BMC Genomics

CrossMark

Self-regulation of functional pathways by motifs inside the disordered tails of beta-catenin

Bi Zhao and Bin Xue[*]

Abstract

Background: Beta-catenin has two major functions: coordinating cell-cell adhesion by interacting with cadherin in cadherin junction formation pathway; and regulating gene expression through Wnt signaling pathway. Accomplishing these two functions requires synergistic action of various sequential regions of the same beta-Catenin molecule, including the N-terminal tail, the middle armadillo domain, and the C-terminal tail. Although the middle armadillo domain is the major functional unit of beta-Catenin, the involvement of tails in the regulation of interaction between beta-Catenin and its partners has been well observed. Nonetheless, the regulatory processes of both tails are still elusive. In addition, it is interesting to note that the three sequential regions have different structural features: The middle armadillo domain is structured, but both N- and C-terminal tails are disordered. This observation leads to another important question on the functions and mechanisms of disordered tails, which is also largely unknown.

Results: In this study, we focused on the characterization of sequential, structural, and functional features of the disordered tails of beta-Catenin. We identified multiple functional motifs and conserved sequence motifs in the disordered tails, discovered the correlation between cancer-associated mutations and functional motifs, explored the abundance of protein intrinsic disorder in the interactomes of beta-Catenin, and elaborated a working model on the regulatory roles of disordered tails in the functional pathways of beta-Catenin.

Conclusion: Disordered tails of beta-Catenin contain multiple functional motifs. These motifs interact with each other and the armadillo domain of beta-catenin to regulate the function of beta-Catenin in both cadherin junction formation pathway and Wnt signaling pathway.

Keywords: Beta-Catenin, Wnt signaling pathway, Cadherin junction formation, Intrinsic disorder, Auto-regulation

* Correspondence: binxue@usf.edu
Department of Cell Biology, Microbiology and Molecular Biology, School of
Natural Sciences and Mathematics, College of Arts and Sciences, University
of South Florida, 4202 E. Fowler Ave, ISA 2015, Tampa 33620, FL, USA

Background

Beta-Catenin is a 92-kDa protein that is composed of two flexible tails at each of the N- and C-termini, and an intermediate structured armadillo domain (ARM) containing 12 repeats of helical segments [1]. The function of beta-Catenin is to regulate cadherin junction formation [2, 3] and to regulate Wnt signaling [4, 5]. Therefore, beta-Catenin plays critical roles in many biological processes, such as embryonic development [6, 7], cell division [8], and maintenance of pluripotency [9]. Disorganized expression of beta-Catenin is associated with many diseases, including cancer [10–12] and cardiovascular diseases [13, 14]. The ARM domain is the major determinant for the function of beta-Catenin. It binds to the cytoplasmatic tail of cadherin at the stage of junction formation. ARM domain also interacts with short motifs on both Axis Inhibition Protein (Axin) and Adenomatous Polyposis Coli tumor repressor protein (APC) in Wnt signaling pathway [15]. The interaction raises the local concentration of beta-Catenin and leads to the phosphorylation, ubiquitination, and degradation of beta-Catenin [16]. Upon the activation of Wnt signaling molecules, beta-Catenin molecules escape from the degradation pathway, accumulate in cytoplasma, and translocate into nucleus. One possible translocation mechanism involves the interaction between the ARM domain and the phenylalanine-glycine(FG)-repeat containing proteins in the Nuclear Pore Complex (NPC) [17, 18]. Once accessing the nucleus, beta-Catenin uses its ARM domain to interact with T Cell Factor/Lymphoid Enhancer Factor (TCF/LEF) family transcription factors to activate downstream gene expression [19–21].

The disordered tails of beta-Catenin were also found to synergistically regulate the function of beta-Catenin in recent studies. The tails regulate the binding between beta-Catenin and cadherin [22]. The N-terminal tail of beta-Catenin is also phosphorylated by the destruction complex, which is formed by Axin, APC, Casein Kinase I alpha (CK1-alpha), Glycogen Synthase Kinase 3β (GSK-3beta), and Protein Phosphatase 2A (PP2A) [23–25]. Phosphorylated beta-Catenin is ubiquitinated and then degraded [26–28]. Wnt signaling molecules outside of the cell are able to prevent the formation of destruction complex and therefore increase the cytoplasmic level of beta-Catenin. In this way, the cytoplasmic level of beta-Catenin is tightly regulated. The flexible tail of beta-Catenin also facilitates its nuclear translocation [29, 30], and recruits factors involved in transcription activation [31].

These discoveries have extended our comprehension on the interaction patterns between beta-Catenin and its partners, and opened a new field on the functional roles of flexible tails. However, the detailed mechanisms associated with these new discoveries are still largely unknown.

In this study, we applied systematic bioinformatics analyses on both N- and C-terminal tails of beta-Catenin, identified multiple functional motifs, characterized the conserved sequential segments, discovered the pattern of interactions between beta-Catenin and its partners, revealed the correlation between cancer-associated mutations and functional motifs, and proposed possible regulatory mechanisms of flexible tails on the function of beta-Catenin. This study increases our knowledgebase on the sequential, structural, and functional features of beta-Catenin tails, facilitates our understanding on the mechanisms and regulatory roles of the terminal tails on the functions of beta-Catenin.

Method

Sequence analysis

(1) Sequence conservation: The FASTA sequences of beta-Catenin from different species were extracted from UniprotKB [32]. The sequences were used as input of different predictors. The sequences were also aligned and uploaded to WebLoGo [33] to generate sequence logo plots to demonstrate the conserved patterns of amino acids. **(2) Mutation analysis:** Amino acid substitutions in human beta-Catenin were retrieved from UniprotKB, COSMIC (Catalogue of Somatic Mutation in Cancer) [34], DMDM (Domain Mapping of Disease Mutations) [35], and BioMuta [36] databases. These four databases have 20, 278, 20, and 138 mutations, respectively. Both of the UniProtKB and DMDM databases contain one unique mutation that is not found in other databases. COSMIC and BioMuta have 128 overlapped mutations. The total number of unique mutations is 281. These mutations were then analyzed in parallel with functional sites/motifs obtained from motif analysis.

Structural analysis

(1) Secondary structure prediction: NetSurfP [37] was used to predict the secondary structure of beta-Catenin. The NetSurfP predictor applied a two-step strategy, of which the first step was used to predict the secondary structure of amino acids in three states (helix, strand, and coil) by using protein sequence as input, and the second step was a filter by taking the predicted secondary structures of the first step over a sliding window as input. In both steps, artificial neural networks were applied to make predictive output based on the input data. **(2) Disorder prediction:** Both PONDR-FIT [38] and PONDR©VLXT [39] were applied to predict per-residue disorder score of beta-catenin. PONDR-FIT is an artificial neural network based meta-predictor and is composed of six individual predictors: PONDR©VLXT, PONDR-VSL2 [40], PONDR-VL3 [41], FoldIndex [42], IUPred [43], and TopIDP [44]. Since the integration of individual predictors in the meta-strategy combines all

the possible correct predictions, the accuracy of meta-predictor is normally improved significantly. In fact, PONDR-FIT is one of the most accurate predictors for protein intrinsic disorder [45]. PONDR©VLXT is highly efficient in identifying hydrophobic motifs within long disordered regions. The input features of PONDR©VLXT include various combinations of amino acids, mainly hydrophobic residues. For this reason, PONDR©VLXT is very sensitive to the local changes of amino acids and can be used to easily detect local hydrophobic clusters. In the outputs of both predictors, residues with score higher than 0.5 were considered to be disordered, residues with score lower than 0.5 were treated as structured. PONDR-FIT was also used to analyze proteins in the primary and the secondary interactomes of beta-Catenin, as well as proteins in the related signaling pathways.

Functional analysis

(1) Binding Motifs: ANCHOR [46], MoRF [47], and MoRFpred [48] were used to predict potential binding motifs inside disordered regions. ANCHOR identifies binding motifs that are dependent on hydrophobic clusters, MoRF is able to detect binding motifs that undergo structural transition from coil to helix, and MoRFpred finds binding motifs that change conformations from coil to helix, strand, and other types of coils. **(2) Linear Motifs:** ELM [49] online server was applied to predict possible functional linear motifs. ELM contains short linear motifs that normally stretch over three to ten residues. The ELM motifs have broad functions, including cleavage, binding, modification, etc. **(3) Protein Interaction Networks and Pathways:** The protein-protein interaction networks were collected from the STRING database [50] by using confidence score 0.85. STRING is one of the most comprehensive databases for protein-protein interactions. Proteins signaling pathways were extracted from KEGG [51], which is a prevailing tool for the scientific community. The gene names/protein IDs from STIRNG and KEGG were also used to extract corresponding protein sequences from UniProtKB. **(4) FATHMM** (Functional Analysis Through Hidden Markov Models) [52]: FATHMM is an advanced high-throughput predictor trained by using hidden Markov model on multiple mutation databases for analyzing the functional impacts of mutations. FATHMM outperformed many other predictors on multiple validation datasets. All the mutations were then classified by FATHMM using two independent variables (criteria): Damaging (or Tolerant) and Cancer-promoting (or non-Cancer-promoting), into four groups: Damaging and Cancer-promoting, Damaging and non-Cancer-promoting, Tolerant and Cancer-promoting, and Tolerant and non-Cancer-promoting.

Results

Disordered tails at both N- and C-termini accommodate predicted binding motifs

PONDR-FIT [38], which is one of the most accurate predictors of protein intrinsic disorder [45], was applied to identify disordered regions of beta-Catenin. The prediction of PONDR-FIT as shown in the upper panel of Fig. 1 demonstrates that: (1) the middle region of human beta-Catenin from ~ AA130 to ~ AA710 is structured; and (2) the N- and C-terminal tails flanking the middle region are disordered. These conclusions are supported by the prediction of PONDR©VLXT [39] predictor as shown in the lower panel of the same figure. The disorder curve of PONDR©VLXT prediction is more fluctuating due to the fact that this predictor is very sensitive to local composition of amino acids. Even though, the disorder scores of the middle region from ~ AA220 to ~ AA600 are much lower than that of the flanking tails, indicating the conformation of the middle region is much more rigid than the flanking tails. In addition to the consistency between different computational predictors, the conclusions on structural flexibility of beta-Catenin are also supported by experimental results. In a recent x-ray crystallography study, the structure of beta-Catenin was resolved for a region spanning from AA138 to AA691 [1], which is similar to the above-mentioned structured region identified by both PONDR-FIT and PONDR©VLXT. This region contains an armadillo domain of which the structure is determined to be twelve repeats of helical segments as shown on the top of Fig. 1. Regions other than this armadillo domain failed to be crystallized, indicating the presence of considerable structural flexibility in these regions. These results come to a consistent conclusion that the armadillo domain has inflexible 3D structure, while both tails are disordered. For the simplicity of discussion in this study, N-terminal and C-terminal tails of human beta-Catenin were defined as AA1-AA150 and AA667-AA781, respectively.

The lower panel of Fig. 1 shows: (1) the disorder prediction of PONDR©VLXT predictor, (2) the secondary structure prediction from NetSurfP [53], which is one of the most accurate and popular predictors of protein secondary structure, and (3) binding motifs identified by Molecular Recognition Feature (MoRF) [47], MoRFpred [48], and ANCHOR [46], respectively. MoRF predicts binding motifs that change their conformations from coils to helices upon binding, MoRFpred identifies binding motifs that transform from coil into one of the three basic secondary structures (helix, strand, and coil), and ANCHOR recognizes interaction motifs dominated by hydrophobic residues. The disorder curves of PONDR©VLXT prediction at both tails display multiple "dip" regions, which indicate hydrophobic clusters of

amino acids inside disordered regions. The segments corresponding to these dips often function as binding motifs. Secondary structure prediction demonstrates that the central ARM domain is composed of many helical segments, the N-terminal tail has multiple scattered helical segments, but the C-terminal tail has only a couple of helical segments in the first ~40 residues and is depleted of helices in the rest ~80 residues. The predicted binding motifs were shown as short bars in the lower panel of Fig. 1. At the N-terminal end (AA1 ~ AA150) of the disorder curve, all three dips (at ~ AA25, ~AA80, and ~ AA130, respectively) are overlapped with binding motifs predicted by at least two predictors. At the C-terminal end (over ~ AA660), the segment corresponding to a composite dip at ~ AA760 was predicted to be a binding motif by three predictors. Another two segments next to the dip at ~ AA720 were predicted to be binding motifs by ANCHOR. The forth segment overlapped with the dip at ~ AA660 was predicted as binding motif by MoRFpred. All the predicted binding motifs, except the last two (~AA720 and ~ AA760) that locate at the very end of C-terminal tail, are also predicted to be helices.

Disordered tails are enriched of linear motifs

Eukaryotic Linear Motifs (ELMs) are short sequence motifs of normally several amino acids, but perform multiple types of functions [49]. By searching ELM database [49], 130 ELMs were identified in beta-Catenin. In which, 73 ELMs locate in the tail regions (Additional file 1: Table S1). Figure 2a shows the density of each type of ELMs in the middle domain and the tail regions of beta-Catenin. It is clear that the densities of different types of functional motifs in both tails are much higher than that in the ARM domain. Both tails are extremely enriched of MOD and LIG motifs. In addition, while only the N-terminal tail has DEG motifs, the C-terminal tail has many more MOD and DOC motifs. To demonstrate the sequential distribution of motifs, the locations of a non-redundant set of ELM motifs in the tail regions were labeled for both N-terminal and C-terminal tails in Fig. 2b and c. At the N-terminal tail, ELM motifs cluster at three locations: ~AA20, ~AA70, and ~ AA120. In the C-terminal tail, the motifs congregate at ~ AA675, ~AA720, and ~ AA760. Clearly, these locations are essentially the same as the dips and binding motifs predicted in Fig. 1.
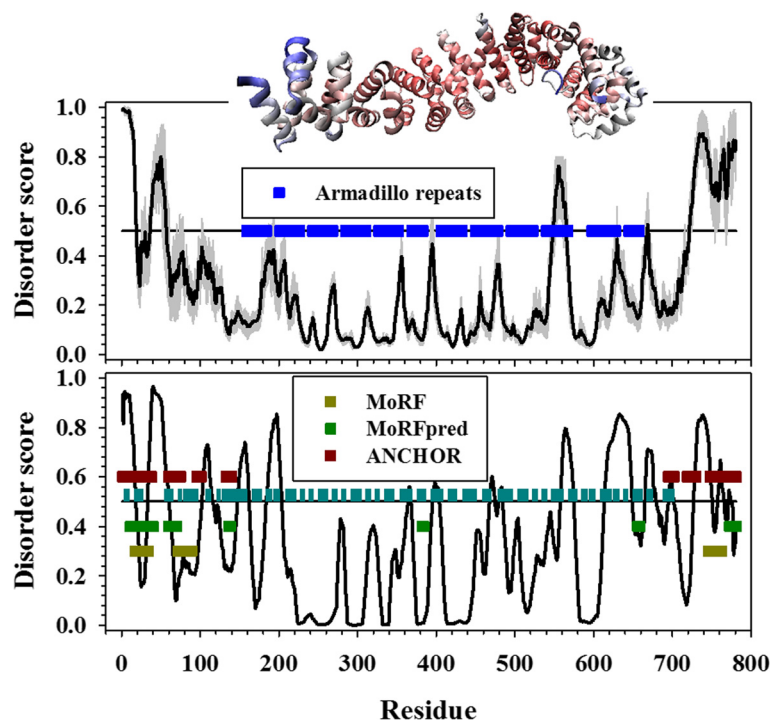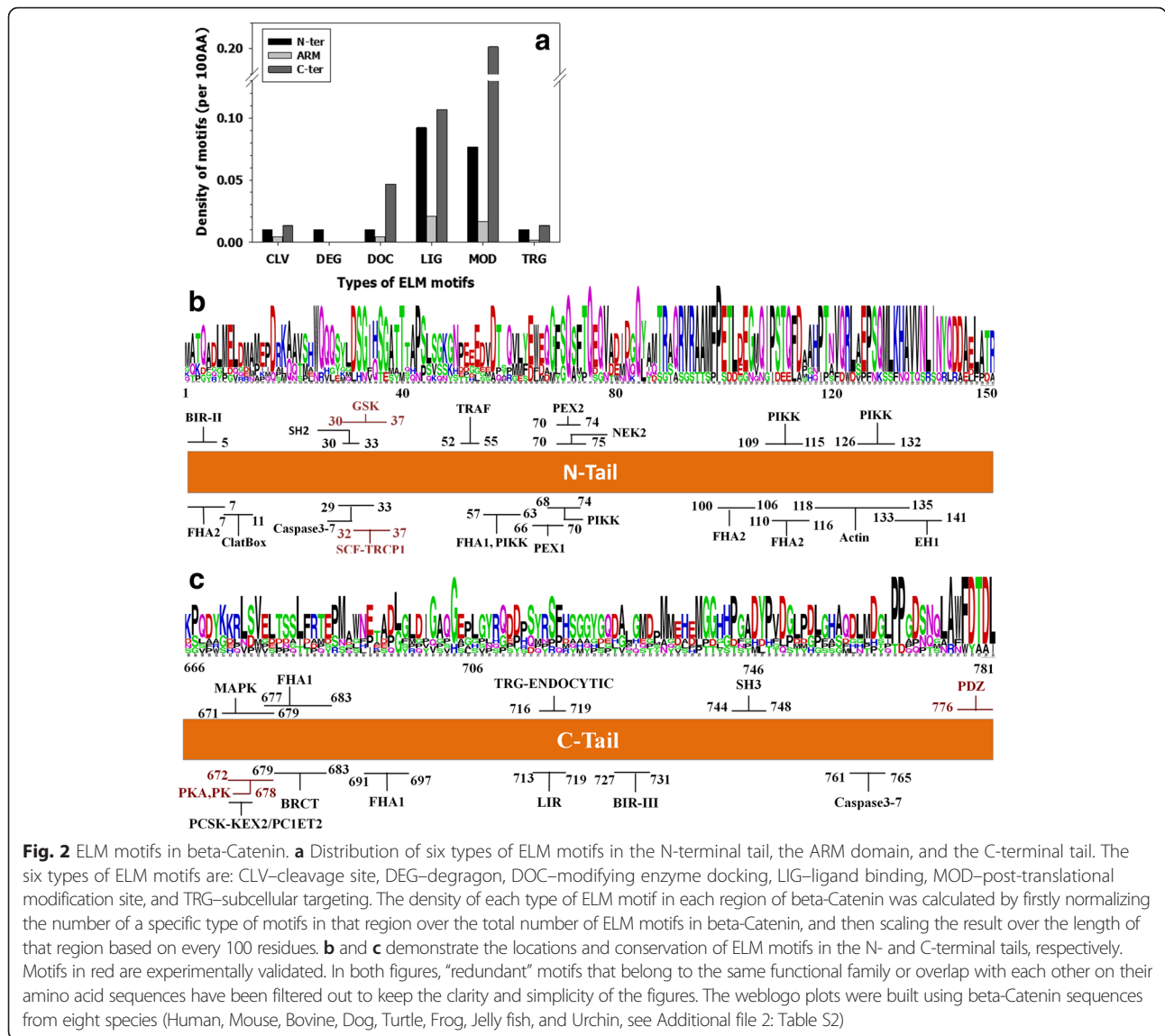

Fig. 1 Predicted intrinsic disorder of and predicted binding motifs in human beta-Catenin (UniProtKB Entry: P35222). X-axis shows sequential indexes of amino acids. Y-axis displays disorder scores predicted by PONDR-FIT (upper panel) and PONDR©VLXT (lower panel). Residues with score higher than 0.5 are disordered, otherwise structured. In addition, residues with score between ~0.3 and 0.5 are assumed to be flexible, and residues with score close to zero are extremely rigid. The gray shadow behind PONDR-FIT prediction is the estimated prediction error from PONDR-FIT predictor. Horizontal blue bars in the upper panel represent the domain structure of twelve armadillo repeats. Dark cyan bars in the middle of lower panel are helices predicted by NetSurP. Binding motifs were predicted by MoRF (dark yellow), ANCHOR (brown), and MoRFpred (dark green) predictors, respectively. The 3D structure for the armadillo domain of human beta-Catenin (PDB id: 1JDH) was placed on the top of the figure

Fig. 2 ELM motifs in beta-Catenin. **a** Distribution of six types of ELM motifs in the N-terminal tail, the ARM domain, and the C-terminal tail. The six types of ELM motifs are: CLV–cleavage site, DEG–degragon, DOC–modifying enzyme docking, LIG–ligand binding, MOD–post-translational modification site, and TRG–subcellular targeting. The density of each type of ELM motif in each region of beta-Catenin was calculated by firstly normalizing the number of a specific type of motifs in that region over the total number of ELM motifs in beta-Catenin, and then scaling the result over the length of that region based on every 100 residues. **b** and **c** demonstrate the locations and conservation of ELM motifs in the N- and C-terminal tails, respectively. Motifs in red are experimentally validated. In both figures, "redundant" motifs that belong to the same functional family or overlap with each other on their amino acid sequences have been filtered out to keep the clarity and simplicity of the figures. The weblogo plots were built using beta-Catenin sequences from eight species (Human, Mouse, Bovine, Dog, Turtle, Frog, Jelly fish, and Urchin, see Additional file 2: Table S2)

Among these predicted motifs, six have been experimentally validated: The MOD_GSK3_1 motif (AA30-37) is phosphorylated by GSK3-beta [54–56]; The ubiquitin ligase binding motif DEG_SCF_TRCP1_1 (AA32-37) recognizes F box in SCF (beta-TrCP1) for subsequent ubiquitination [57]; DOC_WW_Pin1_4 (AA243-248) on beta-Catenin interacts with prolyl somerase Pin1 [53]; MOD_PKA_1 motifs (AA549-555 and AA672-678) are phosphorylated by PKA kinase [58]; and the PDZ-binding motif LIG_PDZ_Class_1 (AA776-781) forms complex with TIP-1 [59].

Weblogo analyses in Fig. 2b and c present multiple sequential patterns and sequence motifs that are conserved cross eight species (Additional file 2: Table S2). Most notably, AA90 ~ AA130 at N-terminal tail and AA774 ~ 781 at the very end of C-terminal tail accommodate highly conserved motifs. Other regions may not contain highly conserved consecutive amino acids sequences, but do have highly conserved patterns of amino acids. To name a few, both AA670 ~ AA680 and AA730 ~ AA740 have conserved hydrophobic patterns. Although being conserved, many of these segments are still lack of functional annotations, such as AA90 ~ AA100, AA140 ~ AA150, AA730 ~ AA740, and AA750 ~ AA760.

N-terminal tail is associated with the majority of sequential variances

Figure 3a shows the sequential distribution of mutations of beta-Catenin collected from four databases, including UniProtKB [32], COSMIC [34], DMDM [35], and BioMuta [36]. Clearly, almost all the deletions and insertions are only observed at the first half of the N-terminal tail (AA1 ~ AA75). The majority of synonymous and
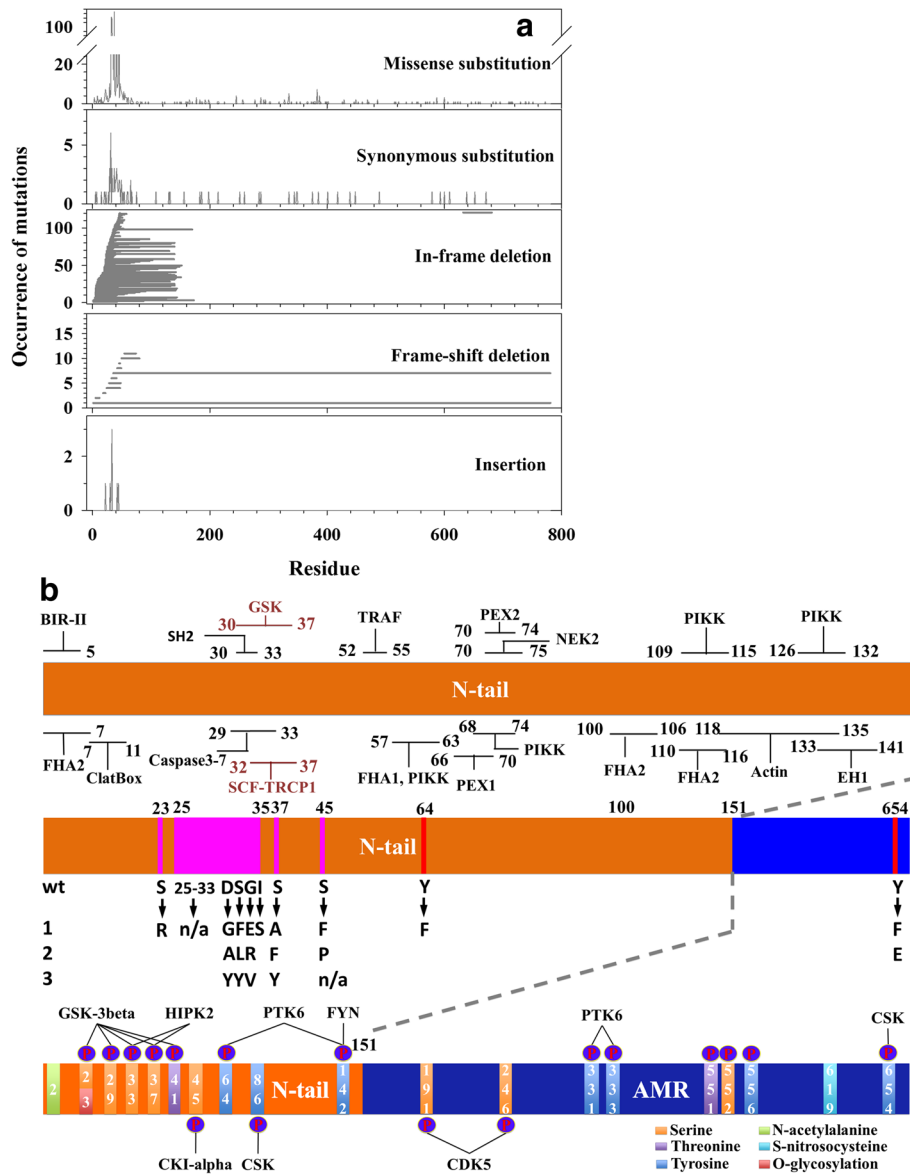
Fig. 3 Sequence variances in disordered tails. Information on the mutation and modification of beta-Catenin was extracted from UniProtKB, COSMIC, BioMute, and DMBM. **a** Occurrences of five different types of mutations. The types of mutations from top to bottom are: (1) missense mutation, (2) synonymous mutation, (3) in-frame deletion, (4) frameshift deletion, and (5) insertion. X-axis shows the amino acid sequence. Y-axis is the occurrences of mutations. The plots in (1), (2), and (5) show the number of point mutations at each site. The lines in (3) and (4) represent deleted segments. **b** Mutations and post-translational modifications in the N-terminal tail of beta-Catenin. The upper inset shows a non-redundant set of functional motifs at the N-terminal tail of beta-Catenin. The middle inset demonstrates mutations from UniProtKB for the N-terminal tail and Tyr-654 near the C-terminal tail. Red bars stand for natural variants and pink bars for mutagenesis sites. The character label "wt" and the numbers at the left-hand side indicate wild type (wt) and mutations (1, 2, and 3). Lower inset shows observed post-translational modifications in both N- and C-terminal tails

missense mutations are also in the first half of the N-terminal tail, although a small portion do spread along the entire sequence. More specifically, over 82 % of the mutations occur in the region from AA1 to AA75, and over 71 % of the mutations present in a smaller segment from AA25 to AA75.

Figure 3b presents the mutations and modifications on both N- and C-terminal tails of beta-Catenin extracted
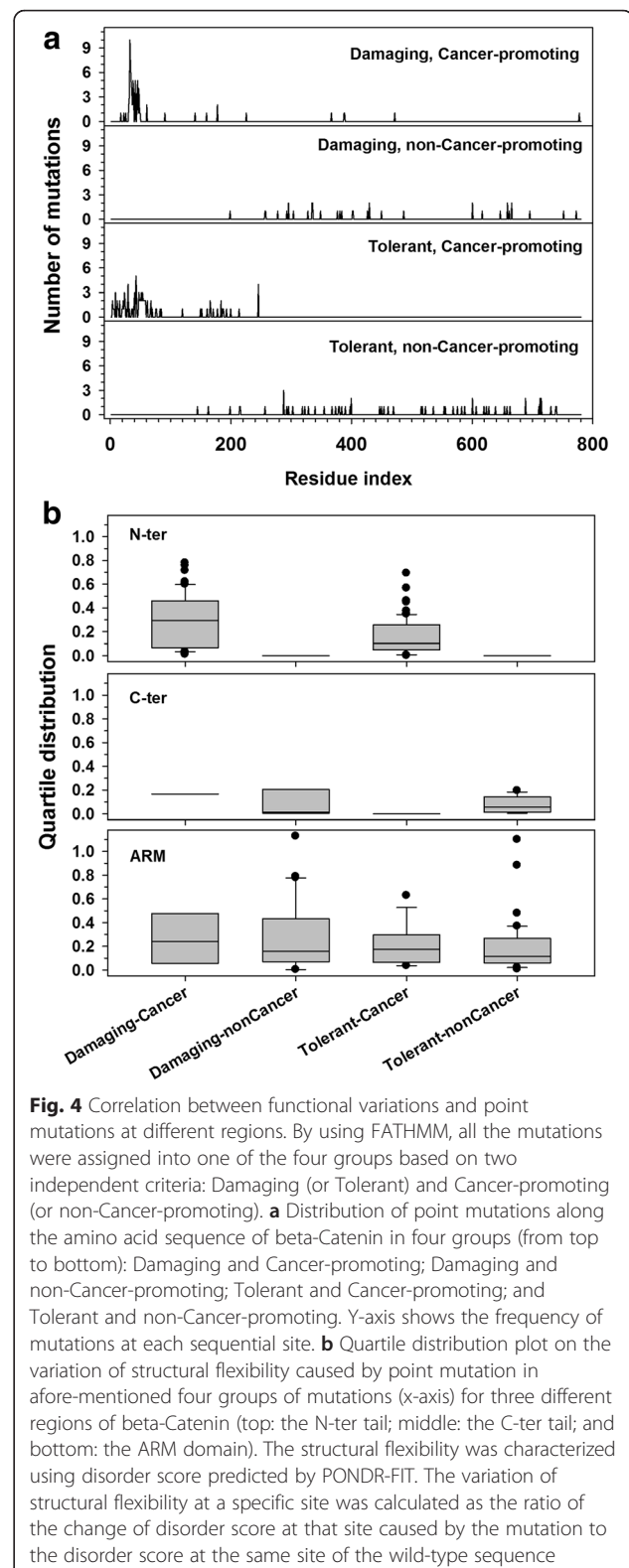
from UniProtKB, together with the functional motifs predicted from ELM. The UniProtKB information on mutation and modification is of high confidence due to the fact that the information is either experimentally validated or observed on their homologous sequences. By comparing the locations of these mutations and modifications, it is clear that most of them are in the region from AA25 ~ AA50. The involved amino acid

residues include Ser-23, 33, 37, 45, Trp-25, Asp-32, Gly-34, Ile-35, and deletion of N'-WQQQSYLDS-C' (AA25-33). These mutations overlap with one CLV, two DEG, one DOC, eight LIG, 12 MOD, and two TRG ELMs (Additional file 1: Table S1). In terms of functional influences, Ser-45, Ser-37, and Ser-33 at the N-terminal tail are three critical residues of which the phosphorylation and subsequent ubiquitination lead to the degradation of beta-Catenin [23, 26–28]. Near the C-terminal tail, several mutations are observed on Tyr-654, which is also a critical residue that uses phosphorylation to regulate the interaction between beta-Catenin and cadherin [22, 60].

FATHMM was used to evaluate the functional impact of point mutations as shown in Fig. 4. Figure 4a shows the sequential distribution of all the point mutations in four groups. Clearly, almost all the cancer-promoting mutations are located in the N-terminal tail of beta-Catenin, and non-cancer-promoting mutations are accommodated in either the ARM domain or the C-terminal tail. Although seeming to be similar, the distributions of mutations in the first group (damaging and cancer-promoting) and the third group (tolerant and cancer-promoting), as well as the distribution of mutations in the second group (damaging and non-cancer-promoting) and the fourth group (tolerant and non-cancer-promoting) are significantly different (P-values of F-test are 2.4e-11 and 1.2e-8, respectively). In addition, the frequencies of mutations at various sites in the N-terminal tail in the first group (damaging and cancer-promoting) are several times higher than the frequencies in other groups and regions. Figure 4b presents the change of structural flexibility caused by point mutation at three different regions in four different groups. Again, from the top and middle panels, it can be seen that the mutations at the N-terminal tail are cancer-promoting and are associated with large variations of structural flexibility. Meanwhile, the mutations at the C-terminal tail are non-cancer-promoting and are demonstrated by very limited variations of structural flexibility. In the bottom panel, it is clear that the damaging associated mutations may have large variations of structural flexibility, but the tolerant-associated mutations have relatively small changes of structural flexibility.

Structural flexibility in the interaction partners of beta-catenin

Beta-Catenin interacts with its partners using its structured ARM domain and/or its flexible tails. Out of a total of 24 PDB structures for complexes formed between beta-Catenin and its partners, five are formed between the tails of beta-Catenin and the structured domains of other proteins, 18 complexes are formed between the structured ARM domain of beta-Catenin

Fig. 4 Correlation between functional variations and point mutations at different regions. By using FATHMM, all the mutations were assigned into one of the four groups based on two independent criteria: Damaging (or Tolerant) and Cancer-promoting (or non-Cancer-promoting). **a** Distribution of point mutations along the amino acid sequence of beta-Catenin in four groups (from top to bottom): Damaging and Cancer-promoting; Damaging and non-Cancer-promoting; Tolerant and Cancer-promoting; and Tolerant and non-Cancer-promoting. Y-axis shows the frequency of mutations at each sequential site. **b** Quartile distribution plot on the variation of structural flexibility caused by point mutation in afore-mentioned four groups of mutations (x-axis) for three different regions of beta-Catenin (top: the N-ter tail; middle: the C-ter tail; and bottom: the ARM domain). The structural flexibility was characterized using disorder score predicted by PONDR-FIT. The variation of structural flexibility at a specific site was calculated as the ratio of the change of disorder score at that site caused by the mutation to the disorder score at the same site of the wild-type sequence
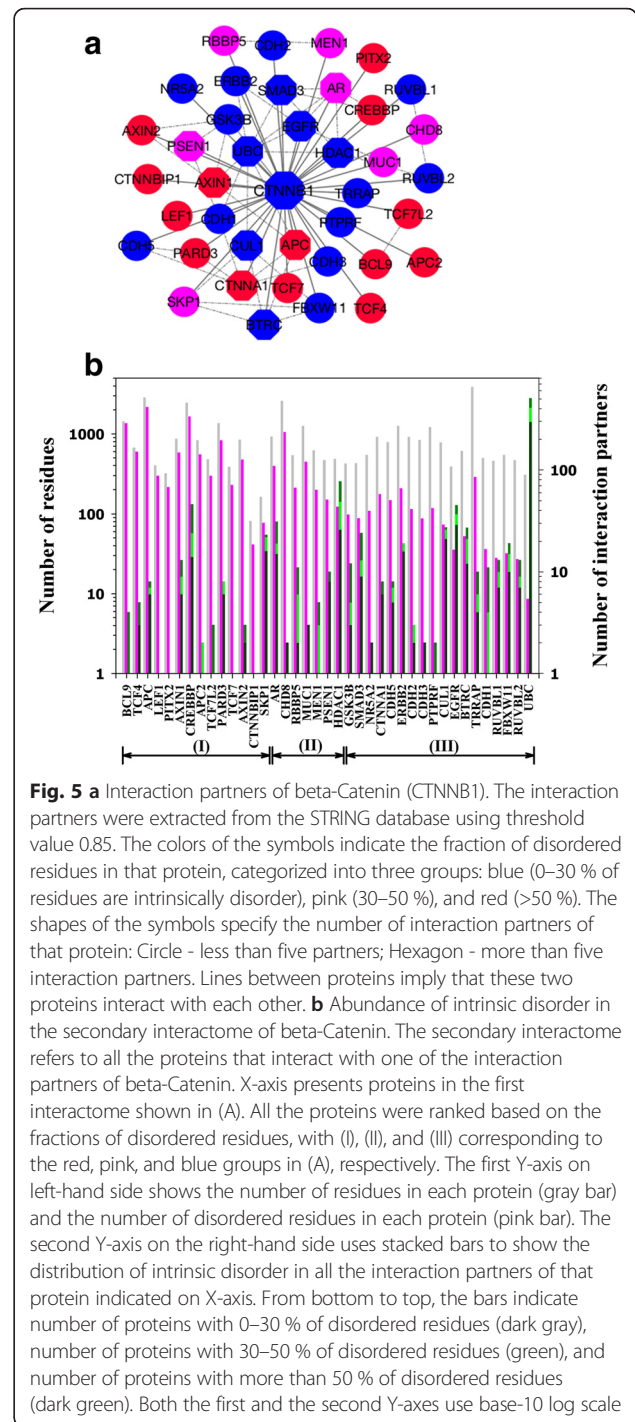
and the short segments of other proteins, the last one complex is formed between the ARM domain of beta-Catenin and a structured domain of another protein

LRH-1 (Additional file 3: Table S3). In the first five complexes, disordered tails of beta-Catenin are involved. In the next 18 complexes, the short segments of the partners are either disordered or inside disordered regions. Therefore, 23 out of 24 interactions between beta-Catenin and its partners involve intrinsically disordered regions.

To determine the abundance of protein intrinsic disorder in the protein interaction networks of beta-Catenin, we retrieved interaction partners of beta-Catenin from the STRING database [50] (Additional file 4: Table S4). STRING is a comprehensive database of protein-protein interactions with multiple choices of functional annotations and source of interactions. Figure 5a illustrates the distribution of protein intrinsic disorder in the interaction network of beta-Catenin. Out of 39 interaction partners of beta-Catenin, 14 proteins were predicted to be disordered, seven were predicted to carry large portion of disordered regions, and 18 were predicted to have small portion of disordered residues and therefore structured. Out of these 18 structured proteins, four cadherin-family members (CDH1, 2, 3, and 5) and a transcription factor LEF-1 use their flexible segments to bind the ARM domain of beta-Catenin (Additional file 3: Table S3). UBC, CUL1, FBXW11, and BTRC are proteins associated with ubiquitin ligase complex, which interacts with the N-terminal tail of beta-Catenin [28]. The other seven proteins are kinases, phosphatases, and deacetylase. These enzymes frequently target amino acids on the tails of beta-Catenin. In fact, as shown in Fig. 3, two groups of amino acids on or near the flexible tails of beta-Catenin undergo critical modifications: one group includes Ser-45, Thr-41, Ser-37, and Ser-33, which can be phosphorylated by CKI-alpha and GSK-3beta; the other group contains Tyr-654 that can be phosphorylated by CSK. The modifications on these two groups of amino acids regulate the function of beta-Catenin in the Wnt signaling pathway and the cadherin junction formation pathway, respectively [22, 23, 26–28, 60]. This observation agrees to the conclusions of our previous studies in which structural flexibility is found to be one of the determinants of amino acid modifications [61, 62].

Figure 5b shows the abundance of protein intrinsic disorder in the secondary interactome of beta-Catenin, which includes all the proteins that interact with the protein(s) in the first interactome of beta-Catenin. All the proteins in the secondary interactome were grouped by their interacting partners in the first interactome. In these groups, 23 have 10 or less partners in the secondary interactome, 15 have more than 10 but less than 100 partners, and only one protein (UBC) has about 500 interaction partners. Altogether, the total number of proteins in the secondary interactome is 1010. To show

Fig. 5 a Interaction partners of beta-Catenin (CTNNB1). The interaction partners were extracted from the STRING database using threshold value 0.85. The colors of the symbols indicate the fraction of disordered residues in that protein, categorized into three groups: blue (0–30 % of residues are intrinsically disorder), pink (30–50 %), and red (>50 %). The shapes of the symbols specify the number of interaction partners of that protein: Circle - less than five partners; Hexagon - more than five interaction partners. Lines between proteins imply that these two proteins interact with each other. **b** Abundance of intrinsic disorder in the secondary interactome of beta-Catenin. The secondary interactome refers to all the proteins that interact with one of the interaction partners of beta-Catenin. X-axis presents proteins in the first interactome shown in (A). All the proteins were ranked based on the fractions of disordered residues, with (I), (II), and (III) corresponding to the red, pink, and blue groups in (A), respectively. The first Y-axis on left-hand side shows the number of residues in each protein (gray bar) and the number of disordered residues in each protein (pink bar). The second Y-axis on the right-hand side uses stacked bars to show the distribution of intrinsic disorder in all the interaction partners of that protein indicated on X-axis. From bottom to top, the bars indicate number of proteins with 0–30 % of disordered residues (dark gray), number of proteins with 30–50 % of disordered residues (green), and number of proteins with more than 50 % of disordered residues (dark green). Both the first and the second Y-axes use base-10 log scale

the abundance of intrinsic disorder, we used both the fraction of disordered residues and the total number of disordered residues because the interaction partners may have large variance in size. In fact, in the first interactome of beta-Catenin, two are short proteins (CTNNBIP1 and SKP1) of about 100 residues, four (APC, CREBBP, CDH8, and TRRAP) are large proteins with more than 2000 residues, and the rest 33 have from hundreds to a little
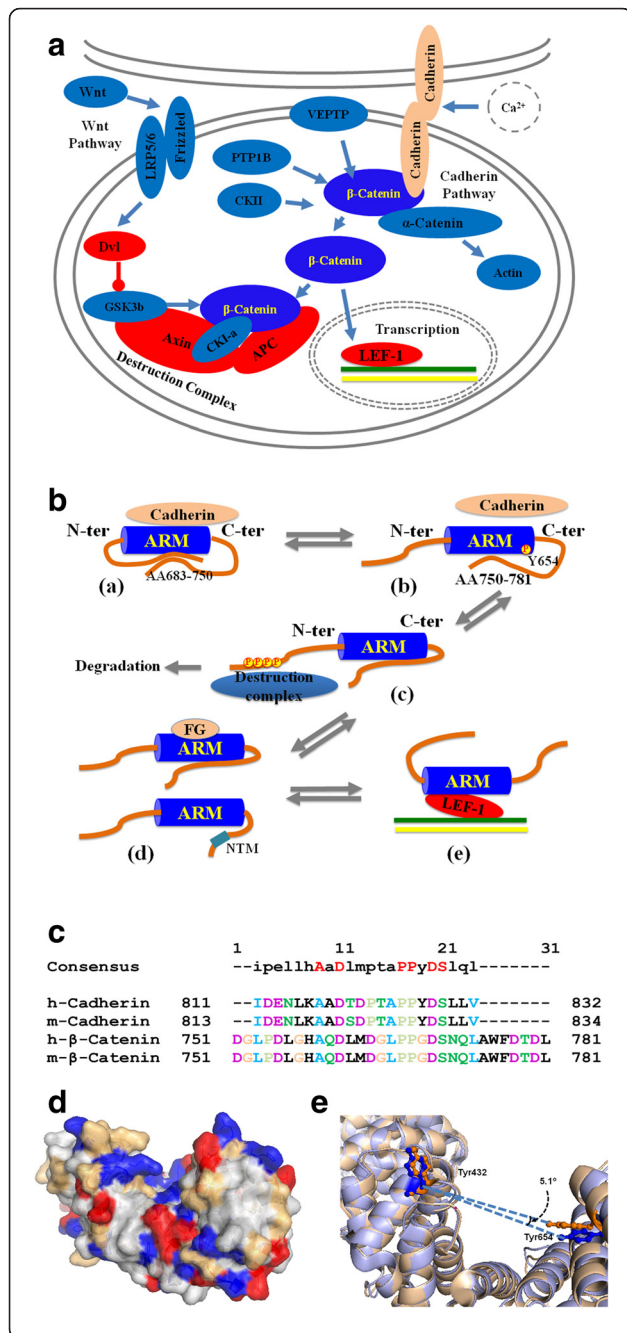
Fig. 6 Disordered tails regulate the function of beta-Catenin in two pathways. (A) Key players in the cadherin junction formation pathway and the Wnt signaling pathway. Proteins in red color are intrinsically disordered, while proteins in blue color are basically structured. Beta-Catenin was shown in dark blue to be differentiated from other proteins. Cadherin was shown in apricot due to its large amount of intrinsically disordered residues. (B) Regulatory roles of disordered tails on the functions of beta-Catenin in both cadherin junction formation and Wnt signaling pathways. ARM is the armadillo domain of beta-Catenin. N-ter and C-ter stand for the N-terminal and C-terminal tails of beta-Catenin. (a) – (e) represent five different functional stages of beta-Catenin. "P" indicates phosphorylation. "FG" stands for phenylalanine-glycine(FG)-repeat containing proteins in the nuclear pore complex. "NTM" refers to nuclear translocation motif. The destruction complex is composed of CKI-alpha, GSK-3beta, Axin, and APC. (C) Conserved sequential patterns among mouse cadherin (UniPortKB Entry: P09803, AA813-834), human cadherin (UniPortKB Entry: P12830, AA811-832), mouse beta-Catenin (UniProtKB Entry: Q02248, AA751-781), and human beta-Catenin (UniPortKB Entry: P35222, AA751-781). (D) Pocket formed by AA400-663 (PDB id: 1JDN). Blue: positive charged residues; Red: negative charge residues; Gray: hydrophobic residues; Light yellow: other residues. (E) Change of the direction of the line formed between the oxygen atom on the side chain of Tyr654 and the Cα atom of Tyr432 in two PDB structures (Blue: PDB 2Z6H; Apricot: PDB 1QZ7). The angle associated with the change of the direction implies the structural variation of the binding pocket. The RMSD calculated from these two structures for the segment from AA576 to AA685 is 0.75 angstroms

more than 1000 residues. After using both measures, it can be seen that most of the interaction partners in the secondary interactome have either higher ratio of disordered residues or higher absolute number of disordered residues. Only in the last five groups on the x-axis (CDH1, RUVBL1, FBWX11, RUVBL2, and UBC), the proteins have smaller portions of interaction partners that have both high fraction and high absolute number of disordered residues. In conclusion, the majority of the proteins in the secondary interactome have significant amount of disordered residues.

Disordered tails regulate the function of beta-Catenin

Beta-catenin is predominantly involved in two pathways: the cadherin junction formation pathway and the Wnt signaling pathway. The major component molecules in these two pathways at different subcellular locations were presented in Fig. 6a. Many proteins associated with the Wnt signaling pathway have large fraction of disordered residues and therefore are labeled in red. Figure 6b describes the functional roles of disordered tails at five representative stages in both of the afore-mentioned pathways. In this figure, stages (a) and (b) correspond to the cadherin junction formation pathway. Stages (c), (d), and (e) are related to the Wnt signaling pathway.

In the cadherin junction formation pathway, cadherin interacts with ARM domain on a region spanning from the 12th to the 7th armadillo repeats [2]. This interaction is enhanced by the association between the N-

terminal tail and the ARM domain of beta-Catenin [22]. A specific C-terminal segment AA683-AA750 also interacts with the N-terminal tail to facilitate the association between the N-terminal tail and the ARM domain [22]. Interestingly, a different C-terminal segment of beta-Catenin competes with cadherin on binding the ARM domain [22]. Figure 6c shows that a C-terminal segment on beta-Catenin starting at Leu-753 is highly conserved to the corresponding binding motif on cadherin. This newly identified location of the C-terminal segment is different from previously reported segment that spans from Gln-760 to the end of beta-Catenin [22]. Besides the competing interactions among different regions, the phosphorylation of Tyr-654 that is close to the C-terminal tail eliminates the interaction between cadherin and the ARM domain of beta-Catenin [22]. These complex regulatory roles of disordered tails of beta-Catenin are illustrated in (a) and (b) of Fig. 6b.

The contacting residues on the interface between human cadherin and human beta-Catenin include: Ile811, Leu815, Asp819, Pro826, Tyr827, Asp828, Ser829, and Leu831 on cadherin, and Ala656, Tyr654, Asn516, Arg474, His470.Tyr469, and Asn426 on beta-Catenin [2]. In which, Tyr-654 on beta-Catenin is critical for the association between cadherin and beta-Catenin. Its phosphorylation disrupts the interaction between cadherin and beta-Catenin [60, 63, 64]. The phosphorylation also decreases the affinity between the ARM domain and the C-terminal tail of beta-Catenin [60]. From a theoretical point of view, phosphorylation on Tyr-654 changes the local polarity of the ARM domain, leading to the interactional and functional modifications of the ARM domain. As shown in Fig. 6d, the interface composed of AA400-663 on the ARM domain is actually a pocket and is enriched of charged and hydrophobic residues, of which the combination is ideal for mediating protein-protein interaction. In addition, structural comparison of the ARM domains in two different structures, which are obtained before and after binding to a motif, shows that the RMSD value associated with the structural change between these two structures for the region from AA576 to AA685 is 0.75 angstroms. The direction of the line formed between the oxygen atom on the side chain of Tyr-654 and the Cα atom of Tyr432 is changed by 5.1° after the motif binding. These results may indicate that the structural variation caused by motif binding is negligible.

Once the interaction between beta-Catenin and its partners is eliminated, the concentration of free beta-Catenin in cytoplasma will be increased. The free beta-Catenin may enter one of two possible processes: (1) the N-terminal tail of free beta-Catenin undergoes phosphorylation and ubiquitination that lead to the degradation of beta-Catenin [1]; (2) free beta-Catenin transports into nucleus to activate gene expression. These two processes comprise the Wnt signaling pathway shown in (c), (d), and (e) of Fig. 6b. At stage (c), CKI-alpha phosphorylates Ser-45 first and then GSK-3beta phosphorylates Thr-41, Ser-37, and Ser-33 successively. Therefore, the N-terminal tail must be free from binding partners for the phosphorylation and subsequent ubiquitination modifications, but the C-terminal tail may or may not be bound to other molecules. The phosphorylation process is facilitated by Axin and APC, which are two large disordered proteins, through forming a destruction complex by binding beta-Catenin, CKI-alpha, and GSK-3beta. The disordered segments in Axin and APC bind to CKI-alpha, GSK-3beta, and the ARM domain of beta-Catenin. The presence of Axin and APC is assumed to increase the local concentration of both enzymes and substrate to accelerate the reaction [16].

When Wnt molecules are present, the destruction complex could not be formed. Therefore, beta-Catenin will accumulate in cytoplasma and transport from cytoplasma into nucleus as shown at stage (d). Different translocation mechanisms have been reported. The translocation can be either independent of the interaction with phenylalanine-glycine(FG)-repeats in the nuclear pore complex [30], or dependent on the interaction between the ARM domain of beta-Catenin and the FG repeats [17, 18]. FG repeats in nuclear pore complexes normally have 200 ~ 400 residues and are intrinsically disordered [65]. In the process of nuclear translocation, both the C-ter tail and the ARM domain of beta-Catenin are required [66]. Once being transported into nucleus, beta-Catenin will interact with TCF/LEF family proteins to trigger the transcription of target genes as shown at stage (e). Although only the ARM domain is required for the association between beta-Catenin and LEF-1 [19–21], both tails of beta-Catenin are observed to promote the transactivation of downstream genes [67, 68].

Discussions

Beta-catenin has two major functions: regulating the formation of cadherin junction and processing Wnt signaling. These functions are performed synergistically by incorporating three sequential regions of beta-Catenin, including the N-terminal tail, the ARM domain, and the C-terminal tail. The ARM domain is composed of 12 helical segments and forms a structured domain, but both N- and C-terminal tails are disordered. In this study, sequential motifs, structural features, mutations, interaction partners, and modifications on both tails were integrated together to characterize the regulatory roles of the disordered tails. This is for the first time that all the five different aspects have been taken into consideration to characterize the regulatory roles of disordered tails of beta-Catenin.

Synergy between ARM and tails

Beta-Catenin interacts with its partners using not only the ARM domain but also the tails. The ARM domain forms several binding pockets, while the tails contain multiple predicted binding motifs and ELM linear motifs. The ARM domain frequently binds to short flexible motifs of the partners, and the disordered tails interact with structured domains of the partners. Therefore, protein intrinsic disorder is very often involved in the interaction between beta-Catenin and its partners. Since the interactions involving disordered regions are often characterized by high-specificity and low-affinity [69], disordered motifs in the signaling pathways facilitate the regulation of the functions.

Functional motifs and sites in tails

Binding motif predictors may be used not only for identifying binding sites, but also for inferring the mechanisms of different binding processes since different predictors are designed for different types of binding motifs. Three binding motif predictors, including MoRF, MoRFpred, and ANCHOR, were used in this study. ANCHOR is able to identify binding motifs composed by mainly hydrophobic residues. MoRF predicts motifs undergoing coil-to-helix transition upon binding to partners. MoRFpred identifies motifs associated with structural transition from coil to all three types of secondary structures. The correlation between binding motifs and secondary structures can be demonstrated by comparing the binding motifs to the results of secondary structure prediction from NetSurfP. As shown in Fig. 1, it is clear that all the predicted binding motifs in the N-terminal tail have helical structures, while the binding motifs at the C-terminal tail are not helices. The structural information of binding motifs can be further used to infer the structural requirements of binding events.

The heterogeneous distribution of the ELM motifs found in the N-terminal and C-terminal tails indicates different functional roles of the tails. All DEG motifs locate in the N-terminal tail. Therefore, the N-terminal tail is the only region responsible for the ubiquitination and degradation of beta-Catenin. Actually, the degradation process of beta-Catenin has been well characterized. The phosphorylation of serine and threonine residues on the N-terminal tail is the prerequisite of subsequent ubiquitination and degradation of beta-Catenin [23, 27]. The C-terminal tail has more DOC and MOD motifs than the N-terminal tail does. The coincident abundance of both DOC and MOD motifs in the C-terminal tail is expected and consistent, since the DOC motifs function in binding modifying enzymes, and the MOD motifs refer to modification sites. Therefore, the C-terminal tail may undergo many more modification events.

The C-terminal tail of beta-Catenin contains two neighboring segments that have contradictory functions. AA683-AA750 interacts with the N-terminal tail to facilitate the binding between the N-terminal tail and the ARM domain, and to further facilitate the interaction between beta-Catenin and cadherin. AA753-AA781, which is right after the previous motif, competing with cadherin on the interaction with the ARM domain. Since these two segments have well conserved sequential patterns and are enriched of functional motifs, it is expected that additional partners may be involved in the regulation of the interaction between these two motifs and the ARM domain.

Tyr-654 is a critical residue in regulating the interaction between the C-terminal tail and the ARM domain of beta-Catenin, as well as the interaction between beta-Catenin and cadherin. The spatial position of Tyr-654 is at the mouth of the binding pocket formed by the 7th to the 12th armadillo repeats. Therefore, when the pocket is occupied by a binding motif, the accessible surface area of Tyr-654 will be decreased significantly. By taking into consideration that the structural variation associated with motif binding is minor, it can be concluded that the phosphorylation of Tyr-654 is unlikely to occur when a binding motif is present, indicating the phosphorylation/dephosphorylation of Tyr-654 has to occur in the unbound state. Therefore, the pre-binding modification of Tyr-654 regulates the interaction between the ARM domain and its binding partners.

Mutations on tails

High substitution rate is often observed in disordered proteins and regions. However, although both of the N- and C-terminal tails of beta-Catenin are disordered, they contain multiple conserved motifs (i.e., AA90 ~ AA130 in the N-terminal tail and AA774 ~ 781 in the C-terminal tail) and conserved sequential patterns (e.g., AA670 ~ AA680 and AA730 ~ AA740). These conserved regions actually overlap with predicted binding motifs and ELM linear motifs, indicating the conservation of functions. Even though, many segments with conserved amino acids or sequential patterns are still lack of functional annotation. AA90 ~ AA100 and AA140 ~ AA150 are two conserved motifs in the N-terminal tail. Both motifs are characterized as composites of hydrophobic residues and charged residues, with positive charged residues in the first motif and negative charged residues in the second motif. AA730 ~ AA740 and AA750 ~ AA760 are another two C-terminal segments with conserved sequential patterns involving hydrophobic, positively charged, and negatively charged residues on each of them.

Almost all the cancer-associated mutations on beta-Catenin are found in the N-terminal tail region, in which, over 70 % of the mutations locate in a short segment of 50 residues (AA25-AA75) at the N-terminal tail. This region accommodates Ser-45, Thr-41, Ser-37, and Ser-33, which are four critical residues in the Wnt signaling pathway. The phosphorylation of these four residues initiates the degradation of beta-Catenin. Therefore, it is conceivable that the mutations in this short region may eliminate the tight control on the degradation of beta-Catenin. Consequently, accumulated beta-Catenin turns on the expression of down-stream genes. In addition, the mutation in this short segment may also affect the function of beta-Catenin in the cadherin function formation pathway. In this pathway, the N-terminal tail of beta-Catenin interacts with the ARM domain to facilitate the interaction between beta-Catenin and cadherin. As shown in Figs. 1, 2b, and Additional file 1: Table S1, the N-terminal tail of beta-Catenin has three candidate binding motifs, locating at ~ AA25, ~AA80, and ~ AA130, respectively. The motif at ~ AA130 is sequentially next to the ARM domain. It is unlikely that this motif is able to "fold back" to interact with the ARM domain. As a result, only the motifs at ~ AA25 and ~ AA80 may perform the function of interacting with the ARM domain. Therefore, the afore-mentioned mutations in the region of AA25-AA75 may also abolish the interaction between the N-terminal tail and the ARM domain.

Conclusions

Disordered tails of beta-Catenin interact with each other and the ARM domain to regulate the function of beta-Catenin synergistically. The tails contain many functional motifs. The types of motifs on both tails are significantly different, indicating divergence of their functional roles. Modifications on or near these motifs further modulate the function of the entire molecule. Mutations in tails are very often associated with diseases. Thorough characterization of the functional motifs in the tails helps further studies on the regulatory mechanisms of beta-Catenin.

Additional files

Additional file 1: Table S1. Eukaryotic linear motifs of beta-Catenin. (PDF 26 kb)

Additional file 2: Table S2. List of beta-Catenin proteins from eight species. (PDF 8 kb)

Additional file 3: Table S3. PDB structures of complexes formed between beta-Catenin and its partners. (PDF 13 kb)

Additional file 4: Table S4. List of proteins in the first and secondary interactome of beta-Catenin. (PDF 198 kb)

Abbreviations

APC, Adenomatous Polyposis Coli tumor repressor protein; ARM, Armadillo, Axin, Axis Inhibition Protein; CK1-alpha, Casein Kinase I alpha; CLV, CLeaVage site; DEG, DEGragon; DOC, modifying enzyme DOCking; ELM, Eukariotic Linear Motif; FG-repeats, phenylalanine-Glycine-repeats; GSK-3beta, Glycogen Synthase Kinase 3β; LIG, LIGand binding; MOD, post-translational MODification; MoRF, Molecular Recognition Feature; PP2A, Protein Phosphatase 2A; TRG, subcellular TaRGeting.

Availability of data and materials
Datasets used in this study can be found in supplementary materials.

Authors' contributions
BZ and BX designed the experiment, analyzed the data, and wrote the manuscript. Both authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Consent for publication
Not applicable.

Ethics approval and consent to participate
Not applicable.

Published: 31 August 2016

References
1. Xing Y, Takemaru K, Liu J, Berndt JD, Zheng JJ, Moon RT, Xu W. Crystal structure of a full-length beta-catenin. Structure. 2008;16(3):478–87.
2. Huber AH, Weis WI. The structure of the beta-catenin/E-cadherin complex and the molecular basis of diverse ligand recognition by beta-catenin. Cell. 2001;105(3):391–402.
3. Brembeck FH, Rosario M, Birchmeier W. Balancing cell adhesion and Wnt signaling, the key role of beta-catenin. Curr Opin Genet Dev. 2006;16(1):51–9.
4. Kramps T, Peter O, Brunner E, Nellen D, Froesch B, Chatterjee S, Murone M, Zullig S, Basler K. Wnt/wingless signaling requires BCL9/legless-mediated recruitment of pygopus to the nuclear beta-catenin-TCF complex. Cell. 2002;109(1):47–60.
5. Xing Y, Clements WK, Kimelman D, Xu W. Crystal structure of a beta-catenin/axin complex suggests a mechanism for the beta-catenin destruction complex. Genes Dev. 2003;17(22):2753–64.
6. Schneider S, Steinbeisser H, Warga RM, Hausen P. Beta-catenin translocation into nuclei demarcates the dorsalizing centers in frog and fish embryos. Mech Dev. 1996;57(2):191–8.
7. Larabell CA, Torres M, Rowning BA, Yost C, Miller JR, Wu M, Kimelman D, Moon RT. Establishment of the dorso-ventral axis in Xenopus embryos is presaged by early asymmetries in beta-catenin that are modulated by the Wnt signaling pathway. J Cell Biol. 1997;136(5):1123–36.
8. Sawa H. Control of cell polarity and asymmetric division in C. elegans. Curr Top Dev Biol. 2012;101:55–76.
9. Sokol SY. Maintaining embryonic stem cell pluripotency with Wnt signaling. Development. 2011;138(20):4341–50.

10. Kobayashi M, Honma T, Matsuda Y, Suzuki Y, Narisawa R, Ajioka Y, Asakura H. Nuclear translocation of beta-catenin in colorectal cancer. Br J Cancer. 2000;82(10):1689–93.

11. Hassanein AM, Glanz SM, Kessler HP, Eskin TA, Liu C. beta-Catenin is expressed aberrantly in tumors expressing shadow cells. Pilomatricoma, craniopharyngioma, and calcifying odontogenic cyst. Am J Clin Pathol. 2003;120(5):732–6.

12. Kypta RM, Waxman J. Wnt/beta-catenin signalling in prostate cancer. Nat Rev Urol. 2012;9(8):418–28.

13. Perriard JC, Hirschy A, Ehler E. Dilated cardiomyopathy: a disease of the intercalated disc? Trends Cardiovasc Med. 2003;13(1):30–8.

14. Mahmoodzadeh S, Eder S, Nordmeyer J, Ehler E, Huber O, Martus P, Weiske J, Pregla R, Hetzer R, Regitz-Zagrosek V. Estrogen receptor alpha up-regulation and redistribution in human heart failure. Faseb J. 2006;20(7):926–34.

15. Xue B, Dunker AK, Uversky VN. The roles of intrinsic disorder in orchestrating the Wnt-pathway. J Biomol Struct Dyn. 2012;29(5):843–61.

16. Xue B, Romero PR, Noutsou M, Maurice MM, Rudiger SG, William Jr AM, Mizianty MJ, Kurgan L, Uversky VN, Dunker AK. Stochastic machines as a colocalization mechanism for scaffold protein function. FEBS Lett. 2013;587(11):1587–91.

17. Fagotto F, Gluck U, Gumbiner BM. Nuclear localization signal-independent and importin/karyopherin-independent nuclear import of beta-catenin. Curr Biol. 1998;8(4):181–90.

18. Sharma M, Jamieson C, Johnson M, Molloy MP, Henderson BR. Specific armadillo repeat sequences facilitate beta-catenin nuclear transport in live cells via direct binding to nucleoporins Nup62, Nup153, and RanBP2/Nup358. J Biol Chem. 2012;287(2):819–31.

19. Behrens J, von Kries JP, Kuhl M, Bruhn L, Wedlich D, Grosschedl R, Birchmeier W. Functional interaction of beta-catenin with the transcription factor LEF-1. Nature. 1996;382(6592):638–42.

20. Huber O, Korn R, McLaughlin J, Ohsugi M, Herrmann BG, Kemler R. Nuclear localization of beta-catenin by interaction with transcription factor LEF-1. Mech Dev. 1996;59(1):3–10.

21. Hsu SC, Galceran J, Grosschedl R. Modulation of transcriptional regulation by LEF-1 in response to Wnt-1 signaling and association with beta-catenin. Mol Cell Biol. 1998;18(8):4807–18.

22. Castano J, Raurell I, Piedra JA, Miravet S, Dunach M, Garcia de Herreros A. Beta-catenin N- and C-terminal tails modulate the coordinated binding of adherens junction proteins to beta-catenin. J Biol Chem. 2002;277(35):31541–50.

23. van Noort M, Meeldijk J, van der Zee R, Destree O, Clevers H. Wnt signaling controls the phosphorylation status of beta-catenin. J Biol Chem. 2002;277(20):17901–5.

24. Clevers H. Wnt/beta-catenin signaling in development and disease. Cell. 2006;127(3):469–80.

25. Klaus A, Birchmeier W. Wnt signalling and its impact on development and cancer. Nat Rev Cancer. 2008;8(5):387–98.

26. Aberle H, Bauer A, Stappert J, Kispert A, Kemler R. beta-catenin is a target for the ubiquitin-proteasome pathway. Embo J. 1997;16(13):3797–804.

27. Orford K, Crockett C, Jensen JP, Weissman AM, Byers SW. Serine phosphorylation-regulated ubiquitination and degradation of beta-catenin. J Biol Chem. 1997;272(40):24735–8.

28. Winer IS, Bommer GT, Gonik N, Fearon ER. Lysine residues Lys-19 and Lys-49 of beta-catenin regulate its levels and function in T cell factor transcriptional activation and neoplastic transformation. J Biol Chem. 2006;281(36):26181–7.

29. Sharma M, Jamieson C, Lui C, Henderson BR. The hydrophobic rich N- and C-terminal tails of beta-catenin facilitate nuclear import of beta-catenin. J Biol Chem. 2014. [Epub ahead of print].

30. Suh EK, Gumbiner BM. Translocation of beta-catenin into the nucleus independent of interactions with FG-rich nucleoporins. Exp Cell Res. 2003;290(2):447–56.

31. Mosimann C, Hausmann G, Basler K. Beta-catenin hits chromatin: regulation of Wnt target gene activation. Nat Rev Mol Cell Biol. 2009;10(4):276–86.

32. UniProtKB. http://www.uniprot.org/. Accessed Feb 2015.

33. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. Genome Res. 2004;14(6):1188–90.

34. Forbes SA, Beare D, Gunasekaran P, Leung K, Bindal N, Boutselakis H, Ding M, Bamford S, Cole C,Ward S et al. COSMIC: exploring the world's knowledge of somatic mutations in human cancer. Nucleic Acids Res. 2015;43(Database issue):D805–11.

35. Peterson TA, Adadey A, Santana-Cruz I, Sun Y, Winder A, Kann MG. DMDM: domain mapping of disease mutations. Bioinformatics. 2010;26(19):2458–9.

36. Wu TJ, Shamsaddini A, Pan Y, Smith K, Crichton DJ, Simonyan V, Mazumder R. A framework for organizing cancer-related variations from existing databases, publications and NGS data using a High-performance Integrated Virtual Environment (HIVE). Database (Oxford). 2014;2014:bau022.

37. Petersen B, Petersen TN, Andersen P, Nielsen M, Lundegaard C. A generic method for assignment of reliability scores applied to solvent accessibility predictions. BMC Struct Biol. 2009;9:51.

38. Xue B, Dunbrack RL, Williams RW, Dunker AK, Uversky VN. PONDR-FIT: a meta-predictor of intrinsically disordered amino acids. Biochim Biophys Acta. 2010;1804(4):996–1010.

39. Romero P, Obradovic Z, Li X, Garner EC, Brown CJ, Dunker AK. Sequence complexity of disordered protein. Proteins. 2001;42(1):38–48.

40. Peng K, Radivojac P, Vucetic S, Dunker AK, Obradovic Z. Length-dependent prediction of protein intrinsic disorder. BMC Bioinformatics. 2006;7:208.

41. Obradovic Z, Peng K, Vucetic S, Radivojac P, Brown CJ, Dunker AK. Predicting intrinsic disorder from amino acid sequence. Proteins. 2003;53 Suppl 6:566–72.

42. Prilusky J, Felder CE, Zeev-Ben-Mordehai T, Rydberg EH, Man O, Beckmann JS, Silman I, Sussman JL. FoldIndex: a simple tool to predict whether a given protein sequence is intrinsically unfolded. Bioinformatics. 2005;21(16):3435–8.

43. Dosztanyi Z, Csizmok V, Tompa P, Simon I. IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. Bioinformatics. 2005;21(16):3433–4.

44. Campen A, Williams RM, Brown CJ, Meng J, Uversky VN, Dunker AK. TOP-IDP-scale: a new amino acid scale measuring propensity for intrinsic disorder. Protein Pept Lett. 2008;15(9):956–63.

45. Li J, Feng Y, Wang X, Liu W, Rong L, Bao J. An Overview of Predictors for Intrinsically Disordered Proteins over 2010-2014. Int J Mol Sci. 2015;16(10):23446–62.

46. Dosztanyi Z, Meszaros B, Simon I. ANCHOR: web server for predicting protein binding regions in disordered proteins. Bioinformatics. 2009;25(20):2745–6.

47. Oldfield CJ, Cheng Y, Cortese MS, Romero P, Uversky VN, Dunker AK. Coupled folding and binding with alpha-helix-forming molecular recognition elements. Biochemistry. 2005;44(37):12454–70.

48. Disfani FM, Hsu WL, Mizianty MJ, Oldfield CJ, Xue B, Dunker AK, Uversky VN, Kurgan L. MoRFpred, a computational tool for sequence-based prediction and characterization of short disorder-to-order transitioning binding regions in proteins. Bioinformatics. 2012;28(12):i75–83.

49. Dinkel H, Van Roey K, Michael S, Davey NE, Weatheritt RJ, Born D, Speck T, Kruger D, Grebnev G, Kuban M et al. The eukaryotic linear motif resource ELM: 10 years and counting. Nucleic Acids Res. 2014;42(Database issue):D259–66.

50. Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, Lin J, Minguez P, Bork P, von Mering C et al. STRING v9.1: protein-protein interaction networks, with increased coverage and integration. Nucleic Acids Res. 2013;41(Database issue):D808–15.

51. Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Res. 2014;42(Database issue):D199–205.

52. Shihab HA, Gough J, Cooper DN, Stenson PD, Barker GL, Edwards KJ, Day IN, Gaunt TR. Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden Markov models. Hum Mutat. 2013;34(1):57–65.

53. Ryo A, Nakamura M, Wulf G, Liou YC, Lu KP. Pin1 regulates turnover and subcellular localization of beta-catenin by inhibiting its interaction with APC. Nat Cell Biol. 2001;3(9):793–801.

54. Hart M, Concordet JP, Lassot I, Albert I, del los Santos R, Durand H, Perret C, Rubinfeld B,Margottin F, Benarous R et al. The F-box protein beta-TrCP associates with phosphorylated beta-catenin and regulates its activity in the cell. Curr Biol. 1999;9(4):207–10.

55. Liu C, Li Y, Semenov M, Han C, Baeg GH, Tan Y, Zhang Z, Lin X, He X. Control of beta-catenin phosphorylation/degradation by a dual-kinase mechanism. Cell. 2002;108(6):837–47.

56. Su Y, Fu C, Ishikawa S, Stella A, Kojima M, Shitoh K, Schreiber EM, Day BW, Liu B. APC is essential for targeting phosphorylated beta-catenin to the SCFbeta-TrCP ubiquitin ligase. Mol Cell. 2008;32(5):652–61.

57. Wu G, Xu G, Schulman BA, Jeffrey PD, Harper JW, Pavletich NP. Structure of a beta-TrCP1-Skp1-beta-catenin complex: destruction motif binding and lysine specificity of the SCF(beta-TrCP1) ubiquitin ligase. Mol Cell. 2003;11(6):1445–56.

58. Taurin S, Sandbo N, Qin Y, Browning D, Dulin NO. Phosphorylation of beta-catenin by cyclic AMP-dependent protein kinase. J Biol Chem. 2006;281(15):9971–6.

59. Zhang J, Yan X, Shi C, Yang X, Guo Y, Tian C, Long J, Shen Y. Structural basis of beta-catenin recognition by Tax-interacting protein-1. J Mol Biol. 2008;384(1):255–63.

60. Piedra J, Martinez D, Castano J, Miravet S, Dunach M, de Herreros AG. Regulation of beta-catenin structure and activity by tyrosine phosphorylation. J Biol Chem. 2001;276(23):20436–43.

61. Xu K, Uversky VN, Xue B. Local flexibility facilitates oxidization of buried methionine residues. Protein Pept Lett. 2012;19(6):688–97.

62. Xue B, Uversky V. Structural characterizations of phosphorylatable residues in transmembrane proteins from Arabidopsis thaliana. Intrinsically Disordered Proteins. 2013;1:1,e25713.

63. Roura S, Miravet S, Piedra J, Garcia de Herreros A, Dunach M. Regulation of E-cadherin/Catenin association by tyrosine phosphorylation. J Biol Chem. 1999;274(51):36734–40.

64. Bonvini P, An WG, Rosolen A, Nguyen P, Trepel J, Garcia de Herreros A, Dunach M, Neckers LM. Geldanamycin abrogates ErbB2 association with proteasome-resistant beta-catenin in melanoma cells, increases beta-catenin-E-cadherin association, and decreases beta-catenin-sensitive transcription. Cancer Res. 2001;61(4):1671–7.

65. Denning DP, Patel SS, Uversky V, Fink AL, Rexach M. Disorder in the nuclear pore complex: the FG repeat regions of nucleoporins are natively unfolded. Proc Natl Acad Sci U S A. 2003;100(5):2450–5.

66. Koike M, Kose S, Furuta M, Taniguchi N, Yokoya F, Yoneda Y, Imamoto N. beta-Catenin shows an overlapping sequence requirement but distinct molecular interactions for its bidirectional passage through nuclear pores. J Biol Chem. 2004;279(32):34038–47.

67. van de Wetering M, Cavallo R, Dooijes D, van Beest M, van Es J, Loureiro J, Ypma A, Hursh D, Jones T, Bejsovec A et al. Armadillo coactivates transcription driven by the product of the Drosophila segment polarity gene dTCF. Cell. 1997;88(6):789–99.

68. Hecht A, Litterst CM, Huber O, Kemler R. Functional characterization of multiple transactivating elements in beta-catenin, some of which interact with the TATA-binding protein in vitro. J Biol Chem. 1999;274(25):18017–25.

69. van der Lee R, Buljan M, Lang B, Weatheritt RJ, Daughdrill GW, Dunker AK, Fuxreiter M, Gough J, Gsponer J, Jones DT et al. Classification of intrinsically disordered regions and proteins. Chem Rev. 2014;114(13):6589–631.