

Drosophila α -Catenin and E-cadherin Bind to Distinct Regions of *Drosophila* Armadillo*

(Received for publication, June 6, 1996, and in revised form, September 30, 1996)

Li-Mei Pai‡, Catherine Kirkpatrick‡, Jason Blanton, Hiroki Oda§, Masatoshi Takeichi§, and Mark Peifer¶

From the Department of Biology, University of North Carolina, Chapel Hill, North Carolina 27599-3280 and the §Department of Biophysics, Kyoto University, Kyoto 606-01, Japan

Adherens junctions are multiprotein complexes mediating cell-cell adhesion and communication. They are organized around a transmembrane cadherin, which binds a set of cytoplasmic proteins required for adhesion and to link the complex to the actin cytoskeleton. Three components of *Drosophila* adherens junctions, analogous to those in vertebrates, have been identified: Armadillo (homolog of β -catenin), *Drosophila* E-cadherin (DE-cadherin), and α -catenin. We carried out the first analysis of the interactions between these proteins using *in vitro* binding assays, the yeast two-hybrid system, and *in vivo* assays. We identified a 76-amino acid region of Armadillo that is necessary and sufficient for binding α -catenin and found that the N-terminal 258 amino acids of α -catenin interact with Armadillo. A large region of Armadillo, spanning six central Armadillo repeats, is required for DE-cadherin binding, whereas only 41 amino acids of the DE-cadherin cytoplasmic tail are sufficient for Armadillo binding. Our data complement and extend results obtained in studies of vertebrate adherens junctions, providing a foundation for understanding how junctional proteins assemble and a basis for interpreting existing mutations and creating new ones.

Cell-cell adhesion and communication are required for cells to form organized tissues. One structure used by cells for these purposes is the adherens junction, found near the apical surface of epithelial cells and found also in other cell types. Adherens junctions mediate calcium-dependent cell-cell adhesion and anchor the actin cytoskeleton (reviewed in Ref. 1). In addition, many signaling molecules are localized to adherens junctions, suggesting a role in the transmission of intercellular signals (reviewed in Ref. 2).

Adherens junctions consist of transmembrane cadherins and a set of cytoplasmic proteins associated with cadherin cytoplasmic domains (reviewed in Refs. 1 and 3). The extracellular domains of cadherins interact homotypically with cadherins of neighboring cells. The cytoplasmic proteins α -catenin, β -catenin, and plakoglobin (or γ -catenin) are required for cadherin adhesive function and anchor the actin cytoskeleton. The Src tyrosine kinase substrate p120^{cas} is also present in adherens junctions (4, 5); its function remains unknown. Changes in tyrosine phosphorylation of β -catenin (reviewed in Ref. 2) and

p120^{cas} (6) correlate with transformation and associated changes in cell adhesion.

To understand the cell biological function of adherens junctions, we must determine how interactions among different adherens junction proteins mediate assembly. β -Catenin and plakoglobin bind directly to the E-cadherin cytoplasmic domain in a mutually exclusive fashion (7, 8). β -Catenin and plakoglobin are 70% identical in amino acid sequence; their central regions, containing ~13 copies of the 42-amino acid Arm¹ repeat (9), are particularly well conserved (~80% amino acid identity). These highly conserved Arm repeats mediate interaction with cadherin (10–12), suggesting that β -catenin and plakoglobin compete for the same binding site. The N-terminal regions of both β -catenin and plakoglobin bind to α -catenin; α -catenin does not bind cadherin directly (11, 13–15). α -Catenin, in turn, links adherens junctions to actin, directly (16) or via α -actinin (17). p120^{cas} also binds directly to E-cadherin (18), but likely to a site distinct from that bound by β -catenin/plakoglobin (4, 5). p120^{cas} does not interact with α -catenin (18), however, and thus does not appear to mediate interaction with actin. The core cadherin-catenin complex forms higher order assemblies such as the zonula adherens. Both E- and N-cadherins dimerize (19, 20), and association with the cytoskeleton may help form larger assemblies.

Adherens junctions were first described in vertebrates, but precisely analogous structures exist in *Drosophila*. The *Drosophila* homolog of β -catenin is Armadillo, first discovered because of its role in transducing the Wingless cell-cell signal (reviewed in Ref. 3). Arm is structurally similar to β -catenin and plakoglobin (it is 73% identical to β -catenin), with 13 Arm repeats (9) flanked by N- and C-terminal regions. *Drosophila* homologs of E-cadherin (DE-cadherin) and α -catenin have been identified (21, 22); no direct homolog of plakoglobin has been found. Both Arm and DE-cadherin are required for proper cell-cell adhesion *in vivo* (23–26).

We undertook a systematic study of the interactions between *Drosophila* Arm and its adherens junction partners, DE-cadherin and α -catenin. We localized binding sites for each of these proteins on Arm using *in vitro* binding assays, the yeast two-hybrid system, and *in vivo* binding assays. We also identified regions of α -catenin and DE-cadherin required for Arm binding. These experiments complement and extend analysis of the vertebrate homologs of Arm. Our parallel *in vivo* studies confirm the importance of these interactions for adherens junction function (27).

* This work was supported by a grant from the March of Dimes and Grant GM47857 from the National Institutes of Health (to M. P.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Contributed equally to this work.

¶ To whom correspondence should be addressed.

¹ The abbreviations used are: Arm, Armadillo; DE-cadherin, *Drosophila* E-cadherin; GST, glutathione S-transferase; PCR, polymerase chain reaction; RIPA, radioimmune precipitation assay; PAGE, polyacrylamide gel electrophoresis; ConA, concanavalin A; APC, adenomatous polyposis coli.

EXPERIMENTAL PROCEDURES

Plasmid Constructions

Further details are available upon request.

GST-Arm Fusions—Arm fragments were PCR-amplified from E9 cDNA (28) with Vent DNA polymerase in 10 cycles. 5'-Primers contain a BamHI site and a methionine codon; 3'-primers contain EcoRI and BamHI sites and a stop codon (Tables I and II show end points of the constructs). PCR products were subcloned into both pBluescript KS⁺ and pLM1; pLM1 was generated by inserting oligomers creating BglII and EcoRI sites in the appropriate reading frame of pGEX2T128/129 (gift of Dr. M. Blonar (29)). The vector/insert junctions were sequenced; many inserts were entirely sequenced. pLM1-N1 and pLM1-N5 were generated by cloning BamHI/partial BclI digestion products of the arm E9 cDNA into the BglII site of pLM1.

PCR-mediated Site-directed Mutagenesis—Mutagenesis was performed as described (30) using pBS-N2 as a template. The fragments from the second PCR were digested with BamHI and subcloned into the pLM1 BglII site. Mutants were screened for fusion protein expression and confirmed by sequencing.

GST- α -Catenin Fusions—The C-terminal two-thirds of α -catenin was subcloned from pBS- α -catenin (21) as a BglII/XbaI fragment, which was ligated with BglII/EcoRI-digested pLM1, followed by filling with Klenow fragment and a second ligation, creating pLM1- α C. The N-terminal end of α -catenin was made by PCR with Vent polymerase in 10 cycles. The 5'-primer contains BglII, EcoRI, and EcoRV sites, while the 3'-primer is just 3' to a BglII site in α -catenin. The PCR product was cut with BglII and cloned into pLM1- α C, creating pLM1- α -catenin. pLM1- α -catenin was generated by eliminating the C-terminal half of α -catenin by EcoRI digestion.

Two-hybrid Plasmids—pCK2 and pCK4 were generated from pBTM116 (gift of P. Bartel and S. Fields) and pACT2 (gift of S. Elledge), respectively, by inserting oligomers creating BamHI and EcoRI sites in the desired reading frame. Arm fragments generated by PCR as described above were cloned into pCK2 and pCK4 as BamHI/EcoRI or BamHI fragments. Mutant Arm fragments were generated by PCR with full-length Arm mutant constructs (27) as templates. The α -catenin N terminus was isolated from pLM1- α -catenin by BglII digestion and cloned into the BamHI site of pCK2 and pCK4. The cytoplasmic domain of DE-cadherin and fragments thereof were amplified by PCR (with primers containing BamHI and EcoRI sites), digested with BamHI and EcoRI, and cloned into pCK4 (see Fig. 11A).

Expression and Purification of GST Fusion Proteins and Bead and Blot Binding Assays

Fusion proteins were expressed in *Escherichia coli* DH5 α . Overnight cultures were diluted 1:10, grown for 1 h, and induced with 0.5 mM isopropyl-1-thio- β -D-galactopyranoside. After 2 h of further growth, cells were lysed in MTPBS (150 mM NaCl, 16 mM Na₂HPO₄, and 4 mM NaH₂PO₄) by sonication (20% output) twice for 30 s. 1% Tween 20 and 1% Triton X-100 were added to lysates, and cell debris was pelleted at 6000 rpm for 10 min. Glutathione-conjugated agarose beads were added, incubated for 30 min at room temperature, and then washed

with MTPBS plus 0.1% Triton X-100 and 1% Tween 20. Wild-type *Drosophila* extracts were made from 0–20-h-old embryos. Embryos were rinsed with 0.1% Triton X-100, dechorionated in 50% bleach for 4 min, rinsed again, and then ground in RIPA buffer (23) or NET (400 mM NaCl, 5 mM EDTA, 50 mM Tris, pH 7.5, and 1% Nonidet P-40), both with 50 μ g/ml phenylmethylsulfonyl fluoride, 1 μ g/ml leupeptin, 1.4 μ g/ml pepstatin, 50 mM NaF, and 1 mM NaVO₃. For bead binding assays, fusion protein bound to glutathione-agarose was mixed with wild-type embryo extract in RIPA buffer or NET at room temperature on a nutator for 2 h, and beads were washed with excess buffer four times for a total of 20 min. Samples were boiled for 5 min in SDS sample buffer and analyzed by SDS-PAGE and immunoblotting. For blot binding assays, fusion proteins were purified on glutathione-agarose, separated by SDS-PAGE, and transferred to nitrocellulose. Blots were stained with Ponceau S to detect total protein and incubated with wild-type embryo extract in RIPA buffer at room temperature for 2 h, followed by washing with excess RIPA buffer five times for a total of 1 h. Blots were immunoblotted with anti- α -catenin antibody and visualized by enhanced chemiluminescence (Amersham Corp.).

Fly Stocks and ConA-Sepharose Fractionation

Wild-type flies were Canton S; *arm*^{H8.6}, *arm*^{XM19}, *arm*^{XP33}, and *arm*^{#2} are described in Ref. 31. *arm*^{DN} was generated as a germ line transformant by injection into *y w* flies. To create pUAST-3armRC, a fragment of Arm containing the repeats and the C terminus (amino acids 128–843) was excised from pLM1-RC with BamHI and EcoRI and ligated into pUAST-3 (a modified version of pUAST (32) with the EcoRI and BamHI sites inverted in the polylinker). ConA-Sepharose fractionation was done as described (23), but using NET instead of RIPA buffer.

Two-hybrid System

We used a version (33) of the yeast two-hybrid system (34) in which one protein is fused to the DNA-binding domain of *E. coli* LexA and its putative partner is fused to the transcriptional activation domain of yeast Gal4p. Interaction of the partners activates two reporter genes, *lacZ*, whose product can be quantitated, and *HIS3*. The yeast strain L40 (*MATa his3 Δ 200 trp1–901 leu2–3,112 ade2 LYS2::(lexAop)_g-HIS3 URA3::(lexAop)_g-lacZ*) (33) was used for all experiments. The strain was transformed (35) simultaneously with two plasmids encoding different fusion proteins (in pCK2 and pCK4); transformants were selected on synthetic complete medium lacking tryptophan and leucine. *HIS3* activation was assayed by spotting dilutions of saturated liquid cultures onto synthetic complete medium lacking tryptophan, leucine, and histidine and containing 25 mM 3-aminotriazole; growth was compared with that of colonies on medium lacking only tryptophan and leucine. In all cases, *HIS3* activation paralleled *lacZ* activation (data not shown).

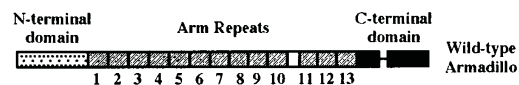


FIG. 1. Arm is a modular protein that can be divided into three regions.

TABLE I

Diagram of portions of Arm tested for association with α -catenin either as GST-Arm fusion proteins or in the two-hybrid system.

Corresponding amino acid positions are indicated on the right. Binding activity is shown as plus and minus symbols. N.D., not determined.

N-terminal domain	Repeat region	C-terminal domain	Amino acid coordinates	α -catenin binding	
				In Vitro	In yeast
		Armadillo	1-843	+	N.D.
	N-terminus 1		1-165	+	+
	N-terminus 2		25-165	+	N.D.
	N-terminus 3		70-165	+	N.D.
	N-terminus 4		90-165	+	+
	N-terminus 5		1-148	-	+
	N-terminus 6		113-165	-	+/-
	Repeats +				
	C-terminal domain		127-843	-	N.D.
	Repeats +				
	N-terminal domain		1-719	+	N.D.
	Repeat 1-13		127-719	N.D.	-
	C-terminal domain		693-843	-	N.D.

Liquid β -galactosidase assays were performed as described (36). Cells from 3 ml of late log phase culture were washed and resuspended in 1 ml of Z buffer (60 mM Na_2HPO_4 , 40 mM NaH_2PO_4 , 10 mM KCl, 1 mM MgSO_4 , and 40 mM 2-mercaptoethanol). 200- μl aliquots of this suspension were mixed with 600 μl of Z buffer, 20 μl of 0.1% SDS, and 3 drops of chloroform. Samples were vortexed for 15 s and incubated at 28 $^\circ\text{C}$ for 10–15 min. Reactions were initiated by adding 160 μl of 4 mg/ml *o*-nitrophenyl- β -D-galactopyranoside (in Z buffer) and stopped when the solution reached an appropriate yellow color ($A_{420} \sim 0.4$ – 0.7) by adding 400 μl of 1 M Na_2CO_3 . β -Galactosidase activity (in Miller units) was calculated as follows: $A_{420} \times 1000$ divided by (the A_{600} of the initial cell suspension \times the volume of suspension used (0.2 ml) \times the time in minutes of color development). All values shown are averages from assays in duplicate or triplicate on cultures of at least six independent transformants. Protein extracts for immunoblotting were made from samples of cultures grown for β -galactosidase assays. Cells from 1.5 ml of culture were resuspended in 50 μl of 2 \times Laemmli buffer, frozen at -70°C , and boiled for 5 min. Immunoblotting was performed using anti-LexA antibody (a gift of Dr. E. Golemis) at 1:5000.

RESULTS

The α -Catenin-binding Site on Armadillo Is at the Junction of the N Terminus and the Repeats—To decipher the biochemical roles of Arm in adherens junctions, we set out to characterize the regions of Arm to which its adherens junction partners, α -catenin and DE-cadherin, bind. We simultaneously took three approaches: *in vivo* in *Drosophila* embryos (Ref. 27; see data below), *in vitro* binding assays, and the yeast two-hybrid system. These approaches were chosen to complement one another, minimizing problems caused by the limitations of each assay.

Arm protein can be divided into three regions (Fig. 1). The central two-thirds of Arm is composed of 13 copies of an imperfect ~ 42 -amino acid repeat, the Arm repeat, with non-repeat amino acids between Arm repeats 10 and 11 (9, 28). There are also regions N- and C-terminal to the repeats. For *in vitro* binding experiments, different parts of Arm were expressed in *E. coli* as GST fusion proteins. Many fusion proteins were generated and tested for their ability to bind α -catenin in two different assays. These assays localized the α -catenin-binding site to the junction of the N terminus and the Arm repeats (Table I).

The first assay used to detect interaction was a bead binding assay, in which purified GST-Arm fusion proteins bound to glutathione-agarose beads (37) were used to recover interacting proteins from wild-type embryo extract. α -Catenin binding was assayed by immunoblotting with anti- α -catenin antibody (21). Neither the Arm repeats nor the C terminus contains strong α -catenin-binding sites (Fig. 2A, upper panel). The smallest region retaining full binding activity is a 76-amino acid region (fusion protein N4, amino acids 90–165) extending from the N-terminal region into Arm repeat 1. The second assay used was a blot binding (far-Western) assay, in which GST-Arm fusion proteins immobilized on nitrocellulose were incubated with wild-type embryo extract (Fig. 3). This assay confirmed that the 76-amino acid region is sufficient for binding and also implied that much of the sequence is necessary; removal of the C-terminal 17 amino acids of this minimal fragment (GST-Arm fusion protein N5) or removal of 23 amino acids from its N terminus (fusion protein N6) eliminates α -catenin binding in this assay (Fig. 3). These results are summarized in Table I.

Similar data were obtained with the yeast two-hybrid system (33, 34). We generated constructs with boundaries identical to those used in the *in vitro* assays. As in those assays, a 76-amino acid fragment at the junction of the N terminus and the Arm repeats of Arm (amino acids 90–165) interacts with α -catenin (Fig. 4, A and B). In this assay, the C-terminal 17 amino acids of this region are not essential for α -catenin interaction (fusion protein N5). We also detected a very weak α -catenin interaction when 23 amino acids were removed from the N terminus of this

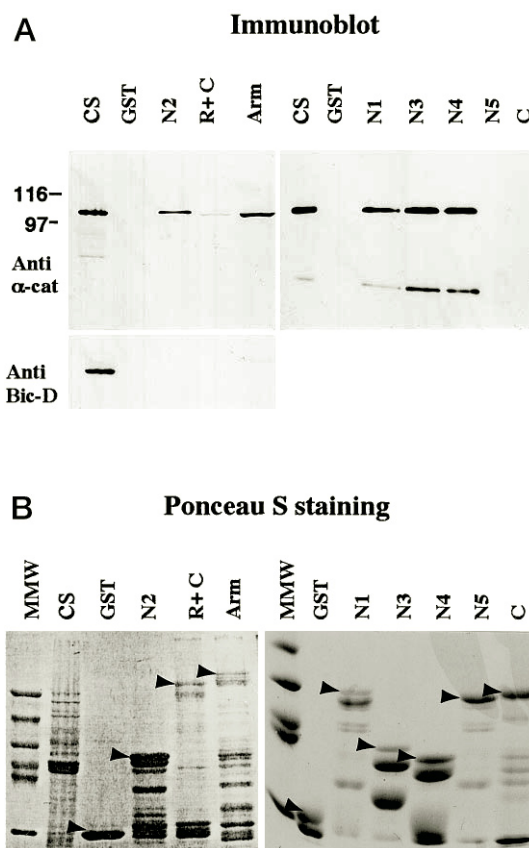


FIG. 2. The α -catenin-binding site on Arm is at the junction of the N terminus and the repeats. A, GST-Arm fusion proteins bound to glutathione-agarose beads were incubated with wild-type *Drosophila* embryo extract in RIPA buffer. Bound proteins were analyzed by SDS-PAGE and immunoblotting with anti- α -catenin antibody (Anti α -cat; upper panel). The same blot was reprobbed with anti-BicD antibody to control for nonspecific retention of extract (lower panel); BicD is a nonjunctional cytoplasmic protein (38). Total embryo extract from wild-type embryos (Canton S (CS)) was used as a positive control. The fusion proteins indicated are diagrammed in Table I. N1 indicates N terminus 1, etc., while GST indicates GST alone. Molecular weight markers are indicated on the left. B, the same blot was stained with Ponceau S to visualize the amount of fusion protein used in each reaction. Full-length fusion proteins are indicated with arrowheads. R+C, fragment of Arm containing the repeats and the C terminus; C, C terminus; MMW, medium molecular weight markers.

region (fusion protein N6). Other fragments of Arm, such as the entire repeat region, do not interact with α -catenin (Fig. 4C).

We also mapped the *in vivo* α -catenin-binding site. We generated a mutant, *arm*^{S14}, with a 39-amino acid deletion in the N terminus (removing amino acids 101–139); these amino acids are within the 76-amino acid region required for α -catenin binding *in vitro*. This mutation was reintroduced into flies and abolishes the ability of Arm to bind to α -catenin *in vivo* (27). In contrast, the most N-terminal portion of Arm is not required for α -catenin binding *in vivo*. We generated and reintroduced *in vivo* a second mutant, *arm*^{S10}, with a 54-amino acid deletion in the N terminus (removing amino acids 34–87); this deletion falls outside the region defined as essential for binding *in vitro*. This mutant protein was tagged with a c-Myc epitope to distinguish it from wild-type endogenous Armadillo. Mutant protein can be specifically immunoprecipitated with anti-c-Myc antibody; α -catenin co-immunoprecipitates with this mutant protein (Fig. 5A).

Point Mutations inside the 76-Amino Acid Region Abolish α -Catenin Binding Activity—The 76-amino acid region sufficient for binding α -catenin *in vitro* is highly conserved between Arm and its vertebrate relatives β -catenin (39) and plakoglobin

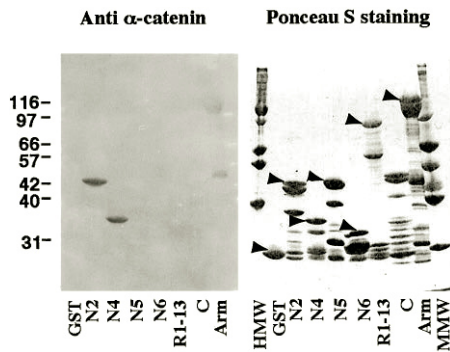


FIG. 3. A 76-amino acid region of Arm is responsible for association with α -catenin. Fusion proteins (diagramed in Table I) were separated by SDS-PAGE and transferred to nitrocellulose, which was stained with Ponceau S to detect the total protein present (right panel), and then incubated with wild-type embryo extract. The blot was then immunoblotted with anti- α -catenin antibody (left panel). Full-length fusion proteins are indicated with arrowheads. Note that only full-length N-terminal fusion proteins bind to α -catenin; C-terminally truncated proteins (generated during purification or in bacteria) fail to bind. Molecular weight markers are shown on the left. *HMW*, high molecular weight markers; *MMW*, medium molecular weight markers; *R1-13*, Arm repeats 1-13; *C*, C terminus.

(40); β -catenin and Arm are 79% identical in this region, while Arm and plakoglobin are 51% identical. To determine whether the entire region or only a subset was required for binding, we generated point mutations using PCR mutagenesis, based on the hypothesis that clustered charged residues would probably be on the surface of the protein. We made seven different mutants with clustered point mutations within fusion protein N2, which has α -catenin binding activity. In the first set of mutants, two to four positively charged amino acids that are conserved among Arm, β -catenin, and plakoglobin were changed to alanines (Fig. 6A). α -Catenin binding activity is abolished in each mutant (Fig. 6B). We also made point mutations in nonconserved residues, replacing Arm residues either with corresponding β -catenin residues or with alanines (Fig. 6A). Both alanine substitution mutants greatly reduced α -catenin binding activity. Putting β -catenin residues at the center of the 76-amino acid region reduced binding, while β -catenin residue substitutions at the N-terminal end of the binding region did not alter binding (Fig. 6B).

The Binding Site for Armadillo on α -Catenin Is in the N-terminal 258 Amino Acids—We generated a full-length GST- α -catenin fusion protein and used it in both bead and blot binding assays. Full-length α -catenin bound Arm in a blot binding assay (Fig. 7 and data not shown), as did proteolytic breakdown products with molecular masses of >25 kDa (data not shown). A smaller GST fusion protein containing the N-terminal half of α -catenin (amino acids 1-526) binds to Arm as well as full-length α -catenin (Fig. 7). We extended these data using the two-hybrid system; in this assay, the N-terminal 258 amino acids of α -catenin bind Arm (100-fold better than the vector control) (Fig. 4).

The *in Vivo* DE-cadherin-binding Site on Armadillo Maps to the Central Repeat Region—Most of the original *in vivo* arm mutations truncate the coding sequence, yielding a series of C-terminally truncated mutant proteins (31). We examined the ability of some of these mutant Arm proteins to associate with DE-cadherin *in vivo*. We also assayed a protein lacking the entire N terminus. We used binding to ConA-Sepharose to measure association with DE-cadherin. Cadherins are glycoproteins that are recognized by this lectin (41); since wild-type Arm associates with DE-cadherin, a fraction of Arm binds to ConA (23). If a mutant protein forms a complex with DE-cadherin, it will bind to ConA-Sepharose and thus be included

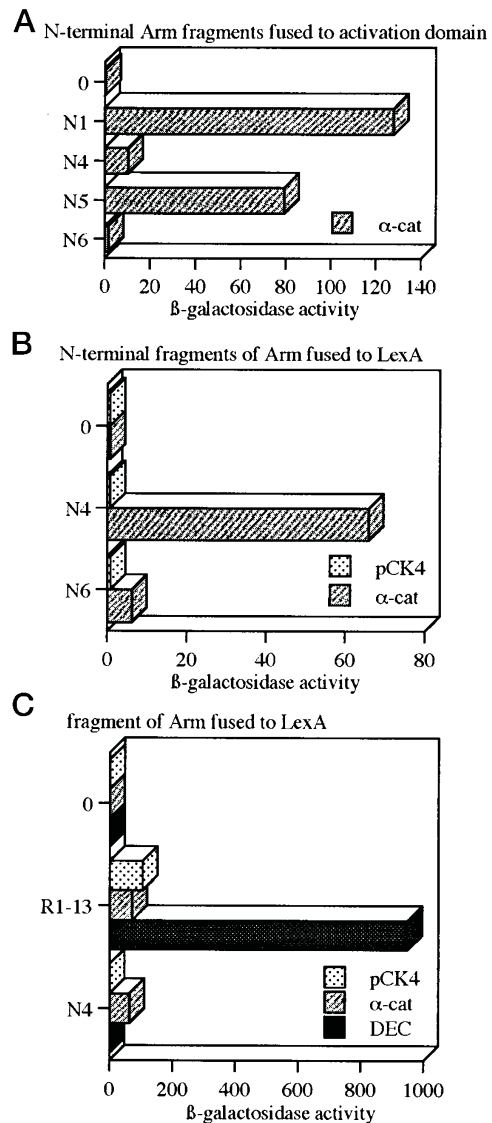


FIG. 4. The N-terminal region of Arm binds to α -catenin, whereas the repeat region binds to DE-cadherin. A, N-terminal fragments of Arm (see Table I) were fused to the Gal4p activation domain and assayed for interaction with a LexA fusion protein containing the N-terminal 258 amino acids of α -catenin (α -cat) by measuring β -galactosidase activity from the *lacZ* reporter gene. Values are in Miller units. 0 indicates a plasmid containing only the Gal4p activation domain. B, fragments N4 and N6 fused to LexA (and a plasmid encoding only LexA (0)) were assayed in combination with pCK4, which encodes the Gal4p activation domain alone, or pCK4- α -catenin, which encodes a fusion of the 258 N-terminal amino acids of α -catenin to Gal4p. (Note that N1 and N5 fused to LexA give strong transcriptional activation; these fragments were therefore assayed only in the target vector.) C, the 76-amino acid N4 fragment or the entire Arm repeat region (*R1-13*) was fused to LexA (in pCK2; 0 indicates a control with no insertion) and assayed for interaction in the two-hybrid system with the N-terminal 258 amino acids of α -catenin (α -cat) or the cytoplasmic domain of DE-cadherin (*DEC*) fused to the Gal4p activation domain in pCK4.

in the bound fraction. In contrast, mutant protein that cannot form a complex will be exclusively in the unbound fraction. Arm^{ΔN}, which lacks the entire N-terminal domain, binds DE-cadherin *in vivo* (Fig. 5B). *arm*^{HS.6} mutant protein, lacking virtually the entire C-terminal domain, also binds DE-cadherin (Fig. 5, C and D). Proteins encoded by *arm*^{XM19}, *arm*^{XP33}, and *arm*^{#2}, with truncations in repeats 13, 10, and 9, respectively, also retain some binding activity (Fig. 5, C and D). Hence, the *in vivo* DE-cadherin-binding site on Arm lies at least in part between Arm repeats 1-9.

The Central-most Arm Repeats Are Required for Binding to DE-cadherin in Vitro and in Yeast—To further narrow down the region of Arm interacting with DE-cadherin *in vitro*, we used GST-Arm fusion proteins containing portions of the Arm repeat region. Both full-length Arm and a fragment containing repeats 1–13 bind strongly to DE-cadherin. Repeats 3–10 and repeats 3–8 retain DE-cadherin binding activity, while the smaller fragment containing repeats 5 and 6 does not bind (Fig. 8).

Regions of Arm were also tested for interaction with DE-cadherin in the two-hybrid system. The N-terminal region of Arm does not interact with the cytoplasmic domain of DE-cadherin (Fig. 4C), but parts of the repeat region do (Fig. 9 and Table II). The smallest fragment of Arm capable of binding DE-cadherin is one containing Arm repeats 3–8 (*R3–8*). Smaller fragments (containing repeats 3–7 or 4–8) do not interact with cadherin, but repeats 4–13 and repeats 1–7 show substantial interaction; repeats 1–6 bind less well. (All LexA-Arm fusion proteins that do not interact with DE-cadherin are expressed in yeast to at least the same level as repeats 1–13, as assayed by Western blotting (data not shown).) We found a similar pattern of interactions between Arm and a vertebrate cadherin (mouse OB-cadherin; a gift of Dr. P. McCrea), except that the vertebrate cadherin shows reduced binding to repeats 1–7 (Fig. 9 and Table II).

We examined the effect of mutations in the Arm repeat region on the interaction with DE-cadherin both in yeast and *in vivo*. Five mutations (diagramed in Fig. 10A) were introduced into the full-length repeat region (repeats 1–13); three were also introduced into the smallest interacting fragment of Arm (repeats 3–8). These mutations (in the context of full-length Arm protein) were also tested *in vivo* (27): mutating repeat 1 or fusing repeats 10 and 11 did not affect DE-cadherin co-immunoprecipitation; removing repeat 5 or 8 greatly reduced cadherin binding; and removing repeats 3–6 eliminated cadherin binding *in vivo*. In the two-hybrid system, we observed similar but weaker effects (Fig. 10B). Mutating repeat 1 or fusing repeats 10 and 11 does not affect the interaction with cadherin, but removing repeats 3–6 eliminates binding. Removing repeat 5 (from repeats 1–13) does not substantially alter DE-cadherin binding, in contrast to the result *in vivo*; the high level of expression of the fusion protein in yeast may compensate for reduced binding affinity. We were not able to assess the effect of removing repeat 8 (from repeats 1–13) because this fusion protein activates the reporters by itself, in the absence of DE-cadherin. Removing one or more repeats from repeats 3–8 eliminates binding to DE-cadherin.

The Armadillo-binding Site on DE-cadherin Lies in a Small Segment of the Cytoplasmic Domain—We used the two-hybrid system to define the portion of DE-cadherin required for interaction with Arm. A series of constructs containing different parts of the DE-cadherin cytoplasmic domain (Fig. 11A) were tested for binding to the full repeat region of Arm. A 41-amino acid fragment (amino acids 1426–1466, DEC6) is sufficient for binding (Fig. 11B). A fragment consisting of amino acids 1350–1446 (DEC3) showed substantial interaction, but a shorter fragment (amino acids 1426–1446, DEC7) and all the fragments beginning at amino acid 1447 (DEC8, DEC9, and DEC10) did not interact with Arm.

DISCUSSION

Armadillo plays a central role in adherens junctions (3), linking the transmembrane adhesive protein DE-cadherin to α -catenin, which connects to the actin cytoskeleton. Depletion of Arm disrupts cell adhesion and actin integrity and thus disrupts organismal development (24, 42). We believe that Arm is a linker, joining together other junctional proteins. To char-

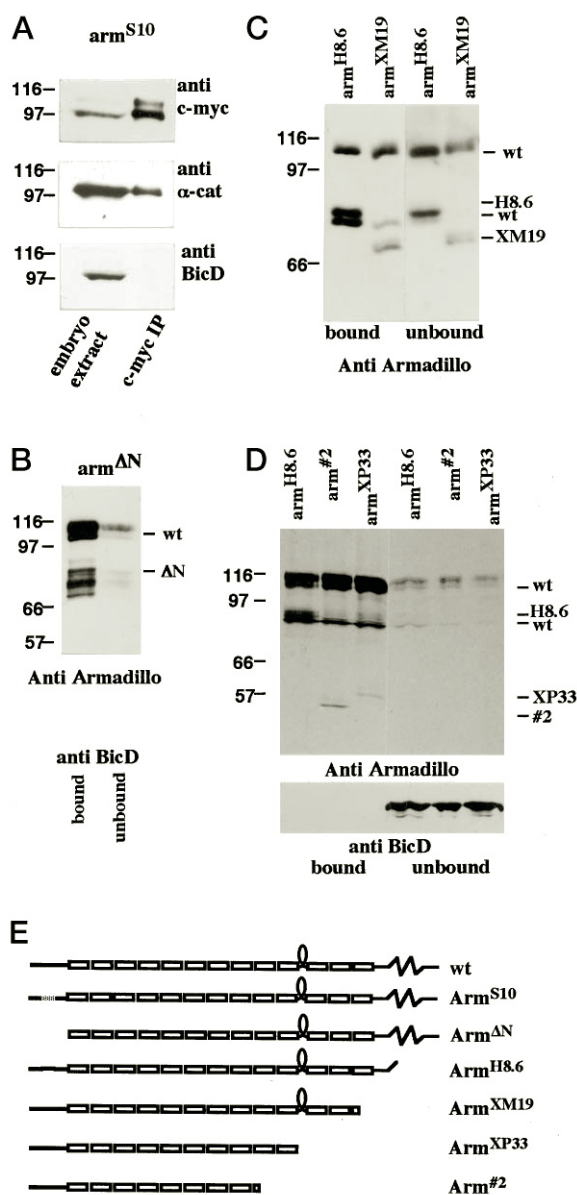













FIG. 5. Association of mutant Arm proteins with α -catenin or DE-cadherin *in vivo*. *A*, *arm*^{S10} mutant protein was immunoprecipitated from embryo extracts (left lane) with anti-c-Myc antibody, and the immunoprecipitate (IP) was analyzed by SDS-PAGE and immunoblotting with anti-c-Myc, anti- α -catenin (*anti* α -cat), and anti-BicD antibodies, respectively. *B–D*, embryo extract made from animals heterozygous for one of a set of different *arm* mutant alleles was incubated with ConA-Sepharose, which interacts with the DE-cadherin-catenin complex. Bound and unbound protein fractions were separated by SDS-PAGE and immunoblotted with anti-Arm antibody. Blots were reprobed with antibody directed against the cytoplasmic protein BicD, which should not bind cadherin, as a negative control. *B*, most *Arm*^{AN} protein is in the bound fraction. *C* and *D*, the C-terminally truncated mutant protein encoded by *arm*^{H8.6} is also found in the bound fraction. Mutant proteins truncated in repeat 13 (*arm*^{XM19}), repeat 10 (*arm*^{XP33}), and repeat 9 (*arm*^{#2}) can also bind to ConA. Proteins derived from different *arm* alleles are labeled on the right; molecular weight markers are on the left. *E*, shown is a diagram of proteins derived from the *arm* mutant alleles used here. *wt*, wild type.

acterize interactions between Arm and its junctional partners, we defined the regions of Arm responsible for each interaction. Both the sequence of Arm (the central two-thirds of the protein is composed of ~13 imperfect 42-amino acid Arm repeats (9, 28)) and the results of previous genetic analysis of mutant Arm proteins (3) suggested that Arm is modular in structure. We thus hypothesized that one might be able to define specific

TABLE II
Summary of Arm/cadherin interactions in the two-hybrid system.

A subset of the repeat region fragments tested for interaction with DE-cadherin (DE-cad) and mouse OB-cadherin (mOB-cad) are diagrammed; amino acids present in each construct are indicated. Binding activities are given as plus and minus symbols. R, Arm repeats.

N-terminal domain	Repeat region	C-terminal domain	Amino acid coordinates	Interaction with		
				DE-cad	mOB-cad	
			Armadillo	1-843		
			R1-13	127-719	++	++
			R1-10	140-596	++	++
			R3-8	214-496	++	++
			R3-7	214-445	-	-
			R4-8	255-496	-	-
			R1-7	127-445	++	+
			R1-6	127-404	+/-	-
			R4-13	255-719	++	++
			R5-13	298-719	-	-
			R7-13	385-719	-	-

many proteins to be tested. However, this assay is performed outside intact cells and involves the fusion of the protein of interest to another protein. The yeast two-hybrid system measures interactions in an intact (although heterologous) cell, but also involves the use of fusion proteins, and no components are purified. Assaying interactions *in vivo* offers by far the most realistic circumstances, but because of the investment involved, one can assay only a small number of mutants.

All three assays agree on the essential results. A region of Arm at the junction of the N terminus and the Arm repeats is both necessary (*in vitro*, in yeast, and in flies) and sufficient (*in vitro* and in yeast) for α -catenin binding, while the central-most Arm repeats (repeats 3–8) are both necessary (*in vitro*, in yeast, and in flies) and sufficient (*in vitro* and in yeast) for DE-cadherin binding.

α -Catenin Binds to Armadillo at the Junction of the N Terminus and the Arm Repeats—A 76-amino acid region at the junction between the N terminus and the Arm repeats (amino acids 90–165) is both necessary and sufficient for interaction with α -catenin *in vitro*. Clustered point mutations throughout the region reduce or abolish binding to α -catenin. Some or all of these mutations may disrupt the structure of this part of Arm; alternatively, mutated residues may reside on and thus disrupt the interaction surface. Regardless, these results suggest that the entire region is required to form either the folded structure or the proper binding surface. Results from the yeast two-hybrid assay were largely consistent with those from *in vitro* binding assays (Fig. 4B).

One discrepancy was noted: in the two-hybrid system, a few amino acids can be removed from the C terminus of this region without eliminating α -catenin binding (amino acids 1–148, N5) (Fig. 4A), while the same alteration blocks α -catenin binding in the *in vitro* assay (Figs. 2 and 3). Perhaps the GST-N5 fusion protein, as purified, lacks C-terminal amino acids due to proteolysis. Our clustered point mutants suggest that amino acids at or near the C terminus of this fusion protein are critical for binding. Further mutational analysis, combined with structural studies, will help sort out which amino acids are critical for forming an appropriately folded domain and which constitute the actual surface involved in binding α -catenin.

We demonstrated the *in vivo* biological relevance of the *in vitro* binding site (Fig. 5) (27). *arm* mutations deleting a region N-terminal to the binding site (Fig. 5) or those deleting various parts of the repeat region (27) do not affect interaction with

α -catenin *in vivo*, while in contrast, an *arm* mutant lacking 39 amino acids (amino acids 101–139) within the minimal *in vitro* binding region blocks interaction with α -catenin *in vivo* (27). This latter mutant is embryonic lethal and completely deficient in adherens junction function (27).

Several studies examined the regions of the vertebrate relatives of Arm, β -catenin and plakoglobin, required for α -catenin binding (11–13, 15). These studies provide an excellent complement to our work. Comparison of these data with our own reveals general agreement as to the site of binding, although differences in the details may reveal subtleties of the *in vivo* situation. Arm, β -catenin, and plakoglobin share substantial amino acid sequence identity through the α -catenin-binding region, and thus, these similarities are not surprising. The most extensive previous study was that of Aberle *et al.* (15), who analyzed the region of plakoglobin involved in α -catenin binding. They found that amino acids 109–137 of plakoglobin (Arm amino acids 125–154) are necessary and sufficient for α -catenin binding (15); point mutations within this region identified amino acids critical to this interaction. In contrast, a larger region of Arm is required for full α -catenin binding; 76 amino acids of Arm are required for strong interaction with α -catenin (Figs. 2 and 3), while only 28 amino acids of plakoglobin are sufficient for binding (15). Our clustered point mutants provide strong support for a more extended binding site. Several mutations outside the minimal region defined by Aberle *et al.* (15) block binding of α -catenin to Arm (Fig. 6). Several possible explanations exist for these differences. First, different assays are employed that may be more or less sensitive. Second, our experiments may identify regions that promote but are not essential for binding. Third, and less likely due to the sequence similarity between plakoglobin and Arm, these proteins may genuinely differ in the details of α -catenin binding.

We also obtained information about the region of α -catenin required to bind Arm. The Arm-binding site on α -catenin maps to its N-terminal third (Figs. 4 and 7). α -Catenin has three blocks of sequence similarity to vinculin, VH1, VH2, and VH3 (43, 44); the Arm-binding site roughly corresponds to VH1, while the actin-binding site of vinculin is found in the C-terminal region near VH3 (43, 44).

The Central Repeats of Armadillo Bind to DE-cadherin—We found that binding of Arm to DE-cadherin *in vivo* does not require the N- or C terminus of Arm, but does require a signif-

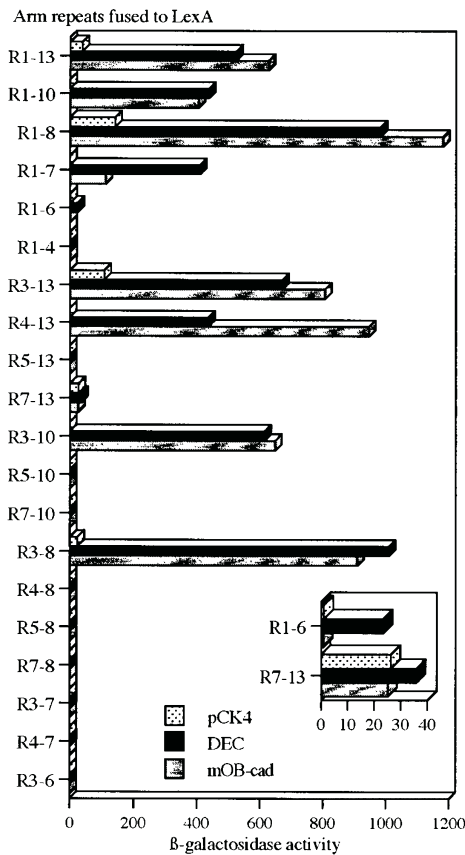


FIG. 9. The central Arm repeats are required for DE-cadherin binding in the two-hybrid system. Different portions of the Arm repeat region (see Table II) were fused to LexA in pCK2 and assayed for interaction with the cytoplasmic domain of DE-cadherin (*DEC*) or with the cytoplasmic domain of mouse OB-cadherin (*mOB-cad*). The control plasmid pCK4 encodes only the Gal4p transcriptional activation domain. Arm fragments named *RX-Y* include all of Arm repeats *X* through *Y*; those ending after repeat 10 include the insert between repeats 10 and 11. Some Arm constructs (such as repeats 1–13) activate the *lacZ* reporter even with pCK4, but give substantially higher β -galactosidase levels when DE-cadherin is present. β -Galactosidase activities are in Miller units. The *inset* shows the data for repeats 1–6 and repeats 7–13 at higher resolution.

inant portion of the central Arm repeat region (Fig. 5) (27). Examination of DE-cadherin binding *in vitro* and in yeast allowed us to further narrow down the region required. A fragment carrying Arm repeats 3–8 is the smallest piece of Arm that interacts strongly with DE-cadherin. Removal of single repeats (repeat 3, 5, or 8) from this minimal fragment abolished interaction with DE-cadherin, although in the context of longer Arm fragments, these particular Arm repeats may not be essential for binding.

We confirmed the *in vivo* relevance of this binding site and compared the effects of mutations in the repeat region both on DE-cadherin binding in yeast and on DE-cadherin binding and adherens junction function in flies. Our two-hybrid data are largely consistent with the mutant phenotypes *in vivo* (27). Mutations in repeat 1 or in repeats 10 and 11 (S6 and S12) do not alter binding of Arm to DE-cadherin in yeast or in flies, while deletion of repeats 3–6 eliminates binding both in yeast and in flies. The one discrepancy involves the deletion of repeat 5 (S5), which reduces but does not eliminate DE-cadherin interaction and adherens junction function in flies (27); in the two-hybrid system, this mutation has little or no effect on binding of repeats 1–13 to DE-cadherin. Perhaps the two-hybrid system, where interacting proteins are expressed at high levels in a foreign environment, is less sensitive to reductions

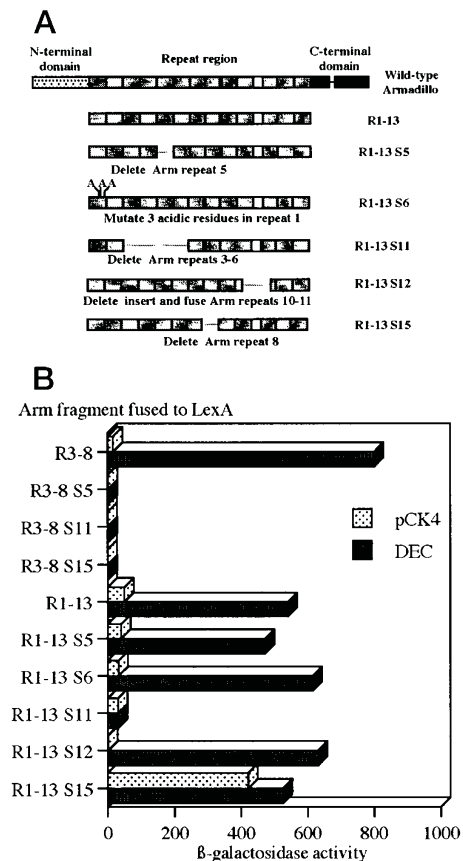


FIG. 10. Some mutations in the central repeats abolish interaction between Arm and DE-cadherin. A, shown is a diagram of the mutations tested and their effect on Arm structure. B, fragments carrying repeats 3–8 (*R3-8*) and repeats 1–13 (*R1-13*) of Arm and mutant forms of each were fused to LexA and tested for interaction with DE-cadherin (*DEC*) in the two-hybrid system. The plasmid pCK4 carries only the transcriptional activation domain. Values are in Miller units.

in the affinity of the interaction.

Our results suggest two models for the nature of the cadherin-binding site on Arm: either it is partially redundant, or the binding site is nonredundant, but must be presented in the context of a minimum of six Arm repeats. Perhaps to form a proper binding site and even for individual repeats to fold into an appropriate tertiary structure, multiple repeats must fold together into a higher order structure. All Arm repeat proteins carry a block of six or more Arm repeats in tandem or near-tandem arrays (9). The DE-cadherin-binding site may span several repeats on the surface of a higher order structure, or it may be localized to a single repeat, but only recognized in the context of that higher order structure.

Our data complement and extend those obtained with the vertebrate relatives of Arm (10–12, 45). All studies agree that the central repeats are involved in binding. Plakoglobin truncated after repeat 8 interacts with N-cadherin, whereas shorter truncations progressively lose interaction (11). A plakoglobin variant lacking repeat 4 shows reduced but detectable binding to E-cadherin (12). *In vivo* deletion analysis of β -catenin implicates the Arm repeats in E-cadherin binding; in one study, interaction of E-cadherin with a β -catenin protein truncated after repeat 10 was not observed (10), while in contrast, an analogous truncation of *Xenopus* β -catenin binds C-cadherin *in vivo*, whereas shorter fragments (repeats 1–9) do not (45). These results are consistent with our *in vivo* results; *arm*^{XP33} mutant protein, truncated in repeat 10, partially colocalizes with DE-cadherin *in vivo* and retains a small amount of adhe-

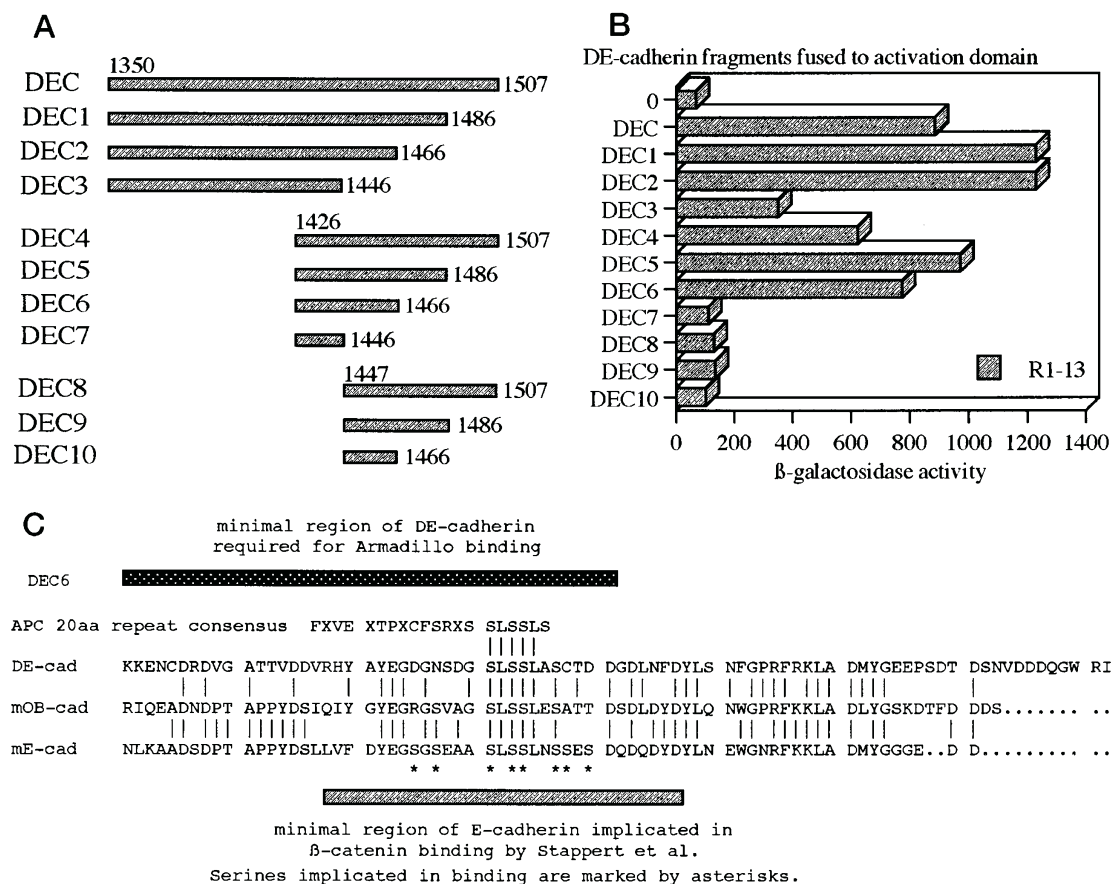


FIG. 11. A small fragment of the DE-cadherin cytoplasmic domain interacts with Arm in the two-hybrid system. *A*, shown is a diagram of the fragments of DE-cadherin tested for Arm binding. The first and last amino acids of each are indicated. *B*, the DE-cadherin fragments shown in *A* were fused to the Gal4p activation domain (in pCK4; 0 indicates pCK4 with no additional sequences inserted) and tested for interaction with Arm repeats 1–13 (R1–13) fused to LexA. β -Galactosidase activities are in Miller units. *C*, shown is the alignment of part of the DE-cadherin cytoplasmic domain (*DE-cad*) with the corresponding regions of mouse OB-cadherin (*mOB-cad*) and mouse E-cadherin (*mE-cad*); the consensus sequence for the 20-amino acid (*aa*) repeats of APC is also shown. Identical amino acids are indicated by lines. The serine residues of mouse E-cadherin mutated by Stappert and Kemler (49) are marked with asterisks. The minimal region of DE-cadherin required for Arm binding and the minimal region implicated in β -catenin binding by Stappert and Kemler (49) are indicated by the boxes above and below the sequence alignment.

rens junction function (24). Together, the data suggest that the central Arm repeats (repeats 3–8) form the core cadherin-binding site, although full interaction *in vivo* may require a more extended region.

In contrast, 41 amino acids of the DE-cadherin cytoplasmic domain are sufficient for Arm binding (Fig. 11). A similar region of vertebrate cadherins is required for β -catenin interaction (14, 46–49); this region of mouse E-cadherin contains eight serine residues, at least some of which are phosphorylated. Mutation of all eight serines to alanines blocks β -catenin binding (49). Stappert and Kemler (49) have suggested that serine phosphorylation may be required for β -catenin binding; five of these serines are conserved in mouse OB-cadherin, and four are also conserved in DE-cadherin (DE-cadherin has a fifth serine, but its position is not strictly conserved) (Fig. 11C). Interestingly, a similar serine cluster is found in the 20-amino acid repeats of the adenomatous polyposis coli (APC) tumor suppressor protein (Fig. 11C) (50). Phosphorylation of the 20-amino acid repeats by the Ser/Thr kinase glycogen synthase kinase 3 β regulates the affinity of APC for β -catenin (51), suggesting that part of the β -catenin binding determinant on its target is phosphoserine or phosphothreonine.

Acknowledgments—We thank M. Blonar for GST vectors; P. McCrea and S. Elledge for two-hybrid vectors; D. DeMarini for help with the two-hybrid system; E. Golemis for anti-LexA antibody; S. Orsulic for *arm*^{S10}; S. Whitfield for photography; and J. Adam, R. Cox, D. DeMarini, D. Kirkpatrick, J. Loureiro, and M. Turnage for reading the

manuscript. We are extremely grateful to J. Shields for making innumerable constructs.

REFERENCES

- Kemler, R. (1993) *Trends Genet.* **9**, 317–321
- Kirkpatrick, C., and Peifer, M. (1995) *Curr. Opin. Genet. & Dev.* **5**, 56–65
- Peifer, M. (1995) *Trends Cell Biol.* **5**, 224–229
- Reynolds, A. B., Daniel, J., McCrea, P., Wheelock, M. J., Wu, J., and Zhang, Z. (1994) *Mol. Cell Biol.* **14**, 8333–8342
- Shibamoto, S., Hayakawa, M., Takehchi, K., Hori, T., Miyazawa, K., Kitamura, N., Johnson, K. R., Wheelock, M. J., Matsuyoshi, N., Takeichi, M., and Ito, F. (1995) *J. Cell Biol.* **128**, 949–957
- Kinch, M. S., Clark, G. J., Der, C. J., and Burridge, K. (1995) *J. Cell Biol.* **130**, 461–471
- Hinck, L., Näthke, I. S., Papkoff, J., and Nelson, W. J. (1994) *J. Cell Biol.* **125**, 1327–1340
- Näthke, I. S., Hinck, L., Swedlow, J. R., Papkoff, J., and Nelson, W. J. (1994) *J. Cell Biol.* **125**, 1341–1352
- Peifer, M., Berg, S., and Reynolds, A. B. (1994) *Cell* **76**, 789–791
- Hülsmen, J., Birchmeier, W., and Behrens, J. (1994) *J. Cell Biol.* **127**, 2061–2069
- Sacco, P. A., McGranahan, T. M., Wheelock, M. J., and Johnson, K. R. (1995) *J. Biol. Chem.* **270**, 20201–20206
- Ozawa, M., Torada, H., and Pedraza, C. (1995) *J. Biochem. (Tokyo)* **118**, 1077–1082
- Aberle, H., Butz, S., Stappert, J., Weissig, H., Kemler, R., and Hoschuetzky, H. (1994) *J. Cell Sci.* **107**, 3655–3663
- Jou, T., Stewart, D. B., Stappert, J., Nelson, W. J., and Marrs, J. A. (1995) *J. Cell Biol.* **92**, 5067–5071
- Aberle, H., Schwartz, H., Hoschuetzky, H., and Kemler, R. (1996) *J. Biol. Chem.* **271**, 1520–1526
- Rimm, D. L., Koslov, E. R., Kebraie, P., Cianci, C. D., and Morrow, J. S. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92**, 8813–8817
- Knudsen, K. A., Soler, A. P., Johnson, K. R., and Wheelock, M. J. (1995) *J. Cell Biol.* **130**, 67–77
- Daniel, J. M., and Reynolds, A. B. (1995) *Mol. Cell Biol.* **15**, 4819–4824
- Shapiro, L., Fannon, A. M., Kwong, P. D., Thompson, A., Lehmann, M. S.,

- Grübel, G., Legrand, J.-F., Als-Nielsen, J., Colman, D. R., and Hendrickson, W. A. (1995) *Nature* **374**, 327–337, and references therein
20. Nagar, B., Overduin, M., Ikura, M., and Rini, J. M. (1996) *Nature* **380**, 360–364, and references therein
21. Oda, H., Uemura, T., Shiomi, K., Nagafuchi, A., Tsukita, S., and Takeichi, M. (1993) *J. Cell Biol.* **121**, 1133–1140
22. Oda, H., Uemura, T., Harada, Y., Iwai, Y., and Takeichi, M. (1994) *Dev. Biol.* **165**, 716–726
23. Peifer, M. (1993) *J. Cell Sci.* **105**, 993–1000
24. Cox, R. T., Kirkpatrick, C., and Peifer, M. (1996) *J. Cell Biol.* **134**, 133–148
25. Uemura, T., Oda, H., Kraut, R., Hatashi, S., Kataoka, Y., and Takeichi, M. (1996) *Genes Dev.* **10**, 659–671
26. Tepass, U., Gruszynski-DeFeo, E., Haaag, T. A., Omatyar, L., Török, T., and Hartenstein, V. (1996) *Genes Dev.* **10**, 672–685
27. Orsulic, S., and Peifer, M. (1996) *J. Cell Biol.* **134**, 1283–1301
28. Riggleman, B., Wieschaus, E., and Schedl, P. (1989) *Genes Dev.* **3**, 96–113
29. Blazar, M. A., and Rutter, W. J. (1992) *Science* **256**, 1014–1018
30. Landt, O., Grunert, H.-P., and Hahn, U. (1990) *Gene (Amst.)* **96**, 125–128
31. Peifer, M., and Wieschaus, E. (1990) *Cell* **63**, 1167–1178
32. Brand, A. H., and Perrimon, N. (1993) *Development (Camb.)* **118**, 401–415
33. Vojtek, A. B., Hollenberg, S. M., and Cooper, J. A. (1993) *Cell* **74**, 205–214
34. Fields, S., and Sternglanz, R. (1994) *Trends Genet.* **10**, 286–291
35. Schiestl, R. H., and Gietz, R. D. (1989) *Curr. Genet.* **16**, 329–346
36. Guarente, L. (1983) *Methods Enzymol.* **101**, 181–191
37. Smith, D. B., and Johnson, K. S. (1988) *Gene (Amst.)* **67**, 31–40
38. Suter, B., and Steward, R. (1991) *Cell* **67**, 917–926
39. McCrea, P., Turck, C., and Gumbiner, B. (1991) *Science* **254**, 1359–1361
40. Cowin, P., Kapprell, H. P., Franke, W. W., Tamkun, J., and Hynes, R. O. (1986) *Cell* **46**, 1063–1073
41. McCrea, P. D., and Gumbiner, B. M. (1991) *J. Biol. Chem.* **266**, 4514–4520
42. Peifer, M., Orsulic, S., Sweeton, D., and Wieschaus, E. (1993) *Development (Camb.)* **118**, 1191–1207
43. Nagafuchi, A., Takeichi, M., and Tsukita, S. (1991) *Cell* **65**, 849–857
44. Herrenknecht, K., Ozawa, M., Eckerskorn, C., Lottspeich, F., and Kemler, R. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88**, 9156–9160
45. Fagotto, F., Funayama, N., Glück, U., and Gumbiner, B. M. (1996) *J. Cell Biol.* **132**, 1105–1114
46. Nagafuchi, A., and Takeichi, M. (1988) *EMBO J.* **7**, 3679–3684
47. Ozawa, M., Baribault, H., and Kemler, R. (1989) *EMBO J.* **8**, 1711–1717
48. Ozawa, M., Ringwald, M., and Kemler, R. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87**, 4246–4250
49. Stappert, J., and Kemler, R. (1994) *Cell Adhesion Commun.* **2**, 319–327
50. Rubinfeld, B., Souza, B., Albert, I., Munemitsu, S., and Polakis, P. (1995) *J. Biol. Chem.* **270**, 5549–5555
51. Rubinfeld, B., Albert, I., Profiri, E., Fiol, C., Munemitsu, S., and Polakis, P. (1996) *Science* **272**, 1023–1026