The structure of a tandem pair of spectrin repeats of plectin reveals a modular organization of the plakin domain.

Arnoud Sonnenberg [1], Ana M Rojas[2] and José M de Pereda [3*]

[1]Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Nethe rlands.

[2]Structural Computational Biology Group, Centro Nacional de Investigaciones Oncológicas, Melchor Fernández Almagro 3, E-28029 Madrid, Spain.

[3]Instituto de Biología Molecular y Celular del Cancer, Consejo Superior de Investigaciones Científicas - University of Salamanca, Campus Unamuno, E-37007 Salamanca, Spain.

* Corresponding author

E-mail address of the corresponding author: pereda@usal.es

Running title: Structure of the plakin domain of plectin

Summary

Plectin is a large and versatile cytoskeletal linker and member of the plakin protein family. Plakins share a conserved region called the plakin domain located near their N-terminus. We have determined the crystal structure of an N-terminal fragment of the plakin domain of plectin to 2.05Å resolution. This region is adjacent to the actin binding domain and is required for efficient binding to the integrin α6β4 in hemidesmosomes. The structure is formed by two spectrin repeats connected by an α-helix that spans these two repeats. While the first repeat is very similar to other known structures, the second repeat is structurally different with a hydrophobic core, narrower than that in canonical spectrin repeats. Sequence analysis of the plakin domain revealed the presence of up to nine consecutive spectrin repeats organized in an array of tandem modules, and a Src-homology 3 domain inserted in the central spectrin repeat. The structure of the plakin domain is reminiscent of the modular organization of members of the spectrin family. The architecture of the plakin domain suggests that it forms an elongated and flexible structure, and provides a novel molecular explanation for the contribution of plectin and other plakins to the elasticity and stability of tissues subjected to mechanical stress, such as the skin and striated muscle.

Keywords: cytoskeleton; hemidesmosome; epithelium; X-ray crystallography

Abbreviations used: ABD, actin binding domain; FnIII, fibronectin type III; BPAG1, bullous pemphigoid antigen 1; MACF, microtubule actin crosslinking factor; rmsd, root mean square deviation; SH3, Src-homology 3; SR, spectrin repeat.

Introduction

Plectin is a large (~500 kDa) and widely expressed cytoskeletal linker[1] that belongs to the plakin protein family[2]. As with other plakins, plectin interconnects filament networks and tethers the cytoskeleton to membrane associated structures involved in cell adhesion. In epithelial tissues, plectin is mostly localized at the basal cell membrane[3] where it is a component of hemidesmosomes[4], which are specialized structures that mediate anchorage of epithelia to the underlying basement membrane by linking the extracellular matrix to the intermediate filament system[5]. In hemidesmosomes, plectin provides a direct link between the integrin $\alpha 6\beta 4$ (a laminin receptor) and the cytokeratin network[6].

Plectin has a multidomain structure that is well suited for its crosslinking functions. It comprises N- and C-terminal domains that contain multiple protein-protein interaction sites, separated by an elongated central rod domain that is predicted to mediate self association via coiled-coil interactions. The N-terminal region contains an actin binding domain (ABD), which is composed of two calponin homology domains, similar to those present in proteins of the spectrin family. In hemidesmosomes, the ABD of plectin binds to the first pair of fibronectin type III (FnIII) domains of the integrin $\beta 4$ cytoplasmic domain[7]. Adjacent to the ABD there is a region of about 1000 residues (the plakin domain) that is conserved among the members of the plakin family. The N-terminal region of the plakin domain of plectin binds to regions of integrin $\beta 4$ downstream of the second FnIII domain[8]. The plakin domain also harbors a binding site for the hemidesmosomal transmembrane protein BP180 (also known as type XVII collagen)[9]. The C-terminal region of plectin contains six plakin repeat domains and harbors binding sites for intermediate filaments[10; 11].

Mutations in the plectin gene result in epidermolysis bullosa simplex (frequently associated with muscular dystrophy), which is a severe skin blistering disorder characterized by fragility at the level of hemidesmosomes[12]. Inactivation of the plectin gene in mice produces a skin blistering which is postnatally lethal[13]. The alterations caused by defects in the plectin gene in humans and mice illustrate the critical role of plectin in maintaining the stability of tissues exposed to strong mechanical stress such as the skin and muscle.

The plakin domain is predicted to have a modular structure made up of α-helices arranged in anti-parallel bundles[14]. Thus, the plakin domain is likely to be flexible, its properties being determined by the structure of each module and by the structure of the linkers that connect these modules. Indeed, electron microscopy of purified plectin molecules suggests that the protein is flexible because the central region adopts many different conformations. This flexibility may facilitate a dynamic response of the protein to mechanical stress. The recent crystal structure of a protease resistant fragment of the plakin domain of BPAG1e was an important breakthrough in our understanding of the structural properties of the plakin domain. It has revealed the presence of a tandem pair of spectrin repeats, while a second pair of spectrin repeats and a Src-homology 3 (SH3) domain have been identified by sequence comparison[15]. Nevertheless, several essential questions remain unanswered. Plectin and other plakins that contain an ABD, such as MACF and the BPAG1 isoforms a, b, and n, have an N-terminal extension of the plakin domain that is contiguous to the ABD and is not present in other plakins. In plectin, this region acts in a synergistic manner with the ABD in binding to β4, and is required for the correct localization of plectin in hemidesmosomes. However, the structure of it is not known. Furthermore, in the current structural model of the plakin domain, some regions have not been assigned to specific

domains. It is uncertain if the plakin domain contains additional modules and if the linkers that connect adjacent modules are disordered and flexible or ordered and inflexible.

Here, we present the crystal structure of the N-terminal fragment of the plakin domain of plectin adjacent to the ABD. This structure is formed by a tandem pair of spectrin repeats connected by a α-helical linker. Moreover, we have identified nine spectrin repeats in the plakin domain of plectin and provide a model of the structure of the plakin domain that accounts for the complete length of this domain. Our results have clear implications for the basis of the elastic properties of the plakin domain and its contribution to tissue integrity.

Results and discussion

Overall structure

We have crystallized the N-terminal region of the plakin domain of human plectin, i.e. residues 300-530 (numbering corresponds to the plectin 1C isoform). In order to improve the crystal quality the two Cys residues, 420 and 435, located in this fragment were replaced by Ala. The mutant protein showed a reduced nucleation rate resulting in larger and better diffracting crystals than the wild type protein. Crystals of the wild type and mutant proteins were isomorphic. The crystal structure was determined using multiple wavelength anomalous diffraction (MAD) methods from a selenomethionine derivative of the C420A/C435A double mutant (Table 1). The structure of the mutant protein was refined against data from a native crystal to 2.05 Å resolution to a final Rwork of 21.0% and an Rfree of 26.1%. The asymmetric unit contains one molecule and the final model includes residues 303-331 and 333-520, 52 solvent molecules, and one molecule of 1,2-propanediol, for which ordered electron density was observed. The model has excellent geometry with main chain torsion angles of all non glycine residues lying in the core regions of the Ramachandran plot as defined in MOLEMAN2[16]. The wild type structure

was refined against data to 2.3 Å resolution to a final Rwork of 21.6% and an Rfree of 26.9%; and it contains the same residues as the mutant structure, but only 16 solvent molecules. The wild type and mutant structures are almost identical with the exception of the mutated residues. The root mean square deviation (rmsd) for all Cα atoms between both structures is 0.21 Å, which is similar to the value of the Cruickshank diffraction precision index[17] for these structures (0.21 Å and 0.16 Å for the wild type and mutant proteins respectively). Thus, we refer to the higher resolution structure of the mutant protein unless otherwise indicated.

The structure is formed by two coiled coils, each of them built up of three α-helices arranged in a left handed supercoil or helical bundle with up-down-up topology. The molecule has an elongated structure with a length of approximately 117 Å and a width of about 27 Å (Figure 1). The first bundle extends from Glu303 to Glu419 and contains helices A, B, and C, while the second bundle expands from Cys420 to Lys520 and contains helices A', B', and C'. The sequences of all helices show the heptad pattern characteristic of left-handed coiled coils, where positions *a* and *d* are preferentially occupied by apolar residues that mediate the inter-helical packing and form a hydrophobic core[18]. Within the group of three-helix left-handed coiled coils the plectin bundles most resemble the spectrin repeat fold (see below). Therefore we refer to the bundles in our structure as the first (SR1) and second (SR2) spectrin repeats of plectin.

Despite sharing an overall common fold, noticeable differences between SR1 and SR2 are observed as expected from the low sequence conservation (9 identical residues). The position of helices A and C in SR1 mostly match helices A' and C' in SR2 despite differences in length, but there are significant deviations between helices B and B' (Figure 2a). Pro357 causes a discontinuity or kink in helix B, commonly present in other spectrin repeats, but not observed in helix B'. As a result B' runs as a single straight helix that

packs against helices A' and C' closer than helix B does against helices A and C. The different packing of the helices in SR1 and SR2 is favored by the nature of the hydrophobic residues mediating the coiled-coil interaction. The core of SR1 contains bulky aromatic residues such as Trp321, Phe329, Phe348, Phe351, and Trp397 and it is wider than the hydrophobic core of SR2, that in the equivalent positions accommodates residues Gln438, Leu446, Ala462, Val465, and Leu508 with smaller side chains (Figure 3).

Helix A' is about two turns longer at its C-terminus than helix A, while the helix B' is about two and a half turns shorter at its N-terminus than helix B. As a consequence loops AB and A'B' run in almost opposite directions along the long axis of the molecule. There are additional differences between the BC and B'C' loops. Helix B has a second kink at Leu373 that makes the last helical turn, not present in helix B', capping one end of the helical bundle. Pro385 to Tyr388 in the BC loop form a type II β-turn that creates a bulge on the side of the repeat necessary to accommodate the additional residues not present in the second repeat.

Comparison with other spectrin repeats

We used the DALI server[19] to search the Protein Data Bank for structural homologues of the plectin repeats. The structure found to be most similar to the SR1 of plectin is the third spectrin repeat of the α-actinin rod[20] (PDB code 1QUU) with a Z score of 14.0, rmsd of 2.0 Å for 114 Cα atoms, and a sequence identity of 18%. Furthermore, spectrin repeat 17 from brain α-spectrin[21] (PDB code 1CUN), repeat 9 of erythroid β-spectrin[22] (PDB code 1S35), repeat 1 of erythroid α-spectrin[23] (PDB code 1OWA) and repeat 1 of BPAG1[15] (PDB code 2IAK), gave similar DALI Z scores which are significantly higher than those of other helical bundles.

The DALI search suggested that the SR1 of plectin is structurally closer to canonical repeats from proteins of the spectrin family, such as α-actinin, than to the repeats of the plakin domain of BPAG1. In order to better understand the structural similarities between the plectin SR1 and other spectrin repeats, pairwise structural superpositions were done using a common set of 80 Cα equivalent positions distributed over the three α-helices of the domain. The structures of the repeats 1 and 3 of α-actinin[24] (PDB entry 1HCI) and repeats 15 and 17 of α-spectrin[21; 25] (PDB codes 1U5P and 1CUN) had the lowest rmsd (1.27 Å, 0.94 Å, 1.08 Å, and 0.93 Å respectively) with respect to the plectin SR1 (figures 2b, 3). In contrast the first and second repeats of BPAG1 had a rmsd of 1.49 Å and 1.55 Å with respect to the SR1 of plectin. The main differences between the SR1 of plectin and the spectrin repeats of BPAG1 are located at the N-terminal half of helix B and at the BC loop. Helix B is interrupted by a kink that occurs at the same position in the SR1 of plectin (see above) and in canonical repeats, such as the SR3 of α-actinin. The SR1 of BPAG1 also presents a discontinuity in helix B, but it is placed one turn towards the C-terminus of the helix with respect to plectin´s SR1, causing a difference in the position of the backbone of the N-terminal half of helix B. This kink is not present in the helix B of the SR2 of BPAG1. The other major structural difference is the BC loop. The poor sequence conservation of this loop between spectrin repeats is the basis of the high conformational variability.

The similarity of the SR1 of plectin with SRs of α-actinin is also at the sequence level. The SR1 and SR3 of α-actinin have the highest sequence identity (20% and 18% respectively) with the SR1 of plectin, which is significantly higher than the sequence identity of the SR1 (11%) and SR2 (5%) of the BPAG1 structure. Thus, the first SR of plectin presents more similarities with repeats of the spectrin family than with repeats of the plakin domain of plakins, both at the level of its primary and tertiary structure.

The DALI search with the SR2 of plectin revealed similarity with repeat 17 of brain α-spectrin[21] (PDB entry 1CUN, Z score of 10.8, and rmsd of 2.2 Å for 99 Cα atoms), repeat 2 of α-actinin[20] (PDB entry 1QUU), repeat 1 of α-spectrin[23] (PDB code 1OWA), and repeat 2 of BPAG1[15] (PDB code 2IAK). Based on pairwise superpositions using a common set of 66 Cα atoms, the SR17 of α-spectrin (rmsd 1.49 Å) is most similar to the SR2 of plectin (Figure 2c). The rmsd between the SR2 of plectin and the first and second repeats of BPAG1 are 1.69 Å and 1.98 Å. In contrast, he rmsd between the first and the second repeat of plectin is 2.36 Å. The main differences between SR2 and other spectrin repeats occur in helix B, which in plectin is packed closer to helices A and C than in other repeats (see above). Overall the SR2 of plectin lacks some of the structural features specific of canonical spectrin repeats and as a consequence it also resembles other 3-helix bundles such as those of Db-STAT and syntaxins.

The sequences of the spectrin repeats of the BPAG1 structure are 50% identical with the equivalent sequence of plectin, suggesting that in plectin this region, which is not present in the structure under study, adopts a similar structure as in BPAG1. Thus, the structural differences between the spectrin repeats of the two plakins mainly reflect the structural diversity within the plakin domain, rather than differences between members of the plakin family. In summary, the first and second repeats of plectin are structurally more similar to the spectrin repeats of other proteins than to each other and to the BPAG1 repeats, suggesting that in plectin these repeats have not arisen from a late duplication event.

Domain-domain organization

Helices C from SR1 and A' from SR2 are fused in a continuous inter-repeat helix 18-turns long that contains the linker region. The helical structure of the linker is a common feature observed in the structures of tandem pairs of spectrin repeats of α-actinin[20; 24], α-spectrin[21],

β-spectrin[22; 25] and BPAG1e[26]. The linker region is characterized by a discontinuity in the heptad pattern; in plectin Leu418 occupies the position *d* of the last heptad repeat of helix C while Leu421 is assigned to the *d* position of an incomplete heptad that initiates helix A'.

The first repeat can be approximately superposed onto the second by a 61 Å translation and a rotation of about 111º along the long axis of the molecule. The translation that relates both repeats is about the length of each single repeat (~64 Å and ~57 Å). Therefore helices A and B of SR1 do not overlap helices B' and C' of SR2 along the longitudinal axis of the molecule. Thus, the repeats are arranged in an extended conformation. The rotation that relates both repeats brings the AB loop in the vicinity of the B'C'. Overall the A and B helices of SR1 and the B' and C' helices of SR2 form a left-handed pseudo-super-helix around the long axis of the molecule.

In order to compare the relative orientation of tandem pairs of spectrin repeats in plectin with those of BPAG1 and spectrins, we have superimposed the N-terminal repeat of each pair with the SR1 of plectin. Spectrin repeats of the α-actinin rod[24] have an extended arrangement similar to the plectin SR1-2 structure (Figure 4a). Nevertheless, the relative rotation of adjacent repeats along the longitudinal axis in plectin is different from that described in α-actinin. The BPAG1 structure[15] reveals differences with plectin both in the relative rotation of the two repeats along the longitudinal axis and in the distance between the repeats along the same axis (Figure 4b). The two repeats of BPAG1 partially overlap along the long axis of the molecule in such a way that helix B of SR1 lies on the opposite side of the inter-domain linker from helices B and C of SR2. Thus, the BPAG1 structure is more tightly compressed along the longitudinal axis of the molecule than the plectin structure. In summary, the relative orientation of the first and second repeats of plectin is

different from the orientation of any other tandem spectrin repeats, and further illustrates the conformational variability of tandem pairs of spectrin repeats.

The plakin domain contains an array of tandem spectrin repeats.

Based on data from sequence analysis the plakin domain is proposed to contain up to six regions termed NN, Z, Y, X, W, and V, each of which build up of α-helices arranged in bundles[14]. The SR2 of plectin approximately corresponds to the NN region, and the two repeats of the BPAG1 structure roughly correspond to the Z and Y regions. Therefore, it is reasonable to assume that the other predicted helical bundles also adopt the spectrin repeat fold. In order to test this hypothesis we have performed a thorough search for spectrin repeats in the plakin domain of plectin by using profile hidden Markov model methods[27] (Figure 5a). Our analysis revealed the presence of eight spectrin repeats in the plakin domain of plectin, and successfully identified the two spectrin repeats present in the crystal structure despite not having included any structural information in our search. Repeats are arranged contiguously in the sequence, with exception of a region of about 85 residues immediately downstream the fifth repeat. A prediction of the secondary structure of this initially unaccounted for region, residues 919-1003, indicates the presence of three α-helices connected by short loops (Figure 5b). These three predicted α-helices exhibit the heptad pattern characteristic of left-handed coiled coil, suggesting that this region adopts a spectrin repeat-like fold, with α-helices shorter than in canonical repeats; hence, we named this region SR6. The linkers between spectrin repeats of the plakin domain of plectin are predicted to be very short; thus helices C and A of adjacent repeats are likely to form a contiguous helix as observed in the crystal structure. The only exception is the linker between SR2 and SR3, which is ~20 residues long and suggests a more flexible inter-repeat organization.

The SR5 contains an insertion of about 80 residues in the BC loop that corresponds to a predicted SH3 domain in BPAG1 and other plakins[15]. Testing this inserted sequence of plectin against the Protein Data Bank using profile-profile alignment and fold recognition algorithms (FFAS[28]) confirmed a strong similarity to SH3 domains (FFAS scores between -29.6 and -37.4), which was further supported by a profile search using hidden markov models (data not shown). The SH3 domain of plectin showed the highest degree of sequence identity, 26%, with the SH3 domain of α-spectrin. Furthermore the localization of the SH3 domain of plectin within the fifth spectrin repeat is equivalent to the insertion of the SH3 domain of α-spectrin in the BC loop of its ninth spectrin repeat[29]. The resemblance in sequence and structural environment of the SH3 domains of α-spectrin with plectin extends the structural similarity to the level of the calponin homology type actin binding domain[30] and between the spectrin repeats, and reinforces the idea of a common ancestral origin of plakins and spectrins. The fact that a Pro-rich sequence, 823PRHPAHPMR831, is present upstream of the SH3 of plectin (Figure 5c); which contains a potential class II SH3 ligand is of interest. The sequential arrangement of a Pro-rich sequence and a SH3 domain is highly reminiscent of the structure of the Tec family of intracellular tyrosine kinases. Intra- and intermolecular interactions between Pro-rich sequences and the adjacent SH3 domain regulate the ability of these proteins to engage with their respective targets[31; 32]. Thus, it is seems likely that the Pro-rich region and SH3 domain of plectin have a regulatory function.

We further investigated the presence of spectrin repeats in the plakin domain of other proteins. Using a profile hidden markov model search we identified in envoplakin, periplakin, BPAG1, and MACF1 six spectrin repeats corresponding to the SR3, SR4, SR5, SR7, SR8, and SR9 in plectin (Figure 5a). Equivalent spectrin repeats with the exception of SR7 were identified in desmoplakin. Comparison of the sequence of the SR6 of plectin

with that of other plakins revealed the presence of highly similar regions downstream SR5 in desmoplakin, BPAG1 and MACF1, that are identical for 32% and 38% to those of plectin; thus these plakins also contain the SR6. In contrast envoplakin and periplakin lack this SR6 region. Our profile-based search also revealed the presence of a repeat equivalent to SR1 in non epithelial isoforms of BPAG1 (such as BPAG1a, BPAG1b, and BPAG1n, that share a common N-terminal region that includes the ABD and plakin domain[33]) and MACF1. Given that the SR1 of plectin is structurally most similar to canonical repeats of spectrins, and that in plakins it always appears in combination with an ABD of the spectrin type, it is likely that the ABD-SR1 pair was simultaneously incorporated into certain plakins from an ancestral spectrin[34]. The region of about 100 residues located immediately downstream of SR1 in BPAG1a/b/n and MACF1 are identical for respectively 29% and 34% to the SR2 of plectin; this is a significant level of similarity and thus a SR1-SR2 tandem is predicted both in BPAG1a/b/n and MACF1. The SR2 of BPAG1a/b/n and MACF1 was not detected in our profile-base search most probably because SR2 is a structurally divergent repeat, as we have observed in our crystal structure. The epithelial isoform BPAG1e lacks the SR1 region, but it does contain the SR2. The main difference between the BPAG1 isoforms and plectin resides in the separation between the SR2 and SR3, in plectin the linker region is ~20 residues long, whereas in BPAG1 isoforms it is ~120 residues long and its C-terminus adopts a loop-like structure[26].

The plakin domain of BPAG1e contains an N-terminal pair of spectrin repeats as shown by x-ray crystallography, and a second pair of repeats and a SH3 domain have been predicted from sequence analysis[26]. Our sequence analysis precisely identifies both the N-terminal pair of spectrin repeats of the BPAG1e structure, equivalent to plectin's SR3 and SR4, and the predicted second pair of repeats, SR8 and SR9. In addition we have identified in BPAG1e another four spectrin repeats (SR2 and SR5-SR7) not detected previously. Our

findings extend the previous model of the molecular organization of the plakin domain of BPAG1[26], revealing a continuous tandem array architecture.

In summary, the plakin domain contains a conserved region built up of seven spectrin repeats organized in a tandem array with an SH3 domain inserted between helices B and C of one of the central repeats (Figure 6a,b). In addition certain plakins contain an N-terminal extension consisting of an additional pair of tandem repeats (SR1-SR2) separated from the SR3 by a region of variable length. The structure of the plakin domain is highly reminiscent of the modular architecture of members of the spectrin family.

Modular organization and mechanical properties of the plakin domain.

Based on the dimensions of the structure of the first pair of spectrin repeats and the predicted presence of nine tandem spectrin repeats in plectin, the plakin domain is estimated to have an elongated rod-like shape with an overall length of about 45 nm, and a thickness of about 2.7 nm. Electron microscopy images of plectin either purified[35] or as part of cell cytoskeletons[36] reveal a dumb bell shape with a central region ~200 nm in length and 2 to 3 nm thick. The estimated elongated shape of the plakin domain is consistent with its localization within the central rope-like structure, predicted to also include the rod domain.

The SR1-SR2 tandem of plectin and other plakins is located adjacent to an ABD, suggesting that they may act as a functional unit. For example, both the ABD and the N-terminal region of the plakin domain of plectin containing the SR1-SR2 tandem harbor binding sites for the cytoplasmic domain of integrin β4. Therefore, the orientation relative to each other of the ABD and the SR1-SR2 is likely to be a determinant for their conjunct function. The SR1 of plectin is separated from the preceding ABD by a short linker, residues 290-303, likely to provide flexibility to the way these domains are relatively

oriented to each other, while limiting the distance at which they are separated from each other. This relative arrangement of the ABD and spectrin repeats arrays is common in members of the spectrin family such as α-actinin, β-spectrin, utrophin and dystrophin[37] (Figure 6c).

The composition of the plakin domain as an array of tandem spectrin repeats has profound implications for its function as it dictates flexibility, and extensibility of the domain. The elastic properties of tandem spectrin repeats are considered to be essential for the function of the proteins that contain such arrays. For example spectrin, whose α and β subunits are paradigms of arrays of tandem repeats, essentially contributes to the stability and elasticity of the membrane associated cytoskeleton (e.g., in erythrocytes), a role that relies on the inherent flexibility of the arrays of spectrin repeats. At the molecular level two models of flexibility of tandem arrays of spectrin repeats have been proposed[21]: a bending model and a conformational rearrangement model.

In the bending model the inter-repeat linkers are involved acting as hinges that allow restricted variations in the way the position of the adjacent repeats are oriented to each other with limited structural changes within each repeat. In order to evaluate the possible flexibility within the first pair of spectrin repeats of plectin, we have predicted protein motions (Figure 7). The main motion involves simultaneous rotation of each spectrin repeat in opposite directions along the long axis of the molecule. Two additional significant motions were predicted, corresponding to two orthogonal "bending" motions of the molecule at the level of the linker region. The expected helical nature of the linkers between other tandem spectrin repeats within the plakin domain suggests that comparable inter-repeat "torsion" and "bending" motions may occur between other spectrin repeats of plectin and other plakins.

The conformational rearrangement model of flexibility implies changes in the distances at which spectrin repeats are separated from each other along the long axis of the molecule or "sliding" of repeats, a movement shown to involve loop-helix transitions at the BC loops[21]. In our crystal structure the first and second spectrin repeats of plectin are in an extended conformation along the long axis of the molecule. Although no experimental data is available to date, it is conceivable for example that the first repeat may slide towards the second; such a movement would not require helix-loop transitions because straightening of the helix B near the BC loop and sliding along helix C would relocate loop BC at the end of the bundle. Interestingly the BPAG1 structure (equivalent to plectin´s SR3-SR4) reveals a shorter arrangement of the repeats along the same axis[26], and illustrates variability in the degree of extension between adjacent spectrin repeats within the plakin domain.

Differences in the number of tandem spectrin repeats and the nature of the linkers of the plakin domain among plakins are likely to determine their specific mechanical properties. For example the two plakins present in the hemidesmosomes, plectin and BPAG1e, have nine and eight repeats respectively; therefore their plakin domains may contribute to the mechanical strength of hemidesmosomes both via specific interaction with other components and by providing specific elastic properties.

Concluding remarks.

In summary, the crystal structure of the first tandem pair of spectrin repeats of plectin described in this study reveals both the unique structural details of each repeat and the organization of the two-repeat module. The structural differences between the spectrin repeats of plectin and BPAG1, which correspond to non-equivalent regions, illustrate variability within the plakin domain. An exhaustive analysis of the plakin domain sequences revealed a modular architecture characterized by a continuous array of tandem

spectrin repeats with an inserted SH3 domain. This structure is very similar to the modular organization of members of the spectrin family, and has profound implications for the role of plakin domains *in vivo*. Plakins interconnect cytoskeletal networks, anchor them to cell adhesion complexes, and contribute to the integrity of tissues subjected to mechanical stress. Tandem arrays of spectrin repeats are flexible structures. Thus, our present model provides a first interpretation at the molecular level of the way the plakin domain contributes directly to the mechanical stability of cells not only by linking cytoskeletal elements but also by providing flexible and deformable connections.

Materials and methods

Protein expression and purification

The cDNA sequence coding for residues 300-530 of human plectin (UniprotKB accession number Q15149-2) was cloned into the pET15b vector (Novagen). The double mutant C420A, C435A was created by site-directed mutagenesis using the QuikChange method (Stratagene, La Jolla CA). Proteins were expressed in *Escherichia coli* strain BL21(DE3) and were purified by Ni-chelating affinity chromatography as described[30]. Proteins were expressed fussed to an N-terminal His-tag that was cleaved in all samples by thrombin digestion, and was removed by a second affinity chromatography. The selenomethionine substituted protein was expressed in the non-auxotrophic strain BL21(DE3) as described [38] and purified as for the unlabelled proteins.

Crystallization and structure determination

Crystals of the wild type protein were grown at room temperature using vapor diffusion methods by mixing a protein solution at 7 mg/ml in 10 mM Tris-HCl (pH 7.5), 50 mM NaCl, 5 mM DTT with an equal volume of mother liquor consisting of 0.1 M citrate-

phosphate (pH 4.6), 10% (v/v) 1,2-propanediol, 5% (w/v) polyethylene glycol 3000, 4% (v/v) glycerol. Crystals of the C420A/C435A mutant were obtained as above but using as mother liquor of 0.1 M citrate-phosphate (pH 4.6), 24% (v/v) 1,2-propanediol, 6% (w/v) polyethylene glycol 3000, 3% (v/v) glycerol. Crystals of the selenomethionine substituted C420A/C435A protein were grown the same way but using as mother liquor 0.1 M citrate-phosphate (pH 4.6), 14% (v/v) 1,2-propanediol, 5% (w/v) polyethylene glycol 3000, 3% (v/v) glycerol. Prior to data collection, all crystals were transferred into a cryoprotectant solution consisting of 0.1 M citrate-phosphate (pH 4.6), 25% (v/v) 1,2-propanediol, 6% (w/v) polyethylene glycol 3000, 10% (v/v) glycerol. All data were collected at 100 K in the beamline BM14 at the ESRF (Grenoble, France). Data were indexed with XDS and reduced with XSCALE[39].

All crystals belong to space group $P2_12_12$ and they contain one molecule in the asymmetric unit corresponding to a solvent content of 46% (Table 1). Multiple wavelength anomalous diffraction (MAD) data from selenomethionine substituted C420A/C435A crystals were used to find the positions of seven Se atoms corresponding to six Met residues (one of them in two conformations); phases were calculated and extended to 2.5 Å resolution with ShelxC/D/E[40; 41] using the HKL2MAP[42] graphical user interface. After phase improvement with DM[43] a readily interpretable map was obtained that allowed automatic building of 149 residues using the helix recognition module in ARP/WARP[44]. The structure was refined against C420A/C435A native data to 2.05 Å resolution using REFMAC5[45]. Manual model building using COOT[46] was alternated with restrained refinement, which at later stages included refinement of four TLS groups optimally identified by the TLS Motion Determination (TLSMD) server[47]. Solvent molecules were built in peaks over 3.5 σ of f_{obs}-f_{calc} maps when reasonable H-bonding pattern was observed. Electron density near Glu431

was interpreted as a molecule of 1,2-propanediol. The wild type structure was refined against data to 2.3 Å resolution in a similar way as for the mutant protein.

Structure comparison and analysis

Superposition of the structures was initially done using DALI Lite[48]. Based on the pre-aligned structures a set of 80 Cα atoms in the first repeat of plectin was identified by visual inspection and was used to superpose other domains onto the SR1 using the LSQKAB[49] program of the CCP4 suite[50]. Superpositions onto the SR2 of plectin were done the same way, but using a set of 66 Cα atoms.

Prediction of protein motions in tandem pairs of spectrin repeats was done using the Dynamite server[51]. The server generated an ensemble of feasible protein conformations by using a non-Newtonian method as implemented in CONCOORD[52], which were analyzed with GROMACS[53] to obtain the principal components of predicted protein motions.

Profile-based sequence analysis and secondary structure prediction

In order to identify spectrin repeats in the plakin domain of plectin and other plakins we applied two different methods to analyze the repeat repertoire of the proteins. Both methods are based on profile searching, which allows detecting remote homologies beyond the reach of alternative sequence comparison methods such as PSI-BLAST. In the first one, that relies exclusively on sequences, a multiple alignment of 300 spectrin repeats was extracted from SMART[54] and PFAM[55] databases. Sequences were curated for redundancy and profiles were built using HMMER[27]. These profiles were used to search customized databases containing spectrin repeats, orthologs and paralogs of the protein of interest. In the second method, the FFAS server[28] was used to create a profile from a sequence of interest that was subsequently compared to profiles derived from the PDB.

Prediction of secondary structure was done using the PredictProtein server[56] and relative correspondence with the real structures was used to establish the boundaries of the repeats.

Protein Data Bank accession numbers

The atomic coordinates and structure factors of the wild type and mutant structures have been deposited in the RCSB Protein Data Bank under ID codes 2ODU and 2ODV.

Acknowledgments

We thank Anastassis Perrakis and Kevin Wilhelmsen for critical comments to the manuscript. We would like to thank the EMBL Grenoble Outstation and in particular Hassan Belrhali and Martin Walsh, for providing support for measurements at BM14 (ESRF) under the European Research Infrastructure Action FP6 program. This work was supported by the Ministerio de Educación y Ciencia and the European Regional Development Fund (grants SAF2003-02509 and BFU2006-01929 to J.M.d.P.), by the Junta de Castilla y León (grant SA041/04 to J.M.d.P.), and by the Marie Curie Action MIRG-CT-2005-016499 to A.M.R.

References

1. Wiche, G. (1998). Role of plectin in cytoskeleton organization and dynamics. *J Cell Sci* **111**, 2477-86.
2. Jefferson, J. J., Leung, C. L. & Liem, R. K. (2004). Plakins: goliaths that link cell junctions and the cytoskeleton. *Nat Rev Mol Cell Biol* **5**, 542-53.

3. Wiche, G., Krepler, R., Artlieb, U., Pytela, R. & Aberer, W. (1984). Identification of plectin in different human cell types and immunolocalization at epithelial basal cell surface membranes. *Exp Cell Res* **155**, 43-9.

4. Hieda, Y., Nishizawa, Y., Uematsu, J. & Owaribe, K. (1992). Identification of a new hemidesmosomal protein, HD1: a major, high molecular mass component of isolated hemidesmosomes. *J Cell Biol* **116**, 1497-506.

5. Litjens, S. H., de Pereda, J. M. & Sonnenberg, A. (2006). Current insights into the formation and breakdown of hemidesmosomes. *Trends Cell Biol* **16**, 376-83.

6. Rezniczek, G. A., de Pereda, J. M., Reipert, S. & Wiche, G. (1998). Linking integrin alpha6beta4-based cell adhesion to the intermediate filament cytoskeleton: direct interaction between the beta4 subunit and plectin at multiple molecular sites. *J Cell Biol* **141**, 209-25.

7. Geerts, D., Fontao, L., Nievers, M. G., Schaapveld, R. Q., Purkis, P. E., Wheeler, G. N., Lane, E. B., Leigh, I. M. & Sonnenberg, A. (1999). Binding of integrin alpha6beta4 to plectin prevents plectin association with F-actin but does not interfere with intermediate filament binding. *J Cell Biol* **147**, 417-34.

8. Koster, J., van Wilpe, S., Kuikman, I., Litjens, S. H. & Sonnenberg, A. (2004). Role of binding of plectin to the integrin beta4 subunit in the assembly of hemidesmosomes. *Mol Biol Cell* **15**, 1211-23.

9. Koster, J., Geerts, D., Favre, B., Borradori, L. & Sonnenberg, A. (2003). Analysis of the interactions between BP180, BP230, plectin and the integrin alpha6beta4 important for hemidesmosome assembly. *J Cell Sci* **116**, 387-99.

10. Nikolic, B., Mac Nulty, E., Mir, B. & Wiche, G. (1996). Basic amino acid residue cluster within nuclear targeting sequence motif is essential for cytoplasmic plectin-vimentin network junctions. *J Cell Biol* **134**, 1455-67.

11. Steinbock, F. A., Nikolic, B., Coulombe, P. A., Fuchs, E., Traub, P. & Wiche, G. (2000). Dose-dependent linkage, assembly inhibition and disassembly of vimentin and cytokeratin 5/14 filaments through plectin's intermediate filament-binding domain. *J Cell Sci* **113**, 483-91.

12. Pfendner, E., Rouan, F. & Uitto, J. (2005). Progress in epidermolysis bullosa: the phenotypic spectrum of plectin mutations. *Exp Dermatol* **14**, 241-9.

13. Andra, K., Lassmann, H., Bittner, R., Shorny, S., Fassler, R., Propst, F. & Wiche, G. (1997). Targeted inactivation of plectin reveals essential function in maintaining the integrity of skin, muscle, and heart cytoarchitecture. *Genes Dev* **11**, 3143-56.

14. Green, K. J., Virata, M. L., Elgart, G. W., Stanley, J. R. & Parry, D. A. (1992). Comparative structural analysis of desmoplakin, bullous pemphigoid antigen and plectin: members of a new gene family involved in organization of intermediate filaments. *Int J Biol Macromol* **14**, 145-53.

15. Jefferson, J. J., Ciatto, C., Shapiro, L. & Liem, R. K. (2006). Structural Analysis of the Plakin Domain of Bullous Pemphigoid Antigen1 (BPAG1) Suggests that Plakins Are Members of the Spectrin Superfamily. *J Mol Biol* **11**, 11.

16. Kleywegt, G. J. & Jones, T. A. (1996). Phi/psi-chology: Ramachandran revisited. *Structure* **4**, 1395-400.

17. Cruickshank, D. W. (1999). Remarks about protein structure precision. *Acta Crystallogr D Biol Crystallogr* **55**, 583-601.

18. Lupas, A. (1996). Coiled coils: new structures and new functions. *Trends Biochem Sci* **21**, 375-82.

19. Holm, L. & Sander, C. (1995). Dali: a network tool for protein structure comparison. *Trends Biochem Sci* **20**, 478-80.

20. Djinovic-Carugo, K., Young, P., Gautel, M. & Saraste, M. (1999). Structure of the alpha-actinin rod: molecular basis for cross-linking of actin filaments. *Cell* **98**, 537-46.

21. Grum, V. L., Li, D., MacDonald, R. I. & Mondragon, A. (1999). Structures of two repeats of spectrin suggest models of flexibility. *Cell* **98**, 523-35.

22. Kusunoki, H., MacDonald, R. I. & Mondragon, A. (2004). Structural insights into the stability and flexibility of unusual erythroid spectrin repeats. *Structure* **12**, 645-56.

23. Park, S., Caffrey, M. S., Johnson, M. E. & Fung, L. W. (2003). Solution structural studies on human erythrocyte alpha-spectrin tetramerization site. *J Biol Chem* **278**, 21837-44.

24. Ylanne, J., Scheffzek, K., Young, P. & Saraste, M. (2001). Crystal structure of the alpha-actinin rod reveals an extensive torsional twist. *Structure* **9**, 597-604.

25. Kusunoki, H., Minasov, G., Macdonald, R. I. & Mondragon, A. (2004). Independent movement, dimerization and stability of tandem repeats of chicken brain alpha-spectrin. *J Mol Biol* **344**, 495-511.

26. Jefferson, J. J., Ciatto, C., Shapiro, L. & Liem, R. K. H. (2006). Structural analysis of the plakin domain of bullous pemphigoid antigen1 (BPAG1) suggests that plakins are members of the spectrin superfamily. *J Mol Biol.*

27. Eddy, S. R. (1998). Profile hidden Markov models. *Bioinformatics* **14**, 755-63.

28. Rychlewski, L., Jaroszewski, L., Li, W. & Godzik, A. (2000). Comparison of sequence profiles. Strategies for structural predictions using sequence information. *Protein Sci* **9**, 232-41.

29. Pascual, J., Castresana, J. & Saraste, M. (1997). Evolution of the spectrin repeat. *Bioessays* **19**, 811-7.

30. Garcia-Alvarez, B., Bobkov, A., Sonnenberg, A. & de Pereda, J. M. (2003). Structural and functional analysis of the actin binding domain of plectin suggests alternative mechanisms for binding to F-actin and to integrin a6b4. *Structure* **11**, 615-625.

31. Andreotti, A. H., Bunnell, S. C., Feng, S., Berg, L. J. & Schreiber, S. L. (1997). Regulatory intramolecular association in a tyrosine kinase of the Tec family. *Nature* **385**, 93-7.

32. Pursglove, S. E., Mulhern, T. D., Mackay, J. P., Hinds, M. G. & Booker, G. W. (2002). The solution structure and intramolecular associations of the Tec kinase SRC homology 3 domain. *J Biol Chem* **277**, 755-62.

33. Leung, C. L., Zheng, M., Prater, S. M. & Liem, R. K. (2001). The BPAG1 locus: Alternative splicing produces multiple isoforms with distinct cytoskeletal linker domains, including predominant isoforms in neurons and muscles. *J Cell Biol* **154**, 691-7.

34. Pascual, J., Castresana, J. & Saraste, M. (1997). Evolution of the spectrin repeat. *Bioessays* **19**, 811-7.

35. Foisner, R. & Wiche, G. (1987). Structure and hydrodynamic properties of plectin molecules. *J Mol Biol* **198**, 515-31.

36. Svitkina, T. M., Verkhovsky, A. B. & Borisy, G. G. (1996). Plectin sidearms mediate interaction of intermediate filaments with microtubules and other components of the cytoskeleton. *J Cell Biol* **135**, 991-1007.

37. Broderick, M. J. & Winder, S. J. (2005). Spectrin, alpha-actinin, and dystrophin. *Adv Protein Chem* **70**, 203-46.

38. Bakolitsa, C., de Pereda, J. M., Bagshaw, C. R., Critchley, D. R. & Liddington, R. C. (1999). Crystal structure of the vinculin tail suggests a pathway for activation. *Cell* **99**, 603-13.

39. Kabsch, W. (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst.* **26**, 795-800.

40. Sheldrick, G. M. (2002). Macromolecular phasing with SHELXE. *Z. Kristallogr.* **217**, 644-650.

41. Sheldrick, G. M. & Schneider, T. R. (1997). SHELXL: High Resolution Refinement. *Methods Enzymol* **277**, 319-343.

42. Pape, T. & Schneider, T. R. (2004). HKL2MAP: a graphical user interface for phasing with SHELX programs. *J. Appl. Cryst.* **37**, 843-844.

43. Cowtan, K. (1994). 'dm': An automated procedure for phase improvement by density modification. *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography* **31**, 34-38.

44. Morris, R. J., Zwart, P. H., Cohen, S., Fernandez, F. J., Kakaris, M., Kirillova, O., Vonrhein, C., Perrakis, A. & Lamzin, V. S. (2004). Breaking good resolutions with ARP/wARP. *J Synchrotron Radiat* **11**, 56-9.

45. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240-55.

46. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-32.

47. Painter, J. & Merritt, E. A. (2006). Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D Biol Crystallogr* **62**, 439-50.

48. Holm, L. & Park, J. (2000). DaliLite workbench for protein structure comparison. *Bioinformatics* **16**, 566-7.

49. Kabsch, W. (1976). A solution for the best rotation to relate two sets of vectors. *Acta Cryst.* **A32**, 922-923.

50. (1994). Collaborative Computational Project, Number 4. *Acta Crystallogr D* **50**, 760-763.

51. Barrett, C. P., Hall, B. A. & Noble, M. E. (2004). Dynamite: a simple way to gain insight into protein motions. *Acta Crystallogr D Biol Crystallogr* **60**, 2280-7.

52. de Groot, B. L., van Aalten, D. M., Scheek, R. M., Amadei, A., Vriend, G. & Berendsen, H. J. (1997). Prediction of protein conformational freedom from distance constraints. *Proteins* **29**, 240-51.

53. Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E. & Berendsen, H. J. (2005). GROMACS: fast, flexible, and free. *J Comput Chem* **26**, 1701-18.

54. Letunic, I., Copley, R. R., Pils, B., Pinkert, S., Schultz, J. & Bork, P. (2006). SMART 5: domains in the context of genomes and networks. *Nucleic Acids Res* **34**, D257-60.

55. Finn, R. D., Mistry, J., Schuster-Bockler, B., Griffiths-Jones, S., Hollich, V., Lassmann, T., Moxon, S., Marshall, M., Khanna, A., Durbin, R., Eddy, S. R., Sonnhammer, E. L. & Bateman, A. (2006). Pfam: clans, web tools and services. *Nucleic Acids Res* **34**, D247-51.

56. Rost, B., Yachdav, G. & Liu, J. (2004). The PredictProtein server. *Nucleic Acids Res* **32**, W321-6.

57. DeLano, W. L. (2002). The PyMOL Molecular Graphics System. *DeLano Scientific, San Carlos, CA, USA.*

58. Diederichs, K. & Karplus, P. A. (1997). Improved R-factors for diffraction data analysis in macromolecular crystallography. *Nat Struct Biol* **4**, 269-75.

Table 1. Summary of crystallographic analysis

Data Collection				
Protein	Plectin (300-530) C420A, C435A			Plectin (300-530)
Space group	$P2_12_12$			$P2_12_12$
Cell dimensions	a = 154.6 Å			a = 154.2 Å
	b = 26.5 Å			b = 26.3 Å
	c = 58.1 Å			c = 58.2 Å
		Se-Met MAD		
Data set	Native	Peak	Remote	Native
Wavelength (Å)	0.9785	0.9785	0.9185	0.9785
Resolution (Å)	2.05 (2.15-2.05)[a]	2.7 (2.8-2.7)[a]	2.5 (2.6-2.5)[a]	2.3 (2.38-2.3)[a]
Unique reflections	15444	12658[b]	15939[b]	11244
Redundancy	14.1 (14.2)[a]	7.8 (7.8)[a]	3.9 (3.8)[a]	8.1 (8.4)[a]
Completeness (%)	97.7 (99.6)[a]	100 (100)[a]	99.9 (100)[a]	99.9 (100)[a]
R_{meas} [c] (%)	5.6 (56.3)[a]	8.3 (47.4)[a]	7.1 (50.8)[a]	7.0 (54.8)[a]
$\langle I/\sigma I \rangle$	31.7 (6.0)[a]	19.5 (5.1)[a]	16.6 (3.4)[a]	22.4 (4.9)[a]
Figure of merit after DM			0.72 (0.59)[a]	

Refinement statistics		
	Plectin (300-530) C420A,C435A	Plectin (300-530)
Resolution range (Å)	39 - 2.05	47 - 2.3
Unique reflections, work/free	14679 / 764	10704 / 533
R work (%)	21.0	21.6
R free [d] (%)	25.6	26.9
Number of residues	217	217
Number of solvent molecules	52	16
Average B value ($Å^2$)		
Wilson plot	34.6	37.8
Protein	47.2[e]	47.6[e]
Solvent	48.1	43.8
1,2-propanediol	50.2	n/a
rmsd bond lengths (Å)	0.017	0.015
rmsd angles (°)	1.545	1.564
Ramachandran plot [f]		
Core regions	203	203
Outliers	0	0

[a] Numbers in parenthesis correspond to the outer resolution shell.

[b] Keeping Bijvoet pairs separate.

[c] R_{meas} is the multiplicity independent R factor as described by Diederichs and Karplus [58].

[d] Calculated using 5% of reflections that were not included in the refinement.

[e] Isotropic equivalent B value.

[f] As defined in the program MOLEMAN2[16].

FIGURE LEGENDS

Figure 1. Crystal structure of the first tandem pair of spectrin repeats of plectin. (a) Ribbon representation of the structure with the equivalent α-helices in each domain shown in the same color. (b) Stereo Cα trace in the same orientation as in (a). The sequence is numbered every 10 residues and the trace is colored as in (a). Molecular figures were produced using PyMOL[57].

Figure 2. Structural comparison of the spectrin repeats of plectin, BPAG1 and spectrins. (a) Cα trace superposition of the first (SR1, blue) and second (SR2, orange) spectrin repeats of plectin. The structures were superposed by aligning the Cα atoms of 66 residues in the three α-helices (rmsd 2.36 Å). The helices are labeled in capitals and the structure of the SR1 is marked every 10 residues. (b) Superposition of the SR1 of plectin with the third repeat of α-actinin (left), and the first (middle) and second (right) repeats of BPAG1. The structures were superposed using a common set of 80 equivalent Cα atoms. The position of the kink in helix B is marked a solid arrow in plectin and α-actinin, and by an empty arrow in the SR1 of BPAG1. (c) Superposition of the SR2 of plectin with the repeat 17 of α-spectrin (left), the first (middle) and second (right) repeats of BPAG1. The structures were superimposed using the same set of equivalent Cα atoms as in (a).

Figure 3. Comparison of the core of the spectrin repeats of plectin, BPAG1 and a**-actinin.** Detailed views of the hydrophobic cores of the SR1 (a) and SR2 (b) of plectin, the

SR3 of α-actinin (c) and the N-terminal repeat of BPAG1 (d). The side chains of equivalent residues that contribute to the interactions between helices are shown. The SR1 of plectin contains bulky hydrophobic residues in the hydrophobic core, similar to the SR3 of α-actinin. In contrast, the hydrophobic core of the SR2 of plectin does not contain aromatic residues. For example the highly conserved Trp in helix A is replaced by Gln438, and the near by position in helix C is occupied by Leu508. The smaller size of the side chains in the core allows helix B to pack closely to helices A and C. The SR1 of BPAG1 has mixed features of the two plectin repeats. BPAG1 contains the aromatic triad W288/F325/Y358 but lacks aromatic residues in the N-terminal half of helix B. All panes are shown in the same orientation as in figure 2. The Cα trace is colored as in figure 1a. Numbering of residues in the BPAG1 structure corresponds to the Uniprot entry Q8WXK8 in accordance with figure 5.

Figure 4. Comparison of the relative orientations of adjacent spectrin repeats in plectin, a-actinin, and BPAG1. (a) Orthogonal views of the SR1-SR2 pair of α-actinin (blue, PDB entry 1HCI) superimposed onto the plectin structure (red). The superposition was done using only the N-terminal repeat of each protein in the calculations. Both structures have an extended arrangement along the long axis of the molecules, but the C-terminal repeat of each pair present a different degree of rotation along the same axis. In the apical view (right side of each panel) only the helices of the C-terminal domains of each pair are shown. (b) Comparison of the BPAG1 structure (green, PDB entry 2IAK) with the plectin one. Superposition was done as in (a) and equivalent points of view are shown. The lateral view reveals the short longitudinal inter-domain translation in BPAG1.

Figure 5. Identification of additional modules in the sequence of the plakin domain. Sequences from the plakin domain of plakins were analyzed. For clarity purposes, only the sequences of human plectin (Accession number Q15149-2), and two other prototypical members of the plakin family BPAG1e (Q8WXK8) and desmoplakin (P15924) are shown. (a) Multiple sequence alignment of the spectrin repeats that constitute the plakin domain. Repeats were identified by doing a exhaustive profile search using hidden markov models (HMMER[27]), with a 300 spectrin repeat profile from the SMART[54] database. The corresponding e-value for each repeat is indicated at the right of the sequences. Repeats within each sequence were numbered using plectin as a reference. Colored boxes localize the structurally determined α-helices in plectin's SR1 and SR2; in the remaining repeats the colored boxes indicates predicted α-helices. The coloring is by secondary structure element as in figure 1a. The heptad pattern observed in the α-helices of the crystal structure is indicated above the alignment. (b) Identification of SR6. Multiple sequence alignment of the region downstream SR5 not identified during the afore-mentioned search. This region includes three predicted α-helices equivalent to helices A, B, and C, of canonical spectrin repeats, that show a heptad pattern as indicated above. The helices are indicated by boxes with coloring as above. (c) Sequence alignment of the SH3 domain inserted in the BC loop of SR5 in plakins with the SH3 domains of α-spectrin (P07751) and the Bruton's tyrosine kinase (BTK, Q06187). The consensus localization of the secondary structure elements observed in the structures of the SH3 domains of α-spectrin (PDB code 1SHG) and BTK (PDB code 1AWW) are indicated under the alignment. The Pro-rich region (PRR) involved in the Tec family SH3 self regulation is underlined in the BTK sequence. The plakin's PRR upstream the SH3 domain is marked on top, with Pro residues highlighted in red boxes.

Figure 6. Module organization of the plakin domain. (a) Schematic representation of the domain organization of full-length plectin. The C-terminal region contains five type B and one type C plakin repeat domains. (b) Detail module organization of the N-terminal region of plectin and other representative plakins including BPAG1a/b/n, BPAG1e, desmoplakin and periplakin. Spectrin repeats (SR) are numbered according to plectin. The repeats of plectin and BPAG1 whose crystal structure is known are shown in dark gray. The ABD is formed by two calponin homology (CH) domains. The position of the six α-helical rich regions (NN, Z, Y, X,W, and V) described in the plakin domain[14] are indicated above the plectin structure. (c) The domain organization of the erythrocyte isoforms of α- and β-spectrin is shown to illustrate the similarities with the plakin domain architecture, noticeably the array organization of tandem SRs, the localization of the ABD preceding the SR1 of β-spectrin, and the insertion of an SH3 domain within the ninth SR of α-spectrin.

Figure 7. Prediction of protein motions in the plectin tandem of spectrin repeats. Schematic representation of the three main suggested relative domain motions, with eigenvector indices of 14.7 (I), 12.1 (II), and 10.0 (III); the next higher motion had an index of 2.5 and was consider not significant. The arrows indicate the suggested rotations, and the approximate rotation axes are shown as dotted lines (I and II) and a spot (III).