PDZ Tandem of Human Syntenin:
Crystal Structure and Functional Properties

Beom Sik Kang,[1,4] David R. Cooper,[1,4] Filip Jelen,[2,4]
Yancho Devedjiev,[1] Urszula Derewenda,[1]
Zbigniew Dauter,[3] Jacek Otlewski,[2]
and Zygmunt S. Derewenda[1,*]
[1]Department of Molecular Physiology and
   Biological Physics and
The Cancer Center
University of Virginia
Charlottesville, Virginia 22908
[2]Laboratory of Protein Engineering
Institute of Biochemistry and Molecular Biology
University of Wroclaw
Tamka 2
50-137 Wroclaw
Poland
[3]Synchrotron Radiation Research Section
Macromolecular Crystallography Laboratory
NCI
Brookhaven National Laboratory
Upton, New York 11973

Summary

Syntenin, a 33 kDa protein, interacts with several cell membrane receptors and with merlin, the product of the causal gene for neurofibromatosis type II. We report a crystal structure of the functional fragment of human syntenin containing two canonical PDZ domains, as well as binding studies for full-length syntenin, the PDZ tandem, and isolated PDZ domains. We show that the functional properties of syntenin are a result of independent interactions with target peptides, and that each domain is able to bind peptides belonging to two different classes: PDZ1 binds peptides from classes I and III, while PDZ2 interacts with classes I and II. The independent binding of merlin by PDZ1 and syndecan-4 by PDZ2 provides direct evidence for the coupling of syndecan-mediated signaling to actin regulation by merlin.

Introduction

Syntenin is a 298 residue long cytosolic protein, originally identified as a molecule linking syndecan-mediated signaling to the cytoskeleton [1]. Subsequently, syntenin was also found to play a role in protein trafficking [2, 3], cell adhesion [4], and activation of the transcription factor Sox4 [5]. Of particular medical interest is the recent report that syntenin is overexpressed in breast and gastric cancer cells and promotes their migration and metastasis [6]. The diverse biological functions of this protein are a result of its interactions with numerous targets. There are currently at least ten putative binding partners reported for syntenin, including IL-5 receptor α subunit (IL5Rα) [5], neuroglian [7], proTGF-α [3], gluta-

mate receptors [8], neurofascin [7], syndecan-4 [1], ephrin B [9, 10], ephrin A7 [9], PTP-η [11], neurexin I [12], and merlin [13]. All the binding partners of syntenin are receptors except for merlin, a cytosolic tumor repressor that is a product of the causal gene for type II neurofibromatosis (NF) [14]. Merlin belongs to the protein 4.1 superfamily, which also includes ezrin, moesin, and radixin, and like its homologs, it binds actin [15].

Based on amino acid sequence analyses, syntenin was predicted to contain a tandem of PDZ domains (PDZ1 and PDZ2) preceded by an N-terminal fragment of 112 amino acids of an unknown structure. PDZ domains are ubiquitous signaling domains, with over 400 distinct copies in the human genome [16, 17], which mediate protein-protein interactions. They may occur in proteins harboring other domains, such as SH2, RGSL, PH, DH, or GUK, but are also found in proteins that contain no other domains: an extreme example, MUPP, is a scaffolding protein with 13 PDZ domains [18]. Through the PDZ domains, signaling proteins bind to receptors, channels, and other targets, often functioning as membrane-associated scaffolds for the assembly of signaling complexes. Finally, PDZ-containing proteins provide a means for subcellular targeting of their partners, as exemplified by the function of Lin-2/CASK [19], Lin-10/MINT1 [20], and GRIP [21, 22].

PDZ domains are structurally conserved modules, about 90 amino acids in length, with a distinct fold of six β strands and two α helices [23, 24]. In most cases, they recognize C-terminal sequence motifs of target proteins and bind these peptides in a pocket between the β2 strand and α2 helix. The PDZ domains are typically grouped into three classes depending on the target tripeptides: class I (-S/T-X-ϕ), class II (-ϕ-X-ϕ), and class III (-D/E-X-ϕ) [17]. Examples outside this paradigm are well documented, and some PDZ domains show degenerate specificity [25]. It has also been reported that interaction between adjacent PDZ domains may modulate peptide binding, further complicating the picture. For example, the PDZ1-PDZ2 tandem within PSD-95 appears to have different binding properties compared to its isolated PDZ domains [16].

The multitude of syntenin's putative partners, which belong to all three classes of target proteins, suggests that its PDZ domains may also exhibit degenerate specificities. Furthermore, it has been reported that the two domains function in a cooperative fashion: for example, isolated PDZ1 and PDZ2 apparently fail to bind merlin- and IL5Rα-derived peptides, whereas binding was reported for full-length protein [5, 13]. A requirement for the complete tandem was also reported for interaction with PTP-η [11] and proTGF-α [3], whereas syndecan-2 was reported to bind to PDZ1-PDZ2 or PDZ2-PDZ2 tandem, but neither the isolated domains nor PDZ1-PDZ1 [12].

In order to explain the molecular basis for the ob-

served properties of syntenin, we have initiated a systematic study aimed at characterizing syntenin's molecular structure and mechanism of action. Here, we report the crystal structure of the intact PDZ tandem, residues 113–273, refined at 1.94 Å resolution. The structure reveals two PDZ domains that have fully solvent-accessible peptide binding pockets. We also report the results of rigorous biophysical binding assays for isolated PDZ domains, the PDZ tandem, and full-length syntenin, with peptides derived from three selected putative binding partners for syntenin: IL5Rα, syndecan-4, and merlin. These data reveal that the binding properties of syntenin are a result of the independent binding events of the two PDZ domains, whose specificities show clear degeneracy. The merlin-derived octapeptide shows the highest affinity for syntenin, and a distinct selectivity for PDZ1. This result reaffirms that merlin is a physiologically important partner for syntenin.

Results and Discussion

Model Quality and Structure Overview

The structure was solved with a three-wavelength MAD experiment, using a SeMet-labeled protein crystal. The model, refined at 1.94 Å resolution to a crystallographic R value of 18.4% (R_{free} 22.7%), contains a noncrystallographic dimer of tandems in the asymmetric unit and a total of 325 residues (Table 1; Figure 1). The only syntenin residue not included in the model is the C-terminal phenylalanine of the second monomer. The refined structure conforms well to standard protein stereochemistry, with rms deviation from ideal bond lengths of 0.015 Å, and with only 2 residues falling into disallowed regions of the Ramachandran plot as judged by Mol-Probity [26]. Only seven side chains are not visible in the σ_A-weighted $2mF_{obs} - DF_{calc}$ electron density map contoured at 1σ (Figure 1C), and their occupancies were set to zero. The average isotropic temperature factor (B) for main chain atoms is 20.6 Å^2, with the highest temperature factors ($\sim$50 Å^2) associated with the linker peptides and the C-terminal end of the α2 helix of PDZ1, all of which are nonetheless clearly visible.

The crystallized fragment of syntenin contains two PDZ domains conjoined by a short linker (Arg193-Pro194-Phe195-Glu196; Figure 1C). Like other domains from this superfamily, the syntenin PDZ modules show a typical fold with two opposing antiparallel β sheets capped by two α helices. Each domain has at least one β strand that is partly contained in both sheets. In the crystal, the two PDZ tandems of syntenin are arranged in a head-to-tail fashion, related by a noncrystallographic dyad, giving the contents of the asymmetric unit the appearance of a four-leaf clover. Interestingly, the linker residue Arg193 forms a salt bridge with Glu240 in PDZ2 and forms a hydrogen bond via its N$^\delta$ with a main chain carboxyl group of Phe154 in PDZ1. This may help to explain why Arg193 falls into a disallowed region of the Ramachandran plot. Superposition of the monomers reveals that there is a slight difference in the angle between the two PDZ domains in the two monomers, explaining why the dimer is noncrystallographic. This suggests that the linker has considerable intrinsic flexibility in solution.

We note that the interaction between the two PDZ domains within a monomer is less extensive than the intermolecular PDZ1-PDZ2 interface. A total of 446 Å^2 of solvent-accessible surface is buried by each intermolecular interaction. Furthermore, this interface is fairly intimate, with a number of hydrogen bonds between the two domains. The few solvent molecules that are at this interface mediate contacts between the two domains. In contrast, there are no direct hydrogen bonds between the PDZ domains within a monomer. Both putative peptide binding grooves of syntenin are located on the same face of the tandem monomer and are completely exposed to the solvent, suggesting that syntenin has two distinct and functional peptide binding sites.

Structure of the PDZ Domains of Syntenin

A structural comparison of the two PDZ domains of syntenin reveals that, in spite of a modest level of sequence identity (26%), they are structurally very similar to each other, with an rms deviation of 1.2 Å on Cα atoms. In both domains, the fragment equivalent to the signature GLGF loop involved in the terminal carboxylate binding deviates from the paradigm by an insertion of a basic residue after the initial glycine (Arg in PDZ1 and His in PDZ2). Such insertions are rarely found in PDZ domains, but they do not seem to disturb the cluster of main chain amides that coordinate the incoming carboxylate of the target peptide. Typically, a Lys or Arg located 4 or 5 residues prior to this loop assists in peptide binding via a water-mediated hydrogen bond. Both of syntenin's PDZ domains have a lysine 4 residues before the initial glycine.

In spite of these similarities, there are some notable differences between the two domains. The most apparent is the length of the β2-β3 loop. When compared to PDZ2, where this loop is shorter than in most other PDZ domains, PDZ1 contains an insertion of 4 residues in the β2-β3 loop (KSIDNGIF versus KN---GK).

Furthermore, in PDZ1, the peptide binding groove is narrower as compared to PDZ2 or other PDZ domains (Figure 2A). This is best illustrated by comparing the distance from the beginning of α2 to the β2 strand of PDZ2, to the corresponding distance in PDZ1, which is 1.8 Å shorter.

The electrostatic potential surrounding the peptide binding groove is another significant difference between the two PDZ domains. The peptide binding surface of PDZ1 is predominately positively charged, surrounded by 3 residues (Lys124, Arg128, and Lys130) from β2 and 2 residues (His175 and Lys179) from α2. Other basic residues flank this region. PDZ2 lacks any clusters of positively charged residues, with His208 as the only charged side chain that extends over the peptide binding groove (Figure 2B).

As a peptide binds to PDZ domains, it mimics an additional antiparallel strand in the sheet containing β2. The position of the β2 strand in both of syntenin's PDZ domains is consistent with this mechanism, with the amino and carboxyl groups of Leu129 and Phe211 available for hydrogen bonding. This type of arrangement dictates that the terminal side chain of the peptide faces the interior of the binding pocket. Both PDZ domains

Table 1. Data Collection and Refinement Statistics

	Edge	Peak	Remote
Data Collection Statistics			
Wavelength	0.97946	0.97900	0.97133
Resolution (Å)	30.0–1.94 (2.01–1.94)[a]	30.0–1.94 (2.01–1.94)[a]	30.0–1.94 (2.01–1.94)[a]
Total reflections	84,171	85,123	86,109
Unique reflections	22,557	22,608	23,095
Redundancy	3.7	3.8	3.7
Completeness (%)	97.4 (75.1)	97.1 (71.9)	98.8 (90.2)
R_{sym} (%)[b]	4.9 (25.9)	6.2 (31.2)	4.8 (31.1)
Average I/σ (I)	20.2	23.1	19.7
Phasing Statistics			
Phasing power,[c] iso/ano	0.37/0.27	0.23/0.35	−/0.31
R_{Cullis},[d] iso/ano	0.43/0.69	0.71/0.57	−/0.66
Overall figure of merit (acentric): 0.68			
Refinement Statistics			
Resolution (Å)	30.0–1.94 (1.99–1.94)		
Reflections (working)	21,926 (1,560)		
Reflections (test)	1,182 (79)		
R_{work} (%)[e]	18.4 (24.0)		
R_{free} (%)[e]	22.7 (25.2)		
Number of waters	254		
Rms deviation from ideal geometry			
Bonds (Å)	0.015		
Angles (°)	1.80		
Average B factor (Å²)			
Main chain	20.6		
Side chain	26.3		
Waters	47.0		

[a] The numbers in parentheses describe the relevant value for the highest resolution shell.
[b] $R_{sym} = \Sigma |I_i - \langle I \rangle| / \Sigma I$, where I_i is the intensity of the i-th observation and $\langle I \rangle$ is the mean intensity of the reflections. The values are for unmerged Friedel pairs.
[c] Phasing power = $\langle |Fh(calc)| / \text{phase-integrated lack of closure}] \rangle$
[d] $R_{Cullis} = \langle \text{phase-integrated lack of closure} \rangle / \langle |Fph - Fp| \rangle$
[e] $R = \Sigma ||F_{obs}| - |F_{calc}|| / \Sigma |F_{obs}|$, crystallographic R factor, and $R_{free} = \Sigma ||F_{obs}| - |F_{calc}|| / \Sigma |F_{obs}|$, where all reflections belong to a test set of randomly selected data.

have a substantial hydrophobic pocket near the binding loop that could accommodate any of the large hydrophobic side chains. The walls and floor of the peptide binding groove in both PDZ domains are lined with hydrophobic residues. Neither PDZ domains has a histidine at the first position of α2, as is found in typical class I PDZ domains. In syntenin, this position is occupied in PDZ1 and PDZ2 by Ser171 and Asp251, respectively. The side chains of both of these residues hydrogen bond to main chain amides at the end of β2. Overall, both of the PDZ domains of syntenin appear to be suitable for peptide binding, although the structural differences suggest diverse specificities.

Stability Studies

Although the crystal structure indicates that both PDZ domains of syntenin are capable of binding peptides, many previously reported binding studies using isolated PDZ1 and PDZ2 domains have failed. In order to better assess the feasibility of performing binding assays with isolated PDZ1 or PDZ2, we conducted stability studies using chemical denaturation monitored by circular dichroism (Figure 3). Full-length syntenin, the PDZ tandem, and isolated PDZ1 (residues 113–193) and PDZ2 (residues 197–273) were used in the assay. Surprisingly,

we found that the isolated PDZ1 and PDZ2 have significantly different stabilities: the free energy of unfolding, ΔG_{un}, for PDZ1 is −3.2 kcal/mol, whereas for PDZ2 it is −4.8 kcal/mol, putting it closer to the average of 5–15 kcal/mol observed for globular proteins. Based on these values, the expected denaturation of the tandem, as simulated by combining single domain transitions, should be less cooperative, with a ΔG_{un} of −2.07 kcal/mol (Figure 3, insert). However, the experimental unfolding of the tandem follows a cooperative, two-state profile, with a ΔG_{un} of −4.1 kcal/mol, suggesting that the domains are associated into a single entity undergoing cooperative denaturation. As already noted, PDZ1 of one molecule interacts with PDZ2 of the adjacent one in the crystal structure. The large buried interface suggests that the interaction could be physiologically relevant, as indeed self-association has been reported for syntenin [7]. It is also possible that the two domains interact in this way within a monomer, and that the crystal structure corresponds to a domain-swapped conformation. Whereas our data are consistent with weak, albeit identifiable domain-domain interactions in syntenin, further work will be required to probe this issue.

Finally, we note that the full-length protein also unfolds in a highly cooperative manner, and shows signifi-
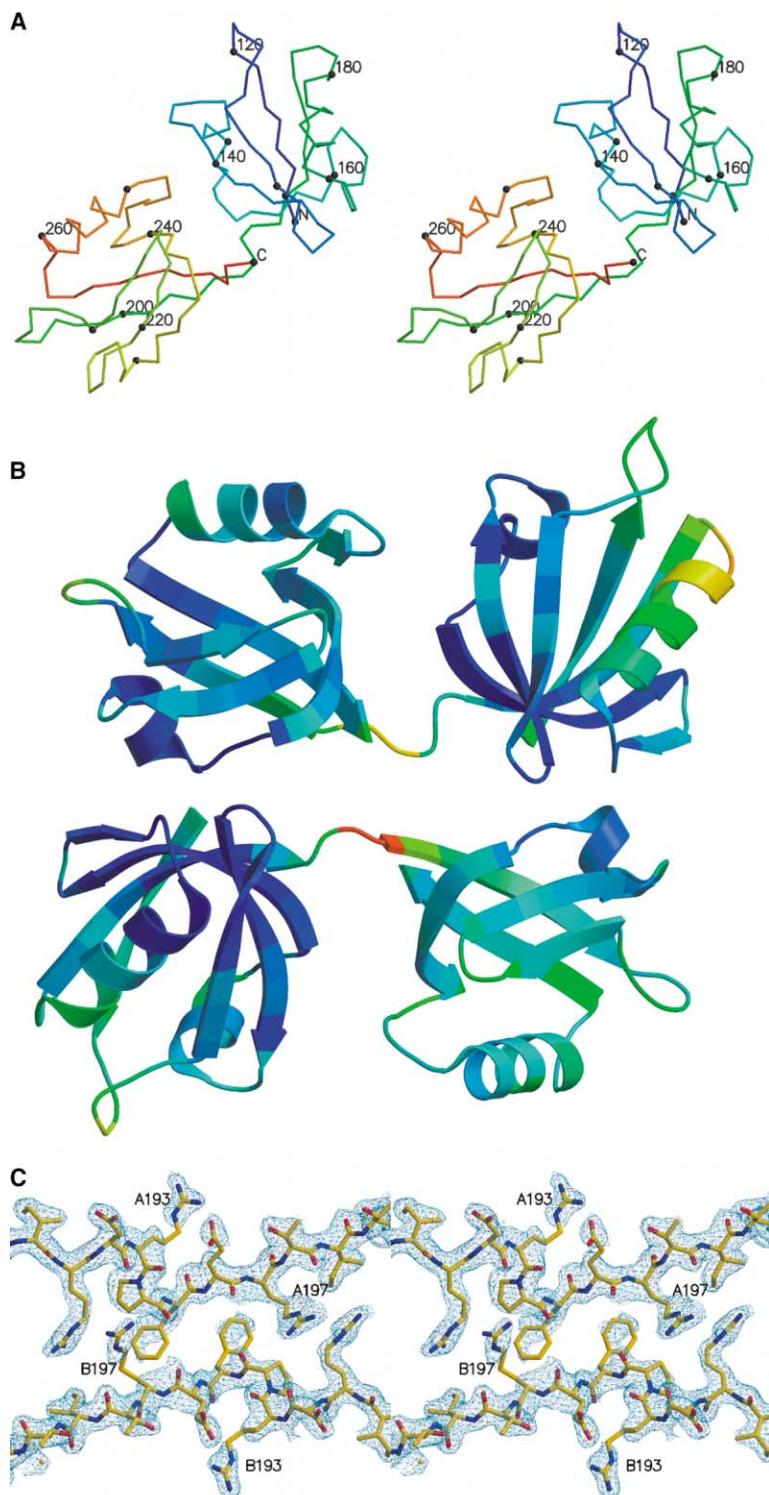
A

B

C

Figure 1. The Structure of the PDZ Tandem of Syntenin

(A) A stereo Cα trace with every tenth α carbon represented as a sphere and every twentieth α carbon labeled, and colored from blue to red as a function of residue number.

(B) Ribbon diagram of the asymmetric unit colored by B factors. B factors are represented with low values (12 Å^2) colored blue and high values (43 Å^2) colored red.

(C) Experimentally determined electron density map of the linker region contoured at 1σ. Residues 189–201 are shown for each monomer. Figures were made using MOLSCRIPT [47] (A and B) and POVSCRIPT+ (http://people.brandeis.edu/~fenn/povscript/) (C) and rendered with RASTER3D.

cantly higher stability (ΔG$_{un}$ of −6.4 kcal/mol) than any of the other constructs. This result may imply that the N-terminal fragment of syntenin is folded and plays a structural role in the protein, possibly interacting with the PDZ domain(s), so that full-length syntenin denatures as a single entity.

Binding Properties of Syntenin and of Isolated PDZ Domains

In the case of full-length syntenin and the PDZ tandem, we were interested in both the affinities and the stoichiometries of peptide binding. With that in mind, we carried out binding assays using isothermal titration calorimetry
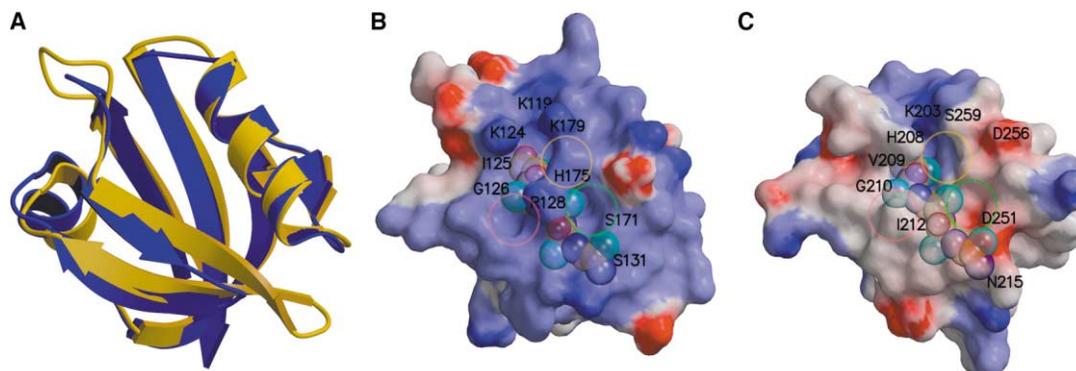
Figure 2. A Comparison of PDZ1 and PDZ2 Domains of Syntenin

(A) Superposition of the two PDZ domains of syntenin. PDZ1 is gold and PDZ2 is blue. The $\alpha2$ helices have been superposed to show the similarity of the fold, yet emphasize the differences of the peptide binding groove. The same orientation is used for all three figures.

(B) The peptide binding surface of PDZ1. The electrostatic potential surface is shown with select residues that surround the peptide binding groove labeled. A superposed C-terminal CRIPT-derived peptide from the structure of PSD-95 (1BE9 [23]) is shown semitransparent, with side chains represented as cyan spheres in the β carbon position. The approximate locations of the P_0, P_{-1}, and P_{-2} binding pockets are indicated by gold, pink, and green circles, respectively.

(C) The peptide binding surface of PDZ2 represented as described in (B). Figures were made using MOLSCRIPT [47] (A) and POVSCRIPT+ (http://people.brandeis.edu/~fenn/povscript/) (B and C) and rendered with RASTER3D [48]. Electrostatic potentials were calculated in GRASP [49].

(ITC). The assays were conducted with the following hexapeptides derived from three putative targets of syntenin, each belonging to one of the three canonical classes: LEDSVF (IL5Rα) representing class I, TNEFYA (syndecan-4) from class II, and AFFEEL (merlin) from class III. We found that all peptides bind to full-length syntenin and to PDZ tandem with dissociation constants (K_d) in the low μM range (Figure 4; Table 2). Interestingly, the IL5Rα-derived peptide shows a stoichiometry of 2:1 for the tandem, but only 1:1 for full-length syntenin, whereas all other measurements indicate equimolar complexes. This is further evidence that suggests a functional role for the N-terminal domain.

To assess whether residues upstream of the C-terminal hexapeptide contribute to binding, we performed the assays with octapeptides for merlin and IL5Rα. The results are very similar for IL5Rα, but the merlin octapeptide binds an order of magnitude more tightly than the corresponding hexapeptide, indicating the functional importance of residues -6 and -7.

The determination of K_d values for isolated domains by ITC proved difficult, because PDZ1 aggregated at the required high concentration and isolated PDZ2 (residues 197–273) was prone to oligomerization (data not shown). To overcome the aggregation problems, we used a fluorescence-based approach using dansylated hexapeptides, allowing for significantly lower protein concentration [27]. We also changed the PDZ2 construct to residues 197–298, which includes syntenin's natural C terminus. To assess whether either the technique or
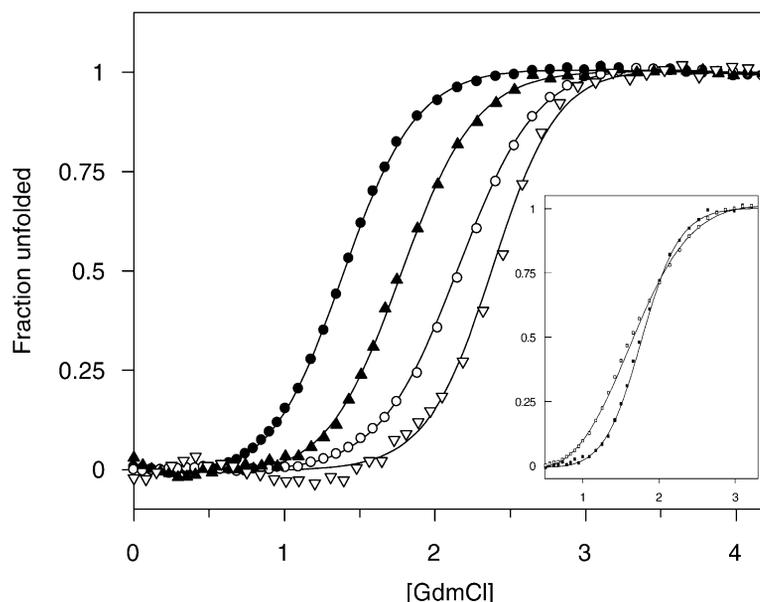

Figure 3. Stability of Syntenin Constructs

GdmCl-induced unfolding of PDZ1 (●), PDZ2 (○), tandem of PDZ domains (▲), and full-length (▽) of syntenin. Measurements were performed in 25 mM Tris, 50 mM NaCl (pH 7.4). Transitions were monitored by the changes of the CD signal at 222 nm. Data were normalized as "fraction unfolded" and fitted to the equation in the text. Insert: combined single domain transitions (□) and tandem of PDZ domains transition (■).
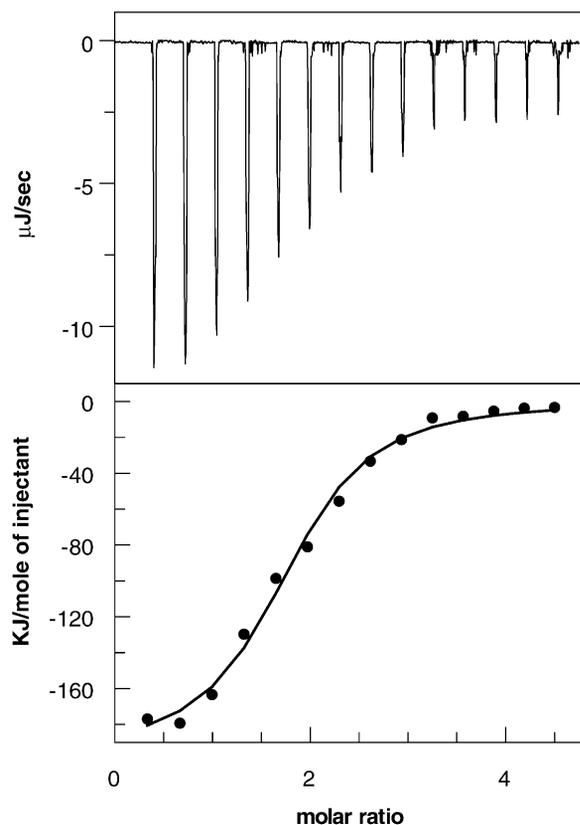
Figure 4. Representative Calorimetric Titration of PDZ Tandem of Syntenin with LEDSVF Peptide

Top: raw heat data corrected for base drift, obtained from 14 consecutive injections of 11.2 mM LEDSVF peptide into a sample cell (1,250 μl) containing 140 μM PDZ tandem of syntenin.
Bottom: the binding isotherm created by plotting the areas under the peaks against the molar ratio of the peptide added to the PDZ tandem present in the cell and the fit line to the model of independent sites. The heats of mixing (dilution) have been subtracted.

peptide dansylation influenced the observed affinities, we included the PDZ tandem in our measurements as a control. We found that the fluorescence data agree well with the calorimetric titrations (Figure 5; Table 3). The IL5Rα peptide binds to both PDZ domains with similar K_d values in the mid-μM range, in agreement
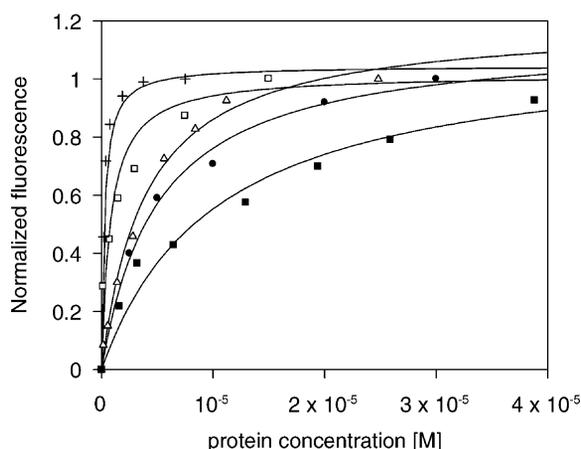

Figure 5. Binding of Dansyl-Labeled Peptides to Syntenin

Binding of dansyl-RVAFFEEL to PDZ1 (+), dansyl-AFFEEL to PDZ1 (●), dansyl-LEDSVF to PDZ1 (■), and PDZ2 (□) and dansyl-TNEFYA to PDZ2 (Δ). Data were normalized as "fraction bound," so that the initial fluorescence was zero and the fluorescence at saturation was equal to unity.

with the 2:1 stoichiometry observed by ITC. The merlin-derived peptide shows no significant affinity toward PDZ2, but binds to PDZ1, to the tandem and the full-length protein, with almost identical K_d values in the sub-μM range. The merlin octapeptide binds to PDZ1 with significantly higher affinity than the hexapeptide, in agreement with the ITC results. The syndecan-4 peptide interacts exclusively with PDZ2, with an affinity virtually identical to those observed for the PDZ tandem and full-length syntenin. This result is again consistent with the 1:1 stoichiometry determined by ITC. However, it is in conflict with the previously reported 2:1 stoichiometry for the whole C-terminal domain of syndecan-2 [12].

It has been suggested previously that the N-terminal fragment of syntenin plays a regulatory function. For example, the association of PTP-η with syntenin was shown to be regulated by tyrosine phosphorylation within this fragment, with phosphorylation preventing the association [11]. This indicates that the N-terminal fragment may, at least under some conditions, regulate the availability of at least one of syntenin's peptide binding pockets. Our data do not provide a clear answer in

Table 2. Isothermal Calorimetry of Syntenin Interactions

	LEDSVF (IL5Rα)	ETLEDSVF (IL5Rα)	AFFEEL (Merlin)	RVAFFEEL (Merlin)	TNEFYA (Syndecan)
PDZ Tandem					
K_d	43.8 μM	32.2 μM	11.6 μM	200 nM	2.9 μM
n	2.26	2.07	0.94	1.12	1.11
ΔH	−2.1 kJ	−7.6 kJ	−7.1 kJ	−3.1 kJ	−5.4 kJ
Full-Length					
K_d	19.5 μM	10.1 μM	8.9 μM	869 nM	2.5 μM
n	1.09	1.12	1.3	1.16	1.14
ΔH	−4.7 kJ	−8.87	−8.9 kJ	−7.3 kJ	−9.8 kJ

Dissociation constants, stoichiometries, and enthalpies for the interactions of the IL5Rα-, merlin-, and syndecan-derived peptides, with the PDZ tandem of syntenin and full-length protein, determined by ITC. Representative data are shown for experiments that were conducted at least twice for each interaction.

Table 3. Dissociation Constants of Syntenin Interactions by Fluorimetric Titrations

	DNS-EDSVF (IL5Rα)	DNS-FFEEL (Merlin)	DNS-VAFFEEL (Merlin)	DNS-NEFYA (Syndecan)
PDZ-PDZ	18.5 μM (±3.2)	6.8 μM (±1.1)	495 nM (±55)	2.1 μM (±0.3)
PDZ1	10.6 μM (±0.8)	5.0 μM (±0.7)	268 nM (±28)	>1 mM
PDZ2	1.9 μM (±0.3)	>0.6 mM	>1 mM	2.3 μM (±0.5)

Dissociation constants for the interactions of the IL5Rα-, merlin-, and syndecan-derived dansylated peptides, with the PDZ tandem of syntenin and isolated PDZ domains, determined by fluorimetric titration. Estimates of error are derived from experimental data.

this regard. The ITC data show that the IL5Rα peptide binds to only one PDZ domain of full-length syntenin, but to both in the tandem. Merlin, which binds only to PDZ1, does so within the context of full-length syntenin as well, suggesting that it is the PDZ2 that is occluded by the N-terminal domain. However, the syndecan-4 peptide is selective for PDZ2 and also binds to full-length syntenin. Further experiments will be necessary to resolve this inconsistency.

Taken together, the structural and binding data indicate that the two domains within the syntenin PDZ tandem function independently. Each domain shows degenerate specificity, so that PDZ1 binds peptides from merlin and IL5Rα, whereas PDZ2 shows affinity toward IL5Rα and syndecan-4. Although our data are internally consistent and reproducible, they are in conflict with some reports in the literature that claim individual PDZ domains are incapable of binding peptides.

Similarities to Other PDZ Domains
As the number of known PDZ domains grows and their importance in a myriad of cellular events becomes evident, numerous attempts have been made to elucidate the factors that govern their specificity. High-resolution crystal structures and solution NMR structures have now been determined for a number of PDZ domains that were classified into distinct groups. It is clear that the overall fold of the domain is well conserved, and the specificity is governed by subtle structural and amino acid sequence variation. The application of generalized rules for governing PDZ domain specificity is complicated by those PDZ domains that show degenerate specificity for more than one archetypal class of peptide. Syntenin is one of the examples of this growing group. Both of syntenin's PDZ domains fit the overall fold of PDZ domains well, with an average rms deviation from the known X-ray structures of 1.4 Å and 1.1 Å for PDZ1 and PDZ2, respectively.

Syntenin, with its tandem PDZ structure, appears to have been well conserved during evolution. The rat and mouse syntenins are virtually identical to the human protein. Recently, the jumbo tiger shrimp *Penaeus monodon* was reported to contain a protein similar to syntenin with extremely high amino acid sequence identity, when compared to the human protein, of 56% and 64%, for PDZ1 and PDZ2 domains, respectively [28]. We conducted a BLAST search of the genome of the malaria vector *Anopheles gambiae*, and found a protein annotated as a syntenin, with 50% amino acid identity to the human molecule. It is noteworthy that both the *Anopheles* and *Penaeus* homologs are far more similar to the human protein than any other PDZ domain in the

human genome (Figure 6). *Ciona intestinalis*, a primitive tunicate with the smallest known genome among *Chordata*, shows the presence of sequences highly similar (50%–60% identity) to the human protein. The high sequence similarity among such diverse species suggests that the molecule predates the appearance of vertebrates.

Syntenin: A Link between Syndecan and the Actin Cytoskeleton
The biological function of syntenin and its domain structure appear to have been stringently guarded by evolutionary mechanisms. The present study strongly supports earlier suggestions that merlin is a physiologically relevant partner for syntenin. Because merlin is an actin binding protein, syntenin may provide another link for syndecan-regulated signaling to the cytoskeleton, with syntenin mediating the colocalization of syndecan, through PDZ2, and merlin, through PDZ1. It will be comparable to the current model of syndecan signaling to actin, through PDZ-containing CASK and protein 4.1 [29] or direct binding to another of the ERM proteins, ezrin [30, 31]. The FERM domain of merlin binds ezrin and could block the interaction of ezrin to actin [31]. This alternate anchoring signal pathway may give clues regarding the involvement of syntenin in metastasis or the tumor suppressor function of merlin. The identification of the syntenin homolog in *Anopheles* prompted us to look in the mosquito's genome for the homologs of merlin and syndecan. We found an annotated merlin homolog with a 57% amino acid identity to the human protein and fully conserved C-terminal RVAFFEEL sequence. Similarly, we found the presence of a syndecan-related protein, with a highly conserved C terminus containing the TNEFYA motif.

Biological Implications

Syntenin is a ubiquitous protein involved in protein targeting and multiprotein assembly, and it is overexpressed in certain cancer cell lines. As inferred from numerous yeast two-hybrid screens and other biochemical assays, syntenin binds biologically important receptors such as IL5Rα and syndecan, as well as the cytosolic actin regulator merlin, which is a tumor suppressor and a product of the causal gene of neurofibromatosis type II. The crystal structure of the biologically functional fragment of syntenin, residues 113–273, solved at 1.94 Å resolution, reveals the presence of two canonical PDZ domains, connected by a 4 residue linker. Both domains appear to be free to interact with target peptides. It is the first crystal structure containing more than one PDZ
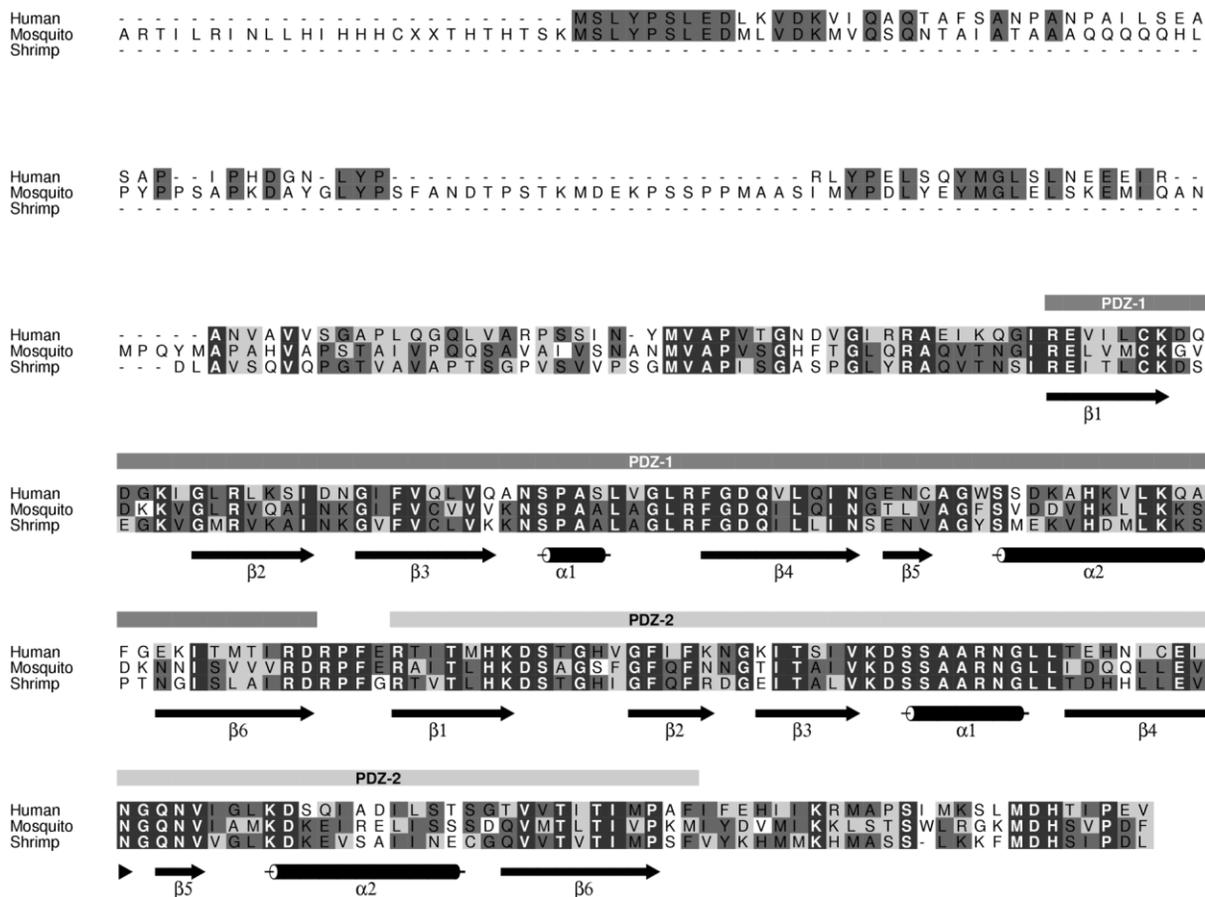
Figure 6. Amino Acid Sequence Alignment of Human Syntenin with the *Anopheles* and *Penaeus* Homologs

The secondary structural elements shown correspond to the PDZ tandem presented in this work.

domain from a single protein. Our binding studies, using stringent biophysical techniques such as isothermal titration calorimetry and fluorimetric titration, show that the properties of the tandem are the sum of the binding properties of the individual domains, with no detectable cooperative effects. Each domain is able to bind peptides belonging to two different classes: PDZ1 binds peptides corresponding to merlin and IL5Rα (classes I and III), whereas PDZ2 interacts with peptides derived from IL5Rα and syndecan-4 (classes I and II). The separate interactions of merlin with PDZ1 and that of syndecan-4 with PDZ2 suggest the physiological coupling of syndecan to merlin through syntenin. Because merlin binds actin, this pathway could be vital for merlin's function as a tumor repressor. The recently completed genome of the malaria vector *Anopheles gambiae* contains homologs of both syntenin and merlin, indicating that this pathway has been conserved during evolution.

Experimental Procedures

Expression and Purification of Protein Samples
A syntenin clone was obtained from American Tissue and Culture Collection (ATCC 72537). The DNA encoding full-length (residues 1–298), PDZ tandem (113–273), PDZ1 (113–193), and PDZ2 (197–273 and 197–298) domains of syntenin were amplified by PCR and cloned into the parallel vector pGST-parallel1 [32], a GST-fusion

protein expression vector containing the recombinant TEV protease (rTEV) cleavage site. The integrity of the insert was verified by direct DNA sequencing. The expression of the proteins was induced by 1 mM IPTG in *E. coli* BL21 strain (Stratagene). The SeMet-labeled PDZ tandem was expressed in the D834 strain (Novagen) with M9 medium with addition of SeMet. The expressed proteins were purified by affinity chromatography using a glutathione-Sepharose 4B (Amersham Pharmacia Biotech). The recombinant protein was subjected to a HiPrep 26/10 desalting column (Amersham Pharmacia Biotech) equilibrated with 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, and was digested using rTEV (Life Technologies) at 10°C in the presence of 0.5 mM EDTA, 1 mM DTT. After complete digestion, the GST tag was removed using a glutathione-Sepharose 4B column. A gel filtration was performed with a Superdex G75 column (Amersham Pharmacia Biotech) equilibrated with 20 mM Tris-HCl (pH 7.5), and the fractions containing the PDZ tandem were collected and concentrated using Centriprep YM10 for crystallization screening. The purified proteins contain an additional five amino acids (GAMDP) at the N terminus due to the cloning procedure.

Crystallization and Data Collection
Crystal Screen (Hampton Research) was used for preliminary screening. Subsequently, crystallization conditions were optimized around 0.1 M sodium acetate (pH 4.6) containing 24% PEG4000 and 0.2 M ammonium acetate. The sitting drop vapor diffusion method was used for all crystallization trials. Drops were formed with 3 μl of protein solution (6 mg/ml) and 3 μl of reservoir buffer, and were overlaid with 15 μl of a 1:1 mixture of silicon and mineral oil. Crystallization trays were stored at 21°C. The best crystals were obtained after microseeding. The crystals used for data collection

were briefly soaked in the crystallization buffer containing 12.5% (v/v) glycerol and frozen by immersion in liquid nitrogen. The structure was solved using a three-wavelength MAD experiment with SeMet-labeled crystals, with 180 images (1° rotation) collected for each wavelength. Data were collected at 0.97946 Å (edge wavelength), 0.97133 Å (remote wavelength), and 0.97900 Å (peak wavelength). The crystals are in space group C2, with unit cell parameters a = 100.7 Å, b = 48.7 Å, c = 74.7 Å, and β = 120.8°. All data were collected at beamline X9B at the NSLS, and processed and scaled using HKL2000 [33]. Data collection statistics are presented in Table 1. The programs SnB [34] and SHELXS [35] were used to locate six of the eight selenium atoms in the asymmetric unit. Phases were generated in SHARP [36] and improved by density modification in SOLOMON [37]. These phases were used as the starting point for automatic model building in ARP/wARP [38]. This generated 275 of the 332 residues in the asymmetric unit. After manually determining which of the resulting polypeptide chains belonged to each monomer in the asymmetric unit, ARP/wARP [38] was used to dock the side chains. Manual model building was performed in O [39]. All but one of the 322 syntenin residues in the asymmetric unit were included in the model, as were 4 N-terminal residues that are an artifact of subcloning. Solvent was added using ARP/wARP [38]. A combination of CNS [40] and REFMAC5 [41] was used to refine this initial model to an R factor of 18.4% and an R_{free} of 22.7%. Maximum likelihood residuals were used throughout the refinement process. TLS refinement [42] and inclusion of experimental phase information [43] from SHARP were included in later stages of refinement to minimize the difference between R_{work} and R_{free}. MOLPROBITY [26] and OOPS2 [44] were used as validation tools during refinement and rebuilding. Refinement data are presented in Table 1.

Calorimetric Binding Assays

Prior to the experiment, the protein solution was extensively dialyzed at 4°C against 25 mM phosphate or 25 mM Tris-HCl, 150 mM NaCl (pH 8.0). The titration was performed using a 4200 isothermal titration calorimeter (CSC). The protein concentration in the sample cell was in the range of 0.1 to 0.3 mM with a cell volume of 1,250 μl. The titrated peptides were dissolved to the concentrations in the range of 3 to 8 mM in dialysis buffer and injected in 5–10 μl aliquots. All experiments were done at 25°C. The titration thermogram was analyzed with BindWorks Applied Thermodynamics software to obtain n, K_d, and ΔH values. Concentration of PDZ tandem and full-length syntenin was estimated using the A_{280} molar absorbance coefficient calculated from the number of Trp and Tyr residues [45]. The concentration of PDZ2 and peptides was estimated using the A_{257} molar absorbance coefficient calculated from the number of Phe residues in the molecules.

Fluorimetric Titrations

Binding of peptides to full-length syntenin and PDZ tandem and isolated domains did not produce detectable change in fluorescence. Therefore, N-terminally dansylated peptides were used to increase sensitivity. The concentration of dansylated peptide was determined using the molar absorbance coefficient of the dansyl group ϵ_{334} = 4,600 $M^{-1}cm^{-1}$. The binding was monitored by following the increase in fluorescence upon titration of a concentrated protein into a 1 cm × 1 cm stirred cell cuvette containing 1.2 ml of 25 mM Tris-HCl, 150 mM NaCl (pH 7.5) and 0.5 μM dansylated peptide. The protein stock concentration was in the range of 1–1.5 mM and the signal was corrected for the dilution factor. Data were fitted to the following equation [46] by nonlinear least squares analysis using the program Grafit 3.01 (Erithacus Software):

$$y = F_o + \frac{(F_{max} - F_o)\frac{x}{K_d}}{1 + \frac{x}{K_d}} \tag{1}$$

where y is the fluorescence signal, x is the concentration of ligand, K_d is the dissociation constant, F_o is the initial fluorescence value, and F_{max} is the fluorescence value at saturation. Experiments were done in duplicate at 21°C using an FP-750 spectrofluorimeter (Jasco)

under following conditions: λ excitation = 335 nm, λ emission = 540 nm, and excitation and emission slit width = 5 nm.

Stability Measurements

Solvent denaturations were performed on a J-715 spectropolarimeter (Jasco) at 21°C with the automatic titrator (Jasco automatic titration system) in 25 mM Tris-HCl, 50 mM NaCl (pH 7.4) and the indicated concentration of guanidinium chloride (GdmCl). The transitions were monitored by the decrease of the CD signal at 222 nm and at 2 nm bandwidth. The apparent free-energy changes in the absence of GdmCl (ΔG_un) were determined by fitting the ellipticity intensity changes at particular concentrations of GdmCl to the equation given elsewhere [46]. Analysis of the data was performed by the program Grafit 3.01 (Erithacus Software). GdmCl concentration was determined by refractometry.

References

1. Grootjans, J.J., Zimmermann, P., Reekmans, G., Smets, A., Degeest, G., Durr, J., and David, G. (1997). Syntenin, a PDZ protein that binds syndecan cytoplasmic domains. Proc. Natl. Acad. Sci. USA *94*, 13683–13688.
2. Fialka, I., Steinlein, P., Ahorn, H., Bock, G., Burbelo, P.D., Haberfellner, M., Lottspeich, F., Paiha, K., Pasquali, C., and Huber, L.A. (1999). Identification of syntenin as a protein of the apical early endocytic compartment in Madin-Darby canine kidney cells. J. Biol. Chem. *274*, 26233–26239.
3. Fernandez-Larrea, J., Merlos-Suarez, A., Urena, J.M., Baselga, J., and Arribas, J. (1999). A role for a PDZ protein in the early secretory pathway for the targeting of proTGF-α to the cell surface. Mol. Cell *3*, 423–433.
4. Zimmermann, P., Meerschaert, K., Reekmans, G., Leenaerts, I., Small, J.V., Vandekerckhove, J., David, G., and Gettemans, J. (2002). PIP₂-PDZ domain binding controls the association of syntenin with the plasma membrane. Mol. Cell *9*, 1215–1225.
5. Geijsen, N., Uings, I.J., Pals, C., Armstrong, J., McKinnon, M., Raaijmakers, J.A., Lammers, J.W., Koenderman, L., and Coffer, P.J. (2001). Cytokine-specific transcriptional regulation through an IL-5Rα interacting protein. Science *293*, 1136–1138.
6. Koo, T.H., Lee, J.J., Kim, E.M., Kim, K.W., Kim, H.D., and Lee, J.H. (2002). Syntenin is overexpressed and promotes cell migration in metastatic human breast and gastric cancer cell lines. Oncogene *21*, 4080–4088.
7. Koroll, M., Rathjen, F.G., and Volkmer, H. (2001). The neural cell recognition molecule neurofascin interacts with syntenin-1 but not with syntenin-2, both of which reveal self-associating activity. J. Biol. Chem. *276*, 10646–10654.
8. Hirbec, H., Perestenko, O., Nishimune, A., Meyer, G., Nakanishi, S., Henley, J.M., and Dev, K.K. (2002). The PDZ proteins PICK1, GRIP, and syntenin bind multiple glutamate receptor subtypes. Analysis of PDZ binding motifs. J. Biol. Chem. *277*, 15221–15224.
9. Torres, R., Firestein, B.L., Dong, H., Staudinger, J., Olson, E.N., Huganir, R.L., Bredt, D.S., Gale, N.W., and Yancopoulos, G.D. (1998). PDZ proteins bind, cluster, and synaptically colocalize with Eph receptors and their ephrin ligands. Neuron *21*, 1453–1463.
10. Lin, D., Gish, G.D., Songyang, Z., and Pawson, T. (1999). The

carboxyl terminus of B class ephrins constitutes a PDZ domain binding motif. J. Biol. Chem. *274*, 3726–3733.

11. Iuliano, R., Trapasso, F., Sama, I., Le Pera, I., Martelli, M.L., Lembo, F., Santoro, M., Viglietto, G., Chiariotti, L., and Fusco, A. (2001). Rat protein tyrosine phosphatase η physically interacts with the PDZ domains of syntenin. FEBS Lett. *500*, 41–44.

12. Grootjans, J.J., Reekmans, G., Ceulemans, H., and David, G. (2000). Syntenin-syndecan binding requires syndecan-synteny and the co-operation of both PDZ domains of syntenin. J. Biol. Chem. *275*, 19933–19941.

13. Jannatipour, M., Dion, P., Khan, S., Jindal, H., Fan, X., Laganiere, J., Chishti, A.H., and Rouleau, G.A. (2001). Schwannomin isoform-1 interacts with syntenin via PDZ domains. J. Biol. Chem. *276*, 33093–33100.

14. Evans, D.G., Sainio, M., and Baser, M.E. (2000). Neurofibromatosis type 2. J. Med. Genet. *37*, 897–904.

15. den Bakker, M.A., Riegman, P.H., Suurmeijer, A.P., Vissers, C.J., Sainio, M., Carpen, O., and Zwarthoff, E.C. (2000). Evidence for a cytoskeleton attachment domain at the N-terminus of the NF2 protein. J. Neurosci. Res. *62*, 764–771.

16. Sheng, M., and Sala, C. (2001). PDZ domains and the organization of supramolecular complexes. Annu. Rev. Neurosci. *24*, 1–29.

17. Hung, A.Y., and Sheng, M. (2002). PDZ domains: structural modules for protein complex assembly. J. Biol. Chem. *277*, 5699–5702.

18. Lee, S.S., Glaunsinger, B., Mantovani, F., Banks, L., and Javier, R.T. (2000). Multi-PDZ domain protein MUPP1 is a cellular target for both adenovirus E4-ORF1 and high-risk papillomavirus type 18 E6 oncoproteins. J. Virol. *74*, 9680–9693.

19. Hsueh, Y.P., Wang, T.F., Yang, F.C., and Sheng, M. (2000). Nuclear translocation and transcription regulation by the membrane-associated guanylate kinase CASK/LIN-2. Nature *404*, 298–302.

20. Okamoto, M., and Sudhof, T.C. (1997). Mints, Munc18-interacting proteins in synaptic vesicle exocytosis. J. Biol. Chem. *272*, 31459–31464.

21. Dong, H., O'Brien, R.J., Fung, E.T., Lanahan, A.A., Worley, P.F., and Huganir, R.L. (1997). GRIP: a synaptic PDZ domain-containing protein that interacts with AMPA receptors. Nature *386*, 279–284.

22. Dong, H., Zhang, P., Song, I., Petralia, R.S., Liao, D., and Huganir, R.L. (1999). Characterization of the glutamate receptor-interacting proteins GRIP1 and GRIP2. J. Neurosci. *19*, 6930–6941.

23. Doyle, D.A., Lee, A., Lewis, J., Kim, E., Sheng, M., and MacKinnon, R. (1996). Crystal structures of a complexed and peptide-free membrane protein-binding domain: molecular basis of peptide recognition by PDZ. Cell *85*, 1067–1076.

24. Morais Cabral, J.H., Petosa, C., Sutcliffe, M.J., Raza, S., Byron, O., Poy, F., Marfatia, S.M., Chishti, A.H., and Liddington, R.C. (1996). Crystal structure of a PDZ domain. Nature *382*, 649–652.

25. Bezprozvanny, I., and Maximov, A. (2001). Classification of PDZ domains. FEBS Lett. *509*, 457–462.

26. Lovell, S.C., Davis, I.W., Arendall, W.B., de Bakker, P.I.W., Word, J.M., Prisant, M.G., Richardson, J.S., and Richardson, D.C. (2002). Structure validation by C-α geometry: φ, ψ and C-β deviation. Proteins *50*, 437–450.

27. Harris, B.Z., Hillier, B.J., and Lim, W.A. (2001). Energetic determinants of internal motif recognition by PDZ domains. Biochemistry *40*, 5921–5930.

28. Bangrak, P., Graidist, P., Chotigeat, W., Supamattaya, K., and Phongdara, A. (2002). A syntenin-like protein with postsynaptic density protein (PDZ) domains produced by black tiger shrimp *Penaeus monodon* in response to white spot syndrome virus infection. Dis. Aquat. Organ. *49*, 19–25.

29. Cohen, A.R., Woods, D.F., Marfatia, S.M., Walther, Z., Chishti, A.H., Anderson, J.M., and Wood, D.F. (1998). Human CASK/LIN-2 binds syndecan-2 and protein 4.1 and localizes to the basolateral membrane of epithelial cells. J. Cell Biol. *142*, 129–138.

30. Granes, F., Urena, J.M., Rocamora, N., and Vilaro, S. (2000). Ezrin links syndecan-2 to the cytoskeleton. J. Cell Sci. *113*, 1267–1276.

31. Nguyen, R., Reczek, D., and Bretscher, A. (2001). Hierarchy of merlin and ezrin N- and C-terminal domain interactions in homo- and heterotypic associations and their relationship to binding of scaffolding proteins ebp50 and e3karp. J. Biol. Chem. *276*, 7621–7629.

32. Sheffield, P., Garrard, S., and Derewenda, Z. (1999). Overcoming expression and purification problems of RhoGDI using a family of "Parallel" expression vectors. Protein Expr. Purif. *15*, 34–39.

33. Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *A276*, 307–326.

34. Weeks, C.M., and Miller, R. (1999). The design and implementation of SnB v2.0. J. Appl. Crystallogr. *32*, 120–124.

35. Sheldrick, G.M., and Gould, R.O. (1995). Structure solution by iterative peaklist optimisation and tangent expansion in space group P1. Acta Crystallogr. *B51*, 423–431.

36. de la Fortelle, E., and Bricogne, G. (1997). Maximum-likelihood heavy atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. Methods Enzymol *276*, 472–494.

37. Abrahams, J.P., and Leslie, A.G.W. (1996). Methods used in the structure determination of bovine mitochondrial F1 ATPase. Acta Crystallogr. *D52*, 30–42.

38. Perrakis, A., Morris, R., and Lamzin, V.S. (1999). Automated protein model building combined with iterative structure refinement. Nat. Struct. Biol. *6*, 458–463.

39. Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for binding protein models in electron density maps and the location of errors in these models. Acta Crystallogr. *A47*, 110–119.

40. Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. *D54*, 905–921.

41. Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. *D53*, 240–255.

42. Winn, M.D., Isupov, M.N., and Murshudov, G.N. (2001). Use of TLS parameters to model anisotropic displacements in macromolecular refinement. Acta Crystallogr. *D57*, 122–133.

43. Pannu, N.S., Murshudov, G.N., Dodson, E.J., and Read, R.J. (1998). Incorporation of prior phase information strengthens maxiumum-likelihood structure refinement. Acta Crystallogr. *D54*, 1285–1294.

44. Kleywegt, G.J. (2000). Validation of protein crystal structures. Acta Crystallogr. *D56*, 249–265.

45. Pace, C.N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995). How to measure and predict the molar absorption coefficient of a protein. Protein Sci. *4*, 2411–2423.

46. Santoro, M.M., and Bolen, D.W. (1992). A test of the linear extrapolation of unfolding free energy changes over an extended denaturant concentration range. Biochemistry *31*, 4901–4907.

47. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J. Appl. Crystallogr. *24*, 946–950.

48. Merritt, E.A., and Bacon, D.J. (1997). Raster3D. Photorealistic molecular graphics. Methods Enzymol. *277*, 505–524.

49. Nicholls, A., Sharp, K., and Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins *11*, 281–296.

Accession Numbers

The coordinates of the syntenin PDZ tandem were deposited in the Protein Data Bank under the ID code 1N99.