Specific recognition of ZNF217 and other zinc-finger proteins at a surface groove of CtBPs

Kate G.R. Quinlan[1], Marco Nardini[2], Alexis Verger[1], Pierangelo Francescato[3], Paul Yaswen[4], Daniela Corda[5], Martino Bolognesi[2] and Merlin Crossley[1]*

[1]*School of Molecular and Microbial Biosciences, G08, University of Sydney, NSW 2006, Australia*

[2]*Department of Biomolecular Sciences and Biotechnology and CNR-INFM, University of Milano, I-20131 Milano, Italy*

[3]*Department of Industrial and Organic Chemistry, University of Milano, I-20131 Milano, Italy*

[4]*Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*

[5]*Department of Cell Biology and Oncology, Consorzio Mario Negri Sud, I-66030 Santa Maria Imbaro, Chieti, Italy*

*Corresponding author: Merlin Crossley
School of Molecular and Microbial Biosciences, G08
University of Sydney
NSW 2006, Australia
Phone: 61-2-9351 2233.
Fax: 61-2-9351 4726.
E-mail address: m.crossley@mmb.usyd.edu.au

Running title: Characterization of a new CtBP interaction motif

Word count (Materials and Methods): 2081

Word count (Introduction, Results and Discussion): 4105

ABSTRACT

Numerous transcription factors recruit C-terminal binding protein (CtBP) co-repressors. We show that the large zinc-finger protein ZNF217 contacts CtBP. ZNF217 is encoded by an oncogene frequently amplified in tumours. ZNF217 contains a typical Pro-X-Asp-Leu-Ser (PXDLS) motif that binds in CtBP's PXDLS-binding cleft. However, ZNF217 also contains a second motif Arg-Arg-Thr (RRT) that binds a separate surface on CtBP. The crystal structure of CtBP bound to a RRTGAPPAL peptide shows that it contacts a surface crevice, distinct from the PXDLS binding cleft. Interestingly, both PXDLS and RRT motifs are also found in other zinc-finger proteins, such as RIZ. Finally, we show that ZNF217 represses several promoters, including one from a known CtBP target gene, and mutations preventing ZNF217's contact with CtBP reduce repression. These results identify a new CtBP interaction motif and establish ZNF217 as transcriptional repressor protein that functions, at least in part, by associating with CtBP.

INTRODUCTION

The C-terminal Binding Proteins (CtBPs) are multi-functional proteins implicated in gene regulation, Golgi maintenance and synaptic ribbon formation (3, 7, 41, 43). They function in gene regulation as transcriptional co-repressors. CtBPs interact with the repression domains of sequence-specific DNA-binding proteins (transcription factors) and recruit a repressor complex that contains histone modifying enzymes, such as histone deacetylases 1 and 2, the histone methyltransferase G9a, and the histone demethylase LSD1 (38-40). Approximately 30 transcription factors that recruit CtBP to gene regulatory elements have been identified. These transcription factors come from diverse families and include proteins with zinc-finger, homeodomain, Ets and Sox type DNA-binding domains. They are united, however, by the fact that they typically contain a Pro-X-Asp-Leu-Ser (PXDLS) or related motif in their repression domains through which they contact CtBP (3, 43).

Crystallographic studies have shown that CtBP is composed of a nucleotide-binding domain, that exhibits homology to dehydrogenase enzymes and includes an extensive dimerization interface and NADH binding motif, and a substrate-binding domain, formed by the N-terminus and part of the C-terminal region (22, 30). The X-ray crystal structure shows that the substrate-binding domain forms CtBP's PXDLS-peptide binding cleft (30). In addition, CtBP contains 80 C-terminal residues recently shown to be intrinsically unstructured (31).

Although the mechanism through which CtBP is recruited by PXDLS partners is well understood, the other CtBP protein contacts remain to be characterized. In an effort to

identify other important contact sites on CtBP, we constructed a CtBP protein with a 'filled' PXDLS cleft. This protein was generated from a fusion gene encoding the well-characterized PXDLS motif found in the transcription factor Basic Krüppel-like Factor (BKLF/KLF3) (42) linked to the 3′ end of the murine *CtBP2* gene. The resulting fusion protein thereby contains a C-terminal tail carrying a PXDLS motif and, since the C-terminus of CtBP is flexible and structurally located near CtBP's PXDLS-binding cleft, we expect this tail to be able to fill the cleft. Indeed, we have found that the linked PXDLS tail does block the binding of additional PXDLS motif partners (data not shown). Importantly, a similar fusion protein, incorporating a point mutation in the PXDLS sequence, does not interfere with the binding of exogenous PXDLS motif partners, arguing against the possibility that the fusion tail is non-specifically impeding access to the PXDLS binding cleft (data not shown).

We used this fusion protein in yeast two-hybrid screens and identified murine Znf217 as a protein partner of CtBP2 that does not depend on the PXDLS cleft for association. Murine Znf217 has not previously been described but, based on homology (Fig. 1) and synteny (10), it appears to be the orthologue of human ZNF217, a recognized oncogene implicated in numerous cancers, most notably breast and colon cancer (47). The human *ZNF217* gene resides on the long arm of chromosome 20, at position q13.2 (5). This region is amplified in up to 40% of breast and 60% of colon cancers (35, 47). The amplification has been shown to correlate with increased ZNF217 protein and poor prognosis. Furthermore, it has been found that over-expression of ZNF217 promotes the immortalization of breast epithelial cells (32), although the precise mechanism through which ZNF217 drives immortalization

89 is not known. Interestingly, human ZNF217 has been found to be present in a number of
90 repression complexes (14, 24, 48), including the CtBP-associated repression complex (39).
91 However, the mechanism through which ZNF217 functions remains unknown.
92
93 Here we show that murine and human ZNF217 directly interact with CtBP. We find that
94 ZNF217 contains a PXDLS motif that binds the CtBP cleft, but also contains a motif that
95 binds elsewhere on CtBP. We map this second motif within ZNF217 to the sequence
96 RRTGCPPAL. Co-crystallization of CtBP with an RRTGAPPAL peptide reveals the
97 location of the second peptide binding site on CtBP.
98
99 We demonstrate that ZNF217 represses transcription driven by a number of promoters, and
100 mutations that prevent it from contacting CtBP impair its ability to repress transcription.
101 This suggests that ZNF217 functions in gene repression by recruiting CtBP and its
102 associated repression complex.
103
104 Our results further indicate that other zinc-finger proteins like RIZ and ZNF516, that also
105 contain a PXDLS and the novel RRT motif, may play direct roles in gene repression
106 through contacting CtBP. We suggest that over-expression of ZNF217 may contribute to
107 tumorigenesis through initiating changes in gene expression profiles.
108

MATERIALS AND METHODS

Plasmid constructs

Full length murine CtBP2 was amplified by PCR and the product was cloned into *Xma* I/*Sal* I pGBT9 (Clontech) vector. This resulting construct, pGBT9-mCtBP2, was used as wild type CtBP2 in the yeast two-hybrid experiments described throughout this manuscript. Secondly, murine BKLF 30-75, containing the ^{61}PVDLT^{65} motif, was amplified by PCR and cloned into the *Not* I/*Sal* I sites of pGBT9-mCtBP2. The resulting yeast two-hybrid constructs expressed BKLF 30-75 fused to the C-terminus of CtBP2. The pGBT9-mCtBP2-BKLF 30-75 construct is referred to as "cleft-filled" CtBP2 throughout this manuscript. The "control ΔDL fusion", which contains a DL to AS mutation in the PVDLT motif in the BKLF 30-75 portion of the fusion, was generated from the pGBT9-mCtBP2-BKLF 30-75 construct by overlap PCR mutagenesis. The mCtBP2, mCtBP2-BKLF 30-75 and mCtBP2-BKLF 30-75 ΔDL inserts were subcloned into the *Xma* I/*Sal* I sites of pGAD10(new) (derived from pGAD10 from Clontech) vector to allow them to be expressed in the yeast two-hybrid system as both Gal4AD and Gal4DBD fusions.

A58E, V72R, E181A and D237A mutations were introduced into mCtBP2 by overlap PCR mutagenesis. *Bgl* II and *Sal* I digested mutant inserts were ligated into the *Bam*H I/*Sal* I sites of pGBT9 and pGAD10(new) vectors to generate pGBT9-mCtBP2-A58E, pGBT9-mCtBP2-V72R, pGAD10(new)-mCtBP2-A58E, pGAD10(new)-mCtBP2-V72R, pGBT9-mCtBP2-E181A/D237A and pGBT9-mCtBP2 A58E/E181A/D237A.

The Gal4DBD (147 amino acids) was amplified by PCR using appropriate primers and was ligated into the *Pst* I/*Not* I sites of pMT3 (derived from pMT2) vector to generate pMT3-Gal4 without a stop codon. A separate pMT3-Gal4 with a stop codon was generated to act as a control in mammalian repression assays. Secondly, wild type mCtBP2, mCtBP2-A58E, mCtBP2-E181A/D237A and mCtBP2-A58E/E181A/D237A mutant inserts were re-amplified by PCR using appropriate primers and cloned into the *Not* I/*Sal* I sites of pMT3-Gal4 without stop, 3′ of the Gal4 gene to generate pMT3-Gal4-mCtBP2, pMT3-Gal4-mCtBP2-A58E, pMT3-Gal4-mCtBP2-E181A/D237A and pMT3-Gal4-mCtBP2-A58E/E181A/D237A.

pMT3-YFP was generated by ligating a *Nsi* I/*Not* I YFP PCR fragment from pEYFP-C1 vector (Clontech) into *Pst* I/*Not* I sites of pMT3 vector (derived from pMT2). mCtBP2, mCtBP2-A58E, mCtBP2-E181A/D237A and mCtBP2-A58E/E181A/D237A were re-amplified by PCR using appropriate primers and cloned into *Not* I/*Sal* I sites of pMT3-YFP. pMT2-HA-mCtBP2 has been previously described (44).

pGAD10-mZnf217 530-932 was isolated from MEL cDNA library with the pGBT9-mCtBP2-BKLF 30-75 bait protein. The NL → AS mutation (ΔDL) was introduced into the putative PXDLS motif, ^{680}PLNLS^{684}, in the pGAD10-mZnf217 530-932 construct using overlap PCR site directed mutagenesis. The mZnf217 530-932 and mZnf217 530-932 ΔDL inserts were liberated from the pGAD10 vector by digestion with *Bam*H I and *Bgl* II and were ligated into the *Bam*H I site of pGBT9(new) vector.

Regions of mZnf217 corresponding to amino acids 548-617, 660-715, 753-794, 869-911 and 686-737 were amplified from pGAD10-mZnf217 530-932 template. Regions of mZnf217 corresponding to amino acids 660-715 ΔDL, 548-715 ΔDL, 753-911, 660-794 ΔDL, 548-794 ΔDL, 660-911 ΔDL, 548-911 ΔDL, 548-775 ΔDL, 548-750 ΔDL, 548-725 ΔDL, 700-911, 725-911 and 730-760 were amplified from the pGAD10-mZnf217 ΔDL template. PCR products were cloned into the *Xma* I/*Bam*H I sites of pGBT9 vector to allow expression of Gal4DBD fusions in yeast. The region of mZnf217 encoding amino acids 700-790 was amplified from pGAD10-mZnf217 530-932 template and the PCR product was cloned into the *Bam*H I/*Pst* I sites of pGBT9 vector.

Triple and single alanine scanning mutations (shown in Fig. 2D) were introduced into pGBT9-mZnf217 700-790 using overlap PCR mutagenesis.

Full length human ZNF217 (1-1048) was amplified by PCR from pLXSN-ZNF217 (32) and cloned into the *Eco*R I site of pGAD10 vector to produce pGAD10-hZNF217 1-1048. An NL → AS mutation (ΔDL) was introduced into the ^{686}PLNLS^{690} motif of pGAD10-hZNF217 1-1048 using overlap PCR mutagenesis to generate pGAD10-hZNF217 1-1048 ΔDL. An RRT → AAA mutation (ΔRRT) was introduced into the ^{752}RRTGCPPAL^{760} motif of pGAD10-hZNF217 1-1048 and pGAD10-hZNF217 1-1048 ΔDL by overlap PCR mutagenesis to produce pGAD10-hZNF217 1-1048 ΔRRT and pGAD10-hZNF217 1-1048 ΔDL ΔRRT.

hZNF217 1-1048, 1-1048 ΔDL, 1-1048 ΔRRT, 1-1048 ΔDL ΔRRT inserts were subcloned into the *Eco*R I site of pMT3-FLAGb vector to generate hZNF217 expression constructs with the FLAG sequence fused to the N-terminus.

Segments of hRIZ1 (amino acids 661-820 containing ^{735}RRTSSPPSS^{743}), mRiz1 (amino acids 772-931 containing ^{858}RRTSSPPSS^{866}) and hZNF516 (amino acids 2381-2580 containing ^{2442}GRTGPPPAL^{2450}) were amplified from human genomic DNA, murine genomic DNA, K562 (human) cDNA library and MEL (murine) cDNA library by PCR and were cloned into the *Eco*R I/*Bam*H I sites of pGBT9 vector. RRT/GRT→AAA mutations were introduced into the putative RRT motifs of hRIZ1 661-820, mRiz1 772-931 and hZNF516 2381-2580 using overlap PCR mutagenesis to generate pGBT9-hRIZ1 661-820 ΔRRT, pGBT9-mRiz1 772-931 ΔRRT and pGBT9-hZNF516 ΔGRT.

Details of primers used in plasmid construction are available on request. The identities of the inserts in each construct were confirmed by automated DNA sequencing.

The firefly luciferase reporter vector pGL2-$(Gal4)_5$-$(LexA)_2$-E1B-Luc and LexA-VP16 mammalian expression plasmid pCMV-LexA (1-202)-VP16 (410-490) were generous gifts from Luke Gaudreau and Mark Ptashne (The Sloan–Kettering Institute, New York, NY). A second firefly luciferase reporter vector containing 5xGal4 binding sites and the TK promoter, pGL2-$(Gal4)_5$-TK-Luc, has been described previously (33). pGL3-human E-cadherin (-427/+53)-luciferase reporter vector was a gift from Stephen Sugrue (Department

of Anatomy and Cell Biology, Harvard Medical School) and has been previously described (1).

Yeast two-hybrid screen and assays

Yeast two-hybrid screens were performed with pGBT9-mCtBP2 BKLF 30-75 as bait and murine erythroleukaemia cell (MEL) and human K562 cell cDNA libraries as described previously (42). For yeast two-hybrid assays, test proteins were expressed in HF7c yeast as either Gal4DBD or Gal4AD fusions. Transformant colonies were selected on Leu/Trp deficient plates and patched onto His/Leu/Trp deficient plates. Growth was scored following 72 hours incubation.

Mammalian cell culture

COS-1 and HEK293 cells were cultured as described previously (1, 34) and transfected, using the transfection reagent FuGENE6 (Roche Diagnostics) following the manufacturer's instructions. *CtBP1*$^{+/-}$*CtBP2*$^{+/-}$ (CtBP+/-) and *CtBP1*$^{-/-}$*CtBP2*$^{-/-}$ (CtBP-/-) cells were a gift from J. Hildebrand and were cultured and transfected as described previously (16).

Co-immunopreciptation experiments

To examine interactions between mCtBP2 and hZNF217 mutants, duplicate 100 mm plates of COS-1 cells were transfected with combinations of 1 μg of pMT2-HA-mCtBP2 and 3 μg of pMT3-FLAGb-hZNF217 wt, ΔDL mutant, ΔRRT mutant and double ΔDL ΔRRT mutant DNA. 48 h following transfection, cells were harvested, duplicates pooled and whole cell protein extracts were prepared (in 50 mM Tris, pH 8.0, 150 mM NaCl, 1% NP-

40, 0.2 mM PMSF, 1 μg/mL aprotinin and 1 μg/mL leupeptin, total volume 500 μL). For the input lanes 20 μL of each extract (10% of the amount used in immunoprecipitation) was mixed with SDS loading dye, boiled and run on an 8% SDS-PAGE gel. Immunoprecipitation was performed with 200 μL of each extract, 10 μL of protein G beads and 7.5 μg of either mouse monoclonal αHA (12CA5 Roche Corporation) or mouse monoclonal αFLAG (Sigma) antibodies to immunoprecipitate HA-mCtBP2 or FLAG-hZNF217 respectively. Following washes, beads were mixed with SDS loading dye, boiled and run on 8% SDS-PAGE gels. Proteins in SDS-PAGE gels were blotted onto nitrocellulose membranes (Western Blot) and immuno-detected with 10 μg of both αHA Ab and αFLAG Ab in 10 mL TBST to detect HA-mCtBP2 and FLAG-hZNF217 respectively. A sheep anti-mouse HRP conjugated secondary Ab (Amersham Bioscience) was used and bands were detected with a Western Lightning Chemiluminescence Reagent Plus (Perkin Elmer Life Sciences) and X-ray film (Eastman Kodak Company). The exposures show the results of a representative experiment.

To examine interactions between hZNF217 and mCtBP2 mutants, 100 mm petri dishes of COS-1 cells were transiently transfected with combinations of 3 μg pMT3-FLAGb-hZNF217 and 250 ng of either pMT3-YFP-mCtBP2, pMT3-YFP-mCtBP2 A58E, pMT3-YFP-mCtBP2 E181A/D237A or pMT3-YFP-mCtBP2 A58E/E181A/D237A. Protein extracts, immunoprecipitation and Western blots were performed as described in the above co-immunoprecipitation methods except that immunoprecipitations were conducted with 10 μg of αHA antibody only and Western blots were immuno-detected using αHA and monoclonal mouse αYFP (BD Living Colors, JL-8, Clontech) antibodies.

Western blots for assessment of protein expression levels

Western blots were performed to confirm equivalent expression of the Gal4-mCtBP2 and FLAG-hZNF217 proteins. 100 mm petri dishes of COS-1 cells were transiently transfected with 4 μg pMT3 alone, pMT3-Gal4-mCtBP2, pMT3-Gal4-mCtBP2 A58E, pMT3-Gal4-mCtBP2 E181A/D237A or pMT3-Gal4-mCtBP2 A58E/E181A/D237A or 3 μg pMT3 alone, pMT3-FLAGb-hZNF217 or pMT3-FLAGb-hZNF217 ΔDL ΔRRT. Cells were incubated for 48 hours following the transfection before cells were harvested and nuclear extracts prepared. Equal amounts of each nuclear extract were run on a 12% SDS PAGE gel, and Western Blots were performed as described above. The Gal4-mCtBP2 was visualized using a mouse monoclonal αCtBP2 antibody (BD Biosciences). The FLAG-hZNF217 was visualized using a mouse monoclonal αFLAG antibody.

Mammalian cell repression assays

To examine CtBP repression of reporter gene expression, 6 well plates of COS-1 cells or CtBP+/- and CtBP-/- cells were transiently transfected. To examine repression of basal expression, the following plasmids were used: 3 μg pGL2-(Gal4)5-TK-Luc reporter, and 50 ng of either pMT3-Gal4-mCtBP2, pMT3-Gal4-mCtBP2-A58E, pMT3-Gal4-mCtBP2-E181A/D237A or pMT3-Gal4-mCtBP2-A58E/E181A/D237A. To examine repression of activated expression, the following plasmids were used: 3 μg pGL2-(Gal4)5-(LexA)2-E1B-Luc reporter, 1 μg pCMV-LexA (1-202)-VP16 (410-490) expression vector and 50 ng of either pMT3-Gal4-mCtBP2, pMT3-Gal4-mCtBP2-A58E, pMT3-Gal4-mCtBP2-E181A/D237A or pMT3-Gal4-mCtBP2-A58E/E181A/D237A. In both experiments, 10 ng of the *Renilla* (R) luciferase vector pRL-Luc (Promega) was co-transfected to allow the

firefly (FF) luciferase measurements to be corrected to control for transfection efficiency. Luciferase activity was measured 48 h post-transfection in a Turner Designs model TD 20/20 luminometer using the dual-luciferase reporter assay system (Promega). Results shown are averaged FF/R luciferase ratios from 4 replicates of a representative experiment.

To examine ZNF217 repression of reporter gene expression, 6 well plates of COS-1 cells were transiently transfected. To examine repression of basal expression, the following plasmids were used: 3 μg pGL2-(Gal4)5-TK-Luc reporter, and 150 ng of either pMT3-FLAGb-hZNF217 or pMT3-FLAGb-hZNF217 ΔDL ΔRRT. To examine repression of activated expression, the following plasmids were used: 3 μg pGL2-(Gal4)5-(LexA)2-E1B-Luc reporter, 1 μg pCMV-LexA (1-202)-VP16 (410-490) expression vector and 150 ng of either pMT3-FLAGb-hZNF217 or pMT3-FLAGb-hZNF217 ΔDL ΔRRT. FF luciferase activity was measured as described above. Results shown are averaged FF luciferase ratios from 4 replicates of a representative experiment.

To examine ZNF217 repression of the *E-cadherin* promoter, 6 well plates of HEK293 cells or CtBP+/- and CtBP-/- cells were transiently transfected with 1 μg of pGL3-E-cad-Luc and 1 μg of various pMT3-FLAG-hZNF217 wild type and mutant derivatives and 10 ng of pRL-Luc. 48 h post transfection, FF and R luciferase activities were quantified as described above. Results shown are averaged FF/R luciferase ratios from 2 replicates of a representative experiment.

Crystallization, structure determination and refinement

289 t-CtBP1-S, bearing a His tag at its N-terminus, was expressed in *E. coli* and purified as
290 described previously (29). Vapor-diffusion co-crystallization experiments on the
291 protein/peptide complex were performed after overnight incubation of t-CtBP1-S (at 10
292 mg/ml concentration) with 10 mM RRTGAPPAL peptide. Bipyramidal-shaped crystals of
293 the t-CtBP1-S/peptide complex grew in a few days using a crystallization solution
294 containing 1.8-2.1 M ammonium formate, 100 mM HEPES, pH 7.5. The crystals belong to
295 the space group *P*6422, with unit cell parameters: $a = b = 89.3$ Å, $c = 162.7$ Å, one
296 molecule per asymmetric unit. A full diffraction data set was collected at 2.85 Å resolution
297 using synchrotron radiation (ID14-EH3 beamline, ESRF, Grenoble, France). All diffraction
298 data were processed using MOSFLM and SCALA (9, 25) (see Table I).
299
300 The RRTGAPPAL peptide was prepared on an Applied Biosystems mod 433A synthesizer
301 according to standard Fmoc (9-fluorenylmethoxycarbonyl) solid-phase synthesis. After
302 purification by preparative RP-HPLC, it was shown to be >95% homogeneous by analytical
303 RP-HPLC. Its identity and molecular weight were confirmed by electrospray ionization
304 mass spectrometry (Finnigan LCQ Advantage) (*m/z*, found 938.3; calcd. for $C_{40}H_{71}N_{15}O_{11}$
305 938.097).
306
307 The structure of the t-CtBP1-S/peptide complex was determined by molecular replacement
308 using the program MolRep (4, 45). The crystal structure of t-CtBP1-S (PDB entry-code
309 1HKU) (30) was used as search model. The structure was then refined using the program
310 REFMAC (28) (rigid body and restrained refinement). After a few cycles of refinement,
311 2Fo-Fc electron density maps showed structural details that allowed unambiguous

312 modeling of the peptide, with the exception of the C-terminal A-L residues, for which poor
313 density was available. As in the case of t-CtBP1-S structure (30), a NAD(H) molecule was
314 found, likely the result of specific uploading during t-CtBP1-S expression/purification (30),
315 tightly bound at the nucleotide-binding domain. The final model contains 331 t-CtBP1-S
316 residues (15-345), 19 water and 1 formate molecules, 1 NAD(H), and 1 RRTGAPPAL
317 peptide molecule (R_{factor} = 22.7% and R_{free} = 27.5%, respectively), with ideal
318 stereochemical parameters (Table I) (8, 23).
319
320 Coordinates and structure factors have been deposited with the Protein Data Bank (2) with
321 accession codes 2HU2 and r2HU2sf, respectively.

RESULTS

Identification of Znf217 as a CtBP partner protein

The CtBP2/BKLF fusion protein, containing residues 30-75 of murine BKLF (encompassing its well-characterized PVDLT CtBP contact site) extending from the CtBP2 C-terminus, was used as a bait in yeast two-hybrid screens. Several positive clones were isolated, including previously reported CtBP partners, Ubc9 (18, 26), HIPK2 (49) and the related protein HIPK1. One clone encoding residues 530 to 932 of murine Znf217 was recovered. This isolate was tested for its ability to interact with normal full length CtBP2 in yeast two-hybrid assays, as both prey and bait (Fig. 2A). Yeast growth was observed in both experiments suggesting that Znf217 is a direct binding partner of CtBP2.

Defining the contact regions in Znf217

We next mapped the domains of Znf217 that contact CtBP2. Znf217 contains 8 classical zinc-fingers. The original cDNA fragment we recovered encodes amino acids 530 to 932 and includes zinc-finger 7. Inspection of this fragment revealed that it also contained the motif PLNLS just upstream of zinc-finger 7. The PLNLS sequence fits the general consensus for PXDLS motifs (3, 43) and is conserved in human ZNF217 (Fig. 1).

First, experiments were carried out to confirm that this PLNLS motif was functional and could slot into the CtBP PXDLS peptide binding cleft. Residues 660 to 715 of Znf217 were amplified and tested for their ability to bind CtBP2 in the yeast two-hybrid assay system. The PLNLS was mutated to PLASS, as the substitution of the central two residues, often DL (but here NL) and referred to as the ΔDL mutation, is known to disrupt binding to the

CtBP PXDLS peptide binding cleft (36, 37). In addition, CtBP derivatives that contain defective PXDLS peptide binding clefts were also tested. Two previously described mutations in the cleft A58E and V72R (30), as well as the 'cleft-filled' mutant, were also tested for their ability to bind the Znf217 PXDLS motif (Fig. 2B). The fragment containing the PLNLS motif was able to interact with wild type CtBP2, the mutation in this motif prevented binding, and the CtBP2 derivatives with defective clefts could not bind this fragment. In summary, Znf217 contains a functional PXDLS motif, as shown in Fig. 1.

Our screen was designed to identify CtBP partners that did not rely on PXDLS motifs for associating with CtBP. To determine if Znf217 did require the PXDLS motif for binding to CtBP, its PLNLS motif was mutated, in the context of a longer fragment of Znf217, residues 530 to 932. We found that this Znf217 fragment retained the ability to interact with CtBP (Fig. 2B). This fragment also retained the ability to bind to the CtBP cleft mutants, A58E and V72R, and the 'cleft-filled' derivative in both orientations in yeast (only one orientation is shown). This result confirmed our expectation that Znf217 was a partner protein that did not rely solely on the CtBP PXDLS peptide binding cleft for contact, and suggested that Znf217 contained a second CtBP contact motif.

We next used deletion analyses on Znf217 fragments containing a mutated PLNLS motif to define the second contact motif (Fig. 2C), and localized it to the region downstream of zinc-finger 7. Further alanine scanning experiments demonstrate that the second contact surface in Znf217 comprises the motif RRTGCPPAL (Fig. 2D). We term this an RRT motif.

We next carried out experiments with full length human ZNF217 and full length CtBP2 to verify the interaction. Full length ZNF217 mutants with defective PXDLS (ΔDL) or RRT (RRTGCPPAL→AAAGCPPAL, ΔRRT) motifs were generated, as well as a double mutant that contained mutations in both motifs. These three mutants were first tested in the yeast two-hybrid assay system. As expected the wild type full length ZNF217 interacted with wild type CtBP2. Additionally, both the single mutants retained the ability to interact. However, the double mutant showed very little CtBP binding (Fig. 3A). This result suggests that the PXDLS and RRT motifs in ZNF217 are the major determinants through which it contacts CtBP2.

We then sought to test whether the protein interactions also occurred in the context of mammalian cells. Epitope tagged FLAG-ZNF217 and HA-CtBP2 were transfected into COS-1 cells and their interaction was monitored using immunoprecipitation. As shown in Fig. 3B, when ZNF217 was recovered with Flag antibody, CtBP2 was efficiently retained as revealed by Western blotting against HA. The PXDLS and RRT ZNF217 mutants and the double mutants were also tested. Both the single ZNF217 mutants bound some CtBP2 (though slightly less than wild type), but the double mutant did not associate with detectable CtBP2. The converse experiment (immunoprecipitating with anti-HA and Western blotting with anti-Flag) was also carried out with similar results, except that in this orientation the reduction in binding brought about by the single mutations was more striking, possibly because the immunoprecipitation or detection of associated proteins by Western blotting was somewhat less efficient in this orientation. Nevertheless, these results confirm the inferences from the yeast two-hybrid assays that full length CtBP2 associates

with full length ZNF217, and that the two motifs, the PXDLS and the RRT motifs, are primarily responsible for the association (Fig. 3C).

RRT motifs occur in several CtBP partner proteins

We searched protein databases to determine whether RRT motifs occur in other proteins. Similar motifs were identified in ZNF516 and RIZ. Both proteins also contained recognizable PXDLS motifs within their sequences (Fig. 4A). We did not identify proteins that contained clear RRT motifs in the absence of the PXDLS motif. Although little is known about ZNF516, it is notable that it was found to co-purify in the repression complex that associates with CtBP in HeLa cells (reported under the name KIA0222) (39). RIZ is a well-studied 8 zinc-finger protein that contains a PR/SET domain and has been reported to possess histone methyltransferase activity (17). It contains two PXDLS motifs and has previously been inferred to be a CtBP partner, although the sites and functional effects of CtBP contact have not been described (15). RIZ also contains two potential RRT motifs. These motifs are conserved in the human and murine forms of RIZ.

In order to test whether the RRT motifs that had been identified by bioinformatics screening were able to physically interact with CtBP, segments of RIZ and ZNF516 were tested for binding to CtBP2 using the yeast two-hybrid system. It was found that the ZNF516 motif and one (but not the other, data not shown) of the RRT motifs in RIZ were able to interact with CtBP (Fig. 4A). Mutations of the RRT sequence abolished the interaction (Fig. 4B), as summarized in Fig. 4C.

Defining the regions in CtBP that contact the RRT motif using X-ray crystallography

To shed more light on the structural bases of the CtBP-ZNF217 interaction, crystallographic evidence was sought on the location of the RRTGCPPAL peptide recognition site on CtBP. To avoid aggregation during crystallization, mutants with the peptide's C residue altered to A or S were tested in yeast two-hybrid assays for binding to CtBP2 (Fig. 2D). Both A and S are tolerated at this amino acid position, so an RRTGAPPAL peptide was synthesized. Co-crystallization experiments were performed by incubating the synthetic peptide with a truncated form of the CtBP1-S isoform (or the short-CtBP1 splice isoform, previously known as CtBP3/BARS). This truncated CtBP1-S (t-CtBP1-S), devoid of 80 C-terminal residues, was successfully used in the past to identify the PXDLS consensus binding site (30). The protein-RRTGAPPAL complex 3D structure was solved by molecular replacement methods using the t-CtBP1-S structure as a starting model (PDB entry-code 1HKU), and refined to 2.85 Å resolution (R_{factor} = 22.7% and R_{free} = 27.5%, respectively; Table I) (8). The crystallized t-CtBP1-S appears as a tight dimer, built across a 2-fold crystallographic symmetry axis, with major packing interactions based on pairing of two nucleotide-binding domains of each monomer, as observed for other t-CtBP1-S crystal forms (Fig. 5 A, B; (30)).

The crystal structure of the protein-RRTGAPPAL complex shows that the consensus peptide binds at a surface cleft mainly defined by the loop connecting helix αC to strand βA, and by helices αF and αG, of the nucleotide-binding domain (Fig. 5 A, C). The bound peptide adopts an extended conformation, antiparallel to the αG helix, burying 146 $Å^2$ of protein surface. Binding of the exogenous peptide is supported by docking of its R1, R2, T3 side-

437 chains into a surface groove lined by CtBP residues Y129, A159, E164, H218, D220, R245,
438 Q246, G247, A248, F249, and R274 (Fig. 5C). The rest of the peptide lies at the protein
439 surface, and shows a kink at residues P6-P7, that locates the peptide C-terminal part next to
440 the last turn of helix αG. The main stabilizing peptide-protein interactions involve two salt
441 bridges (R1-D220, and R2-E164), hydrogen bonds in the residue pairs R1-H217, R2-G247,
442 T3-D220, T3-R245, G4-Q246, and G4-Q246, and intermolecular van der Waals contacts at
443 P6 and P7 residues (Fig. 5C). Interestingly, all protein residues involved in peptide-
444 binding/recognition are conserved within the CtBP family, except for the conservative
445 substitution of H218→Q in the CtBP2 sequence.

446

447 Overlay of t-CtBP1-S and t-CtBP1-S/peptide complex 3D-structures yields a r.m.s.
448 deviation of 0.45 Å, indicating that binding of the consensus peptide RRTGAPPAL is not
449 associated with significant tertiary/quaternary structure modifications. Only local side chain
450 conformational changes are induced by peptide binding. Among these, we notice the
451 substitution of the R245 guanidinio group of t-CtBP1-S with the guanidino head of R1,
452 from the peptide, which thus replaces the intramolecular salt bridge R245-D220 with the
453 intermolecular R1-D220 ion pair. The consensus RRTGAPPAL binding site has no direct
454 contact with the NAD(H) binding region (about 27 Å apart), although both are hosted in the
455 nucleotide-binding domain, nor with the previously identified PXDLS binding site (about
456 53 Å apart). The latter is localized at the N-terminal region of the substrate-binding domain,
457 and on the opposite face of the t-CtBP1-S subunit (Fig. 5 A, B). It is, however, worth
458 noting that in the t-CtBP1-S dimeric assembly, where the two substrate-binding domains lie
459 at opposite poles, the PXDLS binding site of one subunit is located on the same dimer face

of the RRT binding site of the opposite subunit (about 30 Å apart). Considering the close proximity of the ^{686}PLNLS^{690} and ^{752}RRTGCPPAL^{760} motifs in the ZNF217 sequence (only 61 amino acids apart), it is possible for ZNF217 to bind across the CtBP dimer, accessing the PXDLS and RRT binding sites on distinct CtBP subunits, respectively (Fig. 5 A, B).

Confirmation of the structural results using mutagenesis

To confirm the inferred location of the RRTGAPPAL binding site, two CtBP1-S residues building up the peptide recognition cleft, E164 and D220, were selected for mutation to alanine. Mutations were also made in the corresponding residues in CtBP2: E181A and D237A. CtBP derivatives containing each mutation, and the two mutations together were generated. In addition, CtBP proteins containing defective PXDLS-binding clefts were further mutated so that they also carried these additional mutations in the putative RRT binding sites. This panel of CtBP mutants was first tested for interaction with ZNF217 using the yeast two-hybrid system (Fig. 6A). Each of the mutations in the CtBP2 and CtBP1-S RRT motif binding clefts was individually sufficient to abrogate binding (only the results from double mutation of two of the amino acids in this cleft for CtBP2 are shown). As expected the CtBP mutants bearing mutations in either the RRT motif contact region or in the PXDLS binding cleft retained the ability to contact ZNF217, however, when both regions were mutated binding was abrogated. Each of the CtBP mutants retained the ability to dimerize with wild type CtBP indicating that these proteins are expressed and properly folded in yeast. This result is consistent with the structural data and confirms the inference that ZNF217 does contact residues E181 and D237 of CtBP2 through its RRT motif (Fig.

6A). Co-immunoprecipitation experiments were then performed and validated the yeast two-hybrid assay results (Fig. 6B). A summary of the interactions is shown in Fig. 6C.

Mutations in the PXDLS and RRT motif binding clefts of CtBP have little effect on its ability to repress transcription

Having generated a CtBP mutant (A58E/E181A/D237A) that was unable to bind to ZNF217, we sought to assess the effect of this mutation on the ability of CtBP to repress transcription. As CtBP cannot bind to DNA directly, the conventional Gal4-DNA-binding domain (Gal4DBD) fusion strategy that is widely used to assess CtBP repression activity was employed (11, 20, 42). The cDNAs encoding wild type CtBP and CtBP with mutations in the PXDLS motif binding cleft (A58E), in the RRT motif binding cleft (E181A/D237A), and at both clefts, were fused to a cDNA encoding the Gal4DBD. These constructs were transfected into COS-1 cells and were shown to be expressed at equivalent levels (Fig. 7 A). Their ability to repress transcription of the firefly luciferase reporter driven by a core TK promoter with 5 Gal4 binding sites, and a LexA-VP16 activated E1B promoter with 5 Gal4 binding sites and 2 LexA binding sites was examined (Fig. 7 B and C). The mutation in the PXDLS binding cleft (A58E) had a modest effect on the ability of CtBP to repress, but additional mutations in the RRT binding cleft (E181A/D237A), or the RRT cleft mutations alone had no discernible effect on activity. To exclude the possibility that the mutant proteins were retaining repression activity by virtue of their ability to dimerize with wild type endogenous CtBP, we repeated the experiments in murine embryonic fibroblasts derived from *CtBP1*$^{-/-}$*/CtBP2*$^{-/-}$ double knockout murine embryos (16). Wild type CtBP2, A58E and E181A/D237A mutants all exhibited strong repression activity in these CtBP-/-

505 cells (Fig. 7D) although the mutants showed a slight reduction in repression. In summary,
506 these results suggest that ZNF217 contact does not make a major contribution to repression
507 by CtBP.
508
509 **ZNF217 is a transcriptional repressor and mutations in the PXDLS and RRT motifs**
510 **of ZNF217 reduce this activity**
511 We next examined whether ZNF217 is able to repress transcription. We also tested ZNF217
512 mutants that cannot bind to CtBP in these assays. The molecular mechanism through which
513 ZNF217 operates has not been determined, but the finding that it associates directly with
514 CtBP2 suggested that it may play a role in gene repression. We therefore examined its
515 activity on a number of test promoters.
516
517 We first tested the ability of ZNF217 to repress the Gal4 site linked TK promoter. Co-
518 transfection of a plasmid encoding wild type ZNF217 resulted in significant repression. We
519 then tested the mutant derivatives of ZNF217. Mutation of both the PXDLS and RRT
520 motifs in ZNF217 significantly reduced repression (Fig. 8B). Taken together these results
521 suggest that ZNF217 can act as a repressor of transcription and that it in part utilizes CtBP
522 to mediate repression. The residual repression observed indicates that it may also have
523 additional mechanisms through which it can repress gene expression. Similar results were
524 obtained when ZNF217 and the mutant were tested against a second promoter containing
525 LexA and Gal4 sites upstream of the Adenovirus E1B promoter driving a luciferase
526 reporter gene (Fig. 8C).
527

We also sought to test a natural CtBP dependent promoter and chose the *E-cadherin* promoter, as it has previously shown to be a CtBP target gene (12, 13, 39). We transfected a ZNF217 encoding plasmid together with the *E-cadherin* promoter driving a luciferase reporter gene and observed that ZNF217 significantly repressed expression of the reporter gene. We also tested the PXDLS mutant, the RRT mutant, and the double mutant. We found that each single mutation modestly reduced repression and that the double mutation more significantly reduced repression (Fig. 8D). These results were similar to those obtained on the viral promoters used above. We also noted that no repression was observed on other promoters, such as the CMV-promoter driven *Renilla* reporter plasmid, showing that ZNF217 does not non-specifically repress all promoters (data not shown). Wild type ZNF217, and ZNF217 with mutations in both the PXDLS motif (ΔDL) and RRT motif (ΔRRT), were expressed at equivalent levels in COS-1 cells (Fig. 8A).

To further assess the contribution of CtBP to ZNF217 repression activity, we repeated the experiments in CtBP+/- and double knockout cells (Fig. 8E). When transfected into CtBP+/- control cells, ZNF217 represses the *E-cadherin* promoter. Again, this repression activity appears to be mediated in part by CtBP, since the double mutant (ΔDL ΔRTT) shows reduced repression. When tested in CtBP-/- cells, the *E-cadherin* reporter is de-repressed and shows high activity (Fig. 8E). Significantly, co-transfection of ZNF217 leads to significant repression even in the CtBP-/- cells. The double mutant ZNF217 retains equivalent repression activity in these cells (Fig. 8E). Taken together the results suggest that recruitment of CtBP enhances repression but that ZNF217 contacts additional partners that can mediate repression in the absence of CtBP.

DISCUSSION

We have shown that the zinc-finger oncoprotein ZNF217 interacts with CtBP utilizing both a conventional PXDLS motif (localizing to a binding cleft in the CtBP substrate-binding domain (30)), and a distinct RRT motif, that binds at a newly defined surface cleft in the CtBP nucleotide-binding domain (Fig. 5). The two peptide recognition sites are physically well separated (53 Å), being roughly at opposite poles of the CtBP subunit. Moreover, based on geometrical considerations, the two sites do not appear to support simultaneous contacts with the same ZNF217 molecule within one CtBP subunit. Rather ZNF217 may bind across the CtBP dimer, contacting surfaces of both substrate- and nucleotide-binding domains from the two protein subunits, thus recognizing both PXDLS and RRT binding sites on distinct CtBP subunits.

ZNF217 is implicated in human cancers and has been shown to contribute to the immortalization of breast epithelial cells in culture (32). Our work suggests that one mechanism by which increased copy number of *ZNF217* contributes to tumorigenesis could be through altering gene expression, for example, *via* increased repression of tumour suppressor gene promoters. We also show that the ability of ZNF217 to repress transcription is partially dependent on its ability to bind to CtBP.

Having established that ZNF217 represses transcription, future research will focus on the full mechanisms by which it mediates repression. It is known that the protein RIZ, which also contains zinc-fingers, and PXDLS and RRT motifs, can bind GC-rich sites in DNA through its zinc-fingers 1-3 (46) and also has been reported to possess histone

574 methyltransferase activity (19). By analogy, ZNF217 may be a sequence-specific DNA-
575 binding protein that recruits CtBP to silence specific genes and the residual repression
576 activity observed when it is unable to recruit CtBP may reflect an additional repression
577 mechanism. However, to date we have not detected direct DNA-binding by ZNF217 (data
578 not shown). The relationship between ZNF217 and the *E-cadherin* promoter and the
579 mechanism through which it may be recruited to the promoter *in vivo* is still under
580 investigation.
581
582 Interestingly, ZNF217 and ZNF516 (recorded as KIA0222) have been found to be present
583 in a number of repression complexes (14, 24, 48), including the CtBP-associated repression
584 complex that exists in HeLa cells (39), consistent with our data that these proteins directly
585 contact CtBP. Relatively few typical sequence-specific transcription factors have been
586 found in these repression complexes. One known DNA-binding protein that has been found
587 in the CtBP repressor complex is the large zinc-finger homeodomain transcription factor
588 ZEB (39). It is possible that ZNF217, RIZ and ZEB can function as conventional
589 transcription factors and also display additional activities allowing them to contribute
590 directly to gene repression. However, our observation that mutations in CtBP that prevent it
591 from binding ZNF217 had little effect on its ability to repress transcription argues against
592 ZNF217 being an essential effector protein in the CtBP repression complex. It should be
593 noted that a slight loss of repression was apparent when the CtBP mutants were tested in
594 CtBP-/- cells (Fig. 7D) so it is possible that ZNF217 makes some contribution to
595 repression. But taken together the results indicate that ZNF217 is not a critical effector of
596 CtBP activity at least in the promoter and cellular contexts tested.

In summary, we have shown that ZNF217 is a direct partner protein contacting CtBP through the known PXDLS motif, but also through a second RRT motif that binds a novel peptide-recognition groove. Other large zinc-finger proteins also contain PXDLS and RRT motifs. We have shown that mutation of these motifs in ZNF217 reduces its ability to repress transcription. These results suggest that one mechanism through which the ZNF217 oncogene may contribute to tumorigenesis is through CtBP-associated repression of transcription.

ACKNOWLEDGEMENTS

We are grateful to L. Gaudreau and M. Ptashne for the pGL2-(Gal4)5-(LexA)2-E1B-Luc and LexA-VP16 mammalian expression plasmids, S. Sugrue for the E-cad-Luc reporter construct, J. Hildebrand for the CtBP knockout cells, C. Cericola and A. Colanzi for their skillful assistance. We thank the staff of the ID14-EH3 beamline at ESRF, Grenoble (France) for data collection facilities and assistance. K.Q. is supported by an Australian Postgraduate Award. This work was supported by NIH grant NHLBI HL073443 and grants from the Australian ARC and NHMRC to M.C., and by the Italian Ministry of University FIRB grants, and by AIRC (Italy) to M.B. and D.C., and by Telethon (Italy) to D.C. M.B. is grateful to CIMAINA (University of Milano) and to Fondazione CARIPLO (Milano, Italy) for continuous support.

615 **REFERENCES**

616 1. **Alpatov, R., G. C. Munguba, P. Caton, J. H. Joo, Y. Shi, Y. Shi, M. E. Hunt,**
617 **and S. P. Sugrue.** 2004. Nuclear speckle-associated protein Pnn/DRS binds to the
618 transcriptional corepressor CtBP and relieves CtBP-mediated repression of the E-
619 cadherin gene. Mol Cell Biol **24:**10223-10235.

620 2. **Berman, H. M., J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I.**
621 **N. Shindyalov, and P. E. Bourne.** 2000. The Protein Data Bank. Nucleic Acids
622 Res **28:**235-242.

623 3. **Chinnadurai, G.** 2002. CtBP, an Unconventional Transcriptional Corepressor in
624 Development and Oncogenesis. Mol Cell **9:**213-224.

625 4. **Collaborative Computational Project, N. C.** 1994. The CCP4 suite. Programs for
626 protein crystallography. Acta Crystallogr. D Biol. Crystallogr. **50:**760-763.

627 5. **Collins, C., J. M. Rommens, D. Kowbel, T. Godfrey, M. Tanner, S. I. Hwang,**
628 **D. Polikoff, G. Nonet, J. Cochran, K. Myambo, K. E. Jay, J. Froula, T.**
629 **Cloutier, W. L. Kuo, P. Yaswen, S. Dairkee, J. Giovanola, G. B. Hutchinson, J.**
630 **Isola, O. P. Kallioniemi, M. Palazzolo, C. Martin, C. Ericsson, D. Pinkel, D.**
631 **Albertson, W. B. Li, and J. W. Gray.** 1998. Positional cloning of ZNF217 and
632 NABC1: genes amplified at 20q13.2 and overexpressed in breast carcinoma. Proc
633 Natl Acad Sci U S A **95:**8703-8708.

634 6. **Comijn, J., G. Berx, P. Vermassen, K. Verschueren, L. van Grunsven, E.**
635 **Bruyneel, M. Mareel, D. Huylebroeck, and F. van Roy.** 2001. The two-handed E
636 box binding zinc finger protein SIP1 downregulates E-cadherin and induces
637 invasion. Mol Cell **7:**1267-1278.

7. **Corda, D., A. Colanzi, and A. Luini.** 2006. The multiple activities of CtBP/BARS proteins: the Golgi view. Trends Cell Biol **16:**167-173.

8. **Engh, R. A., and R. Huber.** 1991. Accurate bond and angle parameters for X-ray protein structure refinement. Acta Crystallogr. **A47:**392-400.

9. **Evans, P. R.** 1993. Proceedings of the CCP4 study weekend on Data collection and processing. CLRC Daresbury Laboratory, UK.**:**114-122.

10. **Ewart-Toland, A., P. Briassouli, J. P. de Koning, J. H. Mao, J. Yuan, F. Chan, L. MacCarthy-Morrogh, B. A. Ponder, H. Nagase, J. Burn, S. Ball, M. Almeida, S. Linardopoulos, and A. Balmain.** 2003. Identification of Stk6/STK15 as a candidate low-penetrance tumor-susceptibility gene in mouse and human. Nat Genet **34:**403-412.

11. **Furusawa, T., H. Moribe, H. Kondoh, and Y. Higashi.** 1999. Identification of CtBP1 and CtBP2 as corepressors of zinc finger-homeodomain factor deltaEF1. Mol Cell Biol **19:**8581-8590.

12. **Grooteclaes, M., Q. Deveraux, J. Hildebrand, Q. Zhang, R. H. Goodman, and S. M. Frisch.** 2003. C-terminal-binding protein corepresses epithelial and proapoptotic gene expression programs. Proc Natl Acad Sci U S A **100:**4568-4573.

13. **Grooteclaes, M. L., and S. M. Frisch.** 2000. Evidence for a function of CtBP in epithelial gene regulation and anoikis. Oncogene **19:**3823-3828.

14. **Hakimi, M. A., Y. Dong, W. S. Lane, D. W. Speicher, and R. Shiekhattar.** 2003. A candidate X-linked mental retardation gene is a component of a new family of histone deacetylase-containing complexes. J Biol Chem **278:**7234-7239.

15. **Hickabottom, M., G. A. Parker, P. Freemont, T. Crook, and M. J. Allday.** 2002. Two nonconsensus sites in the Epstein-Barr virus oncoprotein EBNA3A cooperate to bind the co-repressor carboxyl-terminal-binding protein (CtBP). J Biol Chem **277:**47197-47204.

16. **Hildebrand, J. D., and P. Soriano.** 2002. Overlapping and Unique Roles for C-Terminal Binding Protein 1 (CtBP1) and CtBP2 during Mouse Development. Mol Cell Biol **22:**5296-5307.

17. **Huang, S., G. Shao, and L. Liu.** 1998. The PR domain of the Rb-binding zinc finger protein RIZ1 is a protein binding interface and is related to the SET domain functioning in chromatin-mediated gene expression. J Biol Chem **273:**15933-15939.

18. **Kagey, M. H., T. A. Melhuish, and D. Wotton.** 2003. The polycomb protein Pc2 is a SUMO E3. Cell **113:**127-137.

19. **Kim, K. C., L. Geng, and S. Huang.** 2003. Inactivation of a histone methyltransferase by mutations in human cancers. Cancer Res **63:**7619-7623.

20. **Koipally, J., and K. Georgopoulos.** 2000. Ikaros interactions with CtBP reveal a repression mechanism that is independent of histone deacetylase activity. J Biol Chem **275:**19594-19602.

21. **Kraulis, P. J.** 1991. MOLSCRIPT: a program to produce both detailed and schenatic plots of protein structures. J. Appl. Crystallogr. **24:**946-950.

22. **Kumar, V., J. E. Carlson, K. A. Ohgi, T. A. Edwards, D. W. Rose, C. R. Escalante, M. G. Rosenfeld, and A. K. Aggarwal.** 2002. Transcription corepressor CtBP is an NAD(+)-regulated dehydrogenase. Mol Cell **10:**857-869.

23. **Laskowski, R. A., M. W. MacArthur, D. S. Moss, and J. M. Thornton.** 1993. PROCHECK, a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. **26:**283-291.

24. **Lee, M. G., C. Wynder, N. Cooch, and R. Shiekhattar.** 2005. An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation. Nature **437:**432-435.

25. **Leslie, A. G. M.** 2003. MOSFLM User guide, Mosflm Version 6.2.3., MRC Laboratory of Molecular Biology, Cambridge, UK.

26. **Lin, X., B. Sun, M. Liang, Y. Y. Liang, A. Gast, J. Hildebrand, F. C. Brunicardi, F. Melchior, and X. H. Feng.** 2003. Opposed regulation of corepressor CtBP by SUMOylation and PDZ binding. Mol Cell **11:**1389-1396.

27. **Merritt, E. A., and M. E. Murphy.** 1994. Raster3D Version 2.0. A program for photorealistic molecular graphics. Acta Crystallogr D Biol Crystallogr **50:**869-873.

28. **Murshudov, G. N., A. A. Vagin, and E. J. Dodson.** 1997. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr **53:**240-255.

29. **Nardini, M., S. Spano, C. Cericola, A. Pesce, G. Damonte, A. Luini, D. Corda, and M. Bolognesi.** 2002. Crystallization and preliminary X-ray diffraction analysis of brefeldin A-ADP ribosylated substrate (BARS). Acta Crystallogr D Biol Crystallogr **58:**1068-1070.

30. **Nardini, M., S. Spano, C. Cericola, A. Pesce, A. Massaro, E. Millo, A. Luini, D. Corda, and M. Bolognesi.** 2003. CtBP/BARS: a dual-function protein involved in transcription co-repression and Golgi membrane fission. Embo J **22:**3122-3130.

31. **Nardini, M., D. Svergun, P. V. Konarev, S. Spano, M. Fasano, C. Bracco, A. Pesce, A. Donadini, C. Cericola, F. Secundo, A. Luini, D. Corda, and M. Bolognesi.** 2006. The C-terminal domain of the transcriptional co-repressor CtBP is intrinsically unstructured. Prot. Sci. **15:**1042-1050.

32. **Nonet, G. H., M. R. Stampfer, K. Chin, J. W. Gray, C. C. Collins, and P. Yaswen.** 2001. The ZNF217 gene amplified in breast cancers promotes immortalization of human mammary epithelial cells. Cancer Res **61:**1250-1254.

33. **Perdomo, J., and M. Crossley.** 2002. The Ikaros family protein Eos associates with C-terminal-binding protein corepressors. Eur J Biochem **269:**5885-5892.

34. **Perdomo, J., A. Verger, J. Turner, and M. Crossley.** 2005. Role for SUMO Modification in Facilitating Transcriptional Repression by BKLF. Mol Cell Biol **25:**1549-1559.

35. **Rooney, P. H., A. Boonsong, M. C. McFadyen, H. L. McLeod, J. Cassidy, S. Curran, and G. I. Murray.** 2004. The candidate oncogene ZNF217 is frequently amplified in colon cancer. J Pathol **204:**282-288.

36. **Schaeper, U., J. M. Boyd, S. Verma, E. Uhlmann, T. Subramanian, and G. Chinnadurai.** 1995. Molecular cloning and characterization of a cellular phosphoprotein that interacts with a conserved C-terminal domain of adenovirus E1A involved in negative modulation of oncogenic transformation. Proc. Natl. Acad. Sci. U S A **92:**10467-10471.

37. **Schaeper, U., T. Subramanian, L. Lim, J. M. Boyd, and G. Chinnadurai.** 1998. Interaction between a cellular protein that binds to the C-terminal region of

adenovirus E1A (CtBP) and a novel cellular protein is disrupted by E1A through a conserved PLDLS motif. J Biol Chem **273:**8549-8552.

38. **Shi, Y., F. Lan, C. Matson, P. Mulligan, J. R. Whetstine, P. A. Cole, R. A. Casero, and Y. Shi.** 2004. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell **119:**941-953.

39. **Shi, Y., J. Sawada, G. Sui, B. Affar el, J. R. Whetstine, F. Lan, H. Ogawa, M. P. Luke, Y. Nakatani, and Y. Shi.** 2003. Coordinated histone modifications mediated by a CtBP co-repressor complex. Nature **422:**735-738.

40. **Shi, Y., and Y. Shi.** 2005. CtBP corepressor complex- a multi-enzyme machinery that coordinates chromatin modifications. CtBP family proteins **LANDES BIOSCIENCES, Georgetown, Texas, USA. Edited by G. Chinnadurai:**http://eurekah.com/abstract.php?chapid=2751&bookid=198&catid=30.

41. **Tom Dieck, S., F. Schmitz, and J. H. Brandstatter.** 2005. CtBPs as synaptic proteins. CtBP family proteins **LANDES BIOSCIENCES, Georgetown, Texas, USA. Edited by G. Chinnadurai:**http://eurekah.com/abstract.php?chapid=2646&bookid=198&catid=30.

42. **Turner, J., and M. Crossley.** 1998. Cloning and characterization of mCtBP2, a co-repressor that associates with basic Kruppel-like factor and other mammalian transcriptional regulators. Embo J **17:**5129-5140.

43. **Turner, J., and M. Crossley.** 2001. The CtBP family: enigmatic and enzymatic transcriptional co-repressors. Bioessays **23:**683-690.

749 44. **Turner, J., H. Nicholas, D. Bishop, J. M. Matthews, and M. Crossley.** 2003. The
750 LIM protein FHL3 binds basic Kruppel-like factor/Kruppel-like factor 3 and its co-
751 repressor C-terminal-binding protein 2. J Biol Chem **278:**12786-12795.
752 45. **Vagin, A., and A. Teplyakov.** 1997. MOLREP: An automated program for
753 molecular replacement. J. Appl. Crystallogr. **30:**1022-1025.
754 46. **Xie, M., G. Shao, I. M. Buyse, and S. Huang.** 1997. Transcriptional repression
755 mediated by the PR domain zinc finger gene RIZ. J Biol Chem **272:**26360-26366.
756 47. **Yaswen, P., and M. R. Stampfer.** 2002. Molecular changes accompanying
757 senescence and immortalization of cultured human mammary epithelial cells. Int J
758 Biochem Cell Biol **34:**1382-1394.
759 48. **You, A., J. K. Tong, C. M. Grozinger, and S. L. Schreiber.** 2001. CoREST is an
760 integral component of the CoREST- human histone deacetylase complex. Proc Natl
761 Acad Sci U S A **98:**1454-1458.
762 49. **Zhang, Q., Y. Yoshimatsu, J. Hildebrand, S. M. Frisch, and R. H. Goodman.**
763 2003. Homeodomain interacting protein kinase 2 promotes apoptosis by
764 downregulating the transcriptional corepressor CtBP. Cell **115:**177-186.
765

FIGURE LEGENDS

Figure 1

The human and murine ZNF217 protein sequences show significant homology.

The sequences of full length human ZNF217 (NM_006526) and murine Znf217 (NM_001033299) proteins are shown. The zinc-finger regions (1-8) are underlined with solid grey lines. The conserved PXDLS motifs and RRT motifs are underlined with grey dashed lines. Residues 530 (lysine) and 932 (glycine) of mZnf217, the first and last amino acids of the yeast two-hybrid screen isolate, are indicated with asterisks.

Figure 2

Znf217 interacts with CtBP2 and the non-PXDLS interaction site was mapped to the motif RRTGXPPXL

A. Yeast two-hybrid assays were performed to examine the interactions between mZnf217 530-932 and CtBP2. These assays were performed with each of the two test proteins fused to the C-terminus of either Gal4AD or Gal4DBD. Growth on -His-Leu-Trp plates (pictured) indicates that the two test proteins interact.

B. The PXDLS motifs in Znf217 530-932 and 660-715 were mutated and both wild type and mutant proteins were tested for binding to wild type CtBP and CtBP with mutations in the PXDLS cleft. Interaction with wild type CtBP (and the BKLF 30-75 ΔDL control protein) only indicates binding which is dependent on a PXDLS motif. Interaction with both wild type and mutant CtBPs indicates binding which is not dependent on a PXDLS motif.

C. Deletion mapping was performed to determine the minimal portion of murine Znf217 capable of interacting with CtBP in a PXDLS motif independent manner. The Gal4DBD-Znf217 proteins are depicted schematically and results of yeast two-hybrid assays with these proteins and Gal4AD-CtBP are shown as either plus (for growth of yeast) or minus. A ΔDL mutation (NL-AS in the PLNLS motif) was introduced into many of the mZnf217 proteins so that only non-PXDLS binding was being examined, and is indicated. The minimal region of mZnf217 required for interaction with CtBP, amino acids 730-760, is indicated by a grey column.

D. Both single and triple mutations were introduced into amino acids 740-760 of Gal4DBD-Znf217 700-790. The mutations in each of the constructs are highlighted within the sequence of Znf217 amino acids 740-760. The results of yeast two-hybrid assays with these mutant Gal4DBD-mZnf217 proteins and Gal4AD-CtBP are shown as either plusses (for relative growth of yeast) or minus. The consensus motif suggested, RRTGXPPXL, is shown below.

Figure 3

Mutation of the PXDLS and RRT motifs of ZNF217 reduce the ability to bind to CtBP and the double mutant has a severe reduction in binding

A. Gal4AD fused wild type hZNF217 and hZNF217 with mutations in the PLNLS motif (ΔDL), RRTGCPPAL motif (ΔRRT) and both motifs (ΔDL ΔRRT) were examined for their ability to interact with Gal4DBD fused wild type and cleft filled (CtBP2-BKLF 30-75) CtBP in yeast two-hybrid assays.

B. Fusion proteins of FLAG and wild type hZNF217 or hZNF217 with mutations in the PLNLS motif motif (ΔDL), RRTGCPPAL motif (ΔRRT) and both motifs (ΔDL ΔRRT) were examined for their ability to interact with HA-CtBP2 in co-immunoprecipitation experiments. COS-1 cells were transfected with the expression vectors indicated and whole cell extracts were immunoprecipitated (IP) separately with both the αFLAG and αHA antibodies. Expression of each of the FLAG fused and HA fused proteins are shown in the top two panels (10% input). FLAG-ZNF217 immunoprecipitated by the αFLAG antibody and the resulting co-immunopreciptated HA-CtBP2 is shown in the middle two panels (IP: αFLAG). HA-CtBP2 immunopreciptated by the αHA antibody and the resulting co-immunoprecipitated FLAG-hZNF217 is shown in the bottom two panels (IP: αHA).

C. A summary diagram combining the results of interaction studies between CtBP and wild type or mutant ZNF217.

Figure 4

RRT motifs are also found in RIZ and ZNF516 and are capable of mediating binding to CtBP

A. mZnf217, hRIZ1, mRiz1 and hZNF516 are large zinc-finger proteins which possess both PXDLS motifs and putative RRT motifs. The features of each protein are shown. The predicted zinc-fingers are shown as arches and numbered, the consensus PXDLS motifs are indicated by hollow rectangles and the motifs are outlined above in black and the putative RRT motifs are indicated by grey filled rectangles and the motifs are outlined above in grey. The PR/SET domains in the RIZ proteins are indicated by wide grey rectangles. The portion of each protein, containing the putative RRT motif, which was tested for interaction

834 with CtBP is indicated below each sequence as a black bar with the numbers of the flanking
835 amino acids indicated.
836 **B.** Segments of ZNF516 and both murine and human RIZ1 with and without mutations in
837 the putative RRT motifs fused to Gal4DBD were tested for their ability to bind to wild type
838 and cleft mutant CtBP fused to Gal4AD in yeast two-hybrid assays.
839 **C.** An alignment is shown of the RRT motifs that have been shown to mediate binding to
840 CtBP. Amino acids within the sequences which are identical to the amino acids in the
841 mZnf217 RRT motif are boxed in grey. The consensus combines information obtained
842 from validated natural RRT motifs and also from mutagenesis studies. The height of each
843 amino acid at each position is representative of the relative frequency in the naturally
844 occurring RRT motif proteins and the tolerance for various amino acids as determined by
845 mutational analysis.
846
847 **Figure 5**
848 **X-ray crystal structure of the RRTGAPPAL peptide bound to t-CtBP1-S**
849 **A.** Ribbon diagram of the t-CtBP1-S dimer. The protein subunits composing the dimer are
850 shown in green and red. The substrate- and the nucleotide-binding domains of each subunit
851 are labeled as SBD and NBD, respectively. The bound NAD(H) and RRTGAPPAL-peptide
852 molecules are shown in ball-and-stick representations (black and magenta, respectively).
853 The PXDLS binding site is reported from the crystal structure of the complex formed by t-
854 CtBP1-S and the PIDLSKK peptide, shown in blue (PDB entry-code 1HL3). (Prepared
855 with MOLSCRIPT (21), and Raster3D (27)).

B. CPK representation of the t-CtBP1-S dimer. In this space filing representation, the molecular complex displayed in panel A has been rotated by about 90° around the vertical axis. In this view the location of the PXDLS and RRTGAPPAL binding sites belonging to different subunits, that fall on the same face of the dimeric assembly, are clearly depicted.
C. Consensus peptide binding site. Stereo view of the consensus RRTGAPPAL peptide (yellow) bound to the t-CtBP1-S nucleotide-binding domain. Salt bridges (black lines) between R1-D220 and between R2-E164 are highlighted. The 2Fo-Fc electron density map at 2.85 Å resolution is shown as a blue grid.

Figure 6

Mutagenesis confirms the RRT contact residues of CtBP

A. Gal4AD-hZNF217 was examined for its ability to interact with Gal4DBD-CtBP2 wild type and with mutations in the PXDLS motif binding cleft (A58E), the newly identified RRT motif binding cleft (E181A/D237A) and with mutations in both clefts, in yeast two-hybrid assays. Interactions between the Gal4DBD CtBP2 mutants and with Gal4AD wild type CtBP were also examined as a positive control for the expression and folding of the CtBP2 mutants in yeast.
B. The ability of FLAG-hZNF217 to interact with YFP-CtBP2 wild type and with mutations in the PXDLS motif binding cleft (A58E), the RRT motif binding cleft (E181A/D237A) and in both clefts in co-immunoprecipitation experiments. COS-1 cells were transfected with the expression vectors indicated above each lane and whole cell extracts of those cells were immunoprecipitated (IP) with αFLAG antibody. Expression of each of the FLAG fused and YFP fused proteins are shown in the top two panels (5%

879 input). FLAG-ZNF217 immunoprecipitated by the αFLAG antibody and the resulting co-
880 immunopreciptated YFP-CtBP2 is shown in the bottom two panels (IP: αFLAG).
881 **C.** A summary diagram combining the results of interaction studies between ZNF217 and
882 wild type or mutant CtBP.
883
884 **Figure 7**
885 **CtBP repression activity does not depend on its ability to bind ZNF217**
886 **A.** A Western blot was performed to examine the expression levels of Gal4DBD fused
887 CtBP2 wild type and A58E, E181A/D237A and A58E/E181A/D237A mutants in
888 transiently transfected COS-1 cells.
889 **B-C**. Gal4DBD-CtBP2 constructs were tested for their ability to repress (B) basal firefly
890 luciferase reporter gene expression from the TK promoter or (C) LexA-VP16 activated
891 firefly luciferase reporter gene expression from the E1B promoter, in COS-1 cells
892 following transient transfection (n=4, ±SD, representative experiment).
893 **D.** Gal4DBD-CtBP2 constructs were tested for their ability to repress basal firefly
894 luciferase reporter gene expression from the TK promoter in CtBP-/- cells following
895 transient transfection (n=2, ±SD, representative experiment).
896
897 **Figure 8**
898 **ZNF217 represses gene transcription and this effect is partially dependent on its**
899 **ability to bind to CtBP**
900 **A.** A Western blot was performed to examine the expression levels of FLAG fused
901 hZNF217 wild type and ΔDL ΔRRT mutant in transiently transfected COS-1 cells.

B-E. FLAG-hZNF217 constructs were tested for their ability to repress (B) firefly luciferase reporter gene expression from the TK promoter in COS-1 cells (n=4, ±SD, representative experiment), (C) LexA-VP16 activated firefly luciferase reporter gene expression from the E1B promoter in COS-1 cells (n=4, ±SD, representative experiment), (D) basal firefly luciferase reporter gene expression from the *E-cadherin* promoter in HEK293 cells (n=2, range of values, representative experiment), (E) basal firefly luciferase reporter gene expression from the *E-cadherin* promoter in CtBP+/- and CtBP-/- cells (n=2, range of values, representative experiment), following transient transfection.

910 **Table I.** Data collection and refinement statistics

	t-CtBP1-S:NAD(H)/RRTGAPPAL
Data collection statistics[a]	
Space group	*P*6422
Unit cell dimensions (Å)	*a*=*b*=89.3, *c*=162.7
Resolution (Å)	2.85
Completeness (%)	99.8 (100)
Multiplicity	9.3 (9.5)
R_{merge}[b](%)	9.4 (39.9)
I/σ(I)	21.1 (4.6)
Refinement statistics and model quality	
R_{factor} (%)	22.7
R_{free}[c] (%)	27.5
No. of residues	331+9 (peptide)
No. of waters	19
No. of formate anion	1
rmsd bond lengths[d] (Å)	0.006
rmsd bond angles (°)	0.95

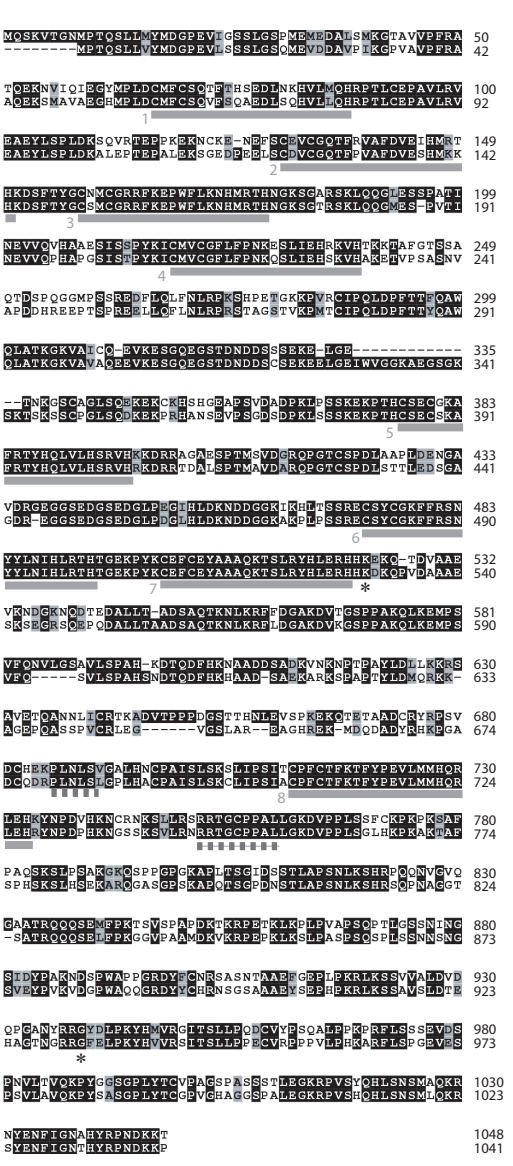
911

912 [a]Values in parentheses are for the highest resolution shell.

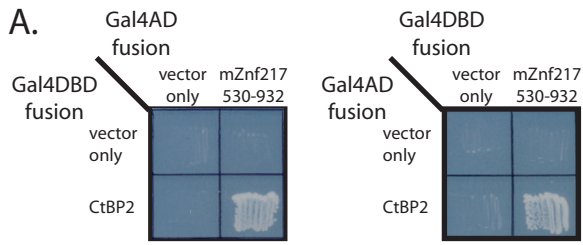
913 bRmerge = $\Sigma h \Sigma i \, |Ihi - < Ih >| / \Sigma h \Sigma i \, Ihi$.

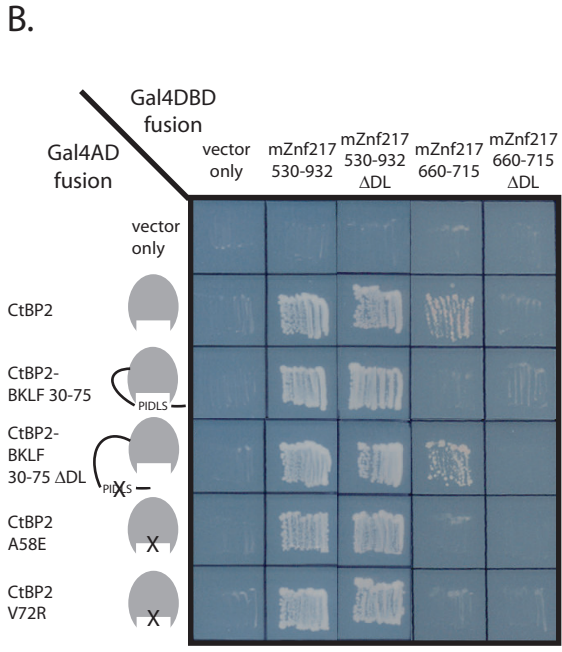
914 [c]R_{free} estimation is based on 10% of data withheld for cross-validation.

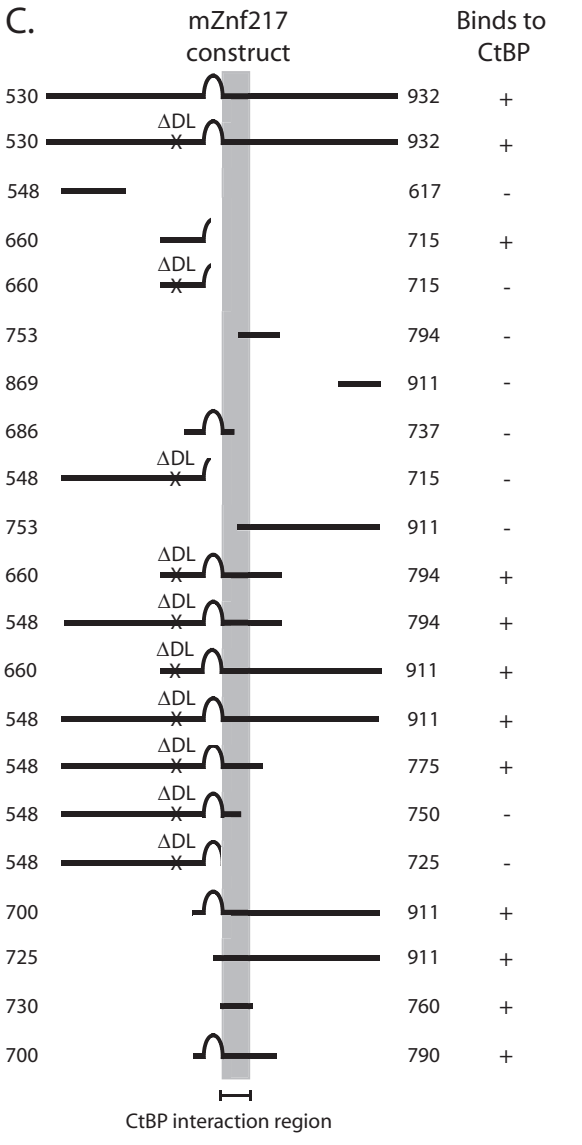
915 [d]The quality of the final model was assessed using the program PROCHECK (23).



Quinlan et al., Figure 1







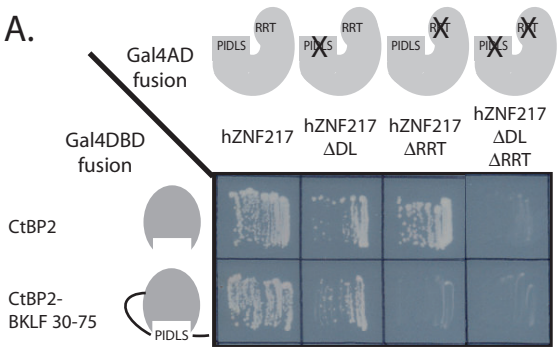
D.

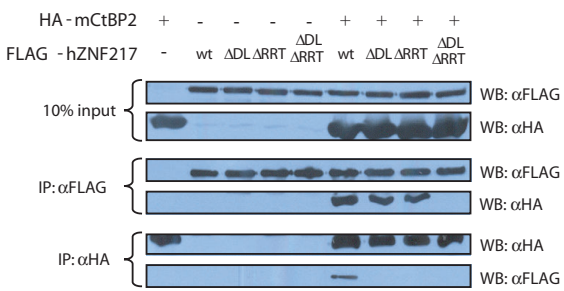
mZnf217 700-790 construct

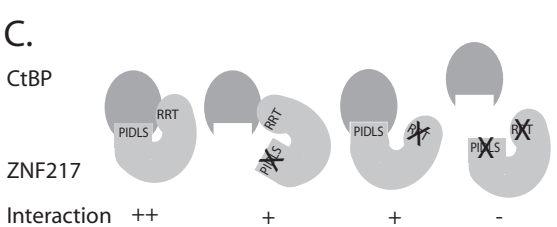
		Binds to CtBP
KSVLRNRRTGCPPALLGKDVP	wt	+++
AAALRNRRTGCPPALLGKDVP	Δ740-742	+++
KSV**AAA**RRTGCPPALLGKDVP	Δ743-745	+++
KSVLRN**AAA**GCPPALLGKDVP	Δ746-748	-
KSVLRNRRT**AAA**PALLGKDVP	Δ749-751	-
KSVLRNRRTGCP**AGA**LGKDVP	Δ752-754	-
KSVLRNRRTGCPPAL**AAA**DVP	Δ755-757	+++
KSVLRNRRTGCPPALLGK**AAA**	Δ758-760	+++
KSVLRN**A**RTGCPPALLGKDVP	Δ746	-
KSVLRNR**A**TGCPPALLGKDVP	Δ747	-
KSVLRNRR**A**GCPPALLGKDVP	Δ748	-
KSVLRNRRT**A**CPPALLGKDVP	Δ749	+
KSVLRNRRTG**A**PPALLGKDVP	Δ750 (C -A)	+++
KSVLRNRRTG**S**PPALLGKDVP	Δ750 (C -S)	+++
KSVLRNRRTGC**A**PALLGKDVP	Δ751	-
KSVLRNRRTGCP**A**ALLGKDVP	Δ752	-
KSVLRNRRTGCPP**G**LLGKDVP	Δ753	++
KSVLRNRRTGCPPA**A**LGKDVP	Δ754	+

RRTGXPPXL

Quinlan et al., Figure 2

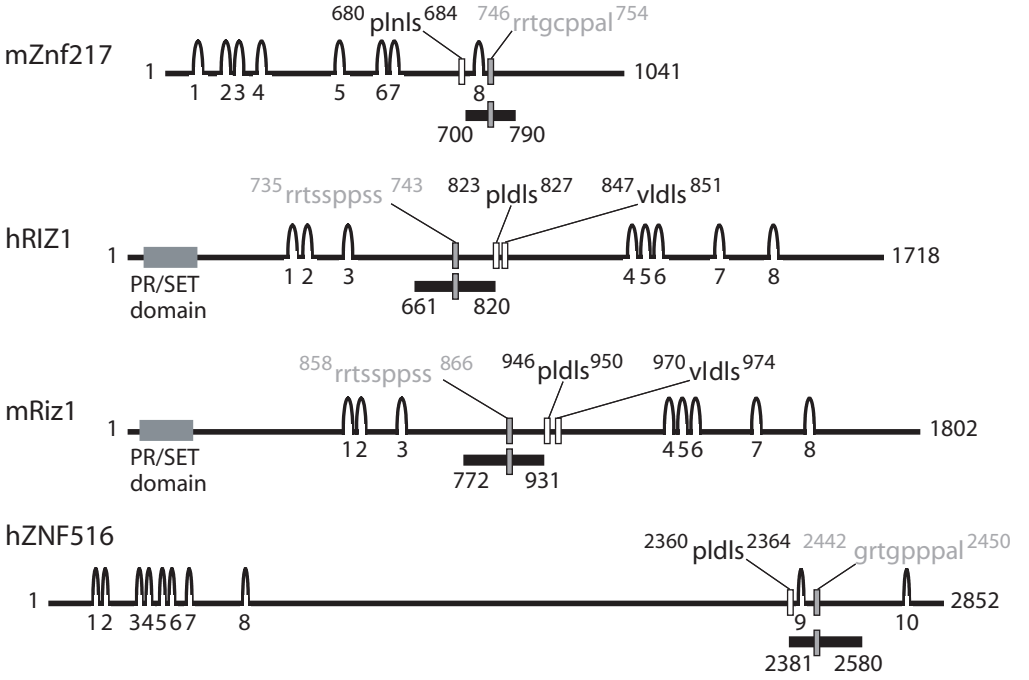






Quinlan et al., Figure 3

A.

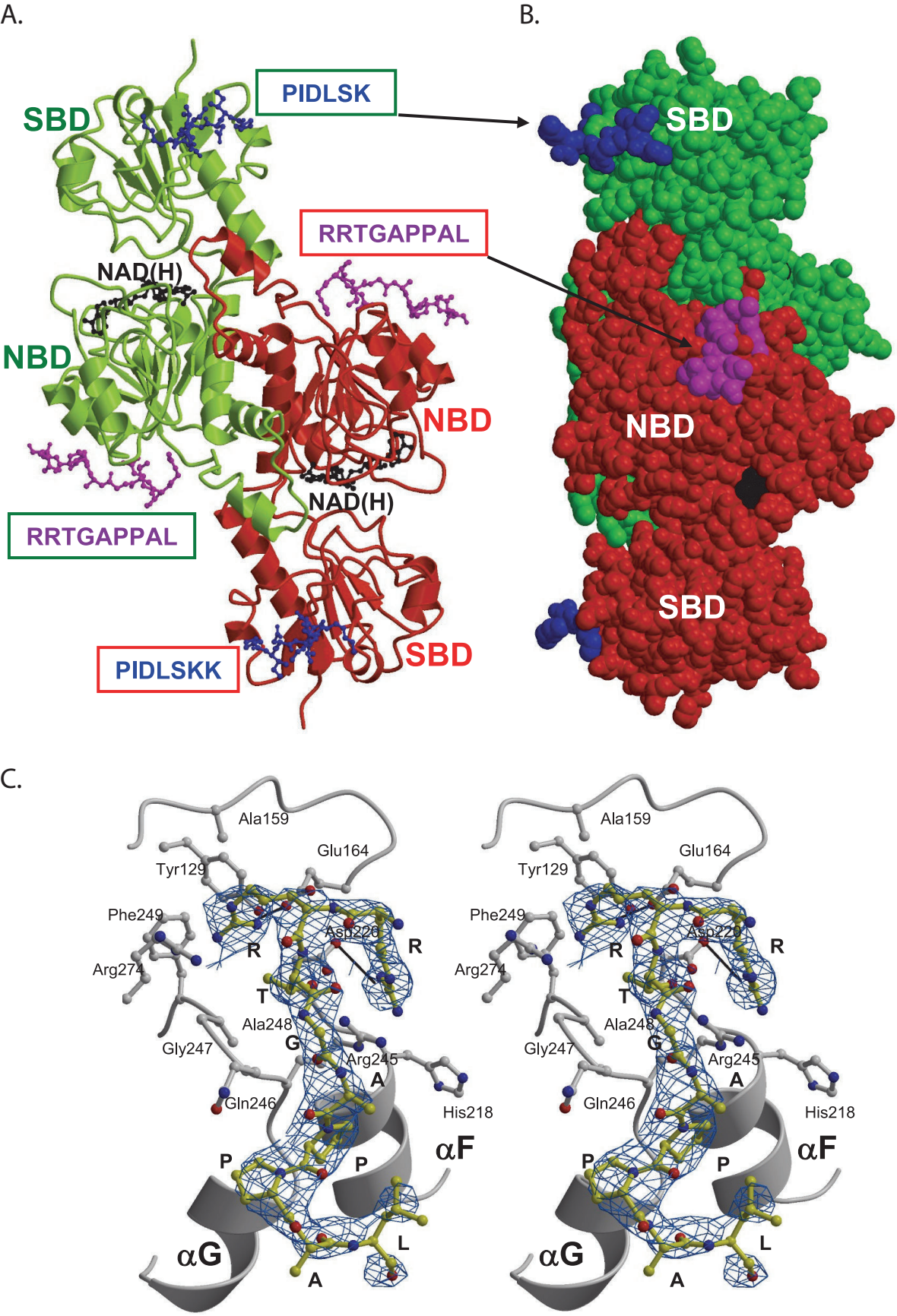


B.

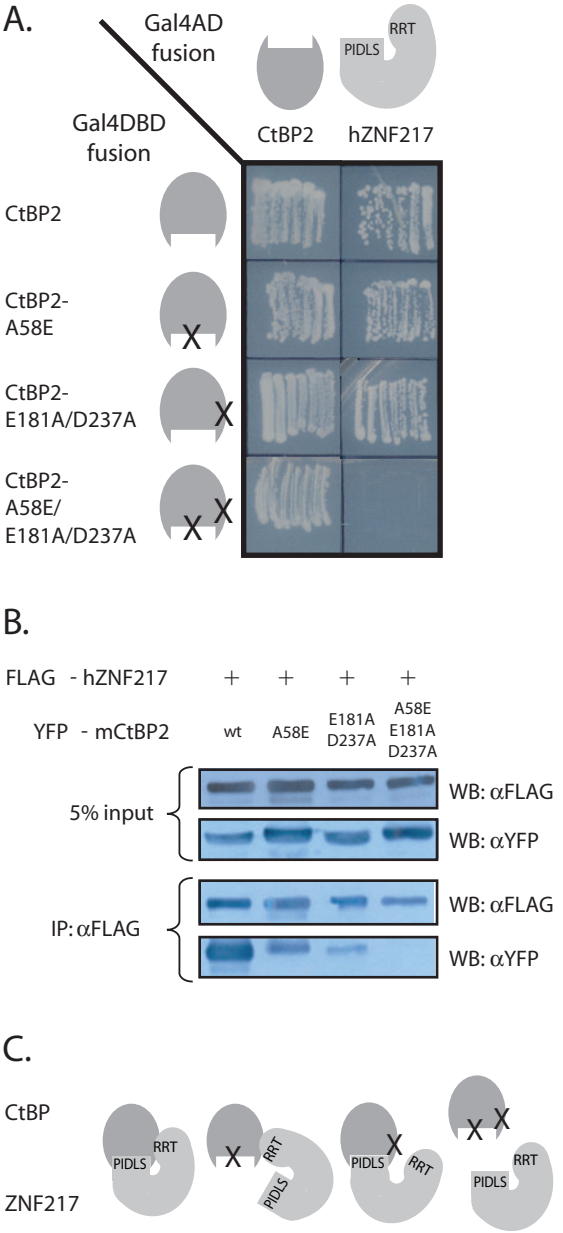
			Binds to wild type CtBP	Binds to PIDLS-pocket mutant CtBP
700 790	746 rrtgcppal 754	mZnf217 700 - 790	+++	+++
700 790	746 aaa gcppal 754	mZnf217 700 - 790 ΔRRT	-	nd
661 820	735 rrtsspp ss 743	hRIZ1 661 - 820	+++	+++
661 820	735 aaa sspps s 743	hRIZ1 661 - 820 ΔRRT	-	-
772 931	858 rrtssppss 866	mRiz1 772 - 931	+++	+++
772 931	858 aaa sppss 866	mRiz1 772 -931 ΔRRT	-	-
2381 2580	2442 grtgpppal 2450	hZNF516 2381 - 2580	+++	+++
2381 2580	2442 aaa gpppal 2450	hZNF516 2381 - 2580 ΔGRT	-	-

C.

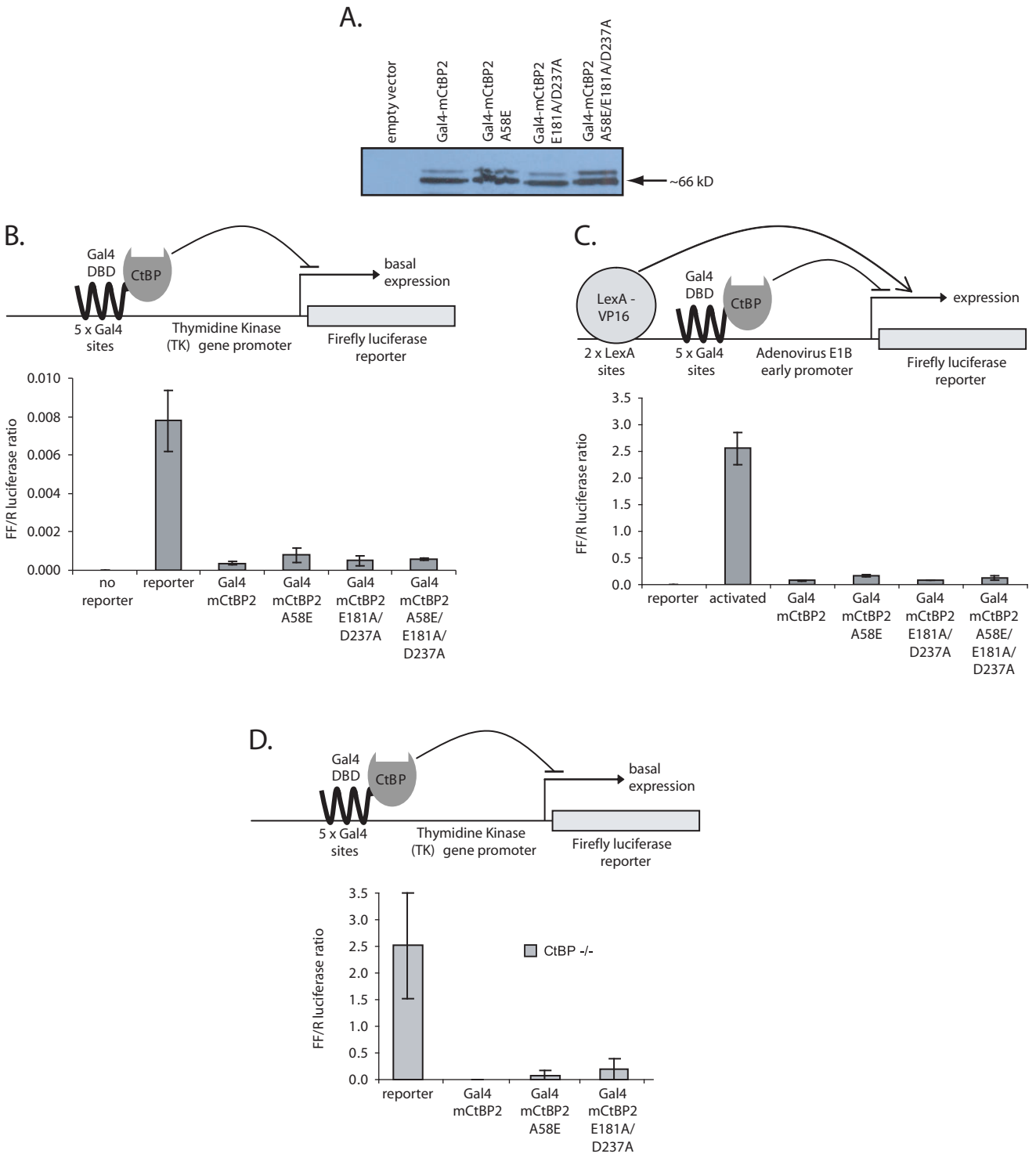
Quinlan et al., Figure 4



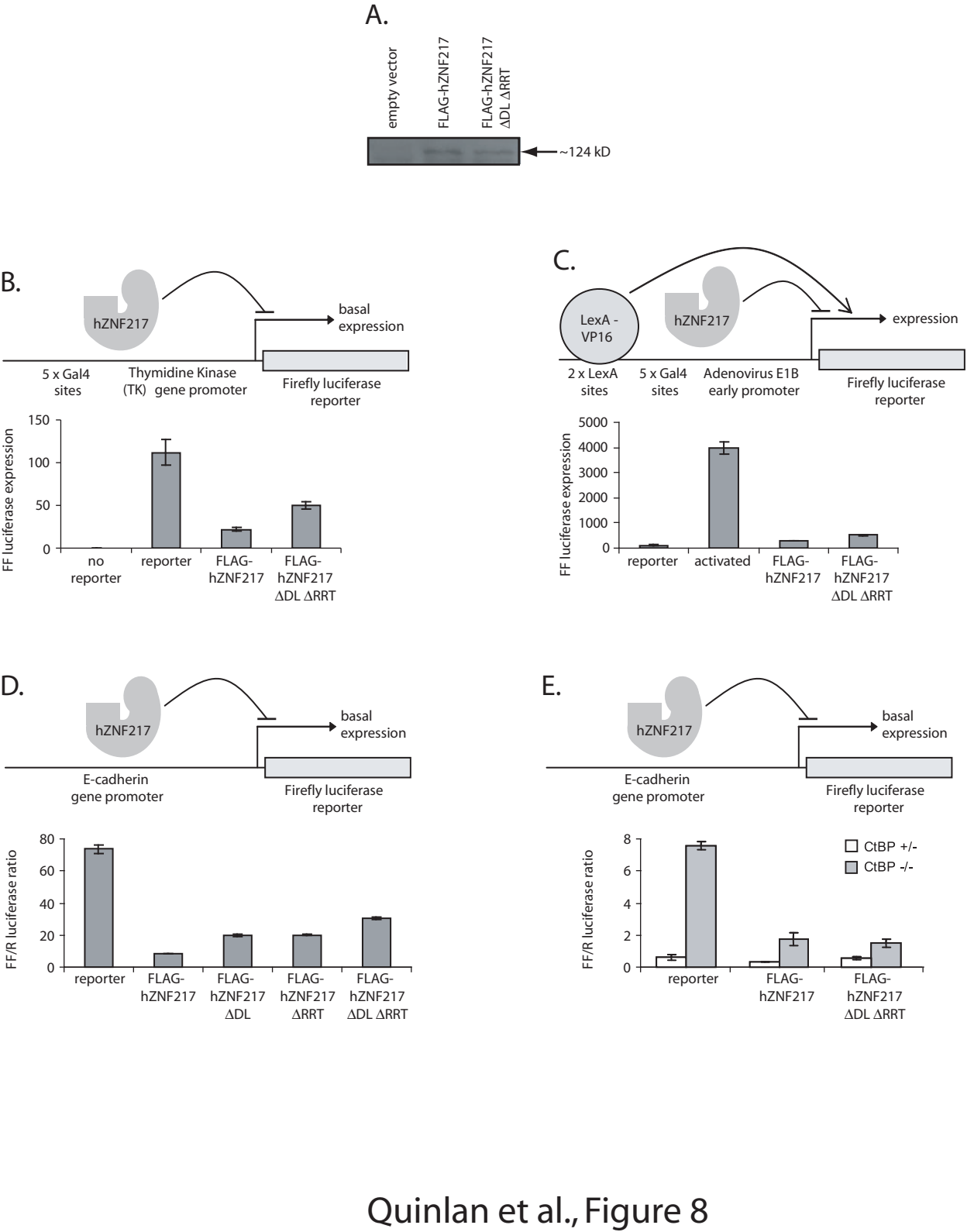
Quinlan et al, Figure 5



Quinlan et al., Figure 6



Quinlan et al., Figure 7



Quinlan et al., Figure 8