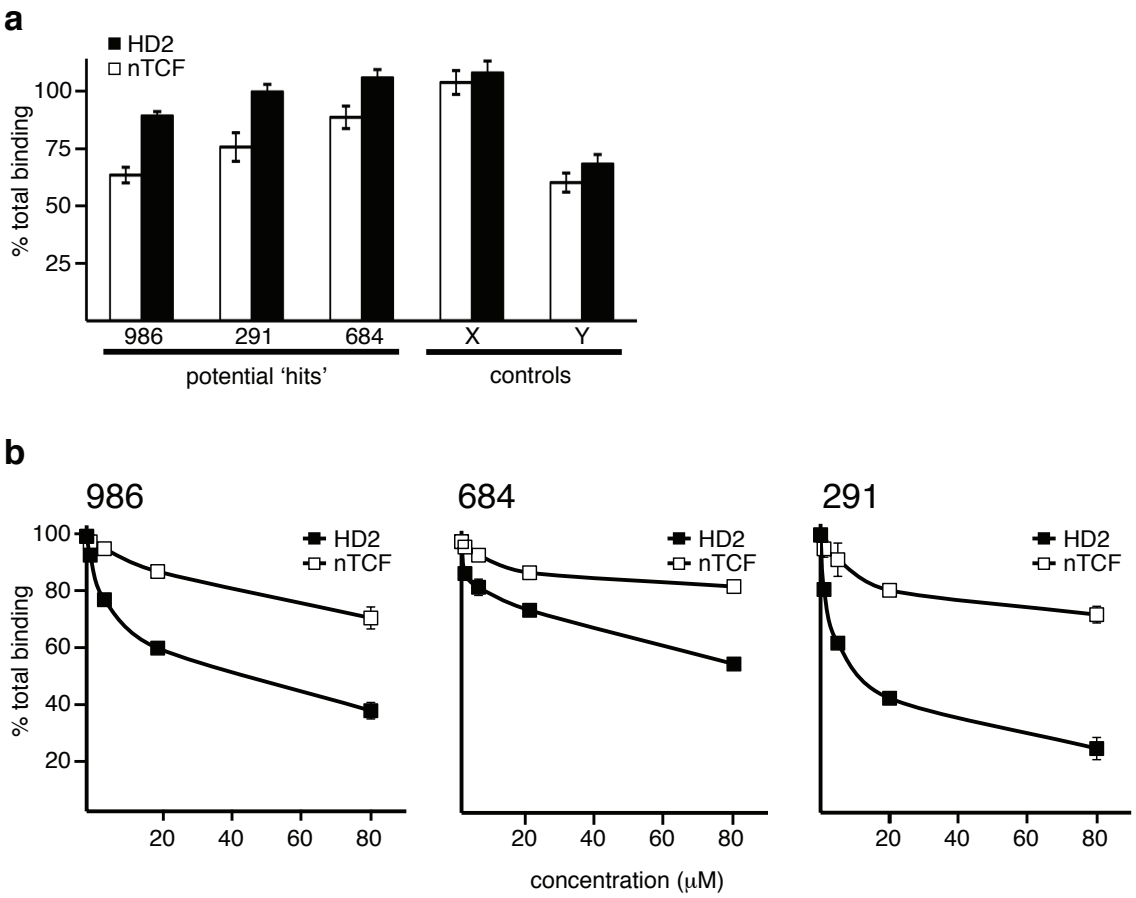
SUPPLEMENTARY INFORMATION

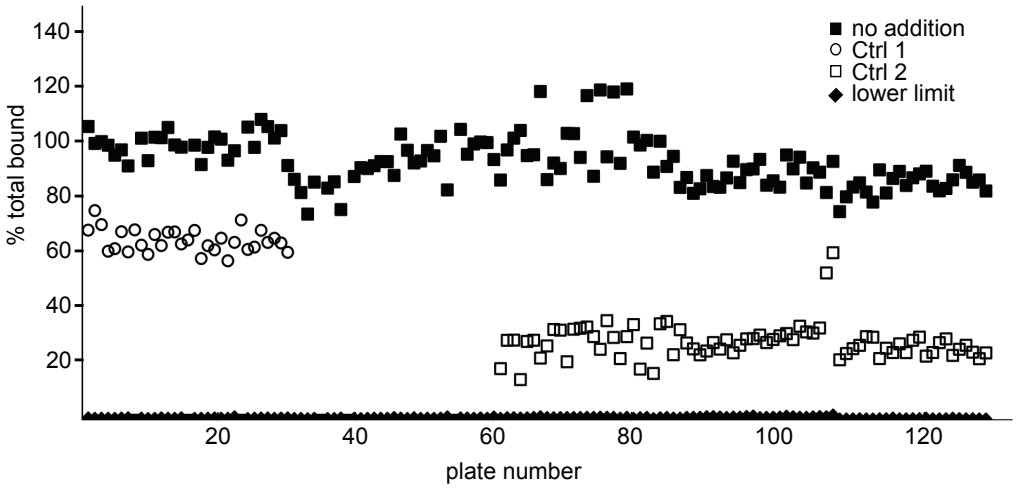
Supplementary Figure S1



***In vitro* inhibitor activities of Phytopure hits**

(a) Inhibition values (percent ARD-HD2 binding after incubation of compound relative to total binding without compound) for 20 μM 986, 291 and 684 in ELISA assays; two compounds (X and Y) are shown as (X) negative in the primary screen, or (Y) positive in both screens; error bars, standard deviations (n=4). **(b)** ELISA assays, showing dose-dependent inhibitions of the ARD-HD2 and ARD-nTCF interactions by 20 μM of compound (986, 291 or 684).

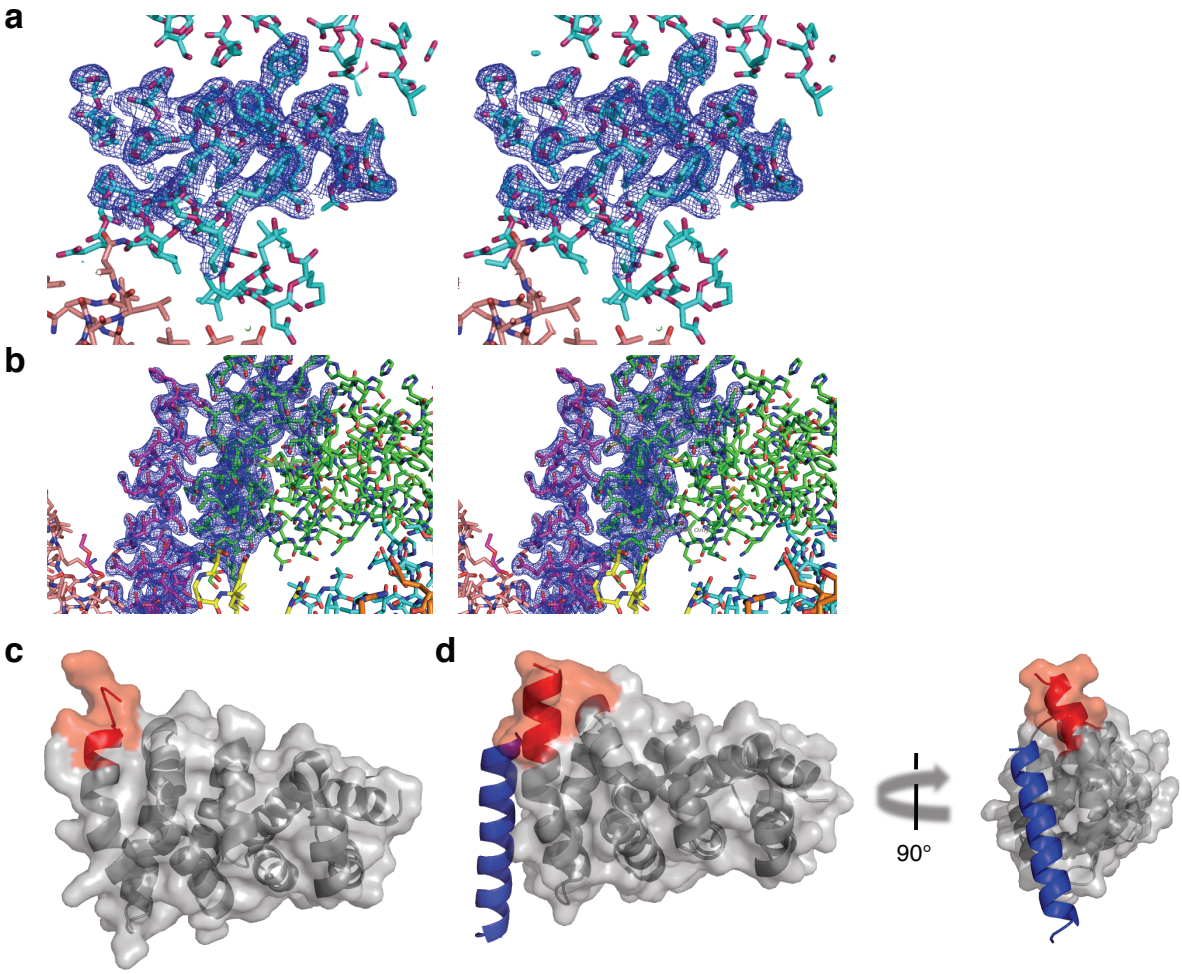
Supplementary Figure S2



Medium throughput screen of the MRCT library

Scan of 135 different 384-well plates, showing upper limit control values derived from wells with no addition of compound (■), and lower limit values (◆) derived from wells coated with GST instead of GST-ARD (unable to bind His-HD2). Two control (Ctrl) compounds from the LOPAC screen (Ctrl 1, aurintricarboxylic acid; Ctrl 2, 6-hydroxy DL DOPA, also known as 2,5-dihydroxy-DL-tyrosine) that scored in the primary and counter screens were used as internal controls in the majority of plates and at 20 μM yielded 30% (Ctrl 1, ○) and 65% (Ctrl 2, □) inhibition of the HD2-ARD interaction, respectively. For automation of this screen, a Biomek Fx automated liquid handling robot (Beckman) coupled to a Biotek plate washer was used. Volumes of protein and compound were 20 μl, and starting concentrations of GST-proteins were 2 μg. Detection of the HRP signal was luminescent-based, developed with 25 μl of POD reagent (Pierce), and quantified on a PHERAstar fluorescence plate reader (BMG Labtech). A hit was defined as a compound that inhibited binding by >60% or >3 standard deviations of the DMSO control, whichever the greater in two independent tests. Hits were re-tested in the counter-screen as described.

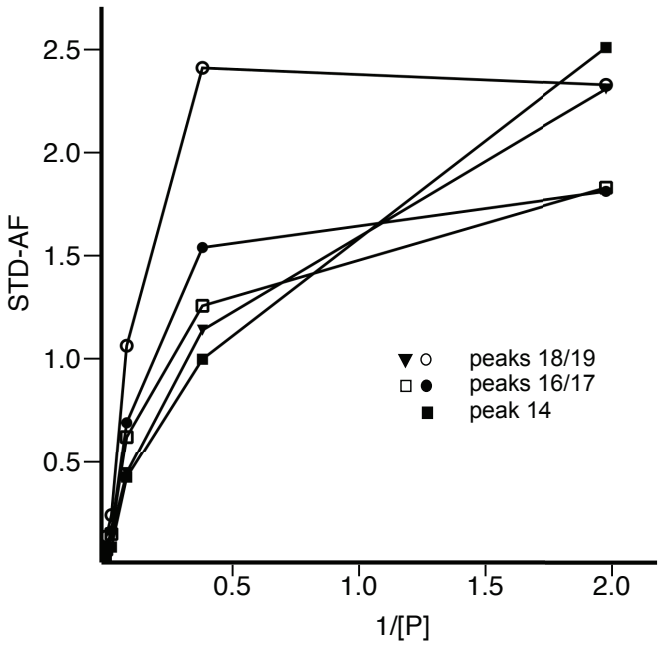
Supplementary Figure S3



X-ray structures of R4 and R4-HD2

(**a, b**) Stereo images of (**a**) 3SLA (R4 structure) and (**b**) 3SL9 (R4-HD2 structure), with electron density maps contoured at 1.0σ, around the selection site, within 1.6 Å of selected atoms based on the 2Fo-Fc map. (**c, d**) Molecular surfaces and superimposed ribbon representations of **(c)** R4 and **(d)** R4-HD2 (images created with PyMOL); grey, R4 (with H1 in red); blue, HD2. Note the HD2-induced conformational change of H1 which is (**c**) largely unstructured in the apo form (i.e. there is no electron density for H1 amino acid side chains, and the observed backbone-derived electron densities are due to artificial crystal contacts), but adopts a helical structure upon HD2 binding (**d**), similar to that previously observed in the ARD-HD2-nTCF complex [18].

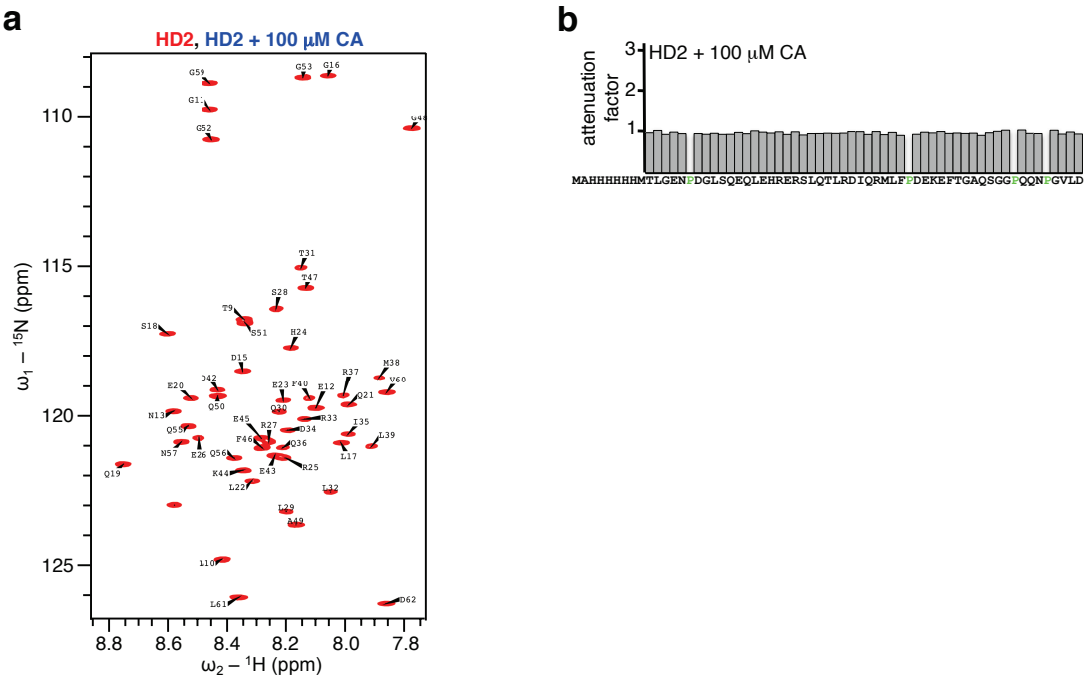
Supplementary Figure S4



Quantitation of CA binding to R4 monitored by STD

Binding isotherms for STD amplification factors versus 1/[P], where STD-AF is defined as $(I_o - I_{STD})/I_o) * ([L]/[P])$ [20]; I_o and I_{sat}, NMR peak heights of individual CA protons (H14, H16-H19; see main **Figure 2d**) measured after 7 seconds off-resonance and on-resonance saturating irradiation, respectively; [P], total protein concentration (0.5-50μM); [L], total ligand concentration (fixed at 50μM); the prochiral methyl group proton resonances of H16/H17 and H18/H19 were not assigned stereospecifically. The plateau of STD-AF at low protein concentrations and the decreasing STD-AF values at higher concentrations indicate that the binding of R4 to CA is saturable.

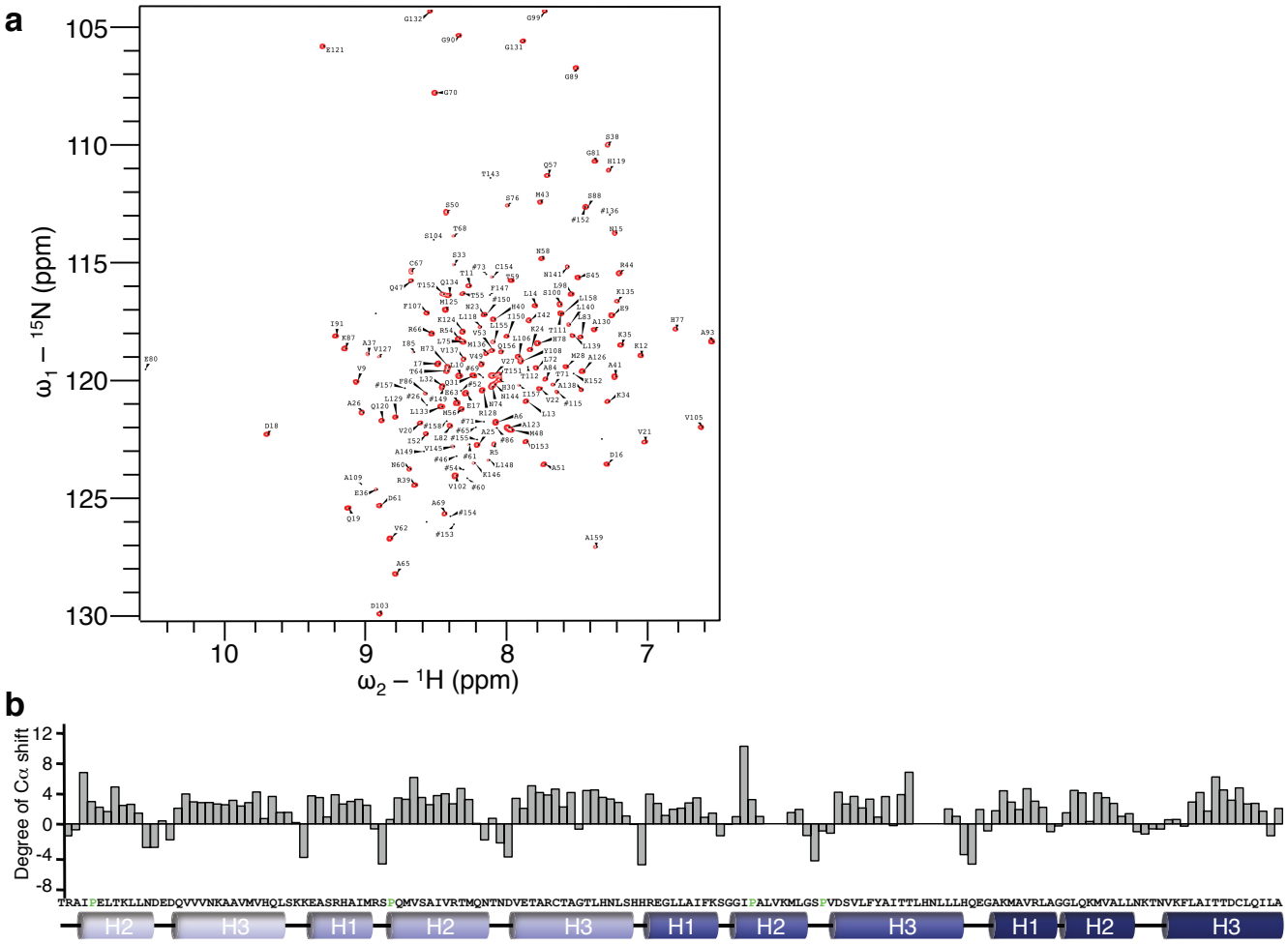
Supplementary Figure S5



The HSQC of HD2 is unperturbed by CA

(a) Overlay of assigned HSQC of ^{15}N-labelled HD2+CA (100 μM each) onto ^{15}N-labelled HD2 (100 μM) in solvent only, as indicated. Assignment of 49 non-proline HD2 residues (plus one residue from the N-terminal tag) was obtained for ^{13}C/^{15}N double-labelled protein at 20˚C, 500 MHz ^{1}H (see also **Supplementary Figure S6**). **(b)** Map of CA-induced relative intensity changes derived from (**a**) onto individual HD2 residues; attenuation factors were calculated as in main **Figure 3**; green, prolines.

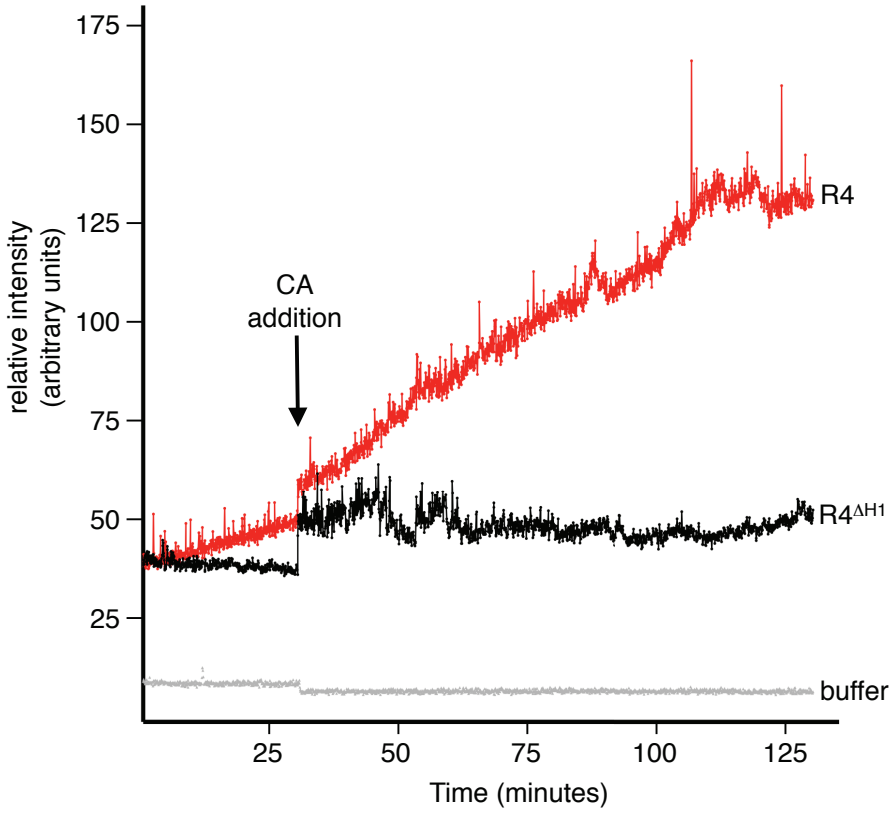
Supplementary Figure S6



Assigned HSQC of R4$^{\Delta H1}$ and secondary structure elements

(**a**) Resonance assignments of 138 non-proline residues were obtained with 300 µM $^{2}H/^{13}C/^{15}N$ triple-labelled R4$^{\Delta H1}$ at 15˚C 600 MHz ^{1}H, using standard TROSY-based triple resonance correlation spectra (trHNCA, trHN(CO)CA, trHN(CA)CB, trHN(COCA)CB, trHNCO & trHN(CA)CO, with unmodified Bruker pulse programs); the remaining 13 residues (in the third repeat, dark blue helices in main **Figure 3c**) remained unassigned due to localised inefficient magnetisation transfer. (**b**) Correlation of resonance assignments with secondary structure likelihood, based on TALOS+ [21], with R4 sequence (underneath) and its α-helices as observed in the crystal structure (see **Supplementary Figure S3**); Cα, α-carbon.

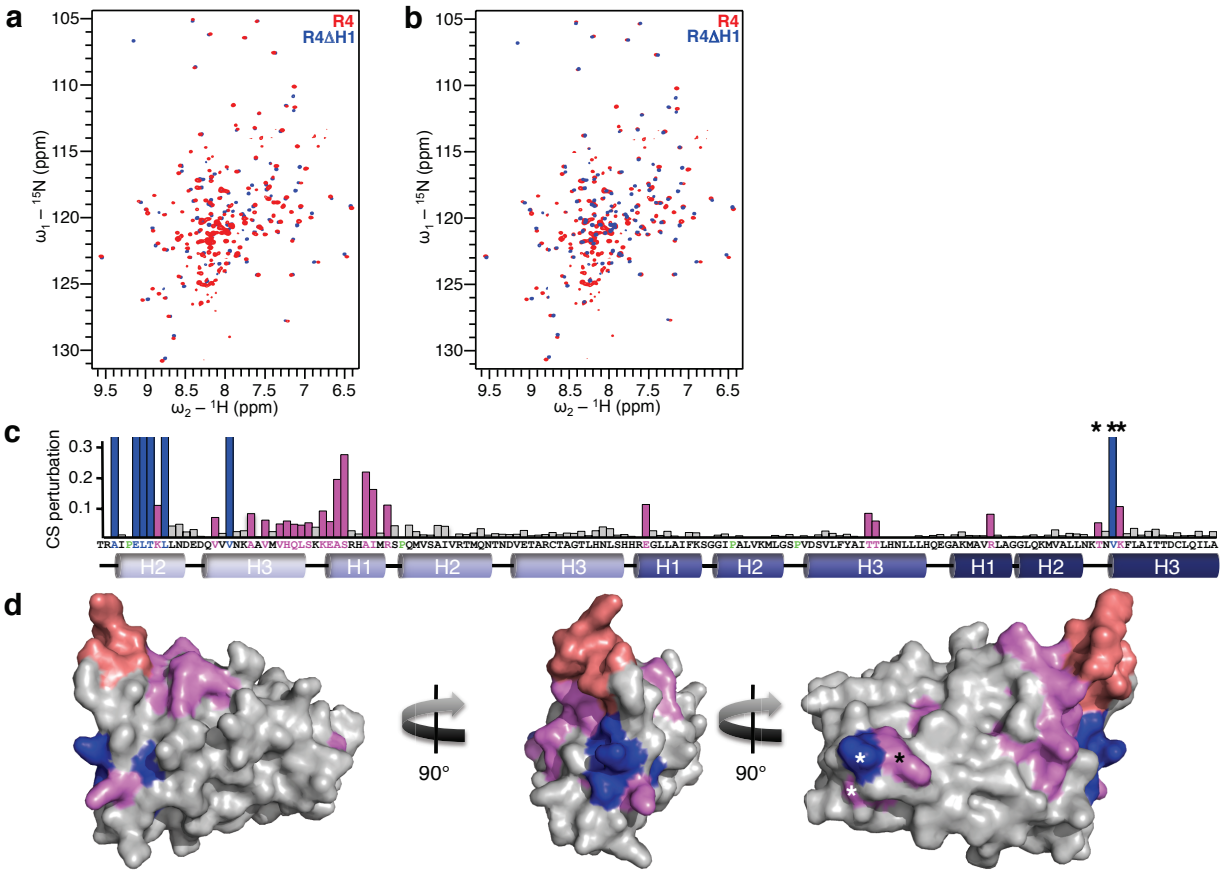
Supplementary Figure S7



CA induces aggregation of R4 but not R4$^{\Delta H1}$

Kinetics of emission intensity changes of scattered light from samples of 25 μM R4 (red) or R4$^{\Delta H1}$ (black), or from buffer alone (grey), based on excitation and emission wavelengths of 350 nM. Right-angle light scattering was recorded (with a Cary Eclipse fluorescence spectrophotometer) for 100 minutes after addition of 25 μM CA, following establishment of baseline light scattering for 30 minutes at 25°C.

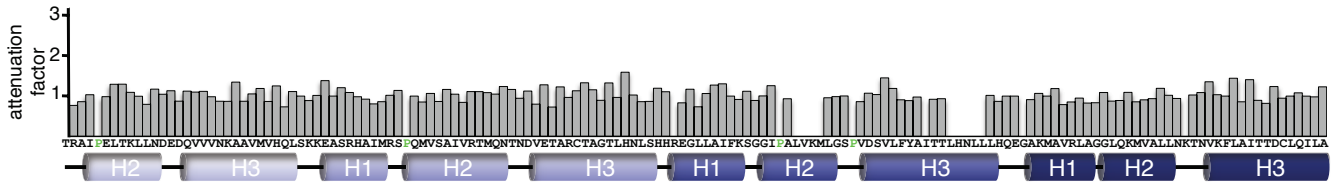
Supplementary Figure S8



HSQC spectral similarity of R4 and R4$^{\Delta H1}$

(a, b) Overlay of TROSY spectra of ^{15}N-labelled **(a)** R4 (red) onto R4$^{\Delta H1}$ (blue; 100 µM each) and (**b)** inverted display. **(c)** Weighted chemical shift (CS) perturbations, indicating shifting of >0.05 ppm (purple) or line broadening (blue) of residues along the amino acid sequence (secondary structure elements underneath). Lack of a bar, proline (green) or unassigned residue (in the third repeat). **(d)** Projection of CS perturbations onto the R4 structure (colour coding as in **c**); the majority of these cluster in repeat 1 and in the first α-helix of repeat 2, which is structurally adjacent to H1 (red), indicative of perturbations experienced by this region by the presence of H1. A small surface-exposed patch (asterisks) revealed by chemical shifts of residues in the third α-helix of repeat 4 suggests that this patch may participate in R4 oligomerisation or aggregation (not observed with R4$^{\Delta H1}$). The prominent valine in this patch (blue) is also solvent-exposed in the structure of the intact ARD [19].

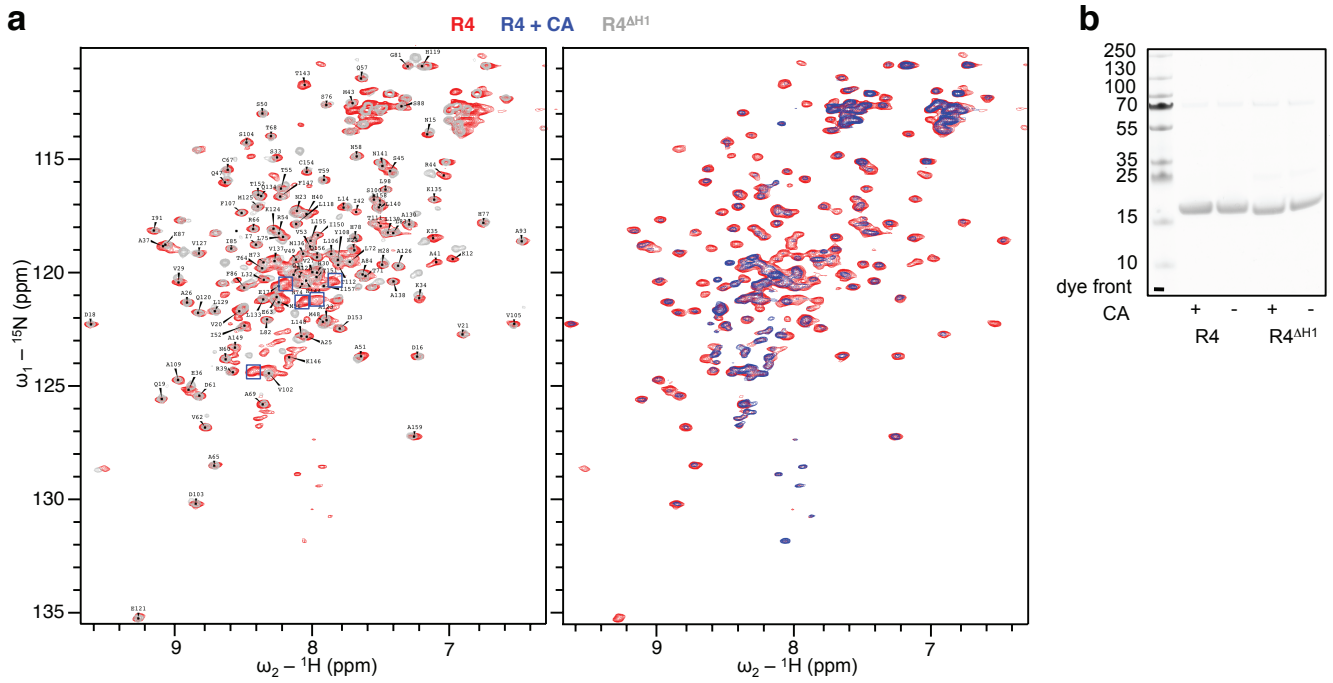
Supplementary Figure S9



R4$^{\Delta H1}$ fails to respond to CA

Attenuation factors (calculated as described in main **Figure 3d**) derived from the HSQC spectral overlays shown in main **Figure 4a** (100 μM ^{15}N-labelled R4$^{\Delta H1}$ +/- 100 μM CA) with regard to individual R4 residues; colour coding as in main **Figure 3d**. No significant changes were observed (attenuation factors ~1 throughout the spectrum).

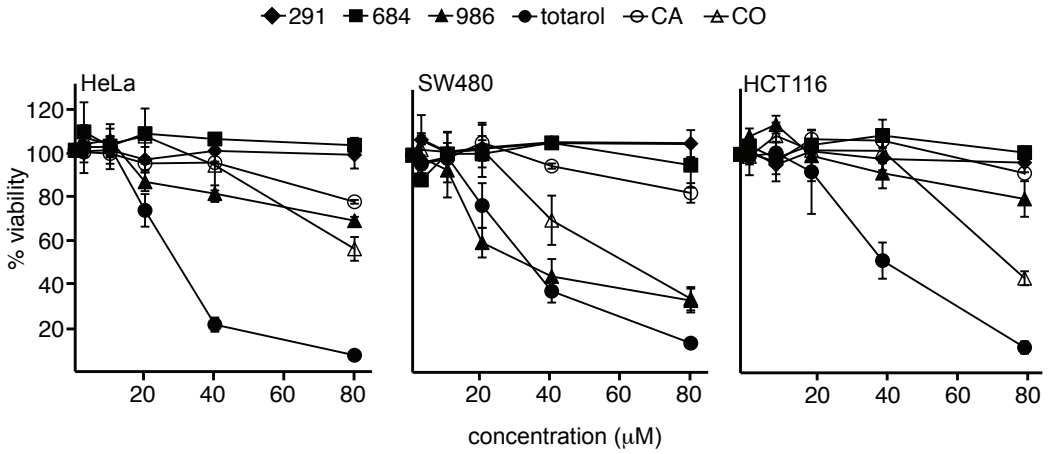
Supplementary Figure S10



Putative H1 residues responding to CA

(a) Double and triple overlays of HSQCs from ^{15}N-labelled R4$^{\Delta H1}$ (shown in main **Figure 4a**, red), R4 and R4+CA (shown in main **Figure 3c**); left-hand panel, grey on red (with annotations of assigned peaks; see **Supplementary Figure S6**); right-hand panel, grey on blue on red; boxed are major unassigned R4-specific peaks (red, but not grey) that are candidates for poorly structured H1 residues (as they appear in the 8.0-8.6 ppm range). A subset of these experience CA-induced line broadenings or slight chemical shift perturbations (compare red and blue in right-hand panel), consistent with direct interactions of these putative H1 residues with CA. The high similarity of the three spectra indicates that the overall R4 fold remains the same after CA addition, arguing against proteolysis as a major factor in the CA response of R4. **(b)** SDS-PAGE of ^{15}N-labelled R4 and R4$^{\Delta H1}$, +/- CA, confirming absence of detectable proteolysis from the NMR samples (faint bands at ~70 kDa represent minor impurities); molecular masses (kDa) are indicated by numbers on the left.

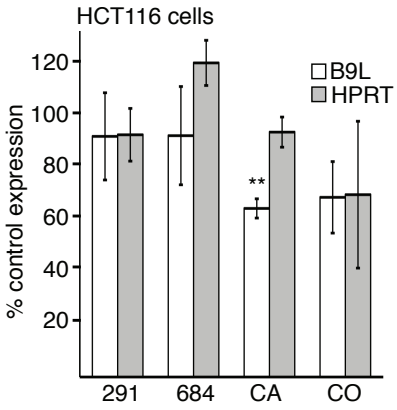
Supplementary Figure S11



Cytotoxicity profiles

HeLa, SW480 and HCT116 cells were seeded in 96-well plates at densities of 5 x 10^3 cells/well and allowed to adhere for 16 hours prior to addition of compounds (at various concentrations, as indicated). After 12 hours, cell numbers were determined with the Cell Counting Kit-8 (Dojindo) according to manufacturer's recommendations, and percent viable cells are shown. Compounds 291, 684 and CA remained non-toxic upon treatment of cells (at 80 μM) for 24 hours; error bars, standard deviations (n=4).

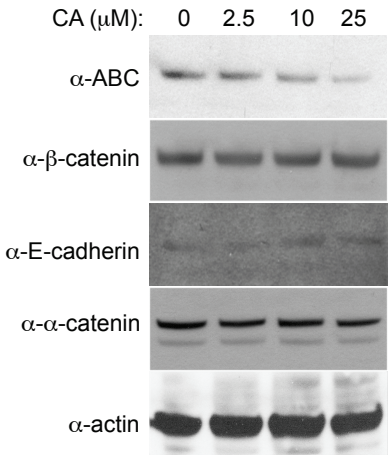
Supplementary Figure S12



Inhibition of *B9L* Wnt target gene expression in HCT116 cells

RT-qPCR analysis of *B9L* and *HPRT* transcript levels, after treatment of HCT116 cells with 25 μM of compounds, as indicated. Reduction of transcript levels relative to DMSO-treated controls are shown at the P<0.001 (**) significance level; error bars, standard deviations (n=4). *AXIN2* expression was not monitored as this gene is known to be hypermethylated in this cell line [54].

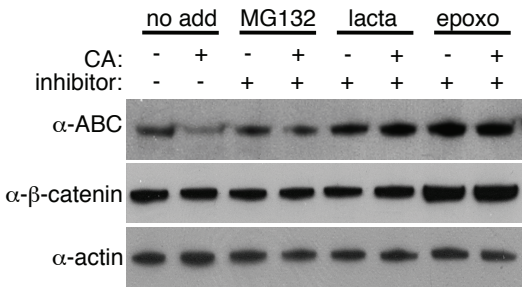
Supplementary Figure S13



Dose-dependent reduction of active β-catenin levels by CA

Western blots of total lysates from SW480 cells exposed for 8 hours to increasing concentrations of CA, probed with antibodies as indicated.

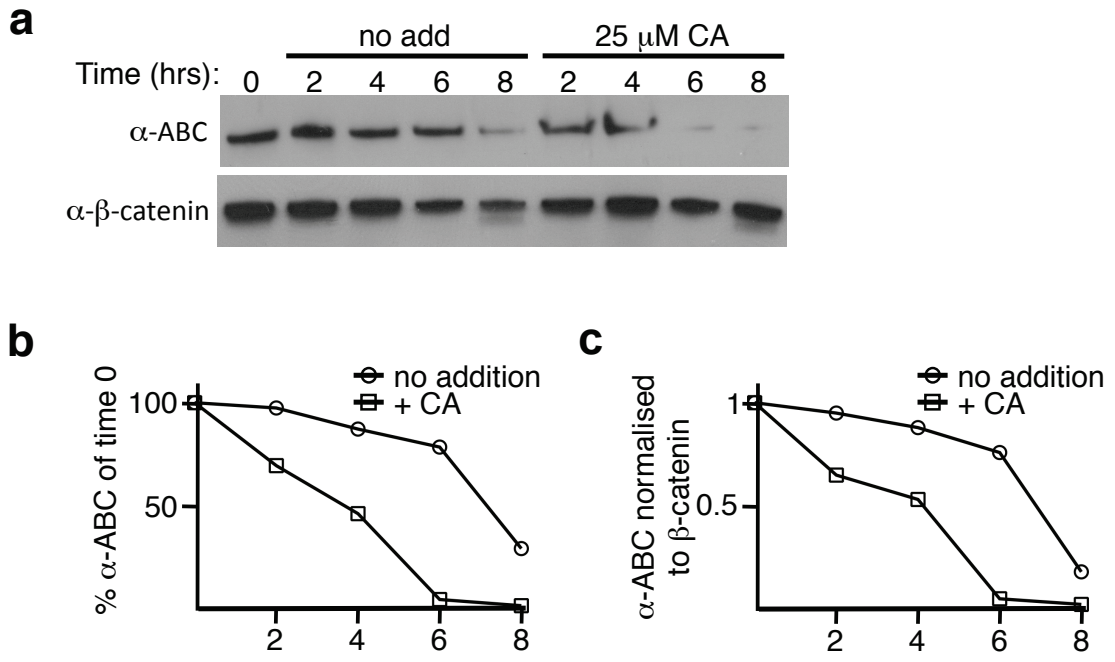
Supplementary Figure S14



Proteasome inhibitors block β-catenin destabilisation by CA

Western blots of total lysates from SW480 cells exposed for 4 hours simultaneously to either DMSO vehicle or 25 μM CA in the presence of 10 μM proteasome inhibitors (lacta, lactacystin; epoxo, epoxomycin), as indicated; Western blots are probed with antibodies as indicated; no add, no addition.

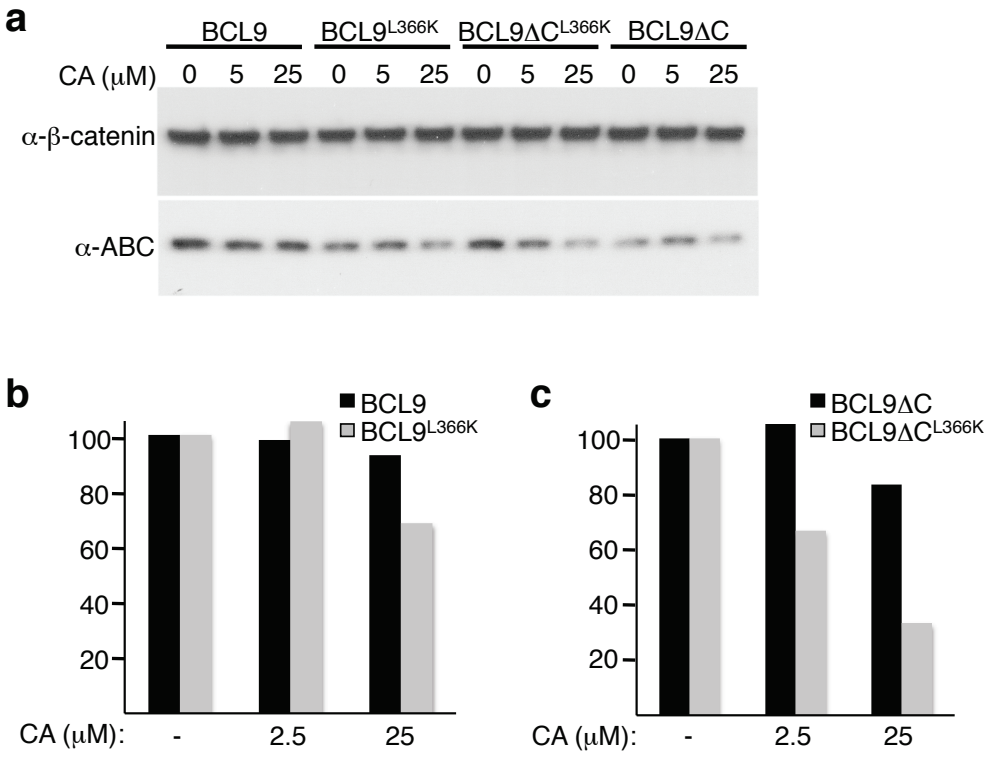
Supplementary Figure S15



CA reduces the half-life of active β-catenin in SW480 cells

(a) Western blots of total lysates from SW480 cells in the presence of 100 μg/ml cycloheximide, mock-treated or treated with 25 μM CA, and probed with antibodies as indicated; no add, no addition. **(b, c)** Densitometry analysis of lanes 2-9 as a percentage of **(b)** total levels of unphosphorylated β-catenin (α-ABC, lane 1 in **a**) at time 0, or (**c**) additionally normalised to total β-catenin levels (α-β-catenin lanes in **a**).

Supplementary Figure S16



BCL9 antagonises the reduction of active β-catenin by CA

(a) Western blots of total lysates from SW480 cells, transfected with wt and mutant FLAG-BCL9 constructs as indicated, treated with increasing concentrations of CA for 8 hours, and probed with antibodies as indicated. **(b, c)** Densitometry analysis of **(b)** lanes 1-6 and **(c)** lanes 7-12; black columns, total β-catenin levels; grey columns, unphosphorylated β-catenin levels.

Supplementary Table S1

statistics	LOPAC screen	Phytopure screen	MRCT Screen
average Z	0.35	0.32	0.70
# of compounds	1280	1280	45'000
percent hit rate	2.5% (32)	3.8% (49)	0.024% (12)
# of discarded hits	32	46	12
% confirmed hits (#)	0% (0)	0.15% (3)	0% (0)

Screen statistics

Discarded hits, hits that scored in the counter-screen; confirmed hits, hits negative in the counter-screen. Due to the low average Z' value, the LOPAC and Phytopure screens were performed twice, which identified the same initial hits. The screen of the MRCT library (containing 45'000 drug-like compounds of <350 MW, <3 H-bond donors, <6 H-bond acceptors, <6 rotatable bonds, cLogP of 1-4 and a polar surface area <120) was performed with a luminescent-based detection method (see **Supplementary Figure S2**), which greatly improved the average Z' value.

Supplementary Table S2

Compound ID	Ki (μM)	
	ARD-HD2	ARD-nTCF
986	4.8 ± 1.0	90 ± 22
291	0.9 ± 0.8	n.d.
684	6.3 ± 2.7	n.d.
CO	8.2 ± 4.3	23 ± 15
CA	3.3 ± 1.8	n.d.

Inhibitor constants

K_i values for inhibition of the ARD-HD2 and ARD-nTCF interactions by the compounds used in this study, with standard deviations (±) indicated; n.d., not detected.

Supplementary Table S3

protein	Kd (μM)	
β-catenin^{ARD}	0.66 ± 0.16	* 0.465 ± 0.035
R4	0.59 ± 0.15	
R4$^{\Delta H1}$	0.35 ± 0.17	

Binding constants of HisS-HD2 to various ARD constructs

K_d values for the binding of fluorescently-labelled HisS-HD2 to ARD, R4 and R4$^{\Delta H1}$, as derived from fluorescence anisotropy measurements, with standard deviations (±) indicated. HisS-HD2 (500 μM) was N-terminally labelled with fluorescein by incubation with 100-fold molar excess of carboxyfluorescein succinimidyl ester reagent in PBS (pH 6.8). After 1 hour, the fluorescently labelled HisS-HD2 was extensively dialysed against PBS (pH 6.8) to remove free label. Fluorescence anisotropy measurements were performed at 20°C with a Cary Eclipse Varian fluorescence spectrophotometer equipped with a Hamilton Microlab M dispenser. Reactions were carried out by titration of His-R4, His-R4$^{\Delta H1}$ or ARD into 1 ml of PBS (pH 6.8) containing 20 nM fluorescently labelled HisS-HD2. * For comparison, a K_d of 0.465 ± 0.035 μM was previously determined (by ITC) for HD2 binding to ARD [18].

Supplementary Table S4

	3SLA	3SL9
Data collection		
Space group	P 41 21 2	C2
Cell dimensions		
a, *b*, *c* (A)	90.79 90.79 364.34	160.89 80.70 87.94
α, β, γ (°)	90.00 90.00 90.00	90.00 102.00 90.00
Resolution	64.51 – 2.51 (2.61 – 2.51)	32.07 – 2.2 (2.32 – 2.2)
R merge	0.08 (0.9)	0.05 (0.25)
$I/\sigma I$	11.1 (1.6)	17.1 (5.3)
Completeness (%)	99.0 (99.3)	97 (97.2)
Redundancy	5.7 (5.4)	3.8 (3.8)
Refinement		
Resolution (Å)	40.71 – 2.5	29.4 – 2.2
Number of reflections	50'801	51'571
R_{work}/R_{free}	0.20/0.26	0.20/0.26
Number of atoms		
Protein	5938	5517
Ligand/ion	0/50	0/144
Water	86	152
B-factors		
Protein	59.46	36.98
Ligand/ion	0/61.14	0/40.31
Water	63.02	43.85
R.m.s. deviations		
Bond lengths (Å)	0.02	0.02
Bond angles (°)	1.9	2.1

	3SLA	3SL9	2GL7	2BCT
3SLA	0			
3SL9	0.41	0		
2GL7	0.43	0.42	0	
2BCT	0.54	0.40	0.49	0

X-ray data collection and refinement statistics

Above, collection and refinement statistic of crystallographic data; data were processed with Mosflm [55] and scaled with Scala [56]; structures were refined with Refmac [57], and the models were updated with Coot [58]; analysis of structures was done with CCP4i programs [59]. *Below*, rmsd values for the core C_α backbones between 3SLA (R4) and 3SL9 (R4-HD2) and the corresponding segments of 2GL7 (ARD-HD2-nTCF) [18] and 2BCT (ARD) [19].

Supplementary Table S5

PDB no.	β-catenin domain	Ligand	Unresolved amino acids in β-catenin (H1, bold)	Reference
2BCT	ARD, 134-671	none	**134-149**, 550-563	Huber et al., 1997
3BCT	ARD, 134-671	none	**134-149**, 550-562	Huber et al., 1997
2Z6H	ARD + C-term, 138-781	none	**138-148**, 551-560, 692-781	Xing et al., 2008
1I7X	ARD, 134-671	E-cadherin, 577-728	134-137, 553-559, 666-670	Huber and Weis, 2001
1M1E	ARD, 134-671	hICAT, 1-81	**134-150**, 550-558	Daniels and Weis, 2002
1LUJ	ARD, 134-664	hICAT, 1-81	**134-150**, 550-562, 664-671	Graham et al., 2002
1JDH	ARD, 134-664	hTCF4, 1-53	**141-150**, 549-559	Graham et al., 2001
1JPW	hARD, 134-668	hTCF4, 8-54	**131-150**, 550-559, 663-670	Poy et al., 2001
2GL7	ARD, 138-686, Y142E	hBCL9, 347-392 and hTCF4, 1-53	137-141, 551-559, 664-686	Sampietro et al., 2006
1DOW	118-151, fused to ligand	α-catenin, 57-261	150, 151	Pokutta and Weis, 2000
1QZ7	hARD, 133-665	XAxin, 435-504	133-141	Xing et al., 2003
1G3J	ARD, 133-664	XTcf3, 1-53	133, 551-559	Graham et al., 2000
1TH1	ARD, 133-665	hAPC, 1362-1540	**133-144,** 553-558	Xing et al., 2004
1T08	ARD, 134-671	hICAT, 1-81 and hAPC, 1484-1498	**134-150**, 551-558	Ha et al., 2004
1V18	ARD, 134-671	Phospho-hAPC, 1484-1498	**134-149**, 551-558	Ha et al., 2004
1JPP	ARD, 134-671	hAPC, 1021-1035	**134-150**, 547-560, 664-671	Eklof Spink et al., 2001
2Z6G	D. rerio, 1-781	none	**138-148**, 551-560, 692-781	Xing et al., 2008
3SLA	ARD 141-306	none	**141-149**	present study
3SL9	ARD 141-306	hBCL9 347-392		present study

Structured β-catenin residues in reported X-ray structures

Listing of the published X-ray structures of vertebrate β-catenin fragments (including residue numbers), with or without ligand [18, 19, 22, 23, 34, 60-68]. Unless otherwise stated, residue numbers refer to the murine proteins; h, human; X, Xenopus. In bold are the H1 residues (plus upstream flanking residues if included) that are not visible in the structures; red indicates structures of ARD complexes in which H1 is helical. Residues 141-150, H1 of ARD repeat 1; residues 550-560, insert in ARD repeat 10; 664-783, C-terminal region.

SUPPLEMENTARY REFERENCES

[54] Koinuma, K. et al. Epigenetic silencing of AXIN2 in colorectal carcinoma with microsatellite instability. *Oncogene* **25**, 139-146 (2006).

[55] Leslie, A. G. The integration of macromolecular diffraction data. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 48-57 (2006).

[56] Evans, P. Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallog.r* **62**, 72-82 (2006).

[57] Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallog.r* **53**, 240-255 (1997).

[58] Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr*. **60**, 2126-2132 (2004).

[59] Potterton, E., Briggs, P., Turkenburg, M. & Dodson, E. A graphical user interface to the CCP4 program suite. *Acta Crystallogr. D Biol. Crystallogr.* **59**, 1131-1137 (2003).

[60] Xing, Y. et al. Crystal structure of a full-length β-catenin. *Structure* **16**, 478-487 (2008).

[61] Daniels, D. & Weis, W. I. ICAT inhibits β-catenin binding to Tcf/Lef-family transcription factors and the general coactivator p300 using independent structural modules. *Mol. Cell* **10**, 573-584 (2002).

[62] Graham, T. A., Clements, W. K., Kimelman, D. & Xu, W. The crystal structure of the β-catenin/ICAT complex reveals the inhibitory mechanism of ICAT. *Mol. Cell* **10**, 563-571 (2002).

[63] Graham, T. A. et al. Crystal structure of a β-catenin/Tcf complex. *Cell* **103**, 885-896 (2000).

[64] Graham, T. A. et al. Tcf4 can specifically recognize β-catenin using alternative conformations. *Nat. Struct. Biol.* **8**, 1048-1052 (2001).

[65] Xing, Y., Clements, W. K., Kimelman, D. & Xu, W. Crystal structure of a β-catenin/axin complex suggests a mechanism for the β-catenin destruction complex. *Genes Dev.* **17**, 2753-2764 (2003).

[66] Xing, Y. et al. Crystal structure of a β-catenin/APC complex reveals a critical role for APC phosphorylation in APC function. *Mol. Cell* **15**, 523-533 (2004).

[67] Ha, N. C. et al. Mechanism of phosphorylation-dependent binding of APC to β-catenin and its role in β-catenin degradation. *Mol. Cell* **15**, 511-521 (2004).

[68] Eklof Spink, K., Fridman, S. G. & Weis, W. I. Molecular mechanisms of β-catenin recognition by adenomatous polyposis coli revealed by the structure of an APC-β-catenin complex. *EMBO J.* **20**, 6203-6212 (2001).