



Structural and mutational analysis reveals that CTNNB1 binds NLSs in a manner distinct from that of its closest armadillo-relative, karyopherin α [☆]



Karuna Ganesh ¹, Febe van Maldegem, Stephanie B. Telerman, Paul Simpson, Christopher M. Johnson, Roger L. Williams, Michael S. Neuberger ², Cristina Rada ^{*}

Medical Research Council Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge CB2 0QH, UK

ARTICLE INFO

Article history:

Received 10 September 2013

Revised 18 October 2013

Accepted 4 November 2013

Available online 20 November 2013

Edited by Gianni Cesareni

Keywords:

Splicing

Nuclear import

Armadillo domain

ABSTRACT

CTNNB1 is a spliceosome-associated protein that binds nuclear localization signals (NLSs) in splice factors CDC5L and Prp31 as well as the antibody diversifying enzyme AID. Here, crystal structures of human CTNNB1 reveal a distinct structure from its closest homologue karyopherin- α . CTNNB1 comprises a HEAT-like domain (including a nuclear export signal), a central armadillo domain, and a coiled-coil C-terminal domain. Structure-guided mutations of the region homologous to the karyopherin- α NLS-binding site fail to disrupt CTNNB1–NLS interactions. Our results identify CTNNB1 as a unique selective NLS-binding protein with striking differences from karyopherin- α s.

Structured summary of protein interactions:

CTNNB1 binds to **CDC5L NLS3** by pull down (View interaction)

CTNNB1 and **CDC5L NLS3** bind by isothermal titration calorimetry (View interaction)

CTNNB1 and **Prp31 NLS** bind by isothermal titration calorimetry (View interaction)

© 2013 The Authors. Published by Elsevier B.V. All rights reserved.

1. Introduction

CTNNB1 is a widely expressed, highly conserved 62 kDa nuclear protein [1] that associates with components of the RNA splicing machinery [2–4]. Mice lacking CTNNB1 die in mid-gestation, whilst CTNNB1-deficient mouse primary B cells and yeast show delayed exit from quiescence [5]. Whilst CTNNB1 lacks significant sequence similarity to other proteins, secondary structural predictions suggest that CTNNB1's central region might comprise an armadillo domain, a multifunctional domain that is found in several unrelated proteins [1,6]. CTNNB1 has been predicted to resemble the karyopherin- α nuclear transport adaptors, in which the armadillo domain surface groove mediates interactions with nuclear localization signals (NLSs) in cargo proteins. In previous work we have shown that like the karyopherin- α s, CTNNB1 binds to NLSs with high selectivity and affinity, in particular the NLS region of essential spliceosome components CDC5L and PRPF31, as well as the ill defined NLS of the antibody diversification enzyme

AID [4]. However, CTNNB1 has an NLS binding specificity distinct from that of karyopherin- α , failing, for example, to bind to the NLS of the SV40 T-antigen [4]. To gain insight into the distinct NLS binding preference of CTNNB1, we solved the crystal structure of CTNNB1, and based on the structure, probed the putative NLS binding region by mutagenesis. The results show a distinctive alpha helical protein, only partially comprised of armadillo repeats, with striking dissimilarities from karyopherin- α .

2. Methods

2.1. Protein expression, purification and crystallization

CTNNB1 and CTNNB1 Δ 76 were expressed and purified as described previously [4]. CTNNB1 Δ 76 mutants were generated by PCR using the primers shown in [Supplementary Table 1](#). Selenomethionine-labeled protein was expressed in B834(DE3) cells as described previously [7] and purified as for unlabeled protein.

Purified recombinant proteins at a concentration of 24 mg ml⁻¹ in crystallization buffer (20 mM Hepes pH 7.5, 50 mM NaCl, 1 M DTT) were tested in crystallization screens with 1536 different conditions in sitting drops (100 nl protein solution + 100 nl reservoir solution) in 96-well MRC crystallization plates using a Cartesian robot (Genomics solutions, Huntingdon, UK). Optimal crystals were grown in sitting drops in 24-well crystallization plates (CrysChem) by mixing 0.5 ml protein solution with 0.5 ml

[☆] This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

^{*} Corresponding author. Fax: +44 0 1223 268305.

E-mail address: car@mrc-lmb.cam.ac.uk (C. Rada).

¹ Present address: Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, USA.

² M.S. Neuberger died 26th October 2013. This article is dedicated to his memory.

reservoir solution. Solutions and cryoprotection details are listed in [Supplementary Table 2](#).

2.2. Data collection

Data were collected at 100 K from cryoprotected crystals frozen in a nitrogen gas stream. MAD datasets from the selenomethionine crystal, CTNNBL1 and CTNNBL1Δ76B (crystal form B of CTNNBL1Δ76) were collected at the Diamond Light Source beamline I02, while CTNNBL1Δ76A (crystal form A of CTNNBL1Δ76) was collected at beamline I04-1. Details of the methods used for Phasing and Model Refinement are described in [Supplementary methods](#). Crystal statistics are given in [Table 1](#).

2.3. Assaying peptide binding

Peptides corresponding to the CDC5L NLS3, SV40 NLS and Prp31 NLS and CTNNBL1Δ76 mutants were dialysed against isothermal titration calorimetry (ITC) buffer (20 mM Hepes pH 7.5, 50 mM NaCl) and concentrated to 0.83 mM (peptide) or 50 μM (protein) for experiments on an ITC₂₀₀ or auto-ITC₂₀₀ calorimeter (Microcal, Inc.). ITC and pulldown assays were performed as described previously [4].

2.4. Protein localization

293T cells transiently transfected with N-terminally GFP-tagged CTNNBL1 fragments were fixed, permeabilized, counterstained with Hoechst 3258 and Alexa-594 conjugated wheatgerm agglutinin, and visualized using confocal microscopy as described [4].

3. Results and discussion

3.1. Crystal structures of human CTNNBL1

We crystallized full-length human 6xHis-CTNNBL1 and N-terminally truncated 6xHis-CTNNBL1Δ76 (to generate a fragment that is necessary and sufficient for NLS-peptide binding

[4]). We were unable to solve the structure using molecular replacement and therefore crystallized selenomethionine-substituted protein for experimental phase determination, solving the structures of CTNNBL1Δ76 and that of full-length CTNNBL1 from crystals with the symmetry of different spacegroups ([Table 1](#)).

The crystals reveal the same overall structure of CTNNBL1. However, despite excellent continuity within the protein chain, we were unable to resolve any density for the N-terminal (76) amino acids in the structure of full length CTNNBL1. The overall structure of CTNNBL1Δ76 reveals a superhelical coiled coil composed of α-helices stacked against each other to form an extended configuration, in agreement with circular dichroism data ([Supplementary Fig. 2B](#)). The structure can be divided into three distinct regions ([Fig. 1A](#)): an N-terminal pre-armadillo (PRE-ARM) domain (residues 76–176); a central compact armadillo (ARM) domain with helices tightly packed against one another to form a solenoid (residues 177–416); and a C-terminal domain (CTD) that projects out from the smooth solenoid of the ARM domain (residues 417–562) ([Fig. 1B](#)).

Residues along the central spine of the solenoid and the CTD have low temperature factors, consistent with their role in stabilizing the core of the molecule, whilst N and C-terminal regions beyond the CTD display higher B-factors for solvent-exposed residues, implying greater flexibility ([Fig. 1C](#)). The solenoid core forms a concave electronegative groove, comprised primarily of acidic amino acids from helices 1 and 3 of the armadillo domain, similar to the acidic concave groove of karyopherin-α [8] and unlike the positively charged groove formed by β-catenin [6,9,10] ([Fig. 1D](#)). In contrast to the negatively charged core, patches of positive charge are located on the periphery.

3.2. The pre-armadillo domain of CTNNBL1 contains a HEAT repeat-like region

Whilst showing no amino acid sequence homology to other armadillo containing proteins, secondary structure predictions have suggested that CTNNBL1 comprises multiple armadillo repeats although the number and extent of these repeats was diffi-

Table 1
Data collection, phasing and refinement statistics.

Crystal name	CTNNBL1Δ76A	CTNNBL1Δ76B	CTNNBL1	SeMetCTNNBL1		
Space group	<i>P</i> ₂ ₁ ₂ ₁	<i>I</i> 222	<i>P</i> ₄ ₃ ₂ ₁ ₂	<i>P</i> ₄ ₃ ₂ ₁ ₂		
Cell dimensions <i>a</i> , <i>b</i> , <i>c</i> (Å)	63.1, 92.7, 121.7	87.4, 92.8, 194.9	66.7, 66.7, 325.9	66.8, 66.8, 328.2		
Wavelength (Å)	0.979500	0.979500	0.979500	0.979700	0.97980	0.96820
Resolution (Å)	63–2.9 (3.08–2.9)	40–3.1 (3.3–3.1)	81–3.0 (3.2–3.0)	47.36–2.32	47.36–2.32	47.46–2.32
<i>R</i> _{merge}	6.8 (49.1)	9.7 (39.8)	11.2 (91.6)	21 (–165)	21.7 (238.7)	2.2 (395.4)
<i>I</i> / <i>σ</i> <i>I</i>	14.6 (3.3)	10.5 (3.6)	16.4 (3.1)	7.3 (–0.1)	7.4 (0.6)	7.7 (0.8)
Completeness (%)	97.6 (99.0)	98.8 (98.4)	99.96 (100)	64.6	62.4	62.2
Redundancy	4.0 (4.1)	5.0 (5.1)	13.4 (13.6)	6.2	6.3	6.3
Refinement						
Resolution (Å)	52.1–2.9	39.9–3.1	65.4–3.0	47.42–3.3		
No. of reflections	15031	13833	14963			
<i>R</i> _{work} / <i>R</i> _{free}	19.2/26.9	21.7/27.9	21.8/29.8			
No. atoms						
Protein	3924	3968	4023			
SO ₄ ³⁻	2	4	0			
Water	5	8	2			
<i>B</i> -Factors	67.9	69.17	90.886			
R.m.s deviations						
Bond lengths (Å)	0.011	0.011	0.009			
Bond angles (°)	1.43	1.74	1.35			

One crystal was used per data set. Values in parentheses are for the highest resolution shell.

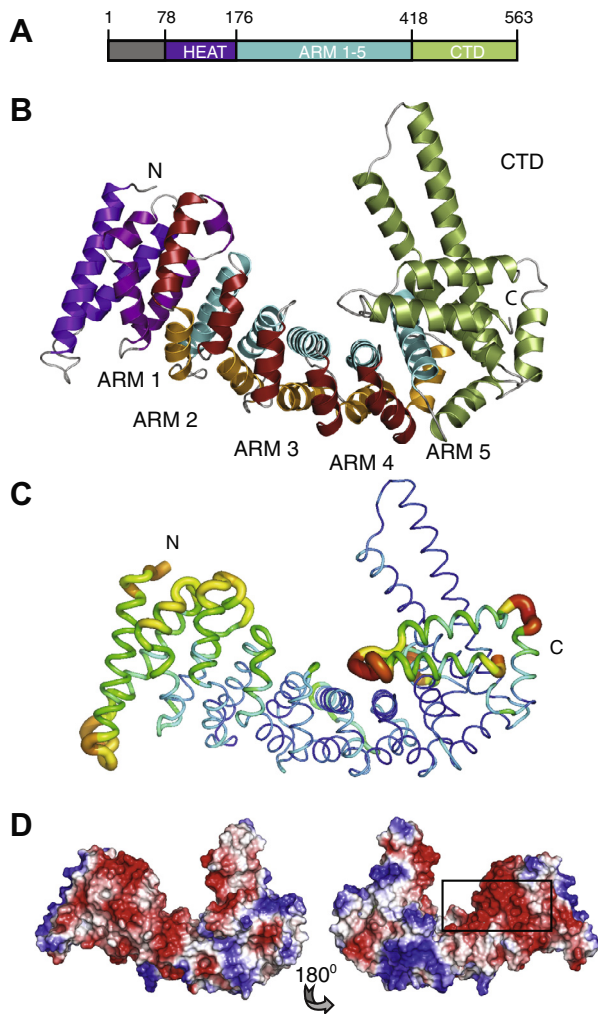


Fig. 1. Structure of CTNNB1. (A) Architecture of CTNNB1 with the N-terminal unresolved region, HEAT-like domain, five armadillo repeats (ARM) and C-terminal domain (CTD). (B) Three-dimensional structure of CTNNB1. The PRE-ARM domain HEAT-like repeats are in pink and purple; ARM repeats indicated with helix 1, helix 2 and helix 3 in red, orange and cyan and the CTD is shown in green. (C) B-factors representation in CTNNB1 crystals. Lower B-factors are blue and are thinner while thicker regions in yellow to red have higher B-factors. (D) Electrostatic surface potential of CTNNB1. (Red, negative; blue, positive). Views rotated by 180°. Acidic groove boxed.

cult to define [1]. The crystal structure shows that the armadillo domain of CTNNB1 begins at E178. Preceding the ARM domain is a region (PRE-ARM domain) comprising 6 α -helices and connecting loops, which do not form the ordered triple helical arrays that constitute armadillo repeats [9, 11–13]. Notably, this region lacks the highly conserved glycines at the junction between the first and second helices that are a hallmark of armadillo repeats [13]. Instead, residues Q79 to T129 form a pair of antiparallel helices with the remaining four helices of the PRE-ARM domain (residues Y134 to E178) resembling an intermediate type of motif that falls between that of a modified armadillo repeat (with the insertion of an extra C-terminal helix) and a HEAT-like repeat (with two kinked helices A and B) (Fig. 2A–C). The HEAT repeat, which is a relative of the armadillo repeat, consists of a tandem array of two antiparallel helices (A and B) instead of the three that comprise the armadillo repeat, although helix A often contains a central proline residue that leads to kinking of this helix [13–15].

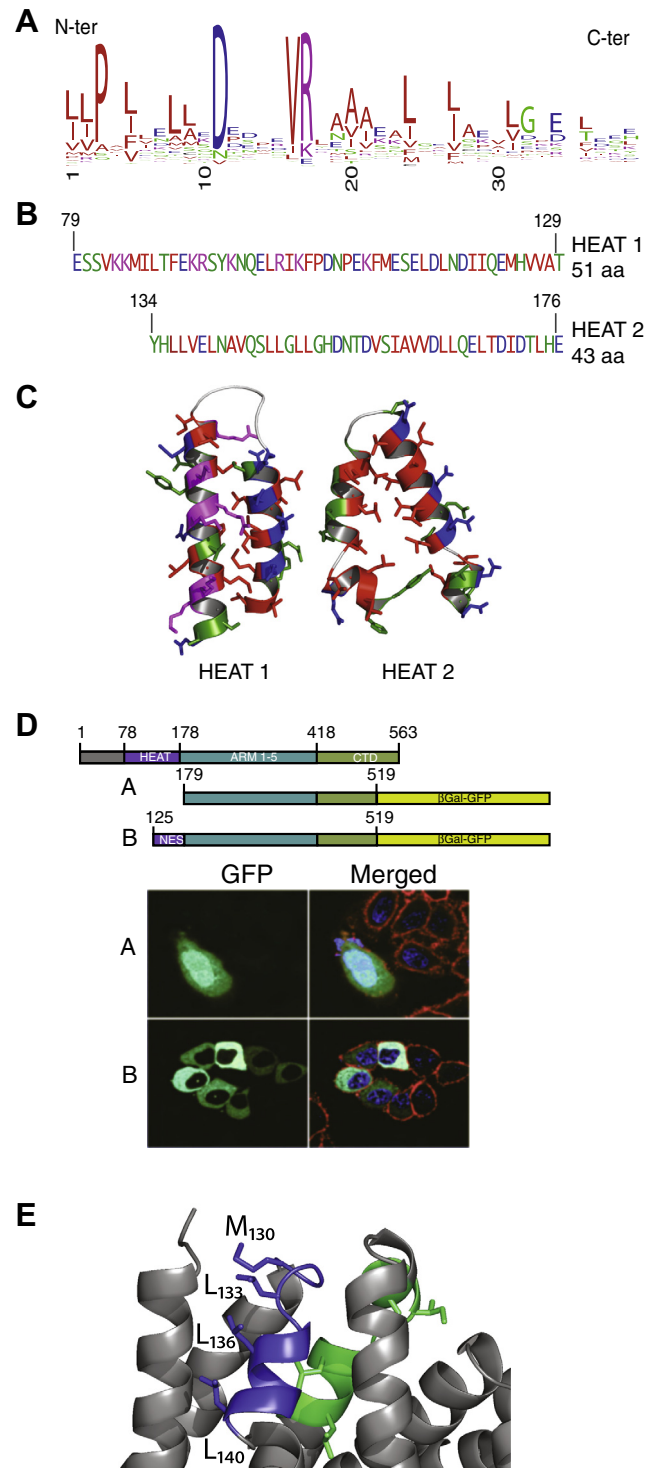


Fig. 2. The PRE-ARM domain of CTNNB1 contains HEAT-like motifs comprising a nuclear export signal. (A) Consensus HEAT repeat (as in Supplementary Fig. 4) (Residues are colored acidic, blue; basic, pink; hydrophobic, red and non-polar, green). (B) Amino acid sequence of HEAT-like regions 1 and 2 of CTNNB1. (C) The structures of the HEAT-like regions are stabilized by conserved hydrophobic residues (colored as in A). (D) The PRE-ARM domain CTNNB1 contains a functional export signal. Cytoplasmic relocalization of a chimera of CTNNB1 residues 179–519- β -galactosidase-GFP in transfected 293T cells by addition of residues 125–179. (E) Detail shows in blue (residues 130–140), the putative functional NES region with the key leucine residues depicted as sticks. The predicted NES in Jabbour et al. is shown in green indicating the leucines buried inside the core of the protein.

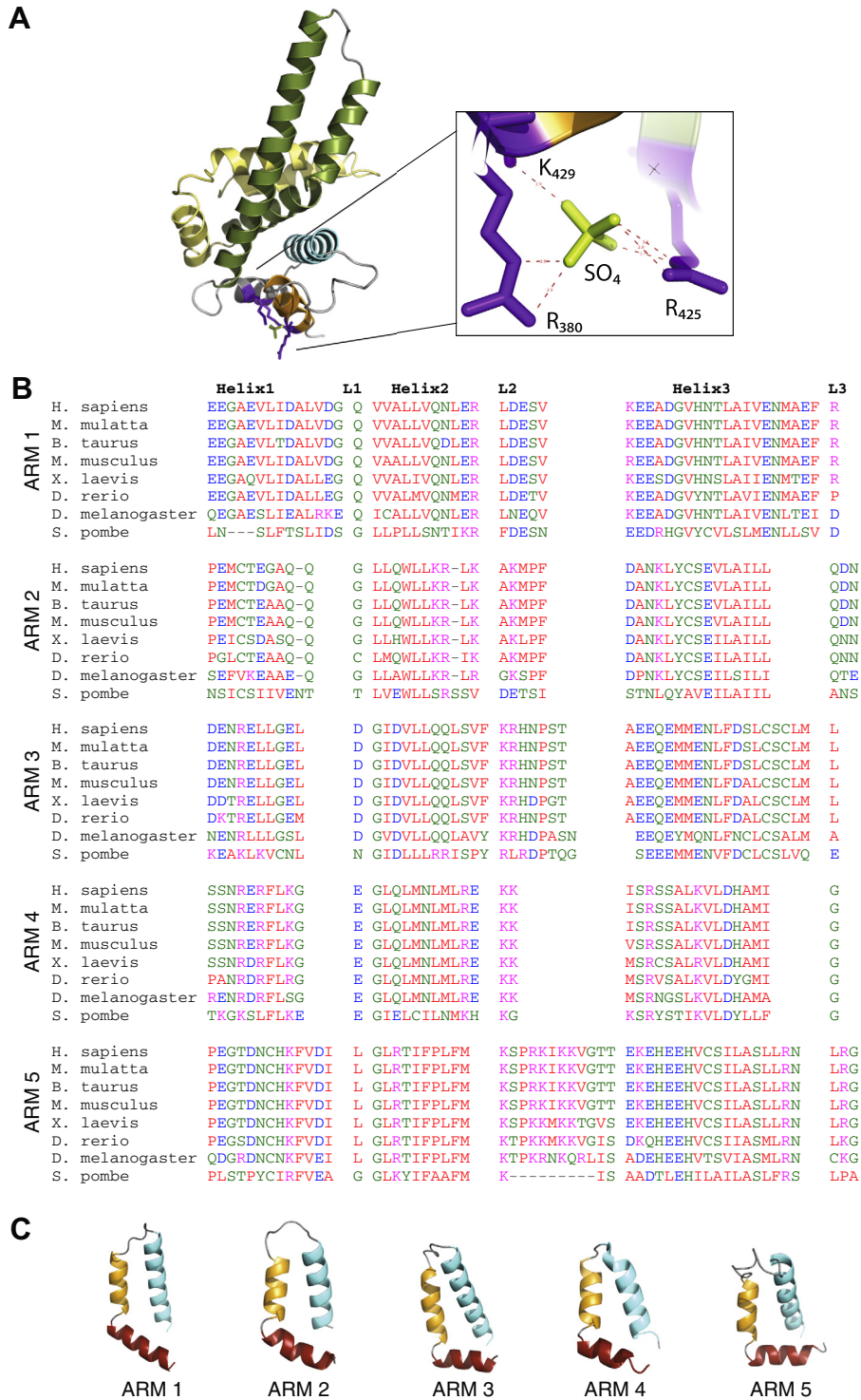


Fig. 3. The rigid CTD of CTNBL1 terminates the conserved ARM domain of CTNBL1. (A) Diagram showing helices 2 and 3 of armadillo repeat 5 (orange and cyan) and the CTD coiled-coil (green) and wing-like region (yellow). Hydrophobic residues stabilizing the base of the CTD are shown in purple. (B) Structure-based multiple sequence alignment of the 5 armadillo repeats of CTNBL1, including *S. pombe*. (C) Structural variation among individual armadillo repeats of CTNBL1. Repeats are oriented vertically along the central axis.

While attempts to narrowly define the HEAT domain have been made (e.g. [13,16]), recent accumulation of structural data has shown that HEAT repeats are highly variable in length, position of hydrophobic residues and the conservation of other residues (such as P3, D11 and R18) (Fig. 2A and Ref. [12]). As in HEAT repeats, the CTNBL1 Δ 76 PRE-ARM helices also contain conserved hydrophobic residues in the buried interaction surface of the

helices where they mediate intra- and inter-repeat packing. Polar residues line the solvent-exposed faces of the amphipathic helices (Fig. 2B and C), and conserved prolines kink the α -helices. Another feature of HEAT repeats is the presence of conserved polar residues near the surface inter-helix loop, which mediate contact between the two helices [13]. In CTNBL1, N96 interacts with S113 and H152, and R100 interacts with E114. Thus the pre-ARM domain

of CTNNB1 is an α -helical region that transitions between a HEAT and an armadillo repeat topology, explaining the difficulties in predicting the precise boundaries of the ARM core domain.

3.3. The HEAT-like region of CTNNB1 contains a functional nuclear export sequence

While CTNNB1 is actively localized to the nucleus [1,4], Jabbour et al. suggested that residues 164–174 (LLQELTDIDTL) could constitute a nuclear export sequence (NES) [1]. The structure reveals that these residues are unlikely to have this role since the leucine residues are buried in the protein core. Nonetheless, addition of residues 125–179 to the armadillo domain of CTNNB1 causes complete cytoplasmic relocalization of a GFP-fusion protein, while the armadillo domain on its own remains entirely nuclear, confirming that the region contains an active NES (Fig. 2D). Leucine-rich NESs are typically located in mobile regions of proteins, close to the surface, at a structural transition between a stacked amphipathic helix and an extended loop [17,18]. On inspection of the structure, there is only one part of the CTNNB1 PRE-ARM domain that meets these criteria: residues 130–140 (MPDLYHLLVEL) have high B factors, and form a transition from a loop to a helix protruding outwards at the surface of CTNNB1 (Fig. 2E). The hydrophobic side chains of residues M130, L133, L136, L140 are exposed on one side of an amphipathic helix, with D132, Y134, E139 on the opposite face. Based on the crystal structure, CTNNB1 residues 130–140 could therefore constitute an active nuclear export sequence.

3.4. The C-terminal domain of CTNNB1 constitutes a novel protein fold containing a coil-coil

The most dramatic feature of the structure is the abrupt termination of the HEAT/ARM solenoid by a rigid extended 34 aa helix and an antiparallel pair of helices 14 and 17 aa long, separated by G489. Despite the presence of the glycine residue, the whole region is quite rigid (as suggested by the B-factors (Fig. 1C), although the 108° angle between the latter two helices forms a kink that makes the coiled-coil region wider at the distal end from the axis of the solenoid. The rigidity of the structure is achieved by a dense cluster of hydrophobic interactions among the last arm motif H3 helix and CTD helices 1, 2, 4 and 5, forming a tightly packed base for the CTD (Fig. 3A). Although residues L519, I526, I531 and I540 are arranged in the heptad pattern typical of leu-ile zipper motifs, the tertiary structure shows that the side-chains point towards the protein core and do not participate in the typical interactions of a Leu-Ile zipper [19], in contrast to prior prediction [1]. The wider distal part of the coiled-coil is less conserved (Supplementary Fig. 1) although an acidic patch is still present in fission yeast *Schizosaccharomyces pombe* (residues D462, E464, E465 in *S. pombe* and D476, E479, E480 in human). The last three helices of the CTD form an additional flexible armadillo-like repeat.

The base of CTD is abutted perpendicularly to the main axis of the solenoid, with a density corresponding to a sulfate ion buried deep within a highly conserved pocket formed by R380, R425 and K429, all making polar contacts. It is possible that the sulfate in the crystals mimic a natural (e.g. phosphorylated) ligand and that this interaction contributes to the stability of the coil-coil region, modulating the position of the CTD and access to the concave surface of the solenoid.

3.5. The core of CTNNB1 comprises a compact array of five armadillo repeats

Although Jabbour et al. had predicted that CTNNB1 might contain a large ARM domain comprising 10 armadillo repeats [1],

analysis of the crystal structure reveals a much smaller ARM domain, with only five armadillo repeats similar to those found in other ARM proteins [12,13] (Fig. 1A). The repeats range in size between 36 and 45 amino acids, and measure about 30 Å in diameter (Fig. 3B). Each armadillo repeat is organized into three alpha helices, with helix 1 arranged perpendicularly to helices 2 and 3, which in turn run antiparallel to one another (Fig. 3C). As in consensus ARM repeats, the helix 1s are short, and helix 3s are longest (up to 20 amino acids) and most highly conserved. The long helix 3s stack against one another to form a concave groove, with an average translational distance between adjacent repeats of 7.9 Å.

The individual ARM repeats in the CTNNB1 structure conform to the overall ARM consensus. They show considerable sequence diversity from repeat to repeat, but the overall array is highly conserved among CTNNB1 homologues (Fig. 3B and C, Supplementary Fig. 2). Each repeat contains at least one conserved glycine at the junction between helices 1 and 2, a hallmark of ARM motifs, with backbone dihedral angles in the α_1 region of the Ramachandran plot that permit the sharp 90° bend between helices 1 and 2 [13]. The loops between helices 2 and 3 are more variable in length, have higher B factors and have the ability to form polar contacts with neighbouring loops (Fig. 3C). With the exception of ARM 5 in lower eukaryotes, the size of both helices and intervening loops is well conserved.

3.6. The concave surface of the CTNNB1 core ARM domain is not essential for NLS binding

The compact core of CTNNB1 ARM domain creates a concave groove similar to karyopherin- α with a linear surface patch of negative charge [8] (Figs. 4A and 1D). In karyopherin- α s, both the N-terminal [2–4] and C-terminal [7,8] ARM repeats are directly involved in binding to bipartite NLSs like nucleoplasmin, whereas the major N-terminal site alone is used in binding the monopartite SV40 large T antigen or Myc NLSs [8,20,21]. Like karyopherin- α s, CTNNB1 also shows selective binding to positively-charged NLSs, exhibiting strong binding to the Prp31 NLS or NLS3 of CDC5L but with no binding to the SV40 NLS or to NLS4 of CDC5L (Fig. 4B). The specific interaction observed by ITC produce complex binding isotherms best described using a two independent site model with K_d 's of ~ 0.1 and 5 μ M, as we have previously reported [4].

In order to probe the NLS binding mode of CTNNB1, we generated multiple mutants of CTNNB1 Δ 76 based on the crystal structure (Table 2 and Fig. 4A). Deletion of the C-terminal wing-like region of CTNNB1, [Δ 521–563], did not affect NLS binding (Fig. 4C) nor the stability or overall folding of the pro-

Table 2
Description of CTNNB1 Δ 76 mutants.

Mutation ^a	Location
WT	CTNNB1 Δ 76
K2: D153 N154 T155 D156	Pre-ARM
K4: D170 D172	Pre-ARM ^b
K6: Q166	Pre-ARM
K5: <u>H216</u> <u>N217</u>	ARM1 Helix3
K7: <u>E223</u> <u>N224</u> <u>Y261</u>	ARM1 & ARM2 Helix3
K8: <u>E264</u> <u>I268</u>	ARM2 Helix3
K9: <u>D315</u> <u>S319</u> <u>H359</u>	ARM3 & ARM4 Helix3
K3: L321 <u>M322</u> <u>H359</u> I362	ARM3 & ARM4 Helix3
Deletion [Δ 521–563]	CTD

Underlined are residues structurally equivalent to the NLS binding WxxxN motif of karyopherin- α as depicted in Fig. 4A.

^a Substitutions to alanine.

^b Residues affecting the acidic pocket in helix 5 of the Pre-ARM.

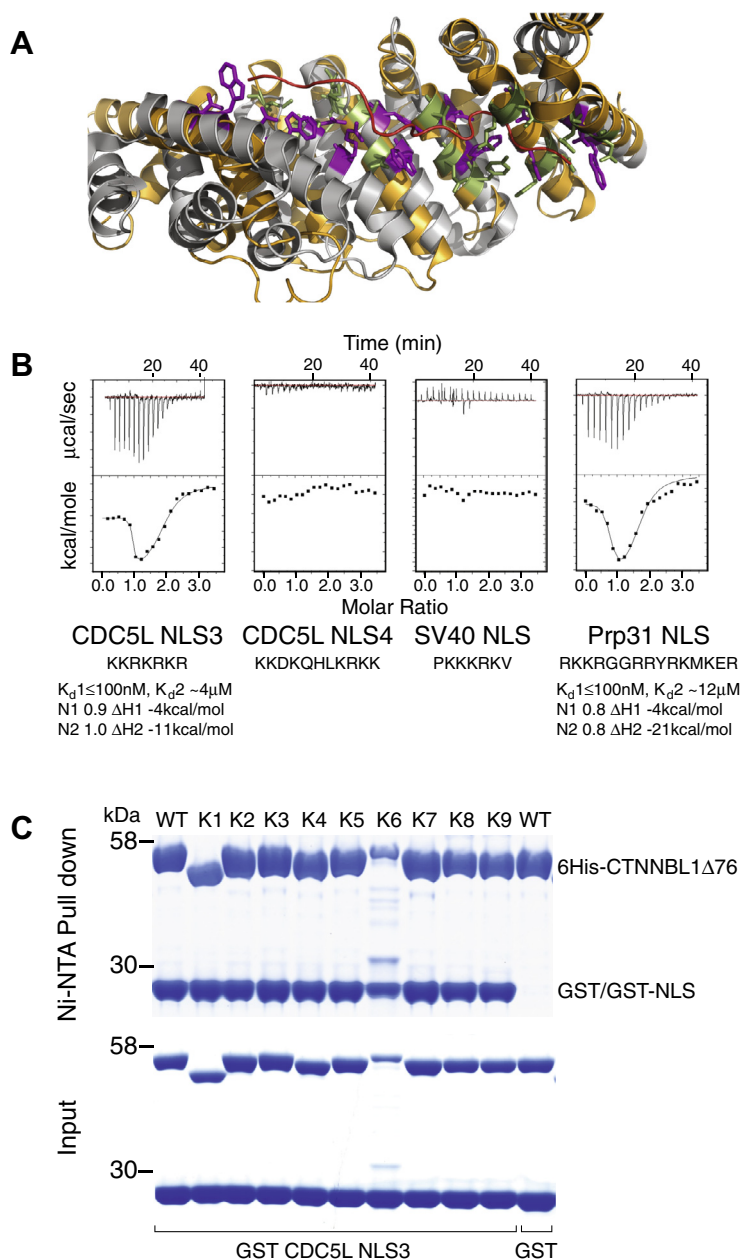


Fig. 4. CTNNB1 binding to NLS peptides. (A) Overlay of mouse karyopherin- α (grey) bound to the NLS of Nucleoplasmin (red) (PDB file 1EJY, [21]) superimposed on the structure of CTNNB1 (orange). Conserved tryptophans and arginines in karyopherin- α that mediate NLS binding are shown in purple with equivalent residues in CTNNB1 in green (residues 519–563 and part of the CTD are excluded for clarity). (B) Selective binding of CTNNB1 Δ 76 to basic peptides of CDC5L-NLS3 and Prp31 compared to GST-CDC5L-NLS3 fusion proteins but not GST alone. Bound proteins were visualized with Coomassie blue following SDS-PAGE. (C) 6His-CTNNB1 Δ 76 and mutants (as in Table 2) interact with GST-CDC5L-NLS3 fusion proteins but not GST alone. Bound proteins were visualized with Coomassie blue following SDS-PAGE.

tein as determined by Circular dichroism (CD) spectroscopy (not shown). Multiple mutants of CTNNB1 were also generated that carried amino acid substitutions in the putative NLS recognition groove (Fig. 4A). Strikingly, all mutants retained the ability to interact with the CDC5L NLS3 as judged by a pull-down assay (Fig. 4C). Indeed, not only was NLS binding retained by these mutants, but the affinity of binding to the CDC5L NLS3 (as well as to the Prp31 NLS) peptide, as monitored by ITC, was little affected by these amino acid changes; the small differences in binding affinity, such as the deletion mutant K1, might reflect a minor influence on the stability of the protein (Supplementary Fig. 5). Thus, CTNNB1 likely recognizes NLS ligands in a manner very different from karyopherin- α .

4. Conclusions

CTNNB1 has a highly compact ARM domain with a novel and rigid C-terminal coiled-coil domain. The N-terminus contains a disordered region and a structured HEAT-like region that includes a functional NES. Despite structural similarities with karyopherin- α , our data suggest a very different NLS binding mode that does not seem to rely on the concave groove formed by the ARM domain. These functional data are in agreement with the lack of conservation of the key tryptophan and asparagine residues that line the groove formed by the ARM helix 3s of karyopherin- α (WxxxN motif, see Fig. 4A and Fig. S4 [8,20,21]) which mediate NLS binding through hydrophobic interactions to the basic side chains and

hydrogen bonding to the main chains of the peptide. The lack of amino acid sequence conservation between karyopherin- α and CTNNB1 therefore reflects a distinct NLS binding mode and may account for the selectivity of the NLS interactions of CTNNB1.

The abbreviated ARM domain of CTNNB1 also precludes space for a potential minor groove as in karyopherins. However, the presence of other patches of negative electrostatic potential, including the extended C terminal helix coil-coil region or the acidic patch between the HEAT-like region and the core ARM domain, opens the possibility of multiple binding sites for NLS containing proteins (such as the NLS region of CDC5L). Indeed, a model docking of one of the interacting partners of CTNNB1, AID, into the independently determined structure of CTNNB1 (reported by Huang et al. while this manuscript was in preparation [22]), suggests that the CTD region might be involved in binding to NLSs. While their model invokes D315, E264 and E223 for the interaction with AID these residues are not required for binding to the spliceosome NLSs. Further functional, structural and biochemical analysis should clarify the distinctive nature of the binding of CTNNB1 to its NLS-containing partners.

PDB references

CTNNB1 Δ 76A 4cb8, CTNNB1 Δ 76B 4cba and CTNNB1 4cb9.

Acknowledgments

Thanks to Alan Warren and Anita Chandra for advice and discussions and to the staff scientist at DSL beamlines I02 and I04-1 for their help. We are grateful for the support from the Leukaemia & Lymphoma Research to F.v.M. and Lucy Cavendish College, Cambridge to S.B.T. This work was supported by the Medical Research Council [MRC reference number MC_U105178806].

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.febslet.2013.11.013>.

References

- [1] Jabbar, L., Welter, J.F., Kollar, J. and Hering, T.M. (2003 Mar) Sequence, gene structure, and expression pattern of CTNNB1, a minor-class intron-containing gene-evidence for a role in apoptosis. *Genomics* 81 (3), 292–303.
- [2] Makarova, O.V., Makarov, E.M., Urlaub, H., Will, C.L., Gentzel, M., Wilm, M. and Lührmann, R. (2004) A subset of human 35S U5 proteins, including Prp19, function prior to catalytic step 1 of splicing. *EMBO J.* 23 (12), 2381–2391.
- [3] Conticello, S.G., Ganesh, K., Xue, K., Lu, M., Rada, C. and Neuberger, M.S. (2008) Interaction between antibody-diversification enzyme AID and spliceosome-associated factor CTNNB1. *Mol. Cell* 31 (4), 474–484.
- [4] Ganesh, K., Adam, S., Taylor, B., Simpson, P., Rada, C. and Neuberger, M. (2011) CTNNB1 is a novel nuclear localization sequence-binding protein that recognizes RNA-splicing factors CDC5L and Prp31. *J. Biol. Chem.* 286 (19), 17091–17102.
- [5] Chandra, A., van Maldegem, F., Andrews, S., Neuberger, M.S. and Rada, C. (2013) Deficiency in spliceosome-associated factor CTNNB1 does not affect ongoing cell cycling but delays exit from quiescence and results in embryonic lethality in mice. *Cell Cycle* 12 (5), 732–742.
- [6] Huber, A.H., Nelson, W.J. and Weis, W.I. (1997) Three-dimensional structure of the armadillo repeat region of beta-catenin. *Cell* 90 (5), 871–882.
- [7] Gill, D.J., Teo, H., Sun, J., Perisic, O., Vepntsev, D.B., Emr, S.D. and Williams, R.L. (2007) Structural insight into the ESCRT-1/II link and its role in MVB trafficking. *EMBO J.* 26 (2), 600–612.
- [8] Conti, E., Uy, M., Leighton, L., Blobel, G. and Kuriyan, J. (1998) Crystallographic analysis of the recognition of a nuclear localization signal by the nuclear import factor karyopherin alpha. *Cell* 94 (2), 193–204.
- [9] Graham, T.A., Weaver, C., Mao, F., Kimelman, D. and Xu, W. (2000) Crystal structure of a beta-catenin/Tcf complex. *Cell* 103 (6), 885–896.
- [10] Huber, A.H. and Weis, W.I. (2001) The structure of the beta-catenin/E-cadherin complex and the molecular basis of diverse ligand recognition by beta-catenin. *Cell* 105 (3), 391–402.
- [11] Groves, M.R. and Barford, D. (1999) Topological characteristics of helical repeat proteins. *Curr. Opin. Struct. Biol.* 9 (3), 383–389.
- [12] Kippert, F. and Gerloff, D.L. (2009) Highly sensitive detection of individual HEAT and ARM repeats with HHpred and COACH. *PLoS One* 4 (9), e7148.
- [13] Andrade, M.A., Petosa, C., O'Donoghue, S.I., Müller, C.W. and Bork, P. (2001) Comparison of ARM and HEAT protein repeats. *J. Mol. Biol.* 309 (1), 1–18.
- [14] Andrade, M.A. and Bork, P. (1995) HEAT repeats in the Huntington's disease protein. *Nat. Genet.* 11 (2), 115–116.
- [15] Groves, M.R., Hanlon, N., Turowski, P., Hemmings, B.A. and Barford, D. (1999) The structure of the protein phosphatase 2A PR65/A subunit reveals the conformation of its 15 tandemly repeated HEAT motifs. *Cell* 96 (1), 99–110.
- [16] Kajava, A.V., Gorbea, C., Ortega, J., Rechsteiner, M. and Steven, A.C. (2004) New HEAT-like repeat motifs in proteins regulating proteasome structure and function. *J. Struct. Biol.* 146 (3), 425–430.
- [17] Dong, X., Biswas, A., Süel, K.E., Jackson, L.K., Martinez, R., Gu, H. and Chook, Y.M. (2009) Structural basis for leucine-rich nuclear export signal recognition by CRM1. *Nature* 458 (7242), 1136–1141.
- [18] la Cour, T., Kiemer, L., Mølgaard, A., Gupta, R., Skriver, K. and Brunak, S. (2004) Analysis and prediction of leucine-rich nuclear export signals. *Protein Eng. Des. Sel.* 17 (6), 527–536.
- [19] Landschulz, W.H., Johnson, P.F. and McKnight, S.L. (1988) The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. *Science* 240 (4860), 1759–1764.
- [20] Conti, E. and Kuriyan, J. (2000) Crystallographic analysis of the specific yet versatile recognition of distinct nuclear localization signals by karyopherin alpha. *Structure* 8 (3), 329–338.
- [21] Fontes, M.R., Teh, T. and Kobe, B. (2000) Structural basis of recognition of monopartite and bipartite nuclear localization sequences by mammalian importin-alpha. *J. Mol. Biol.* 297 (5), 1183–1194.
- [22] Huang, X., Wang, G., Wu, Y. and Du, Z. (2013) The structure of full-length human CTNNB1 reveals a distinct member of the armadillo-repeat protein family. *Acta Crystallogr. D Biol. Crystallogr.* 69 (Pt 8), 1598–1608.