PSG) module of ZO1 com-plexed to C-terminal peptides derived from either the cell adhesion proteins JAM-A [107] or the gap junction protein Connexin45 (Cx45) [68]. Both studies agree on the overall extended shape of the PSG module (PDZ3 does not contact the GK domain). An additional C-terminal α helix (hereafter called α3) of PDZ3 is located in the linker region between the PDZ3 and SH3 domain. Residues of the α2 and α3 helix as well as of the β2–β3 loop of PDZ3 interact with residues from the PDZ-SH3 linker and the SH3 domain. In both crystals, the peptides only partially insert into the binding pocket of PDZ3. The backbone of the peptides exhibit a shift towards the β2 strand, probably forced to this different conformation by the β2–β3 loop of the SH3 domain that is located in front of the peptide binding pocket (Fig. 4B) [107,68]. In both structures the PDZ3-SH3 linker inserts between the β2–β3 loop and the SH3 domain, making the loop inaccessible for interaction with peptide residues. Interestingly, no binding could be observed between JAM-A and the single PDZ3 domain [107] and Cx45 exhibited a 9-fold reduced binding affinity to the single PDZ3 in comparison to the affinity obtained for the PSG module [68]. Similar observations for CRIPT and Neuroligin peptides were independently obtained for the equivalent PSG module of DLG (Disks large) from *D. melanogaster* [108] (Table 2). This difference in affinity might mostly be due to a hydrophobic residue located at p-2 of the peptides (this residue is a Phe in JAM-A and a Val in Cx45) that inserts into a hydrophobic pocket formed by residues from both the α2 helix of PDZ3 and the β2–β3 loop of the SH3 domain of ZO1 [107,68].

The side chains of PDZ3 exhibit almost identical conformations in the two complexes except of R436 from the β2 strand. The aliphatic part of the side chain of R436 establishes hydrophobic contacts with Leu and Trp at p-1 in JAM-A and Cx45, respectively. However, in the PSG-Cx45 complex, the side chain of W_{-1} of Cx45, being bulkier than L_{-1} of JAM-A, displaces the side chain of R436 of PSG, which consequently occupies space that is used by S_{-3} in the PSG-JAM-A complex (Fig. 4B). This in turn displaces S_{-3} and more upstream residues of the Cx45 peptide further away from the PDZ domain as compared to the equivalent residues from the JAM-A peptide. This may be the main reason for the differences in backbone conformation observed for the peptides in the two crystals (Fig. 4B). Based on these observations, one would expect that the JAM-A peptide binds stronger to the PSG module than the Cx45 peptide. Surprisingly, the contrary is the case (Table 2). In both peptides, residues at p−3 and p−4 are very similar being either serine or threonine, while residues further upstream were not observed to significantly contribute to the binding. Hence, the difference in affinity can only be explained from sequence differences at the last three peptide positions being either VWI for Cx45 or FLV for JAM-A. It seems likely that residues W_{-1} (based on previous observations [109]) and Ile at p0 (preference of Leu or Ile over Val at p0 [19]) are the main contributors to the higher affinity of Cx45 to PDZ3. Based on these observations, we speculate that despite of binding with less affinity, JAM-A seems to bind with higher specificity to the PSG module than Cx45.

In summary, the structures of these two complexes revealed very interesting findings. To our knowledge, they are the first to provide atomic details for the direct influence that non-PDZ domains can have on the peptide binding of their PDZ neighbours. They also show that, whenever possible, investigations of small protein fragments such as single PDZs should be complemented by and compared to investigations of larger protein constructs (comprising multiple domains) or even full length proteins. They also serve as a nice example that shows how peptide residues can influence each other's binding to the PDZ domain. Such "cooperative" effects are excluded from most current PDZ–peptide interaction prediction models, which for complexity reasons assume independence of the peptide residues.

4.5. Influence on PDZ–peptide binding from distal domains

A fourth type of sequence context has not been discussed in this review, namely regions in protein sequences that are not in the neighbourhood of PDZs but still influence their peptide binding behaviour. We are convinced that numerous such examples exist, but our current knowledge on this subject is very limited due to a current lack of biophysical methods that allow studying larger protein fragments or even full length proteins at the molecular level. The works of McCann et al. [102], who used single molecule FRET to study the tandem PDZ12 in full length DLG4 as well as of Pan et al. [68] and Nomme et al. [107] who published the structure of the PSG module of ZO1, are promising steps towards bridging this gap.

5. Concluding remarks

This review has been focussed on the influence of sequence context on PDZ–peptide interactions given the vast amount of data available for the PDZ domain family. Other types of globular domain-linear motif interactions are much less studied and our knowledge on sequence context in these systems is sparse. However, we think that it is very likely that sequence context will have

similar importance on linear motif binding in other systems as has been demonstrated for the PDZ domain family. A few studies underpin this speculation (WW domain family [110], SH3 domain family [111]).

To which extent does sequence context influence specificity of PDZ–PBM interactions? An interaction between two proteins is specific when their mutual binding affinity is significantly higher than the affinities of their interactions with most other proteins. Therefore, in principle the specificity of a given PDZ–PBM interaction can only be assayed by comparing its binding affinity to the binding affinities of this particular PDZ domain towards a variety of other PBMs, and/or to the binding affinities of this particular PBM towards a variety of other PDZ domains. Most of the studies discussed here or elsewhere [5,33] show that sequence context influences binding affinity, yet they did not perform the comparative studies to address interaction specificity. There is a need for more studies including different interaction partners and protein sequences of various lengths [26] to better understand how sequence context influences specificity of domain–SLiM interactions in cell signalling processes.

Acknowledgements

We thank Bruno Kieffer and Yves Nomine for fruitful discussions on PDZ domain extensions. We especially thank Robert Weatheritt and Toby Gibson for very constructive comments on the manuscript. This work was supported by CNRS, University of Strasbourg and Association de Recherche contre le Cancer (ARC, grant 3171). KL was supported by the "Région Alsace", ARC, and the Collège Doctoral Européen de Strasbourg. SC was supported by grants from the Agence Nationale de la Recherche (ANR-MIME-2007, project EPI-HPV-3D) and the National Institute of Health (NIH, grant R01CA134737).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.febslet.2012.03.056.

References

[1] Diella, F., Haslam, N., Chica, C., Budd, A., Michael, S., Brown, N.P., Trave, G. and Gibson, T.J. (2008) Understanding eukaryotic linear motifs and their role in cell signaling and regulation. Front Biosci 13, 6580–6603.
[2] Davey, N.E., Roey, K.V., Weatheritt, R.J., Toedt, G., Uyar, B., Altenberg, B., Budd, A., Diella, F., Dinkel, H. and Gibson, T.J. (2012) Attributes of short linear motifs. Mol Biosyst 8 (1), 268–281.
[3] Ponting, C.P. (1997) Evidence for PDZ domains in bacteria, yeast, and plants. Protein Sci 6 (2), 464–468.
[4] Bilder, D. (2001) PDZ proteins and polarity: functions from the fly. Trends Genet 17 (9), 511–519.
[5] Feng, W. and Zhang, M. (2009) Organization and dynamics of PDZ-domain-related supramod-ules in the postsynaptic density. Nat Rev Neurosci 10 (2), 87–99.
[6] Reiners, J., Nagel-Wolfrum, K., Jürgens, K., Märker, T. and Wolfrum, U. (2006) Molecular basis of human Usher syndrome: deciphering the meshes of the Usher protein network provides insights into the pathomechanisms of the Usher disease. Exp Eye Res 83 (1), 97–119.
[7] Liu, W., Wen, W., Wei, Z., Yu, J., Ye, F., Liu, C.-H., Hardie, R.C. and Zhang, M. (2011) The INAD scaffold is a dynamic, redox-regulated modulator of signaling in the Drosophila eye. Cell 145 (7), 1088–1101.
[8] Iden, S. and Collard, J.G. (2008) Crosstalk between small GTPases and polarity proteins in cell polarization. Nat Rev Mol Cell Biol 9 (11), 846–859.
[9] Roh, M.H. and Margolis, B. (2003) Composition and function of PDZ protein complexes during cell polarization. Am J Physiol Renal Physiol 285 (3), F377–F387.
[10] Harris, B.Z. and Lim, W.A. (2001) Mechanism and role of PDZ domains in signaling complex assembly. J Cell Sci 114 (Pt 18), 3219–3231.
[11] C. Nourry, S.G.N. Grant, J.-P. Borg, PDZ domain proteins: plug and play!, Sci STKE 2003 (179) (2003) RE7. doi:10.1126/stke.2003.179.re7.
[12] Cabral, J.H.M., Petosa, C., Sutcliffe, M.J., Raza, S., Byron, O., Poy, F., Marfatia, S.M., Chishti, A.H. and Liddington, R.C. (1996) Crystal structure of a PDZ domain. Nature 382 (6592), 649–652.
[13] Hillier, B.J., Christopherson, K.S., Prehoda, K.E., Bredt, D.S. and Lim, W.A. (1999) Unexpected modes of PDZ domain scaffolding revealed by structure of nNOS-syntrophin complex. Science 284 (5415), 812–815.
[14] Lenfant, N., Polanowska, J., Bamps, S., Omi, S., Borg, J.-P. and Reboul, J. (2010) A genomewide study of PDZ-domain interactions in C. elegans reveals a high frequency of non-canonical binding. BMC Genomics 11, 671.
[15] Gallardo, R., Ivarsson, Y., Schymkowitz, J., Rousseau, F. and Zimmermann, P. (2010) Structural diversity of PDZ-lipid interactions. Chembiochem 11 (4), 456–467.
[16] Songyang, Z., Fanning, A.S., Fu, C., Xu, J., Marfatia, S.M., Chishti, A.H., Crompton, A., Chan, A.C., Anderson, J.M. and Cantley, L.C. (1997) Recognition of unique carboxyl-terminal motifs by distinct PDZ domains. Science 275 (5296), 73–77.
[17] Stricker, N.L., Christopherson, K.S., Yi, B.A., Schatz, P.J., Raab, R.W., Dawes, G., Bassett, D.E., Bredt, D.S. and Li, M. (1997) PDZ domain of neuronal nitric oxide synthase recognizes novel C-terminal peptide sequences. Nat Biotechnol 15 (4), 336342.
[18] Schleinkofer, K., Wiedemann, U., Otte, L., Wang, T., Krause, G., Oschkinat, H. and Wade, R.C. (2004) Comparative structural and energetic analysis of WW domain-peptide interactions. J Mol Biol 344 (3), 865–881.
[19] Tonikian, R., Zhang, Y., Sazinsky, S.L., Currell, B., Yeh, J.-H., Reva, B., Held, H.A., Appleton, A., Evangelista, M., Wu, Y., Xin, X., Chan, A.C., Seshagiri, S., Lasky, L.A., Sander, C., Boone, C., Bader, G.D. and Sidhu, S.S. (2008) A specificity map for the PDZ domain family. PLoS Biol 6 (9), e239.
[20] Linding, R., Jensen, L.J., Pasculescu, A., Olhovsky, M., Colwill, K., Bork, P., Yaffe, M.B. and Pawson, T. (2008) NetworKIN: a resource for exploring cellular phospho-rylation networks. Nucleic Acids Res 36 (Database issue), D695–D699.
[21] Tonikian, R., Xin, X., Toret, C.P., Gfeller, D., Landgraf, C., Panni, S., Paoluzi, S., Castagnoli, L., Currell, B., Seshagiri, S., Yu, H., Winsor, B., Vidal, M., Ger- stein, M.B., Bader, G.D., Volkmer, R., Cesareni, G., Drubin, D.G., Kim, P.M. and Sidhu, S.S. (2009) Boone, Bayesian modeling of the yeast SH3 domain interactome predicts spa-tiotemporal dynamics of endocytosis proteins. PLoS Biol 7 (10), e1000218.
[22] Panni, S., Montecchi-Palazzi, L., Kiemer, L., Cabibbo, A., Paoluzi, S., Santonico, E., Landgraf, C., Volkmer-Engert, R., Bachi, A., Castagnoli, L. and Cesareni, G. (2011) Combining peptide recognition specificity and context information for the prediction of the 14-3-3-mediated interactome in S. cerevisiae and H. sapiens. Proteomics 11 (1), 128–143.
[23] Stein, A. and Aloy, P. (2008) Contextual specificity in peptide-mediated protein interactions. PLoS One 3 (7), e2524.
[24] Chica, C., Diella, F. and Gibson, T.J. (2009) Evidence for the concerted evolution between short linear protein motifs and their flanking regions. PLoS One 4 (7), e6052.
[25] Gibson, T.J. (2009) Cell regulation: determined to signal discrete cooperation. Trends Biochem Sci 34 (10), 471–482.
[26] Luck, K., Fournane, S., Kieffer, B., Masson, M., Nominé, Y. and Travé, G. (2011) Putting into practice domain-linear motif interaction predictions for exploration of protein networks. PLoS One 6 (11), e25376.
[27] Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N. and Bourne, P.E. (2000) The protein data bank. Nucleic Acids Res 28 (1), 235–242.
[28] Letunic, I., Doerks, T. and Bork, P. (2012) SMART 7: recent updates to the protein domain annotation resource. Nucleic Acids Res 40 (Database issue), D302–D305.
[29] Spaller, M.R. and globally, Act (2006) think locally: systems biology addresses the PDZ domain. ACS Chem Biol 1 (4), 207–210.
[30] Bhattacharyya, R.P., Remnyi, A., Yeh, B.J. and Lim, W.A. (2006) Domains, motifs, and scaffolds: the role of modular interactions in the evolution and wiring of cell signaling circuits. Annu Rev Biochem 75, 655–680.
[31] Javier, R.T. and Rice, A.P. (2011) Emerging theme: cellular PDZ proteins as common targets of pathogenic viruses. J Virol 85 (22), 11544–11556.
[32] Giallourakis, C., Cao, Z., Green, T., Wachtel, H., Xie, X., Lopez-Illasaca, M., Daly, M., Rioux, J. and Xavier, R. (2006) A molecular-properties-based approach to understanding PDZ domain proteins and PDZ ligands. Genome Res 16 (8), 1056–1072.
[33] Wang, C.K., Pan, L., Chen, J. and Zhang, M. (2010) Extensions of PDZ domains as important structural and functional elements. Protein Cell 1 (8), 737–751.
[34] Velthuis, A.J.W.T., Sakalis, P.A., Fowler, D.A. and Bagowski, C.P. (2011) Genome-wide analysis of PDZ domain binding reveals inherent functional overlap within the PDZ interaction network. PLoS One 6 (1), e16047.
[35] Chen, J.R., Chang, B.H., Allen, J.E., Stiffler, M.A. and MacBeath, G. (2008) Predicting PDZ domain-peptide interactions from primary sequences. Nat Biotechnol 26 (9), 1041–1045.
[36] Hui, S. and Bader, G.D. (2010) Proteome scanning to predict PDZ domain interactions using support vector machines. BMC Bioinformatics 11 (1), 507.
[37] Smith, C.A. and Kortemme, T. (2010) Structure-based prediction of the peptide sequence space recognized by natural and synthetic PDZ domains. J Mol Biol 402 (2), 460–474.
[38] Gerek, Z.N. and Ozkan, S.B. (2010) A flexible docking scheme to explore the binding selectivity of PDZ domains. Protein Sci 19 (5), 914–928.

[39] Karthikeyan, S., Leung, T. and Ladias, J.A. (2001) Structural basis of the Na+/H+ exchanger regulatory factor PDZ1 interaction with the carboxyl-terminal region of the cystic fibrosis transmembrane conductance regulator. J Biol Chem 276 (23), 19683–19686.
[40] Sheng, M. and Sala, C. (2001) PDZ domains and the organization of supramolecular complexes. Annu Rev Neurosci 24, 1–29.
[41] Lim, I.A., Hall, D.D. and Hell, J.W. (2002) Selectivity and promiscuity of the first and second PDZ domains of PSD-95 and synapse-associated protein 102. J Biol Chem 277 (24), 21697–21711.
[42] Préhaud, C., Wolff, N., Terrien, E., Lafage, M., Mégret, F., Babault, N., Cordier, F., Tan, G.S., Maitrepierre, E., Ménager, P., Chopy, D., Hoos, S., England, P., Delepierre, M., Schnell, M.J., Buc, H. and Lafon, M. (2010) Attenuation of rabies virulence: takeover by the cytoplasmic domain of its envelope protein. Sci Signal 3 (105), ra5.
[43] Babault, N., Cordier, F., Lafage, M., Cockburn, J., Haouz, A., Préhaud, C., Rey, F.A., Delepierre, M., Buc, H., Lafon, M. and Wolff, N. (2011) Peptides targeting the PDZ domain of PTPN4 are efficient inducers of glioblastoma cell death. Structure 19 (10), 1518–1524.
[44] Zhang, J., Yan, X., Shi, C., Yang, X., Guo, Y., Tian, C., Long, J. and Shen, Y. (2008) Structural basis of beta-catenin recognition by Tax-interacting protein-1. J Mol Biol 384 (1), 255–263.
[45] Banerjee, M., Huang, C., Marquez, J. and Mohanty, S. (2008) Probing the structure and function of human glutaminase-interacting protein: a possible target for drug design. Biochemistry 47 (35), 9208–9219.
[46] Yan, X., Zhou, H., Zhang, J., Shi, C., Xie, X., Wu, Y., Tian, C., Shen, Y. and Long, J. (2009) Molecular mechanism of inward rectifier potassium channel 2.3 regulation by tax-interacting protein-1. J Mol Biol 392 (4), 967–976.
[47] Im, Y.J., Kang, G.B., Lee, J.H., Park, K.R., Song, H.E., Kim, E., Song, W.K., Park, D. and Eom, S.H. (2010) Structural basis for asymmetric association of the betaPIX coiled coil and shank PDZ. J Mol Biol 397 (2), 457–466.
[48] Balana, B., Maslennikov, I., Kwiatkowski, W., Stern, K.M., Bahima, L., Choe, S. and Slesinger, P.A. (2011) Mechanism underlying selective regulation of G protein-gated inwardly rectifying potassium channels by the psychostimulant-sensitive sorting nexin 27. Proc Natl Acad Sci U S A 108 (14), 5831–5836.
[49] Mostarda, S., Gfeller, D. and Rao, F. (2012) Beyond the Binding Site: The Role of the ?2 - ?3 Loop and Extra-Domain Structures in PDZ Domains. PLoS Comput Biol 8 (3), e1002429.
[50] Zhang, Y., Dasgupta, J., Ma, R.Z., Banks, L., Thomas, M. and Chen, X.S. (2007) Structures of a human papillomavirus (HPV) E6 polypeptide bound to MAGUK proteins: mechanisms of targeting tumor suppressors by a high-risk HPV oncoprotein. J Virol 81 (7), 3618–3626.
[51] Fournane, S., Charbonnier, S., Chapelle, A., Kieffer, B., Orfanoudakis, G., Trav, G., Masson, M. and Nomin, Y. (2011) Surface plasmon resonance analysis of the binding of high-risk mucosal HPV E6 oncoproteins to the PDZ1 domain of the tight junction protein MAGI-1. J Mol Recognit 24 (4), 511–523.
[52] Charbonnier, S., Nominé, Y., Ramirez, J., Luck, K., Chapelle, A., Stote, R.H., Travé, G., Kieffer, B. and Atkinson, R.A. (2011) The structural and dynamic response of MAGI-1 PDZ1 with noncanonical domain boundaries to the binding of human pa-pillomavirus E6. J Mol Biol 406 (5), 745–763.
[53] Zhang, Z., Li, H., Chen, L., Lu, X., Zhang, J., Xu, P., Lin, K. and Wu, G. (2011) Molecular basis for the recognition of adenomatous polyposis coli by the Discs Large 1 protein. PLoS One 6 (8), e23507.
[54] Liu, Y., Henry, G.D., Hegde, R.S. and Baleja, J.D. (2007) Solution structure of the hDlg/SAP97 PDZ2 domain and its mechanism of interaction with HPV- 18 papillomavirus E6 protein. Biochemistry 46 (38), 10864–10874.
[55] Wang, L., Piserchio, A. and Mierke, D.F. (2005) Structural characterization of the inter-molecular interactions of synapse-associated protein-97 with the NR2B subunit of N-methyl-D-aspartate receptors. J Biol Chem 280 (29), 26992–26996.
[56] Madsen, K.L., Beuming, T., Niv, M.Y., Chang, C.-W., Dev, K.K., We- instein, H. and Gether, U. (2005) Molecular determinants for the complex binding specificity of the PDZ domain in PICK1. J Biol Chem 280 (21), 20539–20548.
[57] Tyler, R.C., Peterson, F.C. and Volkman, B.F. (2010) Distal interactions within the par3-VE- cadherin complex. Biochemistry 49 (5), 951–957.
[58] Feng, W., Wu, H., Chan, L.-N. and Zhang, M. (2008) Par-3-mediated junctional localization of the lipid phosphatase PTEN is required for cell polarity establishment. J Biol Chem 283 (34), 23440–23449.
[59] Skelton, N.J., Koehler, M.F.T., Zobel, K., Wong, W.L., Yeh, S., Pisabarro, M.T., Yin, J.P., Lasky, L.A. and Sidhu, S.S. (2003) Origins of PDZ domain ligand specificity. Structure determination and mutagenesis of the Erbin PDZ domain. J Biol Chem 278 (9), 7645–7654.
[60] Appleton, B.A., Zhang, Y., Wu, P., Yin, J.P., Hunziker, W., Skelton, N.J., Sidhu, S.S. and Wiesmann, C. (2006) Comparative structural analysis of the Erbin PDZ domain and the first PDZ domain of ZO-1. Insights into determinants of PDZ domain specificity. J Biol Chem 281 (31), 22312–22320.
[61] Birrane, G., Chung, J. and Ladias, J.A.A. (2003) Novel mode of ligand recognition by the Erbin PDZ domain. J Biol Chem 278 (3), 1399–1402.
[62] Kozlov, G., Gehring, K. and Ekiel, I. (2000) Solution structure of the PDZ2 domain from human phosphatase hPTP1E and its interactions with C-terminal peptides from the Fas receptor. Biochemistry 39 (10), 2572–2580.
[63] Kachel, N., Erdmann, K.S., Kremer, W., Wolff, P., Gronwald, W., Heumann, R. and Kalbitzer, H.R. (2003) Structure determination and ligand interactions of the PDZ2b domain of PTP-Bas (hPTP1E): splicing-induced modulation of ligand specificity. J Mol Biol 334 (1), 143–155.
[64] Walma, T., Aelen, J., Nabuurs, S.B., Oostendorp, M., van den Berk, L., Hendriks, W. and Vuister, G.W. (2004) A closed binding pocket and global destabilization modify the binding properties of an alternatively spliced form of the second PDZ domain of PTP-BL. Structure 12 (1), 11–20.
[65] Sierralta, J. and Mendoza, C. (2004) PDZ-containing proteins: alternative splicing as a source of functional diversity. Brain Res Brain Res Rev 47 (1-3), 105–115.
[66] Yan, J., Pan, L., Chen, X., Wu, L. and Zhang, M. (2010) The structure of the harmonin/sans complex reveals an unexpected interaction mode of the two Usher syndrome proteins. Proc Natl Acad Sci U S A 107 (9), 4040–4045.
[67] Pan, L., Yan, J., Wu, L. and Zhang, M. (2009) Assembling stable hair cell tip link complex via multidentate interactions between harmonin and cadherin 23. Proc Natl Acad Sci U S A 106 (14), 5575–5580.
[68] Pan, L., Chen, J., Yu, J., Yu, H. and Zhang, M. (2011) The structure of the PDZ3-SH3-GuK tandem of ZO-1 protein suggests a supramodular organization of the membrane-associated guanylate kinase (MAGUK) family scaffold protein core. J Biol Chem 286 (46), 40069–40074.
[69] Chen, J., Pan, L., Wei, Z., Zhao, Y. and Zhang, M. (2008) Domain-swapped dimerization of ZO-1 PDZ2 generates specific and regulatory connexin43-binding sites. EMBO J 27 (15), 2113–2123.
[70] Wu, J., Yang, Y., Zhang, J., Ji, P., Du, W., Jiang, P., Xie, D., Huang, H., Wu, M., Zhang, G., Wu, J. and Shi, Y. (2007) Domain-swapped dimerization of the second PDZ domain of ZO2 may provide a structural basis for the polymerization of claudins. J Biol Chem 282 (49), 35988–35999.
[71] Niraula, T.N., Yoneyama, M., Koshiba, S., Inoue, M., Kigawa, T. and Yokoyama, S. (2008) Riken Structural Genomics/Proteomics Initiative. Solution structure of PDZ domain of Rho GTPase Activating Protein 21.
[72] Akiva, E., Friedlander, G., Itzhaki, Z. and Margalit, H. (2012) A dynamic view of domain-motif interactions. PLoS Comput Biol 8 (1), e1002341.
[73] Punta, M., Coggill, P.C., Eberhardt, R.Y., Mistry, J., Tate, J., Boursnell, C., Pang, N., Forslund, K., Ceric, G., Clements, J., Heger, A., Holm, L., Sonnhammer, E.L.L., Eddy, S.R., Bateman, A. and Finn, R.D. (2012) The Pfam protein families database. Nucleic Acids Res 40 (Database issue), D290–D301.
[74] Zhang, Y., Appleton, B.A., Wu, P., Wiesmann, C. and Sidhu, S.S. (2007) Structural and functional analysis of the ligand specificity of the HtrA2/Omi PDZ domain. Protein Sci 16 (8), 1738–1750.
[75] Korotkov, K.V., Krumm, B., Bagdasarian, M. and Hol, W.G.J. (2006) Structural and functional studies of EpsC, a crucial component of the type 2 secretion system from Vibrio cholerae. J Mol Biol 363 (2), 311–321.
[76] Bhattacharya, S., Dai, Z., Li, J., Baxter, S., Callaway, D.J.E., Cowburn, D. and Bu, Z. (2010) A conformational switch in the scaffolding protein NHERF1 controls au- toinhibition and complex formation. J Biol Chem 285 (13), 9981–9994.
[77] Charbonnier, S., Stier, G., Orfanoudakis, G., Kieffer, B., Atkinson, R.A. and Trav, G. (2008) Defining the minimal interacting regions of the tight junction protein MAGI-1 and HPV16 E6 oncoprotein for solution structure studies. Protein Expr Purif 60 (1), 64–73.
[78] Petit, C.M., Zhang, J., Sapienza, P.J., Fuentes, E.J. and Lee, A.L. (2009) Hidden dynamic allostery in a PDZ domain. Proc Natl Acad Sci U S A 106 (43), 18249–18254.
[79] Ballif, B.A., Carey, G.R., Sunyaev, S.R. and Gygi, S.P. (2008) Large-scale identification and evolution indexing of tyrosine phosphorylation sites from murine brain. J Proteome Res 7 (1), 311–318.
[80] Zhang, J., Petit, C.M., King, D.S. and Lee, A.L. (2011) Phosphorylation of a PDZ domain extension modulates binding affinity and interdomain interactions in postsynaptic density-95 (PSD-95) protein, a membrane-associated guanylate kinase (MAGUK). J Biol Chem 286 (48), 41776–41785.
[81] Consortium, U. (2010) The Universal Protein Resource (UniProt) in 2010. Nucleic Acids Res 38 (Database issue), D142–D148.
[82] Dinkel, H., Chica, C., Via, A., Gould, C.M., Jensen, L.J., Gibson, T.J. and Diella, F. (2011) Phospho.ELM: a database of phosphorylation sites-update 2011. Nucleic Acids Res 39 (Database issue), D261–D267.
[83] Dephoure, N., Zhou, C., Villn, J., Beausoleil, S.A., Bakalarski, C.E., Elledge, S.J. and Gygi, S.P. (2008) A quantitative atlas of mitotic phosphorylation. Proc Natl Acad Sci U S A 105 (31), 10762–10767.
[84] Dinkel, H., Michael, S., Weatheritt, R.J., Davey, N.E., Roey, K.V., Altenberg, B., Toedt, G., Uyar, B., Seiler, M., Budd, A., Jdicke, L., Dammert, M.A., Schroeter, C., Hammer, M., Schmidt, T., Jehl, P., McGuigan, C., Dymecka, M., Chica, C., Luck, K., Via, A., Chatr-Aryamontri, A., Haslam, N., Grebnev, G., Edwards, R.J., Steinmetz, M.O., Meiselbach, H., Diella, F. and Gibson, T.J. (2012) ELM-the database of eu-karyotic linear motifs. Nucleic Acids Res 40 (Database issue), D242–D251.
[85] Round, J.L., Tomassian, T., Zhang, M., Patel, V., Schoenberger, S.P. and Miceli, M.C. (2005) Dlgh1 coordinates actin polymerization, synaptic T cell receptor and lipid raft aggregation, and effector function in T cells. J Exp Med 201 (3), 419–430.
[86] Garrard, S.M., Capaldo, C.T., Gao, L., Rosen, M.K., Macara, I.G. and Tomchick, D.R. (2003) Structure of Cdc42 in a complex with the GTPase-binding domain of the cell polarity protein, Par6. EMBO J 22 (5), 1125–1133.
[87] Peterson, F.C., Penkert, R.R., Volkman, B.F. and Prehoda, K.E. (2004) Cdc42 regulates the Par-6 PDZ domain through an allosteric CRIB-PDZ transition. Mol Cell 13 (5), 665–676.
[88] Whitney, D.S., Peterson, F.C. and Volkman, B.F. (2011) A conformational switch in the CRIB-PDZ module of Par-6. Structure 19 (11), 1711–1722.
[89] Mishra, P., Socolich, M., Wall, M.A., Graves, J., Wang, Z. and Ranganathan, R. (2007) Dynamic scaffolding in a G protein-coupled signaling system. Cell 131 (1), 80–92.

[90] Feng, W., Shi, Y., Li, M. and Zhang, M. (2003) Tandem PDZ repeats in glutamate receptor-interacting proteins have a novel mode of PDZ domain-mediated target binding. Nat Struct Biol 10 (11), 972–978.
[91] Dong, H., O'Brien, R.J., Fung, E.T., Lanahan, A.A., Worley, P.F. and Huganir, R.L. (1997) GRIP: a synaptic PDZ domain-containing protein that interacts with AMPA receptors. Nature 386 (6622), 279–284.
[92] K. Inoue, C. Kurosaki, F. Hayashi, S. Yokoyama, Riken Structural Ge-nomics/Proteomics Initiative. Solution structure of the first PDZ domain of InaD-like protein (2006). .
[93] E. Papagrigoriou, C. Gileadi, C. Phillips, J. Elkins, C. Johansson, E. Salah, P. Savitsky, G. Berridge, D. Doyle, Structural Genomics Consortium. The crystal structure of the 1st PDZ domain of MPDZ. (2006).
[94] K. Yamada, N. Nameki, K. Saito, S. Koshiba, M. Inoue, T. Kigawa, S. Yokoyama, Riken Structural Genomics/Proteomics Initiative. Solution structure of the third PDZ domain of mouse harmonin (2004). .
[95] X.-R. Qin, F. Hayashi, S. Yokoyama, Riken Structural Genomics/Proteomics Initiative. Solution structure of the PDZ domain of Pals1 protein. (2005). .
[96] K. Inoue, T. Nagashima, K. Izumi, F. Hayashi, S. Yokoyama, Riken Structural Ge-nomics/Proteomics Initiative. Solution structure of the 7th PDZ domain of InaD-like protein. (2006).
[97] C. Zhao, T. Kigawa, N. Tochio, S. Koshiba, M. Inoue, S. Yokoyama, Riken Structural Genomics/Proteomics Initiative. Solution structure of the first PDZ domain of Human Atrophin-1 Interacting Protein 1 (KIAA0705 protein). (2003).
[98] W. Ohashi, H. Hirota, T. Yamazaki, Y. Muto, S. Yokoyama, Riken Structural Genomics/Proteomics Initiative. Solution structure of RSGI RUH-003, a PDZ domain of hypothetical KIAA0559 protein from human cDNA. (2004).
[99] Long, J.-F., Tochio, H., Wang, P., Fan, J.-S., Sala, C., Niethammer, M., Sheng, M. and Zhang, M. (2003) Supramodular structure and synergistic target binding of the N-terminal tandem PDZ domains of PSD-95. J Mol Biol 327 (1), 203–214.
[100] Sainlos, M., Tigaret, C., Poujol, C., Olivier, N.B., Bard, L., Breillat, C., Thi- olon, K., Choquet, D. and Imperiali, B. (2011) Biomimetic divalent ligands for the acute disruption of synaptic AMPAR stabilization. Nat Chem Biol 7 (2), 81–91.
[101] Bach, A., Clausen, B.H., Møller, M., Vestergaard, B., Chi, C.N., Round, A., Sørensen, P.L., Nissen, K.B., Kastrup, J.S., Gajhede, M., Jemth, P., Kris-tensen, A.S., Lundström, P., Lambertsen, K.L. and Strømgaard, K. (2012) A high-affinity, dimeric inhibitor of PSD-95 bivalently interacts with PDZ1-2 and protects against is- chemic brain damage. Proc Natl Acad Sci U S A 109 (9), 3317–3322.
[102] McCann, J.J., Zheng, L., Chiantia, S. and Bowen, M.E. (2011) Domain orientation in the N-Terminal PDZ tandem from PSD-95 is maintained in the full-length protein. Structure 19 (6), 810–820.
[103] Wang, W., Weng, J., Zhang, X., Liu, M. and Zhang, M. (2009) Creating conformational entropy by increasing interdomain mobility in ligand binding regulation: a revisit to N- terminal tandem PDZ domains of PSD-95. J Am Chem Soc 131 (2), 787–796.
[104] Li, J., Poulikakos, P.I., Dai, Z., Testa, J.R., Callaway, D.J.E. and Bu, Z. (2007) Protein kinase C phosphorylation disrupts Na+/H+ exchanger regulatory factor 1 autoinhibition and promotes cystic fibrosis transmembrane conductance regulator macromolecular assembly. J Biol Chem 282 (37), 27086–27099.
[105] Long, J.-F., Feng, W., Wang, R., Chan, L.-N., Ip, F.C.F., Xia, J., Ip, N.Y. and Zhang, M. (2005) Autoinhibition of X11/Mint scaffold proteins revealed by the closed conformation of the PDZ tandem. Nat Struct Mol Biol 12 (8), 722–728.
[106] Li, J., Dai, Z., Jana, D., Callaway, D.J.E. and Bu, Z. (2005) Ezrin controls the macromolecu-lar complexes formed between an adapter protein Na+/H+ exchanger regulatory factor and the cystic fibrosis transmembrane conductance regulator. J Biol Chem 280 (45), 37634–37643.
[107] Nomme, J., Fanning, A.S., Caffrey, M., Lye, M.F., Anderson, J.M. and Lavie, A. (2011) The Src Homology 3 domain is required for junctional adhesion molecule binding to the third PDZ domain of the scaffolding protein ZO-1. J Biol Chem 286 (50), 43352–43360.
[108] Qian, Y. and Prehoda, K.E. (2006) Interdomain interactions in the tumor suppressor discs large regulate binding to the synaptic protein gukholder. J Biol Chem 281 (47), 35757–35763.
[109] Luck, K. and Travé, G. (2011) Phage display can select over-hydrophobic sequences that may impair prediction of natural domain-peptide interactions. Bioinformatics 27 (7), 899–902.
[110] Fidan, Z., Younis, A., Schmieder, P. and Volkmer, R. (2011) Chemical synthesis of the third WW domain of TCERG 1 by native chemical ligation. J Pept Sci 17 (9), 644–649.
[111] Bauer, F., Schweimer, K., Meiselbach, H., Hoffmann, S., Rosch, P. and Sticht, H. (2005) Structural characterization of Lyn-SH3 domain in complex with a herpesviral protein reveals an extended recognition motif that enhances binding affinity. Protein Sci 14 (10), 2487–2498.
[112] Chen, Q., Niu, X., Xu, Y., Wu, J. and Shi, Y. (2007) Solution structure and backbone dynamics of the AF-6 PDZ domain/Bcr peptide complex. Protein Sci 16 (6), 10531062.
[113] W.L. DeLano, The PyMOL molecular graphics system (2002). URL http://www.pymol.org.
[114] Durney, M.A., Birrane, G., Anklin, C., Soni, A. and Ladias, J.A.A. (2009) Solution structure of the human Tax-interacting protein-1. J Biomol NMR 45 (3), 329–334.
[115] Im, Y.J., Lee, J.H., Park, S.H., Park, S.J., Rho, S.-H., Kang, G.B., Kim, E. and Eom, S.H. (2003) Crystal structure of the Shank PDZ-ligand complex reveals a class I PDZ interaction and a novel PDZ-PDZ dimerization. J Biol Chem 278 (48), 48099–48104.
[116] Katoh, K. and Toh, H. (2007) PartTree: an algorithm to build an approximate tree from a large number of unaligned sequences. Bioinformatics 23 (3), 372–374.
[117] Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M. and Barton, G.J. (2009) Jalview version 2-a multiple sequence alignment editor and analysis workbench. Bioinformatics 25 (9), 1189–1191.