Cell, Vol. 90, 871–882, September 5, 1997, Copyright ©1997 by Cell Press

Three-Dimensional Structure of the Armadillo Repeat Region of β-Catenin

Andrew H. Huber,* W. James Nelson,† and William I. Weis*‡
*Department of Structural Biology and
†Department of Molecular and Cellular Physiology
Stanford University School of Medicine
Stanford, California 94305

Summary

β-catenin is essential for cadherin-based cell adhesion and Wnt/Wingless growth factor signaling. In these roles, it binds to cadherins, Tcf-family transcription factors, and the tumor suppressor gene product Adenomatous Polyposis Coli (APC). A core region of β-catenin, composed of 12 copies of a 42 amino acid sequence motif known as an armadillo repeat, mediates these interactions. The three-dimensional structure of a protease-resistant fragment of β-catenin containing the armadillo repeat region has been determined. The 12 repeats form a superhelix of helices that features a long, positively charged groove. Although unrelated in sequence, the β-catenin binding regions of cadherins, Tcfs, and APC are acidic and are proposed to interact with this groove.

Introduction

The morphogenesis of solid tissues and the subsequent maintenance of tissue integrity are related processes, with cues that determine cell fate resulting in the formation of tissue-specific patterns and structures. The cytosolic protein β-catenin has essential roles in both of these processes. Assembly of cells into solid tissues requires expression of cadherin cell adhesion molecules (Takeichi, 1991). The extracellular domain of cadherins provides specificity for homotypic cell–cell adhesion, but stable adhesion requires linkage of cadherins to the cytoskeleton (Nagafuchi and Takeichi, 1988; Ozawa et al., 1990). In the adherens junction, cadherins are linked to the actin cytoskeleton by cytosolic proteins, termed catenins (Ozawa et al., 1989). β-catenin binds to the highly conserved cytoplasmic domain of cadherins and to α-catenin (Aberle et al., 1994), which binds to actin (Rimm et al., 1995). The cadherin–catenin complex is a target of regulatory signals that govern cellular adhesiveness and mobility, including tyrosine phosphorylation (Kinch et al., 1995) by growth factor receptors (Hoschuetzky et al., 1994). The cadherin–catenin complex is also present in Drosophila (Peifer et al., 1993), with the product of the segment polarity gene *armadillo* the ortholog of β-catenin.

In addition to its structural role in cellular junctions, β-catenin/armadillo is a critical component of the Wnt/Wingless growth factor signaling pathway that governs cell fate choice in early embryogenesis. In the absence of Wnt/Wingless signal, a nonjunctional pool of β-catenin/armadillo is turned over rapidly (van Leeuwen et al., 1994) in a process that requires the serine/threonine kinase, glycogen synthase kinase-3β/Zeste-white3 (GSK-3). Wnt signaling appears to reduce the activity of GSK-3, which results in an increase in the cytoplasmic pool of β-catenin (Peifer et al., 1994b; He et al., 1995; Yost et al., 1996). Under these conditions, β-catenin forms a complex with members of the Tcf/LEF-1 family of transcription factors that enters the nucleus (Behrens et al., 1996; Huber et al., 1996; Molenaar et al., 1996), binds to target DNA sites, and activates gene expression (Brunner et al., 1997; Riese et al., 1997; van de Wetering et al., 1997). Tcfs are "architectural" transcription factors that introduce bends into DNA but are weak transcriptional activators by themselves (Werner and Burley, 1997). Significantly, binding to β-catenin changes the DNA bending properties of LEF-1 (Behrens et al., 1996). Moreover, the COOH-terminal domain of β-catenin can activate reporter genes when fused to DNA-binding domains of transcription factors (van de Wetering et al., 1997). Thus, β-catenin appears to function as a transcriptional coactivator.

In mammalian cells, β-catenin interacts with the tumor suppressor gene product Adenomatous Polyposis Coli (APC) (Rubinfeld et al., 1993; Su et al., 1993). Mutations in APC are associated with familial and sporadic colorectal cancer (Kinzler and Vogelstein, 1996). APC appears to be critical for the rapid turnover of β-catenin (Munemitsu et al., 1995), and phosphorylation of APC by GSK-3 appears to enhance the interaction of APC and β-catenin (Rubinfeld et al., 1996). Cells that lack APC, or which contain APC mutants lacking the complete β-catenin binding site, display elevated levels of cytosolic β-catenin (Munemitsu et al., 1995) and constitutive transcriptional activation by the β-catenin–Tcf complex (Korinek et al., 1997; Morin et al., 1997; Rubinfeld et al., 1997). These observations indicate that loss of regulation of β-catenin levels is an important step in oncogenic transformation. Moreover, APC localizes to sites involved in active cell migration (Näthke et al., 1996), and the association of β-catenin with APC alters cell adhesiveness (Barth et al., 1997), suggesting that β-catenin–APC interactions are important in regulating migration and adhesion.

The primary structure of β-catenin consists of an NH_2-terminal portion of approximately 130 amino acids, a central region of 550 amino acids that contains 12 imperfect sequence repeats of 42 amino acids known as armadillo (arm) repeats (Peifer et al., 1994a), and a COOH-terminal region of 100 amino acids. These three regions have distinct charge distributions: the calculated pI's of murine β-catenin are 4.4 for the NH_2-terminal domain (residues 1–140), 8.3 for the arm repeats (residues 141–664), and 4.3 for the COOH-terminal domain (residues 665–781). Deletion mutagenesis studies have mapped the binding of cadherins (Hülsken et al., 1994; Orsulic and Peifer, 1996), APC (Hülsken et al., 1994; Rubinfeld et al., 1995), Tcf-family transcription factors (Behrens et

‡To whom correspondence should be addressed.

al., 1996; Molenaar et al., 1996), and the EGF receptor tyrosine kinase domain (Hoschuetzky et al., 1994) to the arm repeat region. The binding site for α-catenin partly overlaps the beginning of this region (Aberle et al., 1996; Pai et al., 1996). The NH_2-terminal region contains consensus phosphorylation sites for GSK-3 (Yost et al., 1996) and tyrosine kinases (A. Barth and W. J. N., unpublished data) and is important for regulating cellular levels of β-catenin (Munemitsu et al., 1996; Morin et al., 1997; Rubinfeld et al., 1997). The acidic COOH-terminal region appears to have the transactivator function required for gene activation by the β-catenin–Tcf complex (van de Wetering et al., 1997).

Here, we describe the three-dimensional structure of the armadillo repeat region of murine β-catenin, designated β59, determined by X-ray crystallography in two crystalline forms at 2.1 and 2.9 Å resolution. The armadillo sequence motif encodes a structural unit that forms an elongated superhelical molecule when tandemly repeated. Features of the superhelical surface suggest possible mechanisms for the interaction of β-catenin with its many binding partners.

Results and Discussion

The Armadillo Repeat Region Is a Single Structural Unit

Limited proteolysis was used to assess whether the distinct regions present in the sequence of β-catenin correspond to compact, discrete elements of structure. Trypsin (Figure 1A) and subtilisin (data not shown) produced similar patterns of digestion, with two fragments of apparent molecular weight 40 and 10 kDa accumulating as limiting digest products. The two tryptic fragments coelute from anion exchange and size exclusion chromatography columns and also copurify on an E-cadherin cytoplasmic domain/glutathione S-transferase fusion protein affinity resin (data not shown). The two fragments were separated from one another by reversed-phase HPLC. Mass spectroscopy and NH_2-terminal sequence analysis showed that the 40 and 10 kDa fragments consist of residues 134–550 and 551–671, respectively. As the fragments stay associated under native conditions and bind to the E-cadherin cytoplasmic domain, these data indicate that trypsinolysis produces a portion of β-catenin spanning residues 134–671 that is nicked at Arg-550 (Figure 1B). Thus, residues 134–671 form a single, functional structural unit that comprises the β-catenin armadillo repeat region as defined by sequence analysis (Peifer et al., 1994a).

Structure Determination

The 59 kDa region of murine β-catenin spanning residues 134–671, designated β59, was cloned and expressed in E. coli. Two crystal forms of β59 were obtained. Form A crystals, grown from a PEG/Urea/Tris solution, diffract to 2.1 Å resolution. Form B crystals, grown from a PEG/NaCl/Tris mixture, diffract weakly to 2.6 Å resolution. A single Form A crystal of selenomethionine-substituted β59 was used for Multiwavelength Anomalous Dispersion (MAD) phasing (Table 1). A molecular model was built into the 2.4 Å MAD phased electron density map (Figure 2) and refined to 2.1 Å, the

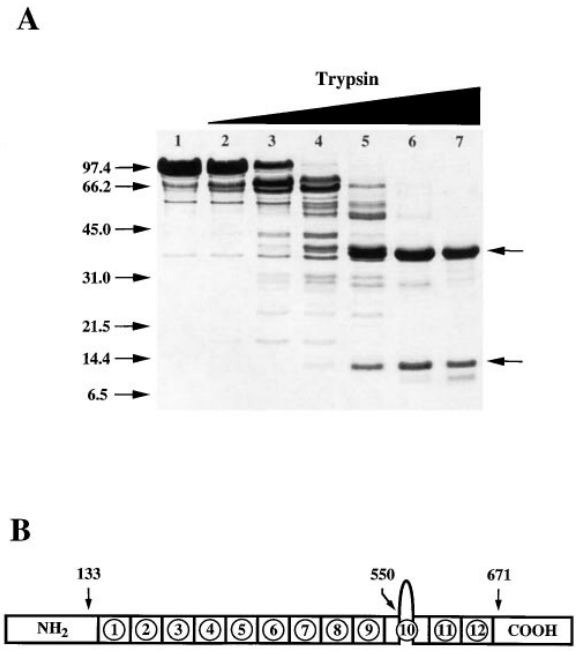

Figure 1. Limited Proteolysis of β-Catenin

(A) SDS–PAGE gel showing digestion of β-catenin with increasing concentrations of trypsin. Positions of molecular weight markers (kDa) are shown on the left, and the positions of tryptic fragments with apparent molecular weights of 40 and 10 kDa are marked with arrows on the right. Increasing trypsin concentration is denoted by the wedge above the gel. Digest mixtures contained 2.0 mg/ml β-catenin in 90 mM Tris–HCl (pH 8.5), 2 mM $CaCl_2$, and 4 mM DTT with 0 μg/ml (lane 1), 0.02 μg/ml (lane 2), 0.08 μg/ml (lane 3), 0.33 μg/ml (lane 4), 1.33 μg/ml (lane 5), 5.33 μg/ml (lane 6), or 21.33 μg/ml trypsin (lane 7). Digests were incubated at 24°C for 20 minutes and stopped with TLCK.

(B) Schematic of β-catenin primary structure showing the location of the cleavage sites that produce the stable products shown in (A). The 12 armadillo repeats of β-catenin are represented with numbered boxes. Numbered arrows indicate the residues NH_2-terminal to the trypsin cleavage sites.

Bragg spacing limit of the λ_1 (high energy remote point) data set (Table 1). The refined Form A model was used to solve the Form B crystal structure at 2.9 Å resolution by molecular replacement. The refined Form B model (Table 1) contains 45 additional residues, mostly at the NH_2 terminus of the molecule, that could not be placed with confidence in the Form A structure due to the poor quality of the electron density in this region of the map. Residues 134–149, 550–562, and 666–671 are not visible in either crystal form. Apart from small differences attributable to packing in the different lattices, the secondary and tertiary structures of β59 seen in the two crystal forms are the same, despite the presence of 2.4 M urea in the Form A crystals.

Structure of β59

The 12 armadillo repeats of β59 form a highly elongated structure composed entirely of α helices and connecting loops. The overall shape of the domain is approximately cylindrical, with a length of 110 Å and an average diameter of 35 Å. Each repeat interacts extensively with its neighbors, packing together to form a single domain

Table 1. Crystallographic Data and Refinement Statistics

A. Data Collection Statistics[a]

Crystal	Wavelength (Å)	Bragg Spacing Limits (Å)	% Complete	R_{sym}[b]	% > 3σ (I)	Average Redundancy
MAD Phasing Data						
Form A	λ_1 = 0.9252 (remote)	45–2.4	99.6 (96.1)	0.054 (0.174)	84.4 (63.7)	2.6 (2.4)
	λ_2 = 0.9757 (peak)	45–2.4	99.6 (96.7)	0.050 (0.163)	85.6 (61.4)	2.5 (2.3)
	λ_3 = 0.9800 (edge)	45–2.4	99.6 (97.0)	0.048 (0.140)	87.2 (68.2)	2.5 (2.3)
	λ_4 = 1.0688 (remote)	45–2.4	99.5 (96.2)	0.043 (0.110)	89.8 (73.5)	2.6 (2.1)
Refinement Data						
Form A	0.9252	10.0–2.1	99.5 (94.5)	0.070 (0.405)	70.5 (29.6)	2.5 (2.0)
Form B	1.5418	40.0–2.9	96.7 (93.4)	0.071 (0.257)	80.4 (57.1)	3.6 (3.4)

B. Phasing Statistics

	Anomalous Diffraction Ratios[c]				Anomalous Scattering Factors[d] (e)		Phasing Power[e]	
Wavelength	λ_1	λ_2	λ_3	λ_4	$\Delta f'$	$\Delta f''$	+ Friedel Mate	− Friedel Mate
λ_1	0.076 (0.050)	0.065	0.069	0.050	−1.8	3.0	1.3	2.0
λ_2		0.089 (0.050)	0.043	0.059	−8.0	4.5	0.8	1.7
λ_3			0.062 (0.047)	0.062	−8.8	1.8	(reference)	1.2
λ_4				0.054 (0.045)	−2.1	0.6	1.5	1.9

MADSYS Phasing Statistics[f]	
$R_{sym}(\lvert{}^{\circ}F_T\rvert)$	0.041
$R_{sym}(\lvert{}^{\circ}F_A\rvert)$	0.463
$\langle\Delta(\Delta\phi)\rangle$	47.1°
$\langle\sigma(\Delta\phi)\rangle$	20.0°

Figure-of-Merit (FOM) from Probabilistic Phasing[g]

Bragg Spacing Limits (Å)	45–4.8	4.8–3.8	3.8–3.3	3.3–3.0	3.0–2.8	2.8–2.6	2.6–2.5	2.5–2.4	Overall
FOM	0.85	0.78	0.73	0.69	0.61	0.55	0.52	0.43	0.65

C. Refinement Statistics

Crystal	Bragg Spacing Limits (Å)	Number of Reflections[h] Working Set (Test Set)	Extent of Models (number of atoms) Amino Acids	Solvent	R_{free}[i]	R_{cryst}[i]	RMSD from Ideality Bonds (Å)	Angles (°)	Temperature Factors (Å²) Bond/Angle RMSD	Average	Overall Anisotropic Tensor B_{11}	B_{22}	B_{33}
Form A	10.0–2.1	59,157 (6,388)	457 (3473)	256 H_2O, 8 urea, 1 Cl^- (288)	0.260	0.229	0.008	1.4	2.0/2.7	31.3	4.1	−6.6	2.5
Form B	40.0–2.9	10,202 (1,178)	502 (3820)	none	0.288	0.211	0.010	1.5	3.7/5.6	26.4	3.5	4.1	−7.5

[a]Values in parentheses are for reflections in the highest resolution bin: 2.49–2.40 Å for the MAD phasing data sets, 2.16–2.10 Å and 3.03–2.90 Å for Form A and Form B refinement data sets, respectively.

[b]$R_{sym} = \Sigma_h \Sigma_i |(I_i(h) - \langle I(h)\rangle| / \Sigma_h \Sigma_i I_i(h)$, where $I_i(h)$ is the i-th measurement of reflection h and $\langle I(h)\rangle$ is the weighted mean of all measurements of h. Bijvoet mates were treated as independent reflections in all Form A data sets.

[c]Anomalous diffraction ratios = $\langle\Delta|F|\rangle/\langle|F|\rangle$, where $\langle\Delta|F|\rangle$ is the rms Bijvoet difference at a single wavelength (diagonal elements) or the rms dispersive difference between two wavelengths (off-diagonal elements). Diagonal elements in parentheses are for centric reflections and provide an estimate of the noise in the anomalous signal; these values would be 0.00 for perfect data.

[d]Refined with MADLSQ (Weis et al., 1991).

[e]Phasing Power = $\langle|F_H|\rangle/E$, where $\langle|F_H|\rangle$ is the rms structure factor amplitude for the anomalous scatterers and E is the estimated lack-of-closure error defined in Burling et al. (1996). Phasing power is listed for each lack-of-closure expression between the reference data set (+ Friedel mate at the edge wavelength) and the + or − Friedel set at each wavelength. Phasing power was calculated using all data between 45 and 2.4 Å.

[f]Data from 45–2.7 Å. $R_{sym}(|F|) = \Sigma_h\Sigma_i||F_i(h)| - |\langle F(h)\rangle||/\Sigma_h\Sigma_i F_i(h)$, where $F_i(h)$ represents the i-th least squares determination of $|{}^{\circ}F_T|$ or $|{}^{\circ}F_A|$. $|{}^{\circ}F_T|$ and $|{}^{\circ}F_A|$ are the structure factor amplitudes for normal scattering from all atoms (Total) and from the Anomalous scatterers, respectively. $\Delta\phi = ({}^{\circ}\phi_T - {}^{\circ}\phi_A)$, where ${}^{\circ}\phi_T$ and ${}^{\circ}\phi_A$ are the phases of ${}^{\circ}F_T$ and ${}^{\circ}F_A$. $\Delta(\Delta\phi)$ is the mean pairwise difference between independent determinations of $\Delta\phi(h)$.

[g]Burling et al., 1996.

[h]The Test Set is a randomly selected subset of the data (10%) that was not used in the refinement of the model. The Working Set comprises the remaining reflections.

[i]$R = \Sigma_h||F_{obs}(h)| - |F_{calc}(h)||/\Sigma_h F_{obs}(h)|$. R_{free} and R_{cryst} were calculated using the Test Set and Working Set reflections, respectively.

with a continuous hydrophobic core. The rotation and translation that relates adjacent repeats produces a right-handed superhelix of helices. The superhelical axis follows a straight course for the first 8 repeats, then bends approximately 50° in the vicinity of repeats 9 and 10 to form a kink in the molecule (Figures 3A and 3B).

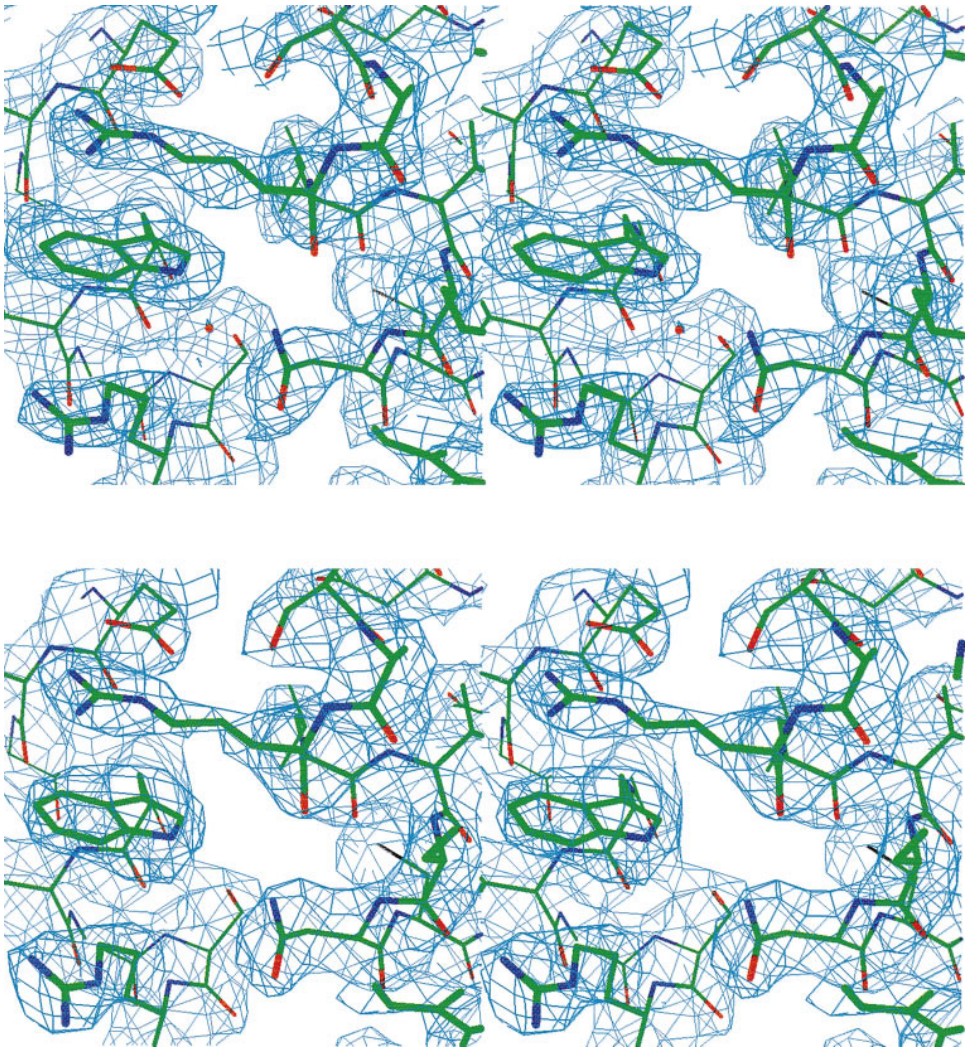

Figure 2. Stereo Views of Electron Density Maps

The region near Trp-338, which interacts with Arg residues in the groove of β59, is shown. The upper panel shows a portion of the refined Form A model in the experimental 2.4 Å MAD-phased map. The lower panel shows the same region of the Form B model in the final 2.9 Å Form B $2F_o - F_c$ map. Both maps are contoured at 1.0 σ.

The arm repeats are structurally very similar, with residues conserved among repeats participating in similar intra- and inter-repeat interactions (Figures 3 and 4). The canonical arm repeat in β59 is 42 amino acids long and contains three α helices: a short helix 1 (H1) of roughly two turns, followed by two larger helices of approximately 2–3 (H2) and 3–4 turns (H3) (Figures 3 and 4). H2 and H3 interact most extensively, packing against one another in an antiparallel orientation. H1 lies roughly perpendicular to the H2–H3 hairpin, contacting the NH_2-terminal end of H2 and the COOH-terminal end of H3. The loop connecting H1 to H2 contains a single residue in a left-handed helical conformation, generally glycine, that reorients the polypeptide backbone by about 90°. The conformations of the loops connecting H2 to H3, and H3 to H1 of the following repeat are also remarkably similar among repeats. An average surface area of 1100 $Å^2$, or approximately 25% of the total helix surface area, is buried in an isolated repeat; 80% of this buried surface is contributed by H2 and H3. In the presence of its neighboring repeats, a single repeat buries an additional 1700 $Å^2$ of surface area, such that 60% of the total helix surface area is buried in β59.

Each armadillo repeat shares extensive interactions with other repeats adjacent in sequence (Figure 4B), and it is unclear whether a single repeat would fold stably in the absence of its neighbors. Hirschl et al. (1996) determined the secondary structure of a putative armadillo repeat peptide from APC. Alignment of the arm repeat region of APC and β-catenin/armadillo (Peifer et al., 1994a) indicates that the peptide starts in the middle of H3 of one repeat and ends in the middle of H3 of the following repeat. Residues corresponding to H1 and a portion of H2, or 38% of the peptide, adopt helical secondary structure, and the remainder of the peptide is unstructured. Because the peptide contains portions of

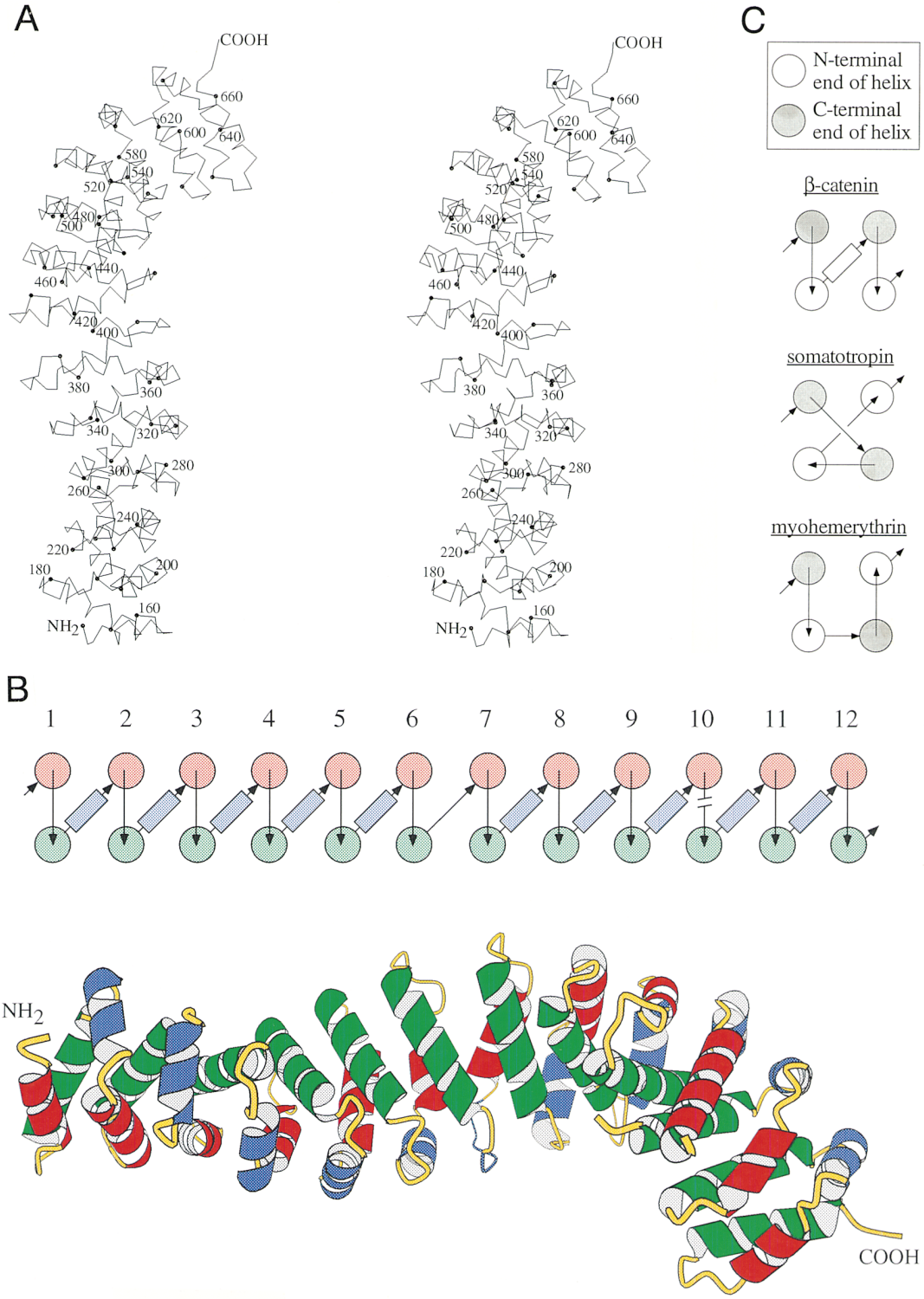

Figure 3. β59 Structure

(A) Stereo Cα trace of the Form B structure. Spheres denote every tenth α-carbon, and every twentieth residue is labeled.

(B) Topology (upper) and ribbon (lower) diagrams of β59. Circles and rectangles represent helices viewed end on and from the side, respectively. The arrows indicate the NH_2- to COOH-terminal direction of the polypeptide. Arrows terminate at the circle perimeter when connected to the end of the helix farthest from the viewer and enter the circle when connected to the end of the helix nearest the viewer. Helices 1, 2, and 3 are colored blue, red, and green, respectively.

(C) Comparison of bundle topology found in β-catenin with those of two representative four-helix bundle proteins, as discussed in Raag et al. (1988). The symbols are the same as in (B). The COOH-terminal end of each helix is shaded.

The Cα trace and the ribbon diagram were made with MOLSCRIPT (Kraulis, 1991).

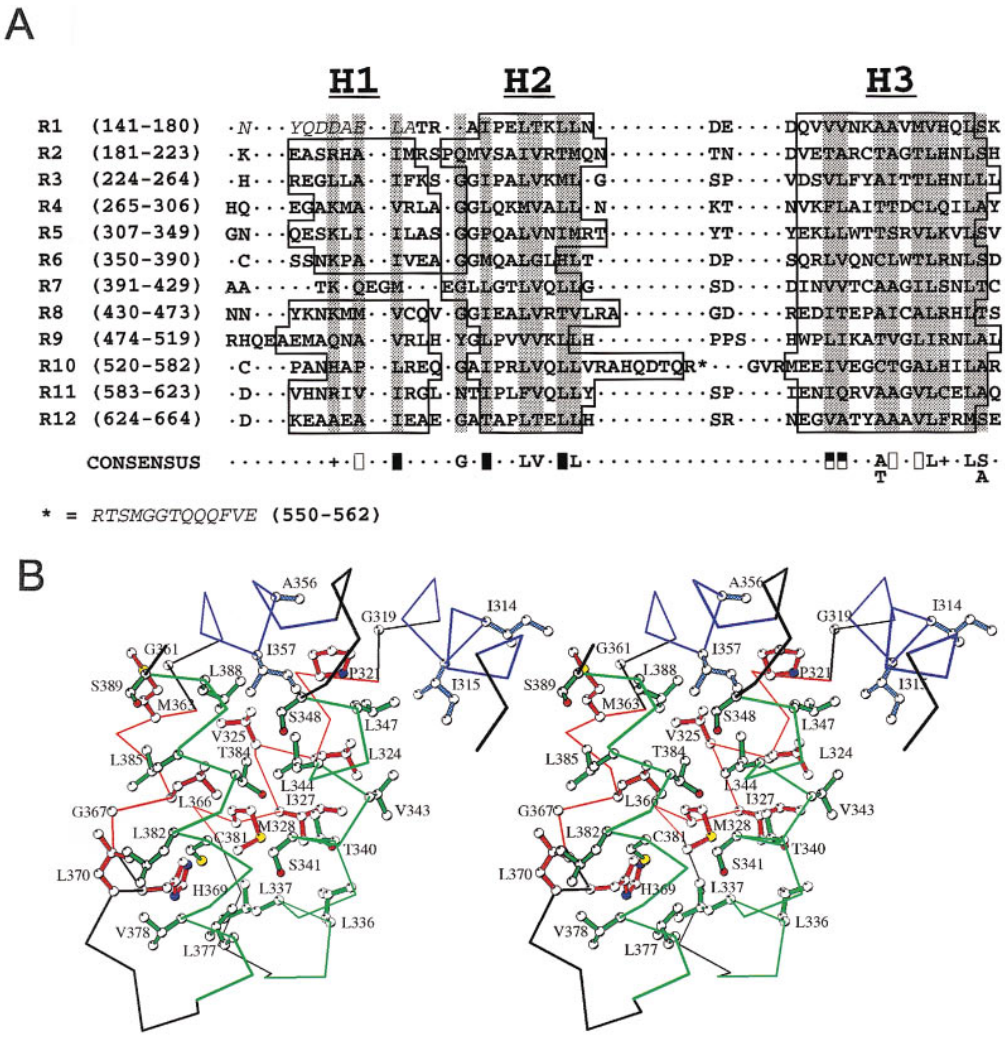

Figure 4. Sequence and Structural Similarities among Repeats

(A) Structure-based sequence alignment of the 12 armadillo repeats of murine β-catenin. Italicized residues are not visible in either crystal form. The repeat numbers and the corresponding ranges of amino acids are shown on the left. The portions of the sequence that form helices H1, H2, and H3 are boxed. Structural positions with strong preferences for a given amino acid or group of amino acids are shaded and listed on the line marked "Consensus" with the following symbols: half-closed box = general hydrophobic (residues in 10 of 12 repeats = A, C, F, I, L, M, P, T, V, or W); open box = small hydrophobic (general hydrophobic with 8 of 12 = A, C, P, V, or T); closed box = large hydrophobic (general hydrophobic with 8 of 12 = F, I, L, M, or W); (+) = basic (8 of 12 = H, K, or R). Consensus positions are assigned specific residue identities if 8 of 12 repeats have the same amino acid, or if two amino acids are present in at least four repeats each.

(B) Stereo view of repeats 5 and 6, showing conserved hydrophobic packing interactions. A Cα trace is shown for residues 307–390. The backbone, and side chains of amino acids in the consensus positions shown in (A), are displayed with blue, red, or green bonds, corresponding to H1, H2, and H3, respectively. The two basic consensus positions ([+] in [A]) have been omitted for clarity. White, blue, red, and yellow spheres represent carbon, nitrogen, oxygen, and sulfur atoms. Amino acids are labeled with single-letter code followed by the residue number. The role of conserved hydrophobic residues in H1 is illustrated by A356, which is near the surface but packs laterally against P321 and other residues from the previous repeat, and I357, which packs against hydrophobic residues of H2 and H3 from both repeats. Note that the exposed hydrophobic residues seen in the figure are actually packed against hydrophobic residues of neighboring repeats. This stereo model was made with MOLSCRIPT (Kraulis, 1991).

two repeats and does not contain a complete H3, no conclusions regarding the tertiary structure or stability of a single repeat can be drawn from these data.

The packing of tandem repeats gives rise to a two-sided spiraling ribbon with each side consisting of parallel H2 or H3 helices (Figures 3A and 3B). The structure resembles a series of four helix bundles with H1 lying diagonally across one end of each bundle. However, there are important topological differences between these "armadillo helix bundles" and the four helix bundles found in other proteins. All directly adjacent helices in a four-helix bundle protein are antiparallel, with the NH_2 and COOH termini located on neighboring helices (Figure 3C). In the armadillo helix bundle, adjacent H2 helices are parallel, as are adjacent H3 helices, whereas H2 and H3 helices are antiparallel (Figures 3A and 3C). This arrangment places the NH_2 and COOH termini on diagonally opposite sides of the bundle, which, when

repeated, produces a structure with a continuously wound topology.

Although the β-catenin arm repeats vary significantly in sequence, the size of the repeat and the character of the hydrophobic core is largely conserved (Figure 4), leading to a surprisingly consistent packing of repeats. The parameters describing the β59 superhelix reflect the rotation/translation relationship produced by the insertion of ridges formed by every third side chain of one helix into grooves created by every fourth side chain of another helix. In β59, this "3 in 4" packing produces an average 30° rotation and 10 Å translation per repeat. Curiously, only two of the five phenylalanine and none of the three tryptophan residues are buried in the hydrophobic core (Figure 4A); the remainder are located at the surface of the molecule and interact with surface residues (e.g., Figure 2), but are not clustered in any one region.

There are two significant deviations from the canonical repeat structure. A loop replaces H1 in repeat 7, but it features the same interactions made by conserved hydrophobic residues present in the H1 helix of other repeats (Figure 4): Gln-395 of the loop packs laterally against H1 and H2 of the previous repeat, and Met-398 of the loop inserts into the hydrophobic core between repeats 6 and 7 (c.f. packing of Ala-356 and Ile-357 in Figure 4B). The other deviation is a 22 amino acid insertion in repeat 10 (Figure 4A). In both crystal forms, the insertion is only partially ordered: the first 6 residues extend H2, and the last 3 lead into H3. The intervening 13 residues are disordered and presumably flexible in both crystal forms, consistent with the observation that Arg-550 of this insert is a site of trypsin cleavage in limited proteolysis experiments (Figure 1).

The repetitive nature of the armadillo domain structure makes it difficult to identify with certainty which structural element marks the "true" beginning of an armadillo repeat. The register presented in Figure 4A is consistent with the extent of the arm repeat region as defined by the pattern of hydrophobic residues in the β-catenin sequence. Similar borders for the arm repeat domain were defined by limited trypsin digestion, which cuts β-catenin ten residues before the putative H1 of repeat 1 and nine residues past H3 of repeat 12 to yield β59 (Figure 1b). However, the portion of β59 preceding residue 150 is not visible in either crystal form, so the present structure provides no evidence that residues 141–149 correspond to an H1 helix in full-length β-catenin.

A right-handed superhelix formed by two layers of parallel helices, analogous to H2 and H3, has been observed in portions of several other proteins: a bacterial muramidase (Thunnissen et al., 1994), lamprey yolk lipovitellin-phosvitin (Raag et al., 1988), and the α-subunit of farnesyl transferase (Park et al., 1997) contain 10, 8, and 7 repeats of two antiparallel helices, respectively. (A similar two-layer topology was predicted for another arm family member, the regulatory A subunit of protein phosphatase 2A [Ruediger et al., 1994].) There are two principal differences between those structures and that of the β-catenin armadillo repeat region. First, the H1 helix seen in β-catenin is not present in these other structures. Second, the superhelical path of the first eight repeats of β59 is very straight, whereas the other structures display significant curvature.

Molecular Flexibility

Elongated molecules composed of tandemly repeated structural units frequently display significant flexibility. However, the extensive interactions observed betweeen each arm repeat and its neighbors contrasts sharply with "beads-on-a-string" structures characteristic of molecules composed of independent folding units such as immunoglobulin domains. The resistance of β59 to proteolysis indicates that it is a compact molecule. The observed structure would appear to allow only limited flexibility, as any large motions would disrupt the single, continuous hydrophobic core. Nonetheless, comparison of β59 in its two crystal forms provides evidence of limited internal motion. The α carbons from the 456 residues ordered in both structures superimpose with an rms deviation of 1.38 Å, whereas superposition of the NH_2- or COOH-terminal halves individually yields rms deviations of 0.52 and 0.48 Å, respectively. When only one half is used for the superposition, the nonsuperimposed halves of the two structures are related by an 11.5° rotation about an axis that passes near the junction between repeats 6 and 7 (Figure 5). Repeat 7 is the only repeat that lacks H1, raising the possibility that H1, when present, adds rigidity to the superhelix. The observed rigid-body motion does not appear to be the result of rotation about a single bond, but rather the product of a number of small changes in the interface between repeats 6 and 7 that do not significantly alter the hydrophobic core. Due to restrictions imposed by packing in the crystal lattice, the full range of motion in solution may be greater than that observed in the crystal structures.

A Potential Binding Surface

The distribution of surface charge in β59 is highly asymmetric, with a large band of positive surface potential extending over the first 10 repeats (Figure 6). Most of the positive charge lies within a shallow groove generated by the right-handed superhelical twist of the molecule. The groove is about 95 Å long and 20 Å wide. H3, the most regular and highly conserved of the three arm repeat helices, defines the groove. The exposed H3 surface of a given repeat defines the floor of the groove, and the NH_2 and COOH termini of the H3s of the flanking repeats define the edges. A significant portion of the positive charge is contributed by the conserved basic position in H3, as well as by residues that reach into the groove from the conserved basic position in H1 (Figure 4A). Most of the basic side chains are well ordered, with many held in position by extensive interactions with other residues (e.g., Figure 2).

Deletion mutagenesis studies have mapped the binding sites for cadherins (Hülsken et al., 1994; Orsulic and Peifer, 1996; Pai et al., 1996), APC (Hülsken et al., 1994; Rubinfeld et al., 1995), and Tcf family members (Behrens et al., 1996; Molenaar et al., 1996) to the arm repeat domain of β-catenin. Attempts to define more closely the interaction sites by deletions within the arm repeat region are difficult to interpret because the assays employed are semiquantitative and binding affinities have not been determined. Moreover, deletions within a single structural domain such as β59 may lead to unanticipated

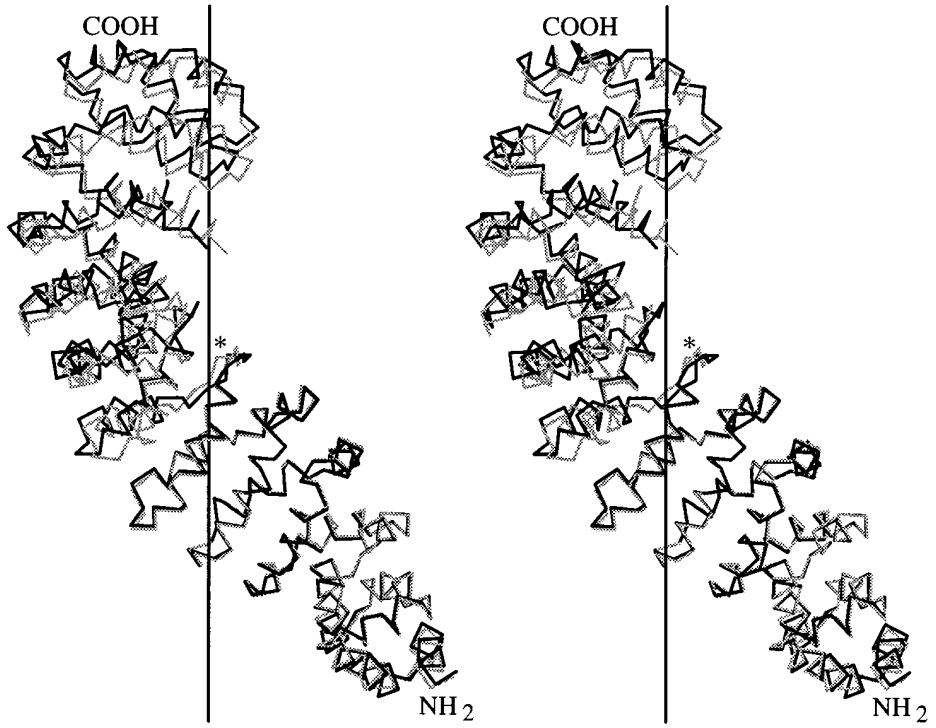

Figure 5. Stereo View Showing Relative Motion of the Two Halves of β59

The Cαs from residues 193–390 from Form A (gray) and Form B (black) have been superimposed. The COOH-terminal portions of the two structures are related by an 11.5° rotation about the axis shown as a straight vertical line. The asterisk marks the middle of the loop that replaces H1 in repeat 7. This figure was made with MOLSCRIPT (Kraulis, 1991).

local changes in the structure that alter the behavior of the mutated protein. Nonetheless, the available data indicate that the different classes of binding partners recognize overlapping portions of β-catenin that span the first 10 repeats. In all cases, a minimum of 6–7 repeats is sufficient for detectable binding, but larger regions give stronger binding. Significantly, E-cadherin and APC form mutually exclusive complexes with β-catenin and appear to compete directly for binding (Hülsken et al., 1994; Rubinfeld et al., 1995).

The regions of cadherins, APC, and Tcf family proteins that interact with β-catenin have also been mapped by deletion mutagenesis. A stretch of 30 amino acids in the cytoplasmic domain of E-cadherin is required for binding to β-catenin (Stappert and Kemler, 1994), and a similar region of Drosophila E-cadherin binds to armadillo (Pai et al., 1996). Several repeats of 15 and 20 amino acids present in APC are required for interaction with β-catenin (Su et al., 1993; Munemitsu et al., 1995). The first 56 amino acids of LEF-1 contain the β-catenin recognition site (Huber et al., 1996), a region highly conserved among Tcf family members (Molenaar et al., 1996; Brunner et al., 1997; van de Wetering et al., 1997).

Although the β-catenin binding regions of cadherins, APC, and Tcf family members show no significant sequence homology, they are all acidic, with calculated pI's of 3.3, 4.1, and 4.0, respectively. The only other common feature of these regions is a short sequence, Ser-Ser-Leu, which may represent a β-catenin binding motif (Molenaar et al., 1996; Pai et al., 1996). We propose that extended, acidic regions of cadherins, APC, and the Tcf-family transcription factors interact with the basic groove of β-catenin. The 95 Å long groove could accommodate a 25–30 residue polypeptide in a fully extended conformation, or larger polypeptides containing bends or local elements of secondary structure. The model is consistent with several observations. First, the cytoplasmic domain of E-cadherin is not structured in solution (A. H. H. et al., unpublished data) and could bind as an extended polypeptide. Unfortunately, nothing is known about the structure of the β-catenin interaction regions of APC or Tcf-family transcription factors. Second, as noted above, deletion mutagenesis studies show a correlation between the number of arm repeats removed and the strength of binding, as would be expected if the interaction surface runs along the groove. Third, phosphorylation of the β-catenin binding region of APC by GSK-3, which would make this region even more acidic, enhances the APC–β-catenin interaction (Rubinfeld et al., 1996). Finally, mutations in the Drosophila Tcf family member Pangolin-1 that diminish, but do not completely abolish, binding to armadillo are changes from glutamate to lysine (Brunner et al., 1997). These results are also consistent with the importance of charge complementarity. Also, an extended region of interaction dominated by electrostatics, rather than close steric complementarity over a more limited area, would be expected to be somewhat tolerant of local disruptions caused by single amino acid substitutions.

The internal rigid-body motion observed between the

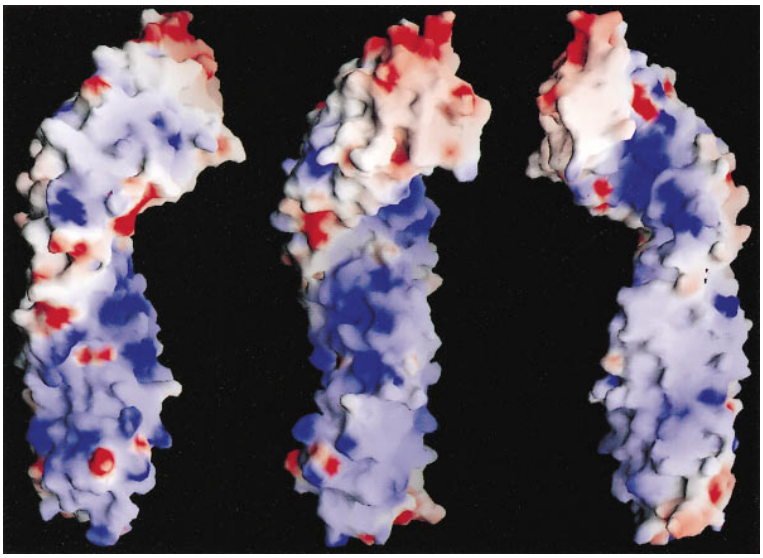

Figure 6. Electrostatic Surface Representation of β59

Blue represents regions of positive potential and red regions of negative potential, at the 10 kT/*e* level. The three views are related by successive 90° rotations about the vertical axis. The NH_2-terminal end of the molecule is at the bottom of the figure. This figure was made with GRASP (Nicholls, 1992).

two crystal forms of β59 produces a local change in the width of the groove near repeat 7. The significance of this motion cannot be assessed in the absence of a β-catenin–ligand complex structure. The two structures may simply represent a "breathing" mode of the molecule. Alternatively, the motion may be evidence of an "induced-fit" type of binding in which β-catenin wraps around its partner. The insert between H2 and H3 of repeat 10 borders the groove and could have a role in ligand binding, as the flexible 13 amino acid loop could reach the groove were it to adopt an extended conformation.

Relationship to Other Armadillo Family Members

Given the continuous packing of armadillo repeats in β59, it is unlikely that a single repeat would form a stable, independently folded structure. Interestingly, no proteins with fewer than six tandem armadillo repeats have been described. As armadillo repeats are largely defined by the conservation of hydrophobic core residues (Figure 4), the repeats found in other proteins (Peifer et al., 1994a; Küssel and Frasch, 1995) are likely to form a superhelix with a shallow groove like that observed in β59. However, the nature of the ligand binding surface is likely to vary greatly among distantly related family members. For example, the armadillo repeat regions of plakoglobin and p120^{cas}, like β-catenin, bind the E-cadherin cytoplasmic domain and are highly basic (calculated pI $\geq$ 8.0), whereas the repeat regions of smgGDS, srp1, and pendulin are mildly acidic.

Ultimately, high resolution cocrystal structures will be required to understand how β-catenin can interact with several classes of proteins and carry out its functions as a structural component of cellular junctions and a signaling molecule. Nonetheless, the remarkable surface profile of β59 observed in the present structure provides a framework for further experiments designed to probe the molecular basis of ligand recognition by β-catenin and other arm repeat proteins.

Experimental Procedures

Purification and Characterization of β-Catenin Tryptic Fragments

Full-length murine β-catenin and the murine E-cadherin cytoplasmic domain (E-cyto) were expressed in E. coli as COOH-terminal fusions to glutathione S-transferase (GST) and were purified from cell lysates by batch affinity purification on glutathione-agarose beads followed by thrombin cleavage, anion exchange, and size exclusion chromatography (A. H. H. et al., unpublished data). A digest mixture containing 2.3 mg/ml β-catenin and either 4.2 μg/ml trypsin (TPCK treated, Worthington Biochemical Corporation) or 8.3 μg/ml subtilisin (Boehringer Mannheim GmbH) in 90 mM Tris–HCl (pH 8.5), 2 mM $CaCl_2$, and 4 mM DTT was incubated at 24°C for 20 minutes. Trypsin was inhibited by adding TLCK to a final concentration of 1.4 mg/ml, and the subtilisin digest was stopped with PMSF at a final concentration of 7.7 mM. Anion exchange chromatography (Mono Q, Pharmacia) was performed by applying the digested sample in 20 mM Tris–HCl (pH 8.5), 1 mM DTT and eluting with a NaCl gradient. A Superdex 200 HR column (Pharmacia) equilibrated with 50 mM Tris–HCl (pH 8.5), 150 mM NaCl, and 2 mM DTT was used for size exclusion chromatography experiments, and a C4 HPLC column (Macrosphere 300, 5 μm particle size, Alltech) was used for reversed-phase protein purification. Anion exchange and size exclusion chromatography steps were carried out at 4°C, and the reversed-phase purification was run at 24°C. Matrix-assisted laser desorption ionization time of flight mass spectroscopy was carried out at the Ericsson Mass Spectrometry Facility, Department of Biochemistry, University of Washington at Seattle, USA, and at Ciphergen Biosystems, Inc., Mountain View, California, USA.

The ability of the digest product to bind to E-cadherin was assessed by incubating a 25 μl volume of either trypsin or subtilisin digest mixture for 80 minutes at 4°C with 1 ml of binding buffer (100 mM Tris–HCl [pH 8.5], 200 mM NaCl, 20 mM EDTA, and 2 mM DTT) containing a 25 μl bed volume of either E-cyto/GST beads or glutathione-agarose beads as a negative control. Bead pellets were washed four times with 1 ml of binding buffer and boiled in reducing SDS–PAGE loading buffer, and the proteins in the supernatant were separated by SDS–PAGE (data not shown).

Expression and Purification of β59

The cDNA sequence encoding residues 134–671 of murine β-catenin (β59) was amplified by polymerase chain reaction and the resulting product subcloned and expressed in E. coli as a COOH-terminal fusion to GST. The last two residues (Gly-Ser) of the thrombin cleavage site linking GST and β59 remain at the NH_2 terminus of β59 after digestion of the fusion protein with thrombin. The fusion protein

was purified from cell lysates on glutathione-agarose beads and cleaved with bovine thrombin, and the soluble β59 was purified by anion exchange chromatography (Mono Q, Pharmacia). For preparation of selenomethionyl-β59, the E. coli methionine auxotroph DL-41 (Hendrickson et al., 1990) was transformed with the expression plasmid and grown in defined medium (Hendrickson et al., 1990) containing the Kao and Michayluk vitamin solution (Sigma, 1× final concentration) and 125 mM seleno-DL-methionine. Amino acid analysis was consistent with 100% substitution of the 16 methionines in β59 with selenomethionine.

Crystallization

Two crystal forms of β59 were obtained using the hanging drop vapor diffusion method and a 7 mg/ml solution of β59 in 10 mM CAPS (pH 10.5), 50 mM NaCl, and 1 mM DTT. Form A crystals ($C222_1$, a = 64.1 Å, b = 102.0 Å, c = 187.0 Å) were obtained from clear drops at 4°C over a well solution of 0.5% (w/v) polyethylene glycol 8000 (PEG 8K), 200 mM Tris–HCl (pH 8.5), and 1.5–2.4 M urea. The crystal of selenomethionyl β59 used for MAD data collection was grown in 2.4 M urea. Form B crystals ($P2_12_12_1$, a = 51.1 Å, b = 75.6 Å, c = 134.4 Å) grew from precipitate at room temperature using a well solution consisting of 10% (w/v) PEG 8K, 200 mM Tris–HCl (pH 8.5), and 875 mM NaCl.

Phasing and Refinement of Form A Crystals

A four wavelength MAD phasing experiment was carried out at beam line 1–5 of the Stanford Synchrotron Radiation Laboratory. Data were measured from a single, cryopreserved crystal of SeMet β59 using a four-circle goniostat and Fuji imaging plates. The crystal was aligned with a major axis coincident with the rotation axis so that Bijvoet pairs could be measured simultaneously. All data were integrated and scaled with DENZO and SCALEPACK (Otwinowski, 1993) (Table 1A). The MADSYS suite of programs (Hendrickson, 1991) was used to extract $|^{\circ}F_A|$, $|^{\circ}F_T|$, and $^{\circ}\varphi_T - {}^{\circ}\varphi_A$ ($\Delta\varphi$) for each reflection and to calculate $^{\circ}\varphi_T$. Six Se sites were located in a $|^{\circ}F_A|$ Patterson map using an automated Patterson search program (RSPS; Knight, 1989) and refined against $|^{\circ}F_A|$ values. Difference Fourier syntheses ($|^{\circ}F_A| - |^{\circ}F_{A|calc}$) were used to locate five additional Se sites. The phases $^{\circ}\varphi_T$ were calculated using both hands of the Se substructure and the correct hand determined by comparison of the resulting $^{\circ}F_T$ electron density maps. Additional refinement of the Se model at 2.7 Å was carried out using a maximum likelihood, pseudo-MIR refinement algorithm (Burling et al., 1996) using the λ_3 (absorption edge) positive Friedel set as a reference. A dispersive difference Fourier map revealed the presence of four more Se sites. A final maximum likelihood refinement of the 15 Se site model (corresponding to 14 of the 16 selenomethionine residues) and phase calculation were performed at 2.4 Å. The resulting figure-of-merit weighted electron density map was readily interpretable (Table 1 and Figure 2). To use the experimental phases as restraints during the initial rounds of refinement, the complex residual target function $(F_{obs} - F_{calc})^2$ was used with $|F_{obs}|$ and φ_{obs} at the reference ("edge") wavelength. Later rounds of refinement employed a maximum likelihood amplitude target function (Pannu and Read, 1996; Adams et al., 1997). The "high energy remote" data set (λ_1) was used to extend the resolution of the structure to 2.1 Å. An overall anisotropic temperature factor tensor (Sheriff and Hendrickson, 1987) was applied throughout the refinement. The final model is composed of amino acids 193–549 and 563–662, and 265 solvent molecules (Table 1C); 95.3% of the amino acids are in the most favorable region of the Ramachandran plot, and none are in outlying regions.

Phasing and Refinement of Form B Crystals

A 2.9 Å resolution native data set (Table 1A) was measured from a single cryopreserved Form B crystal using a rotating anode source and a Rigaku RAXIS-IIc imaging plate detector. The Form B crystal structure was solved by molecular replacement using the refined SeMet β59 structure as a search model. A cross rotation function was calculated using data from 10.0–4.0 Å and Patterson vectors between 5.0 and 45.0 Å. The highest solution and next independent solution were 2.6 and 2.4 standard deviations (σ) above the mean, respectively. Patterson correlation refinement (Brünger, 1990) was used to improve the rotation function solution. The highest and next independent translation function solutions were 8.2 σ and 5.6 σ above the mean, respectively. Rigid-body refinement produced a final molecular replacement solution with an R value of 49.9% for all data with $|F| > 2\sigma(|F|)$ from 10.0–4.0 Å resolution. Manual rebuilding of the molecular replacement solution was followed by simulated annealing refinement to 2.9 Å and several subsequent rounds of conjugate gradient minimization. A maximum likelihood target function and an overall anisotropic temperature factor tensor were used throughout. Refinement of individual atomic temperature factors and the application of a bulk solvent correction (Jiang and Brünger, 1994) were carried out late in the refinement. The bulk solvent correction produced some improvements in ambiguous sections of the electron density map. (The correction was not applied in the Form A crystal structure refinement, as it did not improve map quality and gave a small increase in R_{free} past 3 Å.) The final model is composed of amino acids 150–549 and 564–665; 90.0% of the residues are in the most favorable region of the Ramachandran plot, and none are in outlying regions.

Coordinates and structure factors for both crystal forms of β59 have been deposited in the Protein Data Bank, Brookhaven, NY, with accession codes 2bct (Form B) and 3bct (Form A).

Acknowledgments

We thank H. Bellamy of SSRL for beamline support; A. Kolatkar, K. Misura, K. Ng, and S. Pokutta for assistance with data collection; S. Fridman and Y. Wang for technical assistance; S. Clark for assistance with mass spectroscopy; A. Brünger for providing the maximum likelihood refinement program and for discussions; and S. Halfon, S. Pokutta, and J. Wedekind for comments on the manuscript. A. H. H. was supported by a fellowship from the Jane Coffin Childs Memorial Fund for Medical Research. This work was supported by the National Institutes of Health (W. J. N.) and the Pew Scholars Program in the Biomedical Sciences (W. I. W.).

Received July 7, 1997; revised July 21, 1997.

References

Aberle, H., Butz, S., Stappert, J., Weissig, H., Kemler, R., and Hoschuetzky, H. (1994). Assembly of the cadherin-catenin complex in vitro with recombinant proteins. J. Cell Sci. *107*, 3655–3663.

Aberle, H., Schwartz, H., Hoschuetzky, H., and Kemler, R. (1996). Single amino acid substitutions in proteins of the *armadillo* gene family abolish their binding to α-catenin. J. Biol. Chem. *271*, 1520–1526.

Adams, P.D., Pannu, N.S., Read, R.J., and Brünger, A.T. (1997). Cross-validated maximum likelihood enhances crystallographic simulated annealing refinement. Proc. Natl. Acad. Sci. USA *94*, 5018–5023.

Barth, A.I.M., Pollack, A.L., Altschuler, Y., Mostov, K.E., and Nelson, W.J. (1997). NH_2-terminal deletion of β-catenin results in stable colocalization of mutant β-catenin with Adenomatous Polyposis Coli protein and altered MDCK cell adhesion. J. Cell Biol. *136*, 693–706.

Behrens, J., von Kries, J.P., Kühl, M., Bruhn, L., Wedlich, D., Grosschedl, R., and Birchmeier, W. (1996). Functional interaction of β-catenin with the transcription factor LEF-1. Nature *382*, 638–642.

Brünger, A.T. (1990). Extension of molecular replacement: a new search strategy based on Patterson correlation refinement. Acta Cryst. *A46*, 46–57.

Brunner, E., Peter, O., Schweizer, L., and Basler, K. (1997). *pangolin* encodes a Lef-1 homologue that acts downstream of Armadillo to transduce the Wingless signal in Drosophila. Nature *385*, 829–833.

Burling, F.T., Weis, W.I., Flaherty, K.M., and Brünger, A.T. (1996). Direct observation of protein solvation and discrete disorder with experimental crystallographic phases. Science *271*, 72–77.

He, X., Saint-Jeannet, J.-P., Woodgett, J.R., Varmus, H.E., and Dawid, I.B. (1995). Glycogen synthase kinase-3 and dorsoventral patterning in Xenopus embryos. Nature *374*, 617–622.

Hendrickson, W.A. (1991). Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. Science *254*, 51–58.

Hendrickson, W.A., Horton, J.R., and LeMaster, D.M. (1990). Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. EMBO J. *9*, 1665–1672.

Hirschl, D., Bayer, P., and Müller, O. (1996). Secondary structure of an armadillo single repeat from the APC protein. FEBS Lett. *383*, 31–36.

Hoschuetzky, H., Aberle, H., and Kemler, R. (1994). β-catenin mediates the interaction of the cadherin-catenin complex with epidermal growth factor receptor. J. Cell. Biol. *127*, 1375–1380.

Huber, O., Korn, R., McLaughlin, J., Ohsugi, M., Herrmann, B.G., and Kemler, R. (1996). Nuclear localization of β-catenin by interaction with transcription factor LEF-1. Mech. Dev. *59*, 3–10.

Hülsken, J., Birchmeier, W., and Behrens, J. (1994). E-cadherin and APC compete for the interaction with β-catenin and the cytoskeleton. J. Cell Biol. *127*, 2061–2069.

Jiang, J.-S., and Brünger, A.T. (1994). Protein hydration observed by X-ray diffraction. J. Mol. Biol. *243*, 100–115.

Kinch, M.S., Clark, G.J., Der, C.J., and Burridge, K. (1995). Tyrosine phosphorylation regulates the adhesions of ras-transformed breast epithelia. J. Cell Biol. *130*, 461–471.

Kinzler, K.W., and Vogelstein, B. (1996). Lessons from hereditary colorectal cancer. Cell *87*, 159–170.

Knight, S. (1989). Ribulose 1,5-bisphosphate carboxylase/oxygenase—a structural study. PhD Thesis, Swedish University of Agricultural Sciences, Uppsala, Sweden.

Korinek, V., Barker, N., Morin, P.J., van Wichen, D., de Weger, R., Kinzler, K.W., Vogelstein, B., and Clevers, H. (1997). Constitutive transcriptional activation by a β-catenin-Tcf complex in $APC^{-/-}$ colon carcinoma. Science *275*, 1784–1787.

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

Küssel, P., and Frasch, M. (1995). Pendulin, a *Drosophila* protein with cell cycle-dependent nuclear localization, is required for normal cell proliferation. J. Cell Biol. *129*, 1491–1507.

Molenaar, M., van de Wetering, M., Oosterwegel, M., Peterson-Maduro, J., Godsave, S., Korinek, V., Roose, J., Destrée, O., and Clevers, H. (1996). XTcf-3 transcription factor mediates β-catenin-induced axis formation in Xenopus embryos. Cell *86*, 391–399.

Morin, P.J., Sparks, A.B., Korinek, V., Barker, N., Clevers, H., Vogelstein, B., and Kinzler, K.W. (1997). Activation of β-catenin—Tcf signaling in colon cancer by mutations in β-catenin or APC. Science *275*, 1787–1790.

Munemitsu, S., Albert, I., Souza, B., Rubinfeld, B., and Polakis, P. (1995). Regulation of intracellular β-catenin levels by the adenomatous polyposis coli (APC) tumor-suppressor protein. Proc. Natl. Acad. Sci. USA *92*, 3046–3050.

Munemitsu, S., Albert, I., Rubinfeld, B., and Polakis, P. (1996). Deletion of an amino-terminal sequence stabilizes β-catenin in vivo and promotes hyperphosphorylation of the adenomatous polyposis coli tumor suppressor protein. Mol. Cell. Biol. *16*, 4088–4094.

Nagafuchi, A., and Takeichi, M. (1988). Cell binding function of E-cadherin is regulated by the cytoplasmic domain. EMBO J. *7*, 3679–3684.

Näthke, I.S., Adams, C.L., Polakis, P., Sellin, J.H., and Nelson, W.J. (1996). The adenomatous polyposis coli tumor suppressor protein localizes to plasma membrane sites involved in active cell migration. J. Cell Biol. *134*, 165–179.

Nicholls, A. (1992). GRASP: Graphical Representation and Analysis of Surface Properties (New York: Columbia University).

Orsulic, S., and Peifer, M. (1996). An in vivo structure-function study of armadillo, the β-catenin homologue, reveals both separate and overlapping regions of the protein required for cell adhesion and for wingless signaling. J. Cell Biol. *134*, 1283–1300.

Otwinowski, Z. (1993). Oscillation reduction program. In Proceedings of the CCP4 Study Weekend: "Data Collection and Processing," 29–30 January 1993, L. Sawyer, N. Isaacs, and S. Bailey, eds. (Daresbury, UK, SERC Daresbury Laboratory), pp. 56–62.

Ozawa, M., Baribault, H., and Kemler, R. (1989). The cytoplasmic domain of the cell adhesion molecule uvomorulin associates with three independent proteins structurally related in different species. EMBO J. *8*, 1711–1717.

Ozawa, M., Ringwald, M., and Kemler, R. (1990). Uvomorulin-catenin complex formation is regulated by a specific domain in the cytoplasmic region of the cell adhesion molecule. Proc. Natl. Acad. Sci. USA *87*, 4246–4250.

Pai, L.-M., Kirkpatrick, C., Blanton, J., Oda, H., Takeichi, M., and Peifer, M. (1996). *Drosophila* α-catenin and E-cadherin bind to distinct regions of *Drosophila* Armadillo. J. Biol. Chem. *271*, 32411–32420.

Pannu, N.S., and Read, R.J. (1996). Improved structure refinement through maximum likelihood. Acta Cryst. *A52*, 659–668.

Park, H.-W., Boduluri, S.R., Moomaw, J.F., Casey, P.J., and Beese, L.S. (1997). Crystal structure of protein farnesyltransferase at 2.25 Angstrom resolution. Science *275*, 1800–1804.

Peifer, M., Orsulic, S., Sweeton, D., and Wieschaus, E. (1993). A role for the Drosophila segment polarity gene *armadillo* in cell adhesion and cytoskeletal integrity during oogenesis. Development *118*, 1191–1207.

Peifer, M., Berg, S., and Reynolds, A.B. (1994a). A repeating amino acid motif shared by proteins with diverse cellular roles. Cell *76*, 789–791.

Peifer, M., Sweeton, D., Casey, M., and Weischaus, E. (1994b). *wingless* signal and Zeste-white 3 kinase trigger opposing changes in the intracellular distribution of Armadillo. Development *120*, 369–380.

Raag, R., Appelt, K., Xuong, N.-H., and Banaszak, L. (1988). Structure of the lamprey yolk lipid-protein complex lipovitellin-phosvitin at 2.8 Å resolution. J. Mol. Biol. *200*, 553–569.

Riese, J., Yu, X., Munnerlyn, A., Eresh, S., Hsu, S.-C., Grosschedl, R., and Bienz, M. (1997). LEF-1, a nuclear factor coordinating signaling inputs from *wingless* and *decapentaplegic*. Cell *88*, 777–787.

Rimm, D.L., Koslov, E.R., Kebriaei, P., Cianci, C.D., and Morrow, J.S. (1995). α_1(E)-catenin is an actin-binding and -bundling protein mediating the attachment of F-actin to the membrane adhesion complex. Proc. Natl. Acad. Sci. USA *92*, 8813–8817.

Rubinfeld, B., Souza, B., Albert, I., Müller, O., Chamberlain, S.H., Masiarz, F.R., Munemitsu, S., and Polakis, P. (1993). Association of the *APC* gene product with β-catenin. Science *262*, 1731–1734.

Rubinfeld, B., Souza, B., Albert, I., Munemitsu, S., and Polakis, P. (1995). The APC protein and E-cadherin form similar but independent complexes with α-catenin, β-catenin, and plakoglobin. J. Biol. Chem. *270*, 5549–5555.

Rubinfeld, B., Albert, I., Porfiri, E., Fiol, C., Munemitsu, S., and Polakis, P. (1996). Binding of GSK3β to the APC-β-catenin complex and regulation of complex assembly. Science *272*, 1023–1026.

Rubinfeld, B., Robbins, P., El-Gamil, M., Albert, I., Porfiri, E., and Polakis, P. (1997). Stabilization of β-catenin by genetic defects in melanoma cell lines. Science *275*, 1790–1792.

Ruediger, R., Hentz, M., Fait, J., Mumby, M., and Walter, G. (1994). Molecular model of the A subunit of protein phosphatase 2A: interaction with other subunits and tumor antigens. J. Virol. *68*, 123–129.

Sheriff, S., and Hendrickson, W.A. (1987). Description of overall anisotropy in diffraction from macromolecular crystals. Acta Cryst. *A43*, 118–121.

Stappert, J., and Kemler, R. (1994). A short core region of E-cadherin is essential for catenin binding and is highly phosphorylated. Cell Adhesion Commun. *2*, 319–327.

Su, L.-K., Vogelstein, B., and Kinzler, K.W. (1993). Association of the APC tumor suppressor protein with catenins. Science *262*, 1734–1737.

Takeichi, M. (1991). Cadherin cell adhesion receptors as a morphogenetic regulator. Science *251*, 1451–1455.

Thunnissen, A.-M.W.H., Dijkstra, A.J., Kalk, K.H., Rozeboom, H.J., Engel, H., Keck, W., and Dijkstra, B.W. (1994). Doughnut-shaped structure of a bacterial muramidase revealed by x-ray crystallography. Nature *467*, 750–753.

van de Wetering, M., Cavallo, R., Booijes, D., van Beest, M., van Es, J., Loureiro, J., Ypma, A., Hursh, D., Jones, T., Bejsovec, A., Peifer, M., Mortin, M., and Clevers, H. (1997). Armadillo coactivates transcription driven by the product of the Drosophila segment polarity gene *dTCF*. Cell *88*, 789–799.

van Leeuwen, F., Samos, C.H., and Nusse, R. (1994). Biological activity of soluble *wingless* protein in cultured Drosophila imaginal disc cells. Nature *368*, 342–344.

Weis, W.I., Kahn, R., Fourme, R., Drickamer, K., and Hendrickson, W.A. (1991). Structure of the calcium-dependent lectin domain from a rat mannose-binding protein determined by MAD phasing. Science *254*, 1608–1615.

Werner, M.H., and Burley, S.K. (1997). Architectural transcription factors: proteins that remodel DNA. Cell *88*, 733–736.

Yost, C., Torres, M., Miller, J.R., Huang, E., Kimelman, D., and Moon, R.T. (1996). The axis-inducing activity, stability, and subcellular distribution of β-catenin is regulated in Xenopus embryos by glycogen synthase kinase 3. Genes Dev. *10*, 1443–1454.