Systematic Peptide Array-based Delineation of the Differential 
β-Catenin Interaction with Tcf4, E-Cadherin, and Adenomatous Polyposis Coli*

Robert Gail†, Ronald Frank§, and Alfred Wittinghofer¶¶

From the †Abteilung strukturelle Biologie, Max-Planck-Institut für molekulare Physiologie, Otto-Hahn-Strasse 11, 44227 Dortmund, Germany and §Abteilung Chemische Biologie, Gesellschaft für Biotechnologische Forschung, Mascheroder Weg 1, 38124 Braunschweig, Germany

Nuclear accumulation of the complex between β-catenin and proteins of the T-cell factor (Tcf) family is a hallmark of many cancers. Targeting this interaction for drug development is complicated by the fact that E-cadherin and adenomatous polyposis coli (APC) bind to overlapping sites on β-catenin. Inhibiting their interactions might actually promote tumor growth. To identify selective β-catenin binding hot spots of Tcf4, E-cadherin, and APC, array technology with peptides of up to 53 amino acids length was used. Interactions were monitored by a quantitative fluorescent readout, which was shown to represent a monitor of true equilibrium binding constants. We identified minimal binding motifs in the β-catenin ligands and showed that most of the 15-mer and 20-mer repeats of APC did not interact, at least when non-phosphorylated, and defined a consensus binding motif also present in APC. We confirmed previously found hot spots and identified new ones. The method allowed us to locate a hydrophobic pocket that was relevant for the Tcf, but not the E-cadherin interaction, and would thus constitute an ideal drug target site.

In humans there are around 35,000 protein-coding genes (1, 2). Alternative splicing and posttranslational modifications further increase the number of different proteins. To understand their function, it is necessary to know the interactions they are involved in. In yeast, with its ~6000 genes (3), different high-throughput methods have defined a total of around 80,000 protein-protein interactions, with the real number probably being somewhat lower (4). These searches have also shown that most of the proteins bind multiple other proteins. This may be coupled to diverse and sometimes opposing biological effects. Many interactions are involved in signal transduction pathways leading to differentiation, growth, or apoptosis.

However, therapeutic intervention with such complexes has proven to be challenging because protein-protein interfaces are rather large and often flat, such that they are difficult to disrupt with small molecules. However, in some cases, it has been shown that this is possible (5–8). Even though a growing number of complex structures are available, the relation between interface features and binding strength is not well understood from a theoretical point of view. No parameter observable in crystal structures is correlated to its importance for the interaction. On the other hand, empirical analysis of interfaces has shown that generally only a few residues, the so-called hot spots, contribute the bulk of the binding free energy (9, 10). A detailed understanding of protein-protein interfaces and the rapid identification of hot spot residues are thus required to understand the thermodynamics and specificity of interactions and to apply such knowledge for the development of specific and potent inhibitory drugs.

β-Catenin is an important paradigm for those studies. It exists in two distinct cellular pools, where it carries out two different functions through its binding to a number of diverse proteins. For its role in adherens junctions at the cell membrane, it interacts with the cytoplasmic tail of E-cadherin and with α-catenin. This is crucial for stable cell adhesion (11). The cytoplasmic β-catenin pool serves as a signaling molecule in the Wnt pathway, which plays an important role in development and tissue maintenance and is often found deregulated in cancers (12, 13). In the absence of an extracellular Wnt signal, the cytoplasmic β-catenin concentration is kept low through proteolysis. For β-catenin degradation, its interaction with APC in a multiprotein complex is essential (14). Wnt stimulation results in stabilization of β-catenin, which then translocates into the nucleus, where its complex with Tcf proteins (Tcf1, LEF1, Tcf3, and Tcf4) transcriptionally activates a number of target genes (15, 16).

β-Catenin accumulates in the nucleus of most colorectal tumors and certain other cancers. This is due to mutation in either β-catenin itself or in negative regulators such as APC (17). The β-catenin/Tcf4 complex plays a central role in oncogenesis (18), and its inhibition by dominant negative Tcf4 stops proliferation of colorectal cancer cells and induces differentiation (19), proving that β-catenin is an important anti-cancer target (20). Recently, compounds inhibiting the association between β-catenin and Tcf4 have been described (7). Targeting β-catenin implies another level of complexity due to its interactions with E-cadherin and APC, which should remain unaffected by potential drugs. Loss of the E-cadherin interaction is causal in the transition from adenomas to invasive carcinomas (21), so that its inhibition might actually promote tumor metastasis. Because the β-catenin/APC interaction regulates the level of β-catenin in healthy cells and its localization (22), interfering with it might cause inappropriate activation of β-catenin signaling. The structures of β-catenin complexes with Tcf (23–25), E-cadherin (26), and APC (27) have shown that these proteins all bind to the superhelical groove in the central armadillo repeat region of

* 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 1754 solely to indicate this fact.

† To whom correspondence should be addressed. E-mail: alfred.wittinghofer@mpi-dortmund.mpg.de.

¶¶ The abbreviations used are: APC, adenomatous polyposis coli; Tcf, T-cell factor; CBD, catenin-binding domain; GST, glutathione S-transferase; GFP, green fluorescent protein; Fmoc, N-(9-fluorenyl)methoxy-carbonyl; ITC, isothermal titration calorimetry.
β-catenin and have partially overlapping binding sites, so that selectivity might be difficult to achieve (Fig. 1). The molecules described by Lepourcelet et al. (7) leave the E-cadherin interaction undisturbed but inhibit APC binding, which indicates they need to be optimized. To accomplish this goal, it is necessary to know the specific and differential binding determinants of the respective interactions.

For β-catenin, hot spots have been defined by introducing alanine mutations in the armadillo domain and measuring the effect on LEF-1, APC, and conductin binding (28). This has identified residues that are equally or differentially important for these interactions. Similarly, hot spots were defined in the catenin-binding domain (CBD) of Tcf (29–31), which is located in the first 50–60 residues (32).

We have analyzed β-catenin binding of Tcf4, E-cadherin, and APC in a systematic and comprehensive manner using peptide arrays and fluorescence-labeled β-catenin. The method allowed us to quickly identify specific and overlapping hot spots for different protein-protein interactions. Peptide array data could be directly correlated to affinity by equilibrium measurements using fluorescence polarization titration. We also further narrowed down the minimal binding fragment of Tcf4 and defined binding regions of E-cadherin. Surprisingly, the peptide array technology also showed that β-catenin binding of APC is not strictly defined by or restricted to its 15-mer and 20-mer repeats and allowed us to more rigorously define β-catenin binding regions in APC.

During preparation of the manuscript, our conclusions about binding motifs in APC were corroborated by x-ray structural studies (33, 34).

**EXPERIMENTAL PROCEDURES**

**Expression and Purification of β-Catenin, Tcf4, and E-Cadherin—** Recombinant protein expression was done in Escherichia coli BL21(DE3) using the vector pGEX-4T1 and TB medium supplemented with 100 mg/liter ampicillin. After induction with 0.3 mM isopropyl-β-D-galactopyranoside, the temperature was lowered from 37 °C to 18 °C for overnight production of the GST fusion proteins. The bacteria were lysed in a Microfluidizer M-110S (Microfluidics, Newton, MA). For human β-catenin, the armadillo domain (residues 151–666) was used. The buffer was 30 mM Tris/HCl, pH 8, 300 mM NaCl, and 3 mM dithioerythritol. For cell lysis, 1 mM phenylmethylsulfonyl fluoride was added. After affinity chromatography on reduced glutathione-Sepharose (Amersham Biosciences), β-catenin was cleaved off the column by thrombin and further purified by size exclusion chromatography on Superdex 75 (Amersham Biosciences). For fluorescent β-catenin
(bcat-GFP), a GST-β-catenin-(151–666)-enhanced GFP fusion protein was produced as described above, and bcat-GFP cleaved off the reduced glutathione column was used directly. A Cys(Ser-Gly)-linker, linker was engineered into the N terminus of the human Tcf4 CBD (residues 1–53) for fluorescence labeling. Point mutations were introduced by QuikChange (Stratagene, Amsterdam, the Netherlands). The buffer was 30 mM Tris/HCl, pH 8.0, 200 mM NaCl, and 3 mM dithioerythritol. For lysis, 5 mM EDTA and 1 mM phenylmethylsulfonyl fluoride were added. The proteins were purified by reduced glutathione affinity columns, thrionin cleavage of the GST tag on the column, and size exclusion chromatography on Superdex 75. Two constructs of mouse Tcf4, amino acids 652–700 and 628–723, were used. Point mutations and an N-terminal linker were introduced as described for E-cadherin. The buffer was 30 mM Tris/HCl, pH 7.5, 200 mM NaCl, and 3 mM dithioerythritol. For lysis, 5 mM EDTA and 1 mM phenylmethylsulfonyl fluoride were added. Purification was done as described for Tcf4.

Synthesis of Peptide Arrays—Membranes with peptide arrays were synthesized by the SPOT technique (35) using an ASP22 automated SPOT synthesizer (Intavit Bioanalytical Instruments AG, Cologne, Germany). Briefly, acid-resistant ACS01 amino-pegylated cellulose membranes (0.6 μmol amino groups/cm²; AIMS Scientific Products, Braunschweig, Germany) were derivatized in a grid of spots of 0.3 cm in diameter, defining the location of the peptides by applying to each 0.1 μl of a solution containing 0.05 M Fmoc-β-alanine and 0.27 M N-acetyl-

amidoamine overnight. Throughout solid phase synthesis, Fmoc-amino acid 1-hydroxybenzotriazole esters were used. After every coupling round, unreacted amino groups were blocked with acetic anhydride, Fmoc-NH2, and the terminal amino functions of the growing peptides were stained with bromphenol blue. Finally, all peptides were acetylated at their N termini, and the side chain protecting groups were cleaved by treatment with 82% trifluoroacetic acid, 3% trisobutylsilane, 5% dichloromethane, and 10% water for 16 h.

Investigation of β-Catenin Binding to Peptide Arrays—After moisturization of peptide membranes, peptide membranes were washed twice with Tris-buffered saline (50 mM Tris/HCl, pH 8.0, 173 mM NaCl, and 2.7 mM KCl) and blocked overnight with 80% T-TBS (Tris-buffered saline + 0.05% Tween 20), 20% Blocking Buffer (Genosys, Deisenhofen, Germany), and 5% sucrose. After washing with T-TBS, 1 μM bcat-GFP in 24 mM Tris/HCl, pH 8.0, 240 mM NaCl, 3 mM dithioerythritol, 20% Genosys Blocking Buffer, and 0.05% Tween 20 was added for 2 h at room temperature. After three washing steps with T-TBS, the amount of bound bcat-GFP was quantified on a LA-5000 imager (FujiFilm, Düsseldorf, Germany) using an excitation wavelength of 473 nm and a 510-nm high pass filter for the emitted fluorescence light. The amount of bound bcat-GFP was measured using a one-site binding model. The amount of bound bcat-GFP was measured using a one-site binding model.

RESULTS

Design and Preparation of Peptide Arrays—In order to get a comprehensive view of the energetics and specificity of the different β-catenin interactions with Tcf4, E-cadherin, and APC, we applied a systematic single amino acid replacement strategy. Synthetic peptide arrays as available by the SPOT technique (36) can easily be assembled in parallel on a microscale. The longest peptide successfully synthesized thus far by this technique consists of 44 residues (37). Because the CBD of Tcf4 contains an additional 9 residues, its synthesis is, by itself, challenging. However, because the respective CBDs are unstructured (27, 31, 38), we reasoned that their synthesis and presentation in a binding assay should not be complicated by secondary structure and folding problems. We were therefore confident in extending the synthetic aim to longer sequences.

By using the peptide companion software (www.5z.com/csp/), we predicted the absence of difficult synthetic steps. To avoid inhibitory molecular crowding due to the large molecular weight of the product, we found it necessary to reduce the concentration of starting amino functions on the membrane by blending the Fmoc-β-alanine with a 10-fold excess of the elongation-incapable N-acetyl-β-alanine (data not shown). The completeness of the coupling reaction was assured by repeating the spotting procedure up to four times, if necessary. Additionally, we have chosen a new, acid-resistant cellulose membrane, which allows an extended treatment with high concentrations of trifluoroacetic acid and thus a complete deprotection of the peptides (39).

Alanine Scan of Tcf4 Binding to β-Catenin—Peptide arrays with individual alanine point mutants of the complete CBD of human Tcf4 (residues 1–53) (24) were incubated with a fusion protein of the armadillo domain of human β-catenin and enhanced GFP (bcat-GFP). The amount of bcat-GFP bound to Tcf4 was quantified with a fluorescence imager. The quantitative evaluation is shown in Fig. 2A. The error bars represent the standard deviation from results obtained with three individually synthesized peptide arrays, showing that interaction measurement by the SPOT method was quite reproducible.

Mutation of Asp16 or Phε21 had the strongest effect on β-catenin binding. Replacement of either residue for Ala resulted in a signal indistinguishable from the negative controls (randomized sequence), indicating almost complete loss of β-catenin binding to Tcf4. At least 50% reduction in signal resulted from the mutations of Glu17, Leu18, Ile19, Asp23, Leu41, or Leu48. The mutations V49A and E51A in the C-terminal part of the Tcf4-CBD resulted in a >30% signal reduction. These data confirm and extend previous hot spot studies (see below), establishing peptide array analysis as a valuable method to study β-catenin interactions. In contrast, the acidic region from residue 24 to residue 29 of Tcf4, which binds an important β-catenin lysine through alternative conformations (24), was of only moderate importance for β-catenin binding. Surprisingly, mutation of either Gly13 or Asn15 led to a consistent 30% increase in the amount of bcat-GFP bound to the membrane spot.

Solution Assay to Characterize Hot Spot Mutants—We wondered whether peptide arrays as a non-equilibrium method reflect true relative affinities. Thus, recombinant wild-type and mutant Tcf4 peptides were prepared with an additional linker and a cysteine at the N terminus, to allow coupling of a fluorophore. The affinity of the peptides was determined by titration with the armadillo domain of human β-catenin and measurement of the increase in fluorescence polarization due to formation of a large molecular weight complex. Representative titration curves are shown in Fig. 3A for the L12A and V49A mutants, and the results are summarized in Table I.

The equilibrium dissociation constant Kd of the wild-type CBD peptide was very low (0.84 nM). The most important β-catenin binding residue in Tcf4 was Asp16, in agreement with previous studies using various qualitative and quantitative methods. Mutation of Asp16 to Ala lowered the affinity 1700-fold, corresponding to a ΔΔG of 4.3 kcal/mol at 20 °C. A strong effect was also observed for Phε21, Leu41, and Leu48, whose substitution each reduced affinity 250–300-fold (3.2–3.3 kcal/mol). A considerable contribution to binding was also made by Glu17, Leu18, and Ile19 (1.5–2.1 kcal/mol). In support of the
reliability of the peptide array analysis, mutation of Gly\textsuperscript{13} slightly increased the affinity.

Isothermal titration calorimetry (ITC) was also employed to measure the interaction between unlabeled recombinant Tcf4 peptides and β-catenin. Here, weak binding mutants were employed because the affinity of the wild-type peptide was too high to be measured accurately by ITC. For the mutants D16A and F21A, a $K_d$ of 0.56 and 0.22 $\mu$M, respectively, was determined (Fig. 3B). This was in good agreement with the polarization data and showed that the fluorescence label did not perturb the interaction.

A comparison between the membrane-bound fluorescence intensity data and the equilibrium binding constants measured in solution showed a good correlation, demonstrating that relative affinities can be correctly assigned by the peptide array method (Fig. 4). The comparison also showed, however, that peptide arrays, due to the washing steps involved, were not capable of measuring true equilibrium constants but more likely reflect the different dissociation rate constants, which are likely to be related to affinities.

The Minimal β-Catenin Binding Region of Tcf4—Structures of Xenopus Tcf3-CBD and human Tcf4-CBD in complex with β-catenin showed that Tcf can adopt different conformations and secondary structures in the binding groove of the armadillo domain (Fig. 1, A and B). In addition to an elongated core structure that makes similar contacts to β-catenin, the three structures showed variations in the β-catenin-Tcf interface (23–25). To further delineate and quantify the contributions of different CBD parts and search for a minimal β-catenin binding sequence, a peptide array of overlapping 15-mer peptides covering the complete CBD of Tcf4 was synthesized. We reasoned that a peptide of 15 amino acids might be of sufficient length to bind to β-catenin (27). Filter binding showed that not all of the 53 residues of the CBD as defined previously have to be present. Whereas residues 7–9 were marginally important for binding, strong (relative) binding encompassed the 15-mers starting from residue 11 to residue 15 (Fig. 2B). 15-mers starting from residue 17 or higher did not show any binding by themselves, although some residues in the C-terminal part of the CBD such as Leu\textsuperscript{41} and Leu\textsuperscript{48} did contribute to the binding energy (see Fig. 2A and Table I). Minimal binding sequences required both of the two most crucial hot spots, Asp\textsuperscript{16} and Phe\textsuperscript{21}, for detectable affinity. Affinity measurements of soluble peptides by fluorescence polarization titration showed that

![Peptide array analysis of the β-catenin-Tcf4 interaction. A, alanine scan. Arrays with all single alanine point mutants of the catenin-binding domain of human Tcf4 (residues 1–53) were incubated with a fluorescent variant of β-catenin (bcat-GFP). Binding was quantified by fluorescence imaging. All signals were normalized to the wild-type signal. □, alanine mutants; ■, negative controls (randomized sequence); □, wild-type peptides. The sequence is indicated above the bars of the alanine mutants. B, scan for a minimal Tcf4 binding domain. A membrane with overlapping 15-mer peptides covering the whole Tcf4-CBD was probed with bcat-GFP as described previously. Different-colored bars are from three different experiments and are normalized to the strongest signal.](image-url)
Tcf4-(13–27) did indeed bind to β-catenin. However, its equilibrium dissociation constant was 3 orders of magnitude higher than that of the complete CBD (Table I). Whereas a C-terminal truncation by 2 residues (Tcf4-(13–25)) did not appreciably change the dissociation constant, removal of residues 13 and 14 (Tcf4-(15–27)) reduced affinity by another order of magnitude.

**Hot Spot Scan of E-Cadherin Binding to β-Catenin**—Having established that peptide arrays can be used to elegantly define minimal binding epitopes and interaction hot spots, we applied the method to other β-catenin binding partners. The complete β-catenin-binding region of E-cadherin, as deduced from visible residues in the complex structure (26), might contain as many as 96 residues (residues 628–723 in mature mouse E-cadherin). Because this exceeds the maximum length of a synthetically accessible peptide, a somewhat shorter construct (residues 652–700) was chosen for a systematic investigation of hot spots. Its central region contains sequences that are highly conserved among classical cadherins (residues 667–673) and also contains the extended region (residues 674–682) that superimposes well with the central core region of Tcf and forms essential contacts with β-catenin (26, 40) (Fig. 1). It comprises the complete minimal binding region of 23–27 residues that has been identified for Drosophila E-cadherin (41) and covers a similar surface of β-catenin as the Tcf4-CBD. To account for the effects of phosphorylation, every serine residue phosphorylated in the β-catenin-E-cadherin complex plus every glycogen synthase kinase-3β and casein kinase II site (Ser684, Ser686, Ser689, Ser691, Ser695, and Ser699) (42) was mimicked by a glutamate (41). As for Tcf, variants of the minimal CBD of E-cadherin (residues 652–700) with all possible single alanine mutations were synthesized on a membrane, and β-catenin binding was quantified as described above.

The two most essential residues for β-catenin binding in E-cadherin-(652–700) are Asp674 and Phe679 because their mutation reduced the fluorescence signal by >90% (Fig. 5). These amino acids are homologous to and superimposable onto Asp16 and Phe21, respectively, in Tcf4 (24–26) (Fig. 1C). Other crucial amino acids in this region are Leu676 and Leu677 (80% reduction), which correspond to Tcf4 Leu18 and Ile19 in the overlay. The E-cadherin region from Tyr681 to phospho-Ser684 (mimicked by a Glu) appeared to be more relevant for β-catenin binding than the corresponding residues 23–26 of Tcf4. In the C terminus of the E-cadherin peptides, no hot spots were detected. In particular, apart from position 684, the mimicked phosphoserines did not have an appreciable influence on binding. Notably, Leu691 and Leu694, which are in the same structural position as the important Tcf4 leucines 41 and 48, were dispensable for β-catenin binding. In contrast to Tcf4, the E-
The β-catenin interaction with Tcf4, E-Cadherin, and APC is mediated by specific binding motifs within these proteins. Peptide arrays were used to identify these motifs.

**Table 1**

| Peptide                  | \(K_d\) (nM) | Fluorophore |
|--------------------------|--------------|-------------|
| Tcf4-(1–53) wild-type    | 0.84 ± 0.18 | Fluorescein |
| Tcf4-(1–53) L12A         | 0.82 ± 0.13 | Fluorescein |
| Tcf4-(1–53) G13A         | 0.52 ± 0.05 | Fluorescein |
| Tcf4-(1–53) D16A         | 1.4 ± 0.09 | Alexa-488   |
| Tcf4-(1–53) E17A         | 21 ± 5.6   | Fluorescein |
| Tcf4-(1–53) L18A         | 12 ± 2.6   | Fluorescein |
| Tcf4-(1–53) I19A         | 31 ± 3.2   | Fluorescein |
| Tcf4-(1–53) F21A         | 220 ± 24   | Fluorescein |
| Tcf4-(1–53) D23A         | 1.8 ± 0.62 | Fluorescein |
| Tcf4-(1–53) N34A         | 0.66 ± 0.31| Fluorescein |
| Tcf4-(1–53) L41A         | 250 ± 12   | Fluorescein |
| Tcf4-(1–53) L48A         | 220 ± 22   | Fluorescein |
| Tcf4-(1–53) V49A         | 7.0 ± 0.73 | Fluorescein |
| Tcf4-(1–53) E51A         | 2.5 ± 0.41 | Fluorescein |
| Tcf4-(13–27)             | 1.4 \(\mu\)M| AMCA        |
| Tcf4-(13–25)             | 3.0 \(\mu\)M| AMCA        |
| Tcf4-(15–27)             | 10 \(\mu\)M | AMCA        |
| E-cadherin-(652–700) wild-type | 1.8 \(\mu\)M| Fluorescein |
| E-cadherin-(652–700) E6  | 100 \(nM\) | Fluorescein |
| E-cadherin-(628–723) E6  | 0.52 ± 0.16| Fluorescein |
| E-cadherin-(628–723) E6 L661A | 1.4 ± 2.6 | Fluorescein |
| E-cadherin-(628–723) E6 D674A | 1900 ± 38 | Fluorescein |
| E-cadherin-(628–723) E6 L676A | 9.2 ± 2.8 | Fluorescein |

a AMCA, 7-amino-4-methylcoumarin-3-acetic acid.

In order to determine the equilibrium binding affinities of the short 20-mer repeats (residues 652–700) versus the long 50-mer peptide array covering the 15- and 20-mer region of human APC, a peptide array with overlapping 50-mer peptides was used. For a first screen, we used a peptide array with overlapping 50-mer peptides covering the 15- and 20-mer region of human APC. For a second screen, we used a peptide array with overlapping 15-mer peptides covering the 15- and 20-mer region of human APC.

**Figure 4.** Comparison of peptide array data and equilibrium affinities. For each Tcf4 and E-Cadherin hot spot residues, affinities (\(K_d\) values) to β-catenin of the alanine point mutants measured by equilibrium fluorescence polarization were plotted against the peptide array fluorescence signal normalized to wild-type intensity.

**Figure 6A.** Comparison of peptide array data and equilibrium affinities. For each Tcf4 and E-Cadherin hot spot residues, affinities (\(K_d\) values) to β-catenin of the alanine point mutants measured by equilibrium fluorescence polarization were plotted against the peptide array fluorescence signal normalized to wild-type intensity.

**Figure 6B.** Comparison of peptide array data and affinity. For each Tcf4 and E-Cadherin hot spot residues, affinities (\(K_d\) values) to β-catenin of the alanine point mutants measured by equilibrium fluorescence polarization were plotted against the peptide array fluorescence signal normalized to wild-type intensity.

motif are required to achieve sub-nanomolar affinity. As in the case of Tcf4, peptide array data were a monitor of equilibrium affinities of wild-type and mutant E-cadherin peptides (Fig. 4).

**β-Catenin Binding Motifs in APC—** APC interacts with β-catenin through a region of >1000 amino acids (43, 44). This region does not show any recognizable compact domains and contains four conserved 15-amino acid repeats (repeats A–D) and seven 20-amino acid repeats (repeats 1–7). These are generally accepted to constitute the β-catenin binding regions (27), which were, however, examined predominantly in the presence of flanking sequences. The peptide array methodology appeared to be a perfect tool to analyze the binding motifs in the >1000-residue APC fragment to see whether all the 15- and 20-mer repeats interact with β-catenin and to delineate the binding requirements.

For a first screen, we used a peptide array with overlapping 50-mer peptides covering the 15- and 20-mer region of human APC (residues 1000–1049, residues 1010–1059, and so forth, until residues 2049–2099). Repeat A (residues 1021–1035), for example, was thus present in the 50-mer peptides starting from amino acid 1000, 1010, and 1020, all of which showed strong β-catenin binding when probed with bcat-GFP (Fig. 6A). In the 50-mers starting from residue 1110 to 1170, at least one of repeats B, C, or D was present. Surprisingly, we found that the presence of 15- or 20-mer repeats did not necessarily coincide with strong binding. The peptide starting from residue 1170, which contained repeat D, showed no detectable binding, but the peptide starting from residue 1100, which did not contain any of repeats A-D, apparently bound to bcat-GFP. Furthermore, we found 50-mer peptides containing the 20-mer repeats 3 and 5 that bound more weakly to β-catenin than adjacent peptides that did not contain the respective repeat. As an example, whereas the 1450–1499 peptide outside repeat 3 bound, the 1490–1539 peptide containing repeat 3 had no appreciable affinity.

For a higher resolution of binding sequences, a peptide array with overlapping 15-mer peptides covering the 15- and 20-mer region of human APC (residues 1001–1214) was used. The highest fluorescence signal was observed for the peptide 1021–1035 (Fig. 6B), which corresponded to the 15-mer repeat A, whose binding to β-catenin has been analyzed structurally (27). Additional significant peaks were found for 15-mers starting with residue...
1135, 1140, 1158, or 1159, which overlapped with but were not identical to repeats B, C, and D.

In order to explain these surprising results and find a common binding principle, we compared the sequences of certified β-catenin binding motifs from Tcf proteins, inhibitor of β-catenin and Tcf, and E-cadherin with the APC sequences and searched for motifs close to repeat 3 (residues 1450–1529) and repeat 5 (residues 1810–1899) that would explain the positive and negative binding data. We indeed found homology between sequences in the repeat 3 and repeat 5 areas and previously established binding sites, which is also present in the 15-mer repeat APC-A (Fig. 6C). Alignment of these sequences allowed us to define a consensus motif D-[K]-[LMPV]-[HILM]-X(2,7)-E. This pattern does not coincide with the conventional repeat alignment. An invariant Asp is followed by a position with selection against positively charged residues, two hydrophobic residues, and any residue. Following this is a position with either Phe or Tyr and, with a spacing of 2–7 arbitrary residues, an invariant Glu (Fig. 6C). Several β-catenin complex structures and various mutagenesis studies including the present one provide a rationale for this particular binding motif (see below). Fig. 6C also shows that the other 15-mer repeats (repeats B–D) and all the actual 20-mer repeats do not contain the consensus motif, in agreement with the binding data shown in Fig. 6, A and B. However, the consensus binding motif is present in sequences immediately N-terminal to the 20-mer repeats 3 and 5, but not in the others.

As before, the sequence requirements for binding of the 15-mer APC repeat A were analyzed by a peptide array alanine scan as described above, and 8 of the 15 amino acids were found to be crucial for binding (Fig. 6D). As expected from sequence alignment (Fig. 6C) and the structural analysis, where the extended regions of the β-catenin binding partners overlapped, Asp1022 (homologous to the essential Asp in Tcf4 and E-cadherin) and Tyr1027 (homologous to Phe) were important binding determinants (Fig. 1C). These residues mark the boundaries of a region in which the conformation of all three peptides bound to β-catenin is virtually identical (24–27) (Fig. 1C). Ile1025 was equally as important as the homologous residues in Tcf4 and E-cadherin, whereas the important Asn1026 had no equally important counterpart. Residues Leu1029, Tyr1031, Asp1033, and Glu1034, which were also apparently required for tight binding, are in a region that allows more conformational freedom for the different β-catenin binding partners. In general, we found that certain residues imperfectly conserved among the 15-amino acid repeats A–D, such as Asp1025 or Ile1025, were much more crucial for the interaction than, for example, the invariant Ser1020, arguing that conservation might not reflect the importance for binding in this case.

**DISCUSSION**

Here, we used peptide arrays for the first complete alanine mutagenesis scan of β-catenin-binding domains from E-cadherin, Tcf4, and a 15-amino acid repeat of APC to determine crucial determinants (hot spots) for the interaction with β-catenin. The peptides covalently bound to a membrane support were synthesized by automated SPOT synthesis and incubated with a fluorescent β-catenin protein. Furthermore, using this method, minimal binding sequences were defined for APC and Tcf4. Quantitative peptide array data correlated well with fluorescence polarization equilibrium titrations of the respective peptides in solution and were again correlated with ITC. The results obtained showed that peptide arrays are a method well suited to identify β-catenin binding hot spots for different ligands.

The affinity we found for the wild-type Tcf peptide was in the same range but somewhat higher than the previously published values obtained by ITC (29). However, an affinity below the nanomolar level might be difficult to determine accurately by ITC because concentrations high above the K_d have to be used in order to measure reasonable enthalpy changes. An IC_50 of 15 nM has been found using an enzyme-linked immunosorbent assay (30) in which GST-Tcf was bound to a solid support, which might not be ideally suited to measure true equilibrium constants. The vital importance of Asp_1022 of Tcf proteins for interaction with β-catenin has been noted previously (28–31), and the structure has shown that this residue contacts one of the two positively “charged buttons” formed by crucial β-catenin lysines, which are the major anchor site for Tcf (23–25). For several other residues, including Phe_21, conflicting results have been presented. Pasolini et al. (29) determined a K_d of 1.9 nM for the F21A mutant by ITC, an affinity even higher than that of the wild-type peptide. In our hands, a consistent reduction of β-catenin affinity was found for this mutant using peptide arrays, fluorescence polarization, and ITC. Moreover,
FIG. 6. Peptide array analysis of the β-catenin–APC interaction. Membranes contained peptide fragments of the human APC protein and were probed with bcat-GFP as described previously. A, overlapping 50-mer peptides of the region encompassing all 15-mer (repeats A–D) and 20-mer repeats (repeats 1–7). 50-mers containing the respective repeats are labeled with a horizontal bar (see “Results” for explanation). B, overlapping 15-mer peptides of the region encompassing 15-mer repeats A–D, which are each indicated with an arrow. C, structural alignment of certified β-catenin binding motifs (+) and the binding region adjacent to repeats 3 and 5. These are contrasted with the APC repeat regions that were defined in this study to fail to bind β-catenin (–). D, alanine scan of the β-catenin–APC-A interaction. Peptides with single alanine point mutants were probed with bcat-GFP as described in the Fig. 2 legend.
for the homologous LEF1, an alanine mutation of this phenylalanine, which is invariant in the Tcf family, completely blocks β-catenin interaction in a quantitative yeast two-hybrid assay and abolishes nuclear translocation of β-catenin (28). In the case of Tcf4, a double mutant of Ile19 and Phe21 to alanine consistently reduces reporter gene transcriptional activation (25). In contrast to Poy et al. (25), we also found an important contribution by Leu18, Ile19, and Asp23, by using both peptide arrays and fluorescence polarization. Supporting our data, a lack of β-catenin nuclear translocation for the LEF1 mutant homologous to Asp23 has been detected (28).

The importance of the extended region of Tcf family members (residues 16–24 in Tcf4) for interaction with β-catenin has been well described (29–31). This is also reflected in the peptide array and fluorescence polarization measurements reported in this work. The importance of Leu41 and Leu48 in the C-terminal helix of Tcf4 has also been noticed previously (29, 31). We found two previously undetected residues, Val49 and Glu51, with significant energetic contribution to the interaction with β-catenin. The slight increase in affinity after mutation of Gly13 to alanine, which has not been described previously, might be due to a reduced conformational flexibility and thus enhanced interaction of the region N-terminal to residue 11, which is invisible in any β-catenin-Xenopus Tcf3 structure (23). Xenopus Tcf3 differs from human Tcf4 in only two positions in the region of interest but parallels the situation with the full-length protein in vivo.

Clustered mutation of several residues in an E-cadherin minimal binding region has been tested for binding. This is in agreement with the finding that deletion of the first 14 amino acids of Tcf4, but not deletion of the first 12 amino acids of Tcf4, abolishes β-catenin binding (24).

For anticancer drugs targeted against the oncogenic β-catenin-Tcf4 complex such as the one found recently by high-throughput screening (7), it is mandatory not to affect the interaction of β-catenin to both E-cadherin and APC because this could potentially lead to metastasis of tumor cells or, even worse, an inappropriate accumulation of β-catenin in healthy cells, respectively. Our comparative studies on the binding of Tcf4, E-cadherin, and APC showed that although the extended region of Tcf4 comprising the residues from Asp16 to Phe21 is a suitable mimic of phosphoserines in E-cadherin (41), had only one of which is a conservative substitution (Glu instead of Asp), the other is a Ser instead of a Gly in a stretch of multiple glycine residues.

For E-cadherin, minimal binding regions of about 25 residues have been defined (40, 41), which overlap with the important residues identified in this study. We showed that mutation of individual glutamates, which have been demonstrated to be a suitable mimic of phosphoserines in E-cadherin (41), had only a small effect on β-catenin binding. This agrees well with previous data (40, 41) that mutation of more than one serine of this cluster is necessary to abolish the interaction in vivo. Clustered mutation of several residues in an E-cadherin minimal binding region has been tested for in vivo β-catenin bind-
14. In vivo, which agrees well with the individual hot spots identified here. We identified some additional hot spots N-terminal to the extended core region; however, the most important beta-catenin binding residues are inside the core, namely, the conserved Asp764 and Phe769.

For APC, the correlation between the 15-mer repeats and phosphate binding has not been unequivocal in previous studies. A 27-amino acid fragment containing the 15-mer repeat A interacts with beta-catenin (43), and the x-ray structure of the beta-catenin/ APC-A complex has been solved (27). However, most of the previous studies addressing binding of individual repeats contained additional flanking sequences. On the other hand, Rubinfield et al. (46) found that, in the context of a longer fragment that bound to beta-catenin, mutation of several invariant serines inside repeat 2 did not inhibit the interaction, arguing against an important role of repeat 2. In the same study, it was shown that a fragment containing repeat 2 (residues 1342–1476) does not bind at all. The isolated unphosphorylated APC-3 does not bind to beta-catenin, suggesting that flanking amino acids are required (47). This is all under the assumption that any of the 15-mer repeats B–D and the 20-mer repeats 1–7 are not structured in the absence of beta-catenin, as shown for 15-mer repeat A (27) and the catenin-binding regions of E-cadherin (38) and Tcf4 (31). Our results showed, however, that this is not the case. Therefore, we cannot exclude, however, that phosphorylation would allow binding, which is modulated by phosphorylation (33, 34).

APC is phosphorylated in vivo by at least two kinases, glycogen synthase kinase-3beta and casein kinase-1e. Beta-catenin binding of an APC fragment covering repeats 2 through 7 is strongly augmented by phosphorylation (46), and mimicking of phosphoryserines by aspartates in repeat 3 increases the affinity of an 88-residue fragment (47). It has been proposed that the 20-mer region of APC can bind to two alternative sites on beta-catenin and that phosphorylation serves as a switch (27, 48). We cannot exclude, however, that phosphorylation would allow binding of the 20-mers that failed to bind in our study with unphosphorylated peptides.

While this work was in preparation, two studies on the structure and biochemistry of 20-mer repeat binding to beta-catenin appeared (33, 34). They collectively show that the previously identified 15- and 20-mer repeats may not define a consensus binding motif, that some of these do not seem to bind at all, and that motifs outside the repeats are necessary for binding, which is modulated by phosphorylation (33, 34).

REFERENCES

1. International Human Genome Sequencing Consortium. (2001) Nature 409, 860–921
2. Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., Gocayne, J. D., Amanatides, P., Ballew, R. M., Husen, D. H., Wortman, J. R., Chen, J., Zhang, Q., Kodira, C. D., Zheng, H. X., Chen, L., Skupski, M., Subramanian, G., Thomas, P. D., Zhang, J., Gabor Miklos, G. L., Nelson, C., Broder, S., Clark, A. G., Nadeau, J. M., McKusick, V. A., Zinder, N., Levine, A. J., Roberts, R. T., Simon, M., Shayman, C., Kucan, J., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., Hannenhalli, S., Kravitz, S., Levy, S., Mabardy, C., Reinert, K., Remington, K., Abu-Threedy, J., Beasley, E., Biddle, K., Bonazzi, V., Brandon, R., Cardill, M., Chandramouliswaran, I., Charlab, R., Chaturvedi, K., Deng, Z., Di Francesco, V., Dunn, P., Edieke, B., Evangelista, C., Gabrielian, A. E., Gan, W., Ge, W., Gong, P., Gu, Z., Guan, P., Heiman, T. J., Higgins, M. E., Ji, R. H., Ke, Z., Ketchum, K. A., Lai, Z., Li, Z., Li, J., Liang, Y., Lin, L., Lu, F., Mclay, G. O., McVean, G. H. M., Nielson, S., Nisnay-Dryan, V. A., Neelam, B., Nusskern, D., Rusch, D. B., Salzberg, S. A., Shue, B., Sun, J., Wang, Z., Wang, A., Wang, X., Wang, J., Wei, M., Wides, R., Xiao, C., Yan, C., Yao, A., Ye, J., Zhan, M., Zhang, W., Zhao, Q., Zheng, L., Zhang, P., Zhang, W., Zhu, S., Zhao, S., Gilbert, D., Haung, S., Spier, G., Carter, C., Cravchik, A., Woodage, T., Ali, F., An, H., Awe, A., Baldwin, D., Baden, H., Barnstard, M., Barrow, I., Beeson, K., Busam, D., Curver, A., Center, A., Cheng, M. L., Curry, L., Danaher, S., Davenport, L., Detter, J., Dietz, S., Deshpande, D., DeSantis, V., Dewees, P., Dicks, H., Gluecksmann, A., Hart, B., Haynes, J., Haynes, C., Heidran, S., Hostin, D., Houck, J., Howland, T., Ibegwam, C., Johnson, J., Kalush, F., Kline, L., Koduru, S., Love, A., May, J., Meade, H., Moiseyev, T., Ruoff, G., Spink, K. E., Fridman, S. G., and Weis, W. I. (2001) J. Biol. Chem. 276, 800–807
3. Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., Galbiati, F., Hehiseel, J. D., Jacob, J., Johnston, M., Louis, E. J., Mewes, H. W., Murakami, Y., Philippeps, P., Tettelin, H., and Oliver, S. G. (1996) Science 274, 546–563
4. von Mering, C., Krause, R., Snel, B., Oliver, S. G., and Bork, P. (2002) Nature Rev. Cell 3, 384–390
5. Tilley, J. W., Chen, L., Fry, D. C., Emerson, S. D., Powers, G. D., Biondi, D., Vondriska, T., Trillos, R., Yeh, J. R., Mennona, P., Kaplan, G., Le Mahies, R. A., Carson, M. H., Han, R. J., Liu, C. M., Palermo, R., and Ju, G. (1997) J. Am. Chem. Soc. 119, 7589–7590
6. Kim, J. S., Crooks, H., Foxworth, A., and Waldman, T. (2002) Mol. Cancer Ther. 1, 1053–1057
7. Behrens, J., Piek, J., Peterse, M., and Sengers, M. (2000) J. Mol. Biol. 303, 839–848
β-Catenin Interaction with Tcf4, E-Cadherin, and APC

306, 1179–1189
32. Roose, J., and Clevers, H. (1999) Biochim. Biophys. Acta 1424, M23–M37
33. Ha, N. C., Tonozuka, T., Stamos, J. L., Choi, H. J., and Weis, W. I. (2004) Mol. Cell 15, 511–521
34. Xing, Y., Clements, W. K., Le Trong, I., Hinds, T. R., Stenkamp, R., Kimelman, D., and Xu, W. (2004) Mol. Cell 15, 523–533
35. Frank, R., and Overwin, H. (1996) Methods Mol. Biol. 66, 149–169
36. Frank, R. (1992) Tetrahedron 48, 9217–9232
37. Teepert, F., Pires, J. R., Landgraf, C., Oschkinat, H., and Schneider-Mergener, J. (2001) Angew. Chem. Int. Ed. Engl. 40, 897–900
38. Huber, A. H., Stewart, D. B., Laurents, D. V., Nelson, W. J., and Weis, W. I. (2001) J. Biol. Chem. 276, 12301–12309
39. Zander, N. (2004) Mol. Divers. 8, 189–195
40. Stappert, J., and Kemler, R. (1994) Cell Adhes. Commun. 2, 319–327
41. Simcha, I., Kirkpatrick, C., Sadot, E., Shtutman, M., Polevoy, G., Geiger, B., Peifer, M., and Ben Ze’ev, A. (2001) Mol. Biol. Cell 12, 1177–1188
42. Lickert, H., Bauer, A., Kemler, R., and Stappert, J. (2000) J. Biol. Chem. 275, 5090–5095
43. Su, L. K., Vogelstein, B., and Kinzler, K. W. (1993) Science 262, 1734–1737
44. Rubinfield, B., Souza, B., Albert, I., Muller, O., Chamberlain, S. H., Masiarz, F. R., Munemitsu, S., and Polakis, P. (1993) Science 262, 1731–1734
45. Shuker, S. B., Hajduk, P. J., Meadows, R. P., and Peaik, S. W. (1996) Science 274, 1531–1534
46. Rubinfeld, B., Albert, I., Perfri, E., Munemitsu, S., and Polakis, P. (1997) Cancer Res. 57, 4624–4630
47. Tickenbrock, L., Kosmeier, K., Rehmann, H., Herrmann, C., and Muller, O. (2003) J. Mol. Biol. 327, 359–367
48. Xing, Y., Clements, W. K., Kimelman, D., and Xu, W. (2003) Genes Dev. 17, 2753–2764