Published in final edited form as:

J Mol Biol. 2011 June 24; 409(5): 800–812. doi:10.1016/j.jmb.2011.04.046.

Crystal structure of a rigid four spectrin repeat fragment of the human desmoplakin plakin domain

Hee-Jung Choi and William I. Weis*

Departments of Structural Biology and Molecular & Cellular Physiology, Stanford University School of Medicine

Summary

The plakin protein family serves to connect cell-cell and cell-matrix adhesion molecules to the intermediate filament cytoskeleton. Desmoplakin is an integral part of desmosomes, where it links desmosomal cadherins to the intermediate filaments. The 1056 amino acid N-terminal region of desmoplakin contains a plakin domain common to members of the plakin family. Plakin domains contain multiple copies of spectrin repeats (SRs). We determined the crystal structure of a fragment of desmoplakin, residues 175–630, consisting of four spectrin repeats and an inserted SH3 domain. The four repeats form an elongated, rigid structure. The SH3 domain is present in a loop between two helices of a spectrin repeat, and interacts extensively with the preceding spectrin repeat in a manner that appears to limit inter-repeat flexibility. The intimate intramolecular association of the SH3 domain with the preceding spectrin repeat is also observed in plectin, another plakin protein, but not in α -spectrin, suggesting that the SH3 domain of plakins contributes to the stability and rigidity of this sub-family of spectrin repeat containing proteins.

Introduction

Desmosomes are epithelial intercellular junctions that link members of the cadherin superfamily of transmembrane adhesion molecules to the intermediate filament (IF) cytoskeleton. The extracellular domains of the desmosomal cadherins, desmogleins and desmocollins, mediate intercellular contacts, and their cytoplasmic regions form part of a multiprotein assembly that includes plakoglobin, plakophilins, and desmoplakin (DP). Plakoglobin and plakophilins interact with the desmosomal cadherins, and with DP^{1; 2; 3; 4}. In turn, DP binds to intermediate filaments, and is thus a key linker protein between the membrane-bound cadherin complex and IFs. Mice lacking DP cannot survive beyond E6.5 and have few desmosomes as compared to wild type⁵. A conditional knockout of DP in skin and in heart resulted in severe epidermal fragility and cardiac abnormalities, respectively, both leading to prenatal lethality^{6; 7}. A number of human genetic disorders caused by DP mutations have been reported. For example, desmoplakin haploinsufficiency results in striate palmoplantar keratoderma, and several missense mutations and nonsense mutations causing C-terminal truncation, are linked to arrhythmogenic right ventricular cardiomyopathy (ARVC)^{8; 9; 10}.

© 2011 Elsevier Ltd. All rights reserved.

*Correspondence: Department of Structural Biology, Stanford University School of Medicine, 299 Campus Drive, Stanford, CA 94305, bill.weis@stanford.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

DP belongs to the plakin protein family of cytolinkers, which includes the hemidesmosomal proteins bullous pemphigoid antigen 1 (BPAG1) and plectin that link extracellular matrix-binding integrins to intermediate filaments, and the desmosomal cytolinkers periplakin and envoplakin present in the cornified envelope of the skin¹¹. The primary structure of DP has three distinct regions: a 1056 amino acid N-terminal domain (DPNT), a 890 residue central coiled-coil dimerization domain, and a 925 residue C-terminal intermediate filament binding domain. DPNT, like most other plakins, contains characteristic plakin domain composed of tandem spectrin-like repeats¹². The C-terminal regions of DP and other plakin family proteins contain variable numbers of IF-binding plakin repeat domains¹³.

Yeast two-hybrid assays and coimmunoprecipitation experiments showed that the N-terminal 584 amino acid region of DP interacts with plakoglobin, which, in turn, binds to the cytoplasmic domain of desmosomal cadherins, desmoglein and desmocollin via its central armadillo repeat domain^{1; 14; 15}. The N-terminal 176 amino acids of DP were also reported to interact with desmocollin in blot overlay assays². Studies with N-terminal deletion mutants of DP determined the N-terminal 86 amino acids of DP are sufficient to target DP to desmosomes². Sequence analysis of DP indicates that the N-terminal 180 amino acid region is largely α -helical, but is distinct from the spectrin-like repeat regions present in the rest of DPNT^{12; 16; 17}. Other plakin family members, including plectin and BPAG1a, contain an actin binding domain (ABD) at their N-termini preceding the plakin domain¹¹. The plakin domain of plectin interacts with the cytoplasmic domain of β -dystroglycan *in vitro*, but no direct interaction partners of the equivalent region of DP have been reported¹⁸.

Spectrin repeats (SRs) consist of three α helices, A, B, and C that form an antiparallel triple helical bundle^{19; 20; 21}. Unlike some multiple repeat proteins in which repeating motifs are connected by flexible linkers, successive spectrin repeats are connected by an α -helical linker, such that helix C of one repeat is continuous with helix A of the next. The inherent flexibility of spectrin repeat regions is thought to be responsible for the elasticity of molecular assemblies containing such proteins. Structural^{22; 23} and theoretical²⁴ studies have shown that the α -helical linker is the weakest point in the structure, as the packing contacts between residues at heptad repeat positions are disrupted in this region, thereby imparting flexibility to the structure. Moreover, conformational rearrangement of spectrin repeats with an elongated helix B and shortened helix C have also been observed²⁰.

Crystal structures of two tandem spectrin repeats of BPAG1e and plectin showed that they share the basic three-helix architecture found in other spectrin repeats^{12; 17; 25}. Based on sequence conservation DPNT was predicted to have 8 spectrin repeats, with the first spectrin repeat of DP equivalent to the third repeat of plectin such that the DPNT plakin domain contains spectrin repeats designated SR3–6, and SR8–9; SR7 found in other plakin domains appears to be absent in desmoplakin, and is replaced by a non-helical connector. Also, sequence analysis suggested that an SH3 domain is inserted between helices B and C of SR5 in DP, plectin and BPAG1 (Fig. 1a)¹².

The SH3 domain mediates protein-protein interactions in a variety of contexts including signal transduction pathways and vesicular trafficking^{26; 27}. The SH3 domain within SR9 of α -spectrin has been shown to be involved in Rac activation in integrin clusters²⁸. In the case of nonerythroid α -spectrin, the SH3 domain directly interacts with Fanconi anemia protein as part of the DNA repair process²⁹. High resolution NMR structures of the isolated α -spectrin SH3 domain and its complex with a proline rich polypeptide showed that it adopts a canonical SH3 fold and binds to peptide ligands using a similar binding interface as that seen in other SH3 domain complexes. The functional role of the SH3 domain of DP is not known, and no binding partners have been reported. Very recently, a structural study of plectin SR4-SR5 containing an SH3 domain showed that the SH3 domain does not contain

the canonical ligand binding interface and instead has intramolecular interactions with hydrophobic residues in SR4, suggesting that the SH3 domain contributes to the structural stability of plakin family ²⁵.

Here we report the crystal structure of a large fragment of DPNT, residues 175–630, consisting of spectrin repeats SR3 to SR6 and the SH3 domain. The SH3 domain is inserted into the B–C loop of SR5 and interacts extensively with SR4. SR6 has a divergent structure relative to the other SRs. The SH3 domain of DP appears to form a tight interface with the preceding SR that would limit inter-repeat flexibility. The intimate intramolecular association of the SH3 domain is also observed in plectin but not in α -spectrin, suggesting that the SH3 domain of plakins contributes to the stability and rigidity of this sub-family of spectrin repeat containing proteins.

Results and Discussion

Overall structure

Secondary structure prediction using PredictProtein ³⁰ of human DPNT suggested a high proportion of α helix for residues 1–630, followed by a non-helical stretch and then another α -helical region. The bacterially-expressed construct comprising residues 1–630, designated DP(1–630), degraded during purification and did not crystallize. To identify a stable fragment more amenable to crystallization, DP(1–630) was subjected to limited proteolysis by trypsin, chymotrypsin or V8 protease. V8 protease did not cleave the construct, whereas trypsin and chymotrypsin each yielded a ~55 kDa stable fragment on SDS-PAGE (data not shown). N-terminal sequencing of these fragments showed that trypsin and chymotrypsin cleaved after Lys167 and Tyr172, respectively, suggesting that this region forms a flexible linker (Fig. 1a). Residues 175–630 were expressed in bacteria, and produced a stable protein that crystallized. Most DP(175–630) crystals diffracted to only 8–10 Å, but one native crystal out of 52 screened diffracted to 2.95 Å. Likewise, one selenomethionyl protein crystal out of 129 screened diffracted to 3.5 Å, and 3.8 Å resolution SAD data set measured from this crystal was used for phasing and initial refinement. The final structure at 2.95 Å resolution was obtained by refinement against the native data. Data collection, phasing and refinement statistics are provided in Table 1.

DP(175–630) is an elongated, rod-shaped molecule with approximate dimensions 180 Å × 40 Å × 40 Å (Fig. 1b). The structure consists of four spectrin repeats (SRs) connected by helical linkers. There is also a single SH3 domain that connects two helices, B and C of SR5 (Fig. 1b). The DP(175–630) crystals contain two molecules in the asymmetric unit, which pack perpendicularly against each other near the middle of the molecule (Fig. 1c). The interface buries only 180 Å² surface area, which corresponds to about 0.7% of free protomer area, consistent with the observation that DP(175–630), as well as DPNT, sizes as a homogeneous monomer on a size exclusion chromatography column (data not shown).

Structural superposition of individual repeats within a monomer reveals that the helical regions of SR3, SR4, and SR5 are very similar, with root-mean-square deviations (rmsds) of 0.8–1.2 Å (Fig. 2a), whereas SR6 shows significant deviations in helices A and B (rmsd of 1.9 Å). These two helices are significantly shorter in SR6 than in the other repeats, and SR6 also is more divergent in sequence: SR3, SR4, and SR5 are 23–27% similar to one another, whereas SR6 is only about 13% similar to the others (Fig. 2b; Table 2). A DALI search with SR6 revealed that SR6 is more similar to other three helix bundle proteins, including Target of Myb1 and the syntaxin Habc domain, than to other spectrin repeat proteins.

The deviation of SR6 from the other repeats correlates with the absence of key residues thought to stabilize the spectrin domain fold: a conserved tryptophan residue in helix A,

another conserved aromatic residue at a heptad *a* position in helix C, and a less conserved hydrophobic residue at the heptad *d* position in helix B interact in the core of most spectrin repeats¹². In DP(175–630), helix C of SR6 has the conserved aromatic residue (Phe612), but the tryptophan in helix A is replaced by aspartic acid (Asp556), which is solvent exposed rather than buried (compare to Trp194 in SR3, Trp285 in SR4, and Tyr391 in SR5; Fig. 2c). Helix A in SR6 is shorter than in the other repeats, and its N-terminal half displays a large deviation relative to the other repeats. The large displacement of the N-terminal part of helix A is caused by the presence of Phe600 in B-C loop, whose side chain points into the hydrophobic core between helix A and B, tightly packing with Trp550 in helix A and Tyr587 and Phe590 in helix B (Fig. 2c). All of these aromatic residues, including Phe600, are conserved in BPAG1 and plectin, suggesting that this aromatic packing network confers a similar SR6 structure in other SR6 containing plakin proteins. For convenience and compatibility with earlier nomenclature we refer to this region as a “spectrin-like repeat” and label it as “SR6”, although it is clear that it lacks features characteristic of other spectrin repeats¹⁷. Indeed, other variations of the spectrin repeat architecture are seen in other spectrin-repeat proteins. For example, plectin SR2 lacks the conserved aromatic residues in helices A and C, such that helix B is closer to helices A and C than in canonical spectrin repeats¹⁷. Repeat 9 of erythroid β -spectrin also lacks these conserved aromatic residues in helices A and C, and two water molecules are recruited to fill up the cavity²².

Inter-repeat flexibility and the role of the SH3 domain

The individual spectrin repeats of the two crystallographically independent molecules superimpose closely, with rmsds between 0.5 and 0.8 Å. However, superposition of the two full monomers (Fig. 1d) gives a larger rmsd of 1.8 Å, indicating significant flexibility despite the fact that the spectrin repeats are connected by continuous helices rather than flexible linkers. When the first repeats of the two molecules are superimposed, significant deviations of the other repeats are evident (Fig. 1d). Inter-repeat flexibility was previously reported in the crystal structure of two tandem repeats of α -spectrin^{20; 22; 23}, with two different sources of flexibility noted. The first is conformational rearrangement, which is observed within one repeat of α -spectrin, where a loop region shifts position such that helix B extends and helix C shortens²⁰. The other type of flexibility arises from different degrees of bending at the linker region. In DP(175–630), the individual repeat of one crystallographically independent molecule aligns well with the corresponding repeat of the other molecule, demonstrating that conformational rearrangements do not contribute to the flexibility. Instead, different bending angles of the linker helices, ranging from 0.3° to 15.8° are observed. Interestingly, the SR5-SR6 di-repeat shows the most flexibility with an rmsd of 1.6 Å, whereas the SR4-SR5 di-repeat, which includes the SH3 domain, shows the least flexibility with an rmsd of 1.0 Å.

As expected¹², the SH3 domain is observed in the loop between helices B and C of SR5. Like other SH3 domains, it consists of 5 β -strands that form two orthogonal beta sheets (Fig. 3a). The first and the second β -strands are connected by the long RT loop. The n-Src loop and the distal loop connect β -strands 2 and 3, and 3 and 4, respectively. Structural alignments of the DP(175–630) SH3 domain with other known SH3 domains (see Methods) give rmsd values between C α atoms of 0.6–0.8 Å, excluding the three variable loops. Several conserved hydrophobic residues, Leu463, Leu484, Trp493, Val495, and Leu510 form the protein core, and other key structural residues, Gly479, Asp480, and Gly500 are conserved in β -turns.

Canonical SH3 domains with specificity for the sequence motif PXXP feature a set of conserved aromatic and proline residues that form the ligand-binding surface. Some of these aromatic residues are not present in the DP(175–630) SH3 domain. The first tyrosine residue in the conserved YDY sequence of the RT loop is replaced with Cys467, while Asp468 and

Tyr469, which are known to be crucial for protein stability, are preserved. Lys492 and Gly509 also replace the conserved tryptophan and tyrosine residues located in β -strand 3 and the region N-terminal to β -strand 5, respectively (Fig. 3b). The absence of these critical ligand-binding residues may be related to the fact that this surface is partially occupied by hydrophobic residues from SR4. Although the SH3 domain and spectrin repeats were suggested to be independent structural units by the fact that SH3 domain is located in a loop of SR5, the structure reveals an intimate association between the SH3 domain and SR4: 23% of the solvent accessible surface area of the SH3 domain (1090 Å²) is buried in the spectrin-SH3 interface. In particular, Trp360, located in helix C of SR4, lies in the canonical SH3 ligand-binding site, where it packs against the conserved Tyr469 (Fig. 3b). This interaction is enabled by the replacement of a conserved tyrosine residue with Gly509, which removes a potential steric clash. The position of Trp360 is further stabilized by a hydrogen bond between the tryptophan indole NH and Glu290. Likewise, Cys367 packs against Cys467 in the RT loop of the SH3 domain. Another important interaction occurs between a hydrophobic surface created by Trp360 and Ile364 in SR4 helix C, Leu294 in SR4 helix A, and Cys367 in SR4 helix C, with Val508 located in the β 4/ β 5 loop of the SH3 domain. As in the case of the Trp360 interaction, the Val508 interaction is further fixed by a hydrogen bond between its main-chain carbonyl oxygen and His371 of SR4.

It was suggested that a proline rich sequence (⁴⁵⁰PRNP⁴⁵³) present N-terminal to the DP SH3 domain would engage in intra- or inter-molecular binding, as shown in the NMR structures of the Itk and Rlk SH3 domains^{17; 31}. The DP(175–630) structure shows no evidence for such interactions. Instead, the N-terminal and the C-terminal linker peptides that connect the SH3 domain to SR5 helices B and C interact with nearby spectrin repeats, thereby orienting the SH3 domain for the extensive contacts described above. Lys449, which is positioned at the N-terminal linker, forms a salt bridge with Asp300 of SR4. Leu448 and Pro514, which are positioned in the N-terminal and the C-terminal linker regions, respectively, pack against Tyr378 and Ile445 in SR5. The fact that these linkers make close contacts with the main body of the structure is consistent with proteolytic sensitivity experiments using trypsin and chymotrypsin, which showed no cleavage in these linkers despite the presence of several basic and aromatic residues (data not shown) (Fig. 2b). Moreover, the orientation of SH3 domain with respect to the nearby spectrin repeats is similar in both crystallographically independent copies, suggesting a rigid interface (Fig. 1d).

No ligands for the SH3 domains of DPNT and other plakin family members have been reported. The DP(175–630) structure suggests that the polyproline-binding surface of the SH3 domain is not accessible to other proteins. One possibility is that the SH3 domain is not stably associated with SR4, such that when it is away from the main body of DPNT it might interact with a ligand in vivo. However, we have found no evidence for the SH3 being flexibly linked to the rest of the protein. Moreover, the isolated DP SH3 domain is unstructured by NMR spectroscopy, indicating that it is unstable in the absence of its interaction with SR4 (M. Overduin, personal communication). We also tested whether the SH3 domain might affect the stability of DP(175–630). When the SH3 domain was deleted from DP(175–630) and replaced with a 6 amino acid linker (Gly-Ser-Gly-Ser-Gly-Ser) to directly connect residues 447 to 516, the mutant did not express as a soluble protein in *E. coli*. This likely arises from the now solvent-exposed hydrophobic residues that normally interact with the SH3 domain, including Leu294, Trp360, and Cys367. Collectively, these observations support the notion that the hydrophobic interface between the spectrin repeats and the SH3 domain is fixed and rigid. The SH3 domain could, however, have non-canonical binding surfaces, as has been observed in other proteins^{32; 33; 34; 35}. In this regard, it is interesting that the exposed portion of the DP SH3 domain displays an acidic patch adjacent to a highly basic surface region on SR4 (Fig. 1b).

Tandem spectrin repeats form elongated, flexible structures that connect different functional regions. For example, the N-terminal 175 amino acids of desmoplakin that precede the spectrin repeat region appear to mediate interactions with desmosomal cadherins, plakoglobin and plakophilins near the plasma membrane, and the regions following DPNT mediate dimerization and IF binding. The elongated structure of DP(175–630) likely serves to impart appropriate mechanical properties to desmosomes. Bending of inter-repeat linker helices is believed to be an important factor in imparting overall flexibility to molecules containing tandem spectrin repeats^{22; 23}. We suggest that the extensive hydrophobic interactions between the SH3 domain of SR5 and the preceding SR4 rigidify the molecule in this region, and could also prevent conformational rearrangement of spectrin repeats seen in the structure of α -spectrin (Grum et al, 1999). This notion correlates with the observation above that the di-repeat SR4-SR5 shows the least flexibility amongst the various tandem pairs of SRs in the DP(175–630) structure.

ARVD/C mutations

Several mutations in DPNT that cause Arrhythmogenic Right Ventricular Dysplasia/Cardiomyopathy (ARVD/C) lie in the region visualized in the crystal structure presented here (Fig. 4). These sites fall into two groups: core residues and surface residues. Core residues Ser299, Asn375, Ile445, and Ser507 are conserved in BPAG1 and plectin, whereas surface residues are not, except for Asn287 and Lys470. Interestingly, the core mutation sites, which likely affect structural stability, cluster near the SH3 domain, supporting the notion that the SH3 domain has a crucial role in maintaining the structural integrity of DPNT. The other six residues lie on the surface of the molecule, both on SRs and the SH3 domain. To date, protein partners have been reported only for the first 175 amino acids of DPNT, but it is possible that the residues corresponding to surface ARVD/C mutations in DP(175–630) mediate interactions with other desmosomal proteins, or perhaps intramolecular interactions with other portions of desmoplakin. In this regard, it is interesting that desmoplakin has a predicted non-helical region in place of SR7¹⁷. It is possible that SRs 8 and 9 following this region are flexibly linked to SR3–6 and that the two regions might interact under some circumstances. Indeed, SAXS analysis of a fragment spanning all SRs of DPNT suggests a major kink following SR6 (M. Overduin, personal communication). Also, the surface of SR4-SR5 of DP(175–630) is very basic (Fig. 1b), and could be a binding site for other proteins.

Comparison with other spectrin repeat-containing structures

Tertiary structures of spectrin repeat protein fragments comprising two to four consecutive spectrin repeats from α -actinin, α -spectrin, β -spectrin, BPAG1e, and plectin are available. Structural superpositions using about 80 core residue C α positions of SR3 of DP were performed: SR3s of BPAG1e and plectin had the lowest rmsd values (1.0 Å), whereas SR1 of plectin showed the most divergence (rmsd = 2.1 Å). Spectrin repeats of α -actinin (SR1), α -spectrin (SR16) and β -spectrin (SR8) had rmsd values of 1.9 Å, 1.4 Å, and 1.5 Å, respectively. Amino acid sequence alignments show that SR3 of DPNT is much more similar to SR3 of plectin and BPAG1 than to other DPNT spectrin repeats or to others in plectin and BPAG1. For example, sequence identities of SR3 of DPNT to SR3 and SR4 of plectin are about 32% and 9%, respectively, and similarly SR4 and SR5 are more similar between proteins than to each other (Table 2). These observations indicate that plakin family members and other spectrin repeat-containing proteins likely diverged from a common ancestor that contained multiple tandem spectrin repeats. Curiously, the structural deviations between repeats within DP(175–630) are relatively small considering the low level of sequence identity and similarity between different repeats (Table 2), suggesting that there are strong constraints on the overall structure, with the few conserved residues dictating very similar relative positions of the three helices.

During the preparation of this manuscript, the crystal structure of the plectin SR4-SR5 containing an SH3 domain was reported²⁵. As expected from amino acid sequence conservation, the intramolecular interface between SR4 and the SH3 domain described here for DP is very similar to that of plectin, indicating that it is a general structural feature of plakin domains (Supp. Fig.1). Moreover, the residues that form the interface between the SH3 domain and SR4 in DP are conserved in BPAG1 and plectin, except for Gly509, which corresponds to Cys882 in plectin and Cys612 in BPAG1. In contrast to DP and plectin, the SH3 domain of α -spectrin repeat 9 retains the key aromatic residues that are important for peptide binding, and the structure of the α -spectrin SH3 domain bound to a polyproline peptide (p41) shows that the interaction mode is similar to other canonical SH3-peptide complexes³⁶ (Fig. 4c).

Comparison of the structures DP(175–630) and plectin reveals one noteworthy difference: the extra helical element found in plectin SR5 in the loop preceding helix B, designated helix B0, is located in a position that could not interact with SR6. As noted by the authors of that study, this region is stabilized by a crystal packing interaction and a calcium ion present in the crystallization cocktail, suggesting that it is an artifact of crystallization (supp. Fig.1). Most of the crystallized constructs of spectrin repeats, including plectin SR4-SR5, have a break in the helical linker that joins tandem SRs. In contrast, the structure of DP(175–630) represents the four intact spectrin repeats with the SH3 domain, flanked by flexible elements at either end.

It is interesting to compare the overall organization of repeats in the known tandem spectrin repeat structures, as the orientation of one spectrin repeat relative to the previous repeat is determined by different degrees of linker helix bending, and/or a change in the phase of the heptad pattern depending on the length of linker helix. Most reported structures have two repeats but three 3-repeat structures (α -spectrin repeats 15–17, β -spectrin repeats 14–16, and erythroid β -spectrin repeats 13–15) and one other 4-repeat structure (α -actinin repeats 1–4) have been determined^{21; 23; 37; 38}. Fig. 5a shows the comparison of domain organization of spectrin repeats in DPNT, BPAG1e (SR3-SR4), plectin (SR1-SR2), erythroid β -spectrin (SR8-SR9), α -spectrin (SR15-SR16-SR17) and α -actinin (SRs1–4)^{12; 17; 21; 22; 23}. In order to compare the relative orientation of the repeats, each N-terminal repeat was aligned with SR3 of DP(175–630).

The difference in relative orientation between the second repeats of BPAG1e and DPNT arises mainly from differences in linker helix bending (Fig. 5b). However, in the case of plectin (SR1-SR2), helix B of SR2 is located at the position of helix C of SR4 in DPNT. This is caused by the presence of 11 more residues in the linker helix, which changes the orientation of hydrophobic core residues on helix A, which changes the positions of helices B and C relative to helix A. In contrast, repeat 9 of erythroid β -spectrin is rotated in the opposite direction such that its helix C adopts a position similar to that helix B in DPNT. Sequence alignment between erythroid β -spectrin and DPNT showed that the linker helix of β -spectrin has 3 more residues. These examples highlight the diverse domain organization possible for tandem spectrin repeats. Interestingly, plectin, BPAG1e, and DPNT have the same number of residues in linker helices connecting repeats 3 and 4, repeats 4 and 5, and repeats 5 and 6, implying that these plakin proteins would have similar overall domain organization with DPNT but simply varied by flexibility of linker helices. These observations are consistent with the notion that the plakin subfamily diverged from a common multi-spectrin repeat protein.

Methods

Protein expression and purification

The cDNA encoding residues 1–630 and residues 175–630 of human desmoplakin was cloned into the pGEX-TEV vector, which produces a tobacco etch virus (TEV) protease cleavable fusion to an N-terminal glutathione-S-transferase affinity tag. Protein was expressed in *Escherichia coli* strain BL21 and purified by glutathione-agarose affinity chromatography. After removal of the tag by TEV protease, the protein was further purified by anion exchange (MonoQ, GE Healthcare) and size exclusion (S200, GE Healthcare) chromatography (all buffers include 25mM Tris-Cl, pH 8.0, 0.5mM EDTA, and 2mM DTT; a NaCl gradient from 50–250 mM was used for MonoQ and 100 mM NaCl was present in the S200 buffer). The selenomethionine-substituted protein was produced using metabolic inhibition of methionine synthesis³⁹ and purified as for the native protein.

Crystallization and structure determination

Crystals of the native protein were grown by vapor diffusion at 20°C by mixing a protein solution at 20 mg ml⁻¹ with an equal volume of reservoir consisting of 0.1 M Tris-Cl pH 8.5, 16% PEG 8000 and 4 mM dithiothreitol. Crystals of the selenomethionyl protein were grown similarly but with a reservoir comprising 50 mM CHES pH 9.25, 14% PEG 8000, and 8 mM TCEP. The crystals were typically 0.3 × 0.3 × 0.5 mm³.

All crystals were flash frozen into liquid nitrogen using perfluoropolyether-X175/08 (Hampton Research) as cryoprotectant. Native and single-wavelength anomalous dispersion (SAD) data sets were measured on beamlines 9–2 and 11–1 of the Stanford Synchrotron Radiation Laboratory, respectively. Data were processed with Mosflm⁴⁰ and Scala⁴¹. The crystals belong to space group C2 and have two molecules in the asymmetric unit. The selenomethionyl SAD data were measured by inverse beam at the Se absorption peak. Of 30 possible Se sites, 24 were found with Shake and Bake⁴² and refined with Sharp⁴³. SAD phases calculated with Solve⁴⁴ were improved by density modification in Resolve⁴⁵. The initial model was built into the density-modified map using Coot⁴⁶. Rigid-body refinement followed by minimization was performed with CNS, giving an R value of 40%. Molecular replacement (Phaser) using this partially refined model was used to phase the native data set; the R value of the top solution was 47%. The structure was refined against the native data to 2.95 Å resolution using Phenix⁴⁷. At later stages of refinement, 4 TLS groups per each molecule, identified by the TLSMD server⁴⁸, were used for TLS refinement.

Structure comparisons

Structural superpositions were performed with Lsqkab program of the CCP4 suite⁴¹. For superpositions of SR4 and SR5 onto SR3, 75 Ca atoms of SR3 were used as a reference set and 51 Ca atoms of SR3 were used for superposition with SR6. In the comparisons with other spectrin repeat structures, each representative spectrin repeat structure, except for α -actinin, was superimposed onto 75 Ca atoms of SR3 of DP(175–630). Superposition of α -actinin was done using 55 Ca atoms instead. The SH3 domain of DP(175–630) was superimposed with three other representative SH3 domain structures: Abl kinase (pdb ID; 1ABO), IL-2-inducible T-cell kinase (pdb ID; 1AWJ), and α spectrin (pdb ID; 2JMA) using 32 Ca atoms in the β strands as a reference set.

Accession numbers

Coordinates and structure factors for DP(175–630) have been deposited in the Protein Data Bank, code 3R6N.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank M. Overduin for communicating results prior to publication. Portions of this research were carried out at the Stanford Synchrotron Radiation Lightsource, a Directorate of SLAC National Accelerator Laboratory and an Office of Science User Facility operated for the U.S. Department of Energy Office of Science by Stanford University. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research, and by the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program (P41RR001209), and the National Institute of General Medical Sciences. This work was supported by National Institutes of Health grant GM56169 to W.I.W.

References

1. Kowalczyk AP, Bornslaeger EA, Borgwardt JE, Palka HL, Bhaliwal AS, Corcoran CM, Denning MF, Green KJ. The amino-terminal domain of desmoplakin binds to plakoglobin and clusters desmosomal cadherin-plakoglobin complexes. *J Cell Biol.* 1997; 139:773–784. [PubMed: 9348293]
2. Smith EA, Fuchs E. Defining the interactions between intermediate filaments and desmosomes. *J Cell Biol.* 1998; 141:1229–1241. [PubMed: 9606214]
3. Chen X, Bonne S, Hatzfeld M, van Roy F, Green KJ. Protein binding and functional characterization of plakophilin 2. Evidence for its diverse roles in desmosomes and beta -catenin signaling. *J Biol Chem.* 2002; 277:10512–22. [PubMed: 11790773]
4. Kowalczyk AP, Hatzfeld M, Bornslaeger EA, Kopp DS, Borgwardt JE, Corcoran CM, Settler A, Green KJ. The head domain of plakophilin-1 binds to desmoplakin and enhances its recruitment to desmosomes. Implications for cutaneous disease. *J Biol Chem.* 1999; 274:18145–8. [PubMed: 10373410]
5. Gallicano GI, Kouklis P, Bauer C, Yin M, Vasioukhin V, Degenstein L, Fuchs E. Desmoplakin is required early in development for assembly of desmosomes and cytoskeletal linkage. *J Cell Biol.* 1998; 143:2009–2022. [PubMed: 9864371]
6. Vasioukhin V, Bowers E, Bauer C, Degenstein L, Fuchs E. Desmoplakin is essential in epidermal sheet formation. *Nat Cell Biol.* 2001; 3:1076–1085. [PubMed: 11781569]
7. Garcia-Gras E, Lombardi R, Giocondo MJ, Willerson JT, Schneider MD, Khoury DS, Marian AJ. Suppression of canonical Wnt/beta-catenin signaling by nuclear plakoglobin recapitulates phenotype of arrhythmogenic right ventricular cardiomyopathy. *J Clin Invest.* 2006; 116:2012–21. [PubMed: 16823493]
8. Armstrong DK, McKenna KE, Purkis PE, Green KJ, Eady RA, Leight IM, Hughes AE. Haploinsufficiency of desmoplakin causes a striate subtype of palmoplantar keratoderma. *Hum Mol Genet.* 1999; 8:143–148. [PubMed: 9887343]
9. Norgett EE, Hatsell SJ, Carvajal-Huerta L, Cabezas JC, Common J, Purkis PE, Whittock N, Leigh IM, Stevens HP, Kelsell DP. Recessive mutation in desmoplakin disrupts desmoplakin-intermediate filament interactions and causes dilated cardiomyopathy, woolly hair and keratoderma. *Hum Mol Genet.* 2000; 9:2761–6. [PubMed: 11063735]
10. Alcalai R, Metzger S, Rosenheck S, Meiner V, Chajek-Shaul T. A recessive mutation in desmoplakin causes arrhythmogenic right ventricular dysplasia, skin disorder, and woolly hair. *J Am Coll Cardiol.* 2003; 42:319–27. [PubMed: 12875771]
11. Jefferson JJ, Leung CL, Liem RK. Plakins: goliaths that link cell junctions and the cytoskeleton. *Nat Rev Mol Cell Biol.* 2004; 5:542–53. [PubMed: 15232572]
12. Jefferson JJ, Ciatto C, Shapiro L, Liem RK. Structural analysis of the plakin domain of bullous pemphigoid antigen1 (BPAG1) suggests that plakins are members of the spectrin superfamily. *J Mol Biol.* 2007; 366:244–57. [PubMed: 17161423]
13. Leung CL, Green KJ, Liem RK. Plakins: a family of versatile cytolinker proteins. *Trends Cell Biol.* 2002; 12:37–45. [PubMed: 11854008]

14. Wahl JK, Sacco PA, McGranahan-Sadler TM, Sauppe LM, Wheelock MJ, Johnson KR. Plakoglobin domains that define its association with the desmosomal cadherins and the classical cadherins: identification of unique and shared domains. *J Cell Sci.* 1996; 109 (Pt 5):1143–54. [PubMed: 8743961]
15. Witcher LL, Collins R, Puttagunta S, Mechanic SE, Munson M, Gumbiner B, Cowin P. Desmosomal cadherin binding domains of plakoglobin. *J Biol Chem.* 1996; 271:10904–9. [PubMed: 8631907]
16. Virata ML, Wagner RM, Parry DA, Green KJ. Molecular structure of the human desmoplakin I and II amino terminus. *Proc Natl Acad Sci U S A.* 1992; 89:544–8. [PubMed: 1731325]
17. Sonnenberg A, Rojas AM, de Pereda JM. The structure of a tandem pair of spectrin repeats of plectin reveals a modular organization of the plakin domain. *J Mol Biol.* 2007; 368:1379–91. [PubMed: 17397861]
18. Reznicek GA, Konieczny P, Nikolic B, Reipert S, Schneller D, Abrahamsberg C, Davies KE, Winder SJ, Wiche G. Plectin 1f scaffolding at the sarcolemma of dystrophic (mdx) muscle fibers through multiple interactions with beta-dystroglycan. *J Cell Biol.* 2007; 176:965–77. [PubMed: 17389230]
19. Yan Y, Winograd E, Viel A, Cronin T, Harrison SC, Branton D. Crystal structure of the repetitive segments of spectrin. *Science.* 1993; 262:2027–30. [PubMed: 8266097]
20. Grum VL, Li D, MacDonald RI, Mondragon A. Structures of two repeats of spectrin suggest models of flexibility. *Cell.* 1999; 98:523–35. [PubMed: 10481916]
21. Djinovic-Carugo K, Young P, Gautel M, Saraste M. Structure of the alpha-actinin rod: molecular basis for cross-linking of actin filaments. *Cell.* 1999; 98:537–46. [PubMed: 10481917]
22. Kusunoki H, MacDonald RI, Mondragon A. Structural insights into the stability and flexibility of unusual erythroid spectrin repeats. *Structure.* 2004; 12:645–56. [PubMed: 15062087]
23. Kusunoki H, Minasov G, Macdonald RI, Mondragon A. Independent movement, dimerization and stability of tandem repeats of chicken brain alpha-spectrin. *J Mol Biol.* 2004; 344:495–511. [PubMed: 15522301]
24. Paramore S, Voth GA. Examining the influence of linkers and tertiary structure in the forced unfolding of multiple-repeat spectrin molecules. *Biophys J.* 2006; 91:3436–45. [PubMed: 16891371]
25. Ortega E, Buey RM, Sonnenberg A, de Pereda JM. The structure of the plakin domain of plectin reveals a non-canonical SH3 domain interacting with its fourth spectrin repeat. *J Biol Chem.* 2011
26. Li SS. Specificity and versatility of SH3 and other proline-recognition domains: structural basis and implications for cellular signal transduction. *Biochem J.* 2005; 390:641–53. [PubMed: 16134966]
27. Mayer BJ. SH3 domains: complexity in moderation. *J Cell Sci.* 2001; 114:1253–63. [PubMed: 11256992]
28. Bialkowska K, Saido TC, Fox JE. SH3 domain of spectrin participates in the activation of Rac in specialized calpain-induced integrin signaling complexes. *J Cell Sci.* 2005; 118:381–95. [PubMed: 15632109]
29. Lefferts JA, Wang C, Sridharan D, Baralt M, Lambert MW. The SH3 domain of alphaII spectrin is a target for the Fanconi anemia protein, FANCG. *Biochemistry.* 2009; 48:254–63. [PubMed: 19102630]
30. Rost B, Yachdav G, Liu J. The PredictProtein server. *Nucleic Acids Research.* 2004; 32:W321–W326. [PubMed: 15215403]
31. Laederach A, Cradic KW, Fulton DB, Andreotti AH. Determinants of intra versus intermolecular self-association within the regulatory domains of Rlk and Itk. *J Mol Biol.* 2003; 329:1011–20. [PubMed: 12798690]
32. Mitin N, Betts L, Yohe ME, Der CJ, Sondek J, Rossman KL. Release of autoinhibition of ASEF by APC leads to CDC42 activation and tumor suppression. *Nat Struct Mol Biol.* 2007; 14:814–23. [PubMed: 17704816]
33. Murayama K, Shirouzu M, Kawasaki Y, Kato-Murayama M, Hanawa-Suetsugu K, Sakamoto A, Katsura Y, Suenaga A, Toyama M, Terada T, Taiji M, Akiyama T, Yokoyama S. Crystal structure

- of the rac activator, Asef, reveals its autoinhibitory mechanism. *J Biol Chem.* 2007; 282:4238–42. [PubMed: 17190834]
34. Kami K, Takeya R, Sumimoto H, Kohda D. Diverse recognition of non-PxxP peptide ligands by the SH3 domains from p67(phox), Grb2 and Pex13p. *EMBO J.* 2002; 21:4268–76. [PubMed: 12169629]
35. Schiller MR, Chakrabarti K, King GF, Schiller NI, Eipper BA, Maciejewski MW. Regulation of RhoGEF activity by intramolecular and intermolecular SH3 domain interactions. *J Biol Chem.* 2006; 281:18774–86. [PubMed: 16644733]
36. Candel AM, Conejero-Lara F, Martinez JC, van Nuland NA, Bruix M. The high-resolution NMR structure of a single-chain chimeric protein mimicking a SH3-peptide complex. *FEBS Lett.* 2007; 581:687–92. [PubMed: 17275816]
37. Davis L, Abdi K, Machius M, Brautigam C, Tomchick DR, Bennett V, Michael P. Localization and structure of the ankyrin-binding site on beta2-spectrin. *J Biol Chem.* 2009; 284:6982–7. [PubMed: 19098307]
38. Ipsaro JJ, Harper SL, Messick TE, Marmorstein R, Mondragon A, Speicher DW. Crystal structure and functional interpretation of the erythrocyte spectrin tetramerization domain complex. *Blood.* 2010; 115:4843–52. [PubMed: 20197550]
39. Yu RC, Jahn R, Brunger AT. NSF N-terminal domain crystal structure: models of NSF function. *Mol Cell.* 1999; 4:97–107. [PubMed: 10445031]
40. Leslie, AGW. Autoindexing of rotation diffracton images and parameter refinement. In: Sawyer, L.; Isaacs, N.; Bailey, S., editors. *Proceedings of the CCP4 Study Weekend: "Data Collection and Processing"*, 29–30 January 1993. SERC Daresbury Laboratory; Daresbury, U.K.: 1993. p. 44–51.
41. Collaborative Computational Project, N. The CCP4 suite: programs for protein crystallography. *Acta Cryst.* 1994; D50:760–763.
42. Weeks CM, DeTitta GT, Hauptman HA, Thuman P, Miller R. Structure solution by minimal-function phase refinement and Fourier filtering. II. Implementation and applications. *Acta Crystallogr A.* 1994; 50 (Pt 2):210–20. [PubMed: 8166952]
43. de La Fortelle E, Bricogne G. Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol.* 1997; 276:472–494.
44. Terwilliger TC, Berendzen J. Automated structure solution for MIR and MAD. *Acta Cryst.* 1999; D55:849–861.
45. Terwilliger TC. Maximum-likelihood density modification. *Acta Crystallogr D Biol Crystallogr.* 2000; 56:965–72. [PubMed: 10944333]
46. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr.* 2004; 60:2126–32. [PubMed: 15572765]
47. Adams PD, Grosse-Kunstleve RW, Hung LW, Ioerger TR, McCoy AJ, Moriarty NW, Read RJ, Sacchettini JC, Sauter NK, Terwilliger TC. PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr.* 2002; 58:1948–54. [PubMed: 12393927]
48. Painter J, Merritt EA. Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D Biol Crystallogr.* 2006; 62:439–50. [PubMed: 16552146]
49. Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG, Thompson JD. Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.* 2003; 31:3497–500. [PubMed: 12824352]
50. Laskowski RA, Moss DS, Thornton JM. Main-chain bond lengths and bond angles in protein structures. *J Mol Biol.* 1993; 231:1049–67. [PubMed: 8515464]

Figure 1. Overall structure of DP(175–630)

(A) Desmoplakin primary structure. Desmoplakin consists of three major regions whose amino acid boundaries are indicated. A close up view of the plakin domain within DPNT is shown, with the colored regions representing the 4 spectrin repeats found in the crystal structure of DP(175–630).

(B) Overall structure of DP(175–630) shown as a ribbon representation and as an electrostatic surface (contoured at ± 5 kT/e), shown in two orientations. The N and C termini and each spectrin repeat are labeled. Each spectrin repeat is colored as shown in panel (A).

(C) Packing of the two copies of DP(175–630) in the asymmetric unit, shown in green and purple.

(D) The two molecules in the asymmetric unit after superposition of their SR3 domains.

Figure 2. Comparison of spectrin repeat structures of desmoplakin

(A) Helical regions of SR4, SR5, and SR6 are superimposed onto corresponding region of SR3. The color scheme is the same as in Figure 1B.

(B) Structure-based sequence alignments of SR3, SR4, SR5 and SR6 of desmoplakin. Secondary structure elements were assigned according to the program PROCHECK⁵⁰. The α helices in the spectrin repeats are enclosed by green colored boxes and the *a* and *d* positions in the heptad patterns are shown above the sequences. The SH3 domain with N- and C-terminal linkers is inserted into the loop between helices B and C of SR5 at the position indicated by the red asterisk. Magenta boxes indicate β -strands in the SH3 domain. Blue and red underlined residues at the N-terminal linker region of the SH3 domain represent potential cleavage sites for trypsin and chymotrypsin, respectively.

(C) Comparison of hydrophobic interactions in SR3 and SR6. Trp194 in SR3 is a conserved aromatic residue position found in the hydrophobic core of most spectrin repeats, but is replaced by Asp556 in SR6, which faces toward the solvent region. In SR6, Phe600, located in the loop between helices B and C of SR6, participates in hydrophobic core interactions.

Figure 3. SH3 domain structure

(A) Overall structure of the SH3 domain of desmoplakin. Beta-strands, RT loop, n-Src loop, and the N and the C termini are labeled.

(B) Comparison of canonical ligand-binding sites of the SH3 domains from desmoplakin and α -spectrin. The SH3 domain of α -spectrin and p41 peptide complex (PDB code 2JMA) (right) is shown in the same orientation as the SH3 domain of desmoplakin (left). SR4 and the SH3 domain of desmoplakin are displayed in the same colors used in Figure 1; the SH3 domain of α -spectrin and its peptide ligand are colored yellow and red, respectively. Amino acid residues that form the canonical binding pocket of the SH3 domain in α -spectrin and those from equivalent positions in the SH3 domain of desmoplakin are labeled, and their side chains are shown as sticks; the C α of glycine 509 is shown as a ball. Side chains of residues in SR4 interacting with the SH3 domain are also shown. The position of tryptophan 360 in SR4 overlaps that of proline in the p41 peptide ligand.

Figure 4. ARVD/C mutations in demoplakin

Residues mutated in ARVD/C are shown as sticks on ribbon representation of DP(175–630) in the same orientation as in the right figure of Figure 1b. Core residues, which contribute the structural integrity of DP(175–630) and surface residues are colored in cyan and red, respectively. A close up view of the SH3 domain and nearby regions is shown at the right.

Figure 5. Structural comparisons of DP(175–630) with other spectrin repeat proteins
 (A) Crystal structures of α -actinin (SRs1–4; PDB code 1HCI), α -spectrin (SR15–SR16–SR17; PDB code 1U4Q), erythroid β -spectrin (SR8–SR9; PDB code 1S35), plectin (SR1–SR2; PDB code 2ODU), and BPAG1e (SR3–SR4; PDB code 2IAK) were used for the comparison with the structure of DP(175–630). The N-terminal repeat of each protein was superimposed onto SR3 of desmoplakin (green). (B) The C-terminal repeat of BPAG1e, plectin, and β -spectrin are shown in apical view when their N-terminal repeats are superimposed into SR3 of DP(175–630). The color scheme is the identical to that in panel (A).

Table 1

Crystallographic statistics

A. Data collection		
Data set	Native	SAD
wavelength (Å)	1.0332	0.9792
Space group	C2	C2
Unit cell parameters		
<i>a</i> (Å)	157.5	168.1
<i>b</i> (Å)	82.5	80.5
<i>c</i> (Å)	140.0	150.3
β (°)	125.5	129.6
Resolution (Å) (last shell)	50–2.95 (3.03–2.95)	50–3.8 (3.9–3.8)
Unique reflections	30845 (2187)	14577 (1089)
Completeness (%)	99.5 (96.7)	95.8 (95.8)
Multiplicity	2.9 (2.6)	5.2 (5.2)
I/σ(I)	9.1 (2.0)	5.3 (1.3)
R _{merge} ^a	0.046 (0.365)	0.084 (0.428)
B. SAD Phasing		
Resolution (Å)		50–4.0
Phasing power ^b		2.55 / 2.58
figure-of-merit		0.35
C. Refinement		
Resolution (Å)	50–2.95	
No. of reflections working set (test set)	30776 (3090)	
No. residues	450 (A) / 447 (B)	
R _{cryst} / R _{free} ^c	0.216 / 0.265	
bond length rmsd from ideal (Å)	0.003	
bond angle rmsd from ideal (°)	0.59	
Temperature factors (Å ²)		
Wilson B from data	95.3	
Average B	97.8 (protein chain A) / 113.3 (protein chain B) / 157.1 (DTT ^d)	
Anisotropic B tensor	B ₁₁ = 6.2 / B ₂₂ = 0.7 / B ₃₃ = -6.9 / B ₁₃ = -10.1	
Ramachandran analysis ^e		
% most favored regions	94.9	
% additionally allowed regions	5.1	
% generously allowed regions	0.0	
% disallowed regions	0.0	

Values in parentheses are for the highest resolution shell. Rmsd, root-mean square deviation.

^aR_{merge}= $\sum_h \sum_i |I_i(h) - \langle I(h) \rangle| / \sum_h \sum_i I_i(h)$, where $I_i(h)$ is the i^{th} measurement of reflection h , and $\langle I(h) \rangle$ is the weighted mean of all measurements of h .

^bPhasing power = $\langle |F_H| \rangle / E$, where $\langle |F_H| \rangle$ is the rms structure factor amplitude for the anomalous scatterers and E is the estimated lack-of-closure error. Overall phasing powers are given separately for the + and – Friedel mates.

^c $R = \sum_h |F_{\text{obs}}(h)| - |F_{\text{calc}}(h)| / \sum_h |F_{\text{obs}}(h)|$. R_{cryst} and R_{free} were calculated using the working and test reflection sets, respectively.

^dThe model includes 2 molecules of reduced and 2 molecules of oxidized dithiothreitol.

^eAs defined in Procheck⁵⁰

Table 2

Sequence and structural similarities amongst SRs of desmoplakin and plectin

Plectin	DPNT	SR3	SR4	SR5
SR3		32[57] / 1.0	10[22] / 1.2	10[20] / 1.3
SR4		9[18] / 1.4	48[75] / 1.0	8[17] / 1.2
SR5		10[24] / 1.3	13[22] / 1.3	33[66] / 1.0

Entries indicate sequence identity (%) [similarity (%)] / rmsd after superposition (Å)

Sequences of each plectin repeat were aligned to each repeat of DPNT with ClustalW²⁴⁹ and sequence identities (%) and similarities (%) are shown. Structural superposition of each plectin repeat was carried out with Lsqkab program of the CCP4 suite⁴¹ using a set of about 75 C α atoms of the helical region in each repeat of DP(175–630) as a reference.