Cloning and characterization of mCtBP2, a co-repressor that associates with basic Krüppel-like factor and other mammalian transcriptional regulators

Jeremy Turner and Merlin Crossley¹

Department of Biochemistry, G08, University of Sydney, NSW, Australia, 2006

¹Corresponding author

e-mail: M.Crossley@biochem.usyd.edu.au

Basic Krüppel-like factor (BKLF) is a zinc finger protein that recognizes CACCC elements in DNA. It is expressed highly in erythroid tissues, the brain and other selected cell types. We have studied the activity of BKLF and found that it is capable of repressing transcription, and have mapped its repression domain to the N-terminus. We carried out a two-hybrid screen against BKLF and isolated a novel clone encoding murine C-terminal-binding protein 2 (mCtBP2). mCtBP2 is related to human CtBP, a cellular protein which binds to a Pro-X-Asp-Leu-Ser motif in the C-terminus of the adenoviral oncoprotein, E1a. We show that mCtBP2 recognizes a related motif in the minimal repression domain of BKLF, and the integrity of this motif is required for repression activity. Moreover, when tethered to a promoter by a heterologous DNA-binding domain, mCtBP2 functions as a potent repressor. Finally, we demonstrate that mCtBP2 also interacts with the mammalian transcription factors Evi-1, AREB6, ZEB and FOG. These results establish a new member of the CtBP family, mCtBP2, as a mammalian co-repressor targeting diverse transcriptional regulators.

Keywords: BKLF/CtBP/mCtBP2/transcriptional repression/zinc finger proteins

Introduction

Basic Krüppel-like factor (BKLF) (Crossley *et al.*, 1996) is a zinc finger protein that belongs to the subfamily of Krüppel-like proteins, which includes erythroid Krüppel-like factor (EKLF) (Miller and Bieker, 1993), lung Krüppel-like factor (LKLF) (Anderson *et al.*, 1995) and gut-enriched Krüppel-like factor (Shields *et al.*, 1996) [also known as endothelial zinc finger protein, EZF (Garrett-Sinha *et al.*, 1996; Yet *et al.*, 1998)]. Members of this subfamily contain three characteristic Cys–Cys: His–His Krüppel-like zinc fingers and recognize CACCC motifs in the promoters and enhancers of various genes. The founding member of the family, EKLF, is expressed in erythroid and mast cells and specifically recognizes the CACCC-box in the β -globin promoter (Miller and Bieker, 1993). Naturally occurring mutations which inhibit the binding of EKLF to this element are associated with clinical β -thalassaemia (Feng *et al.*, 1994), and knockout studies in mice have demonstrated that EKLF is required

for transcription of the β -globin gene (Nuez *et al.*, 1995; Perkins *et al.*, 1995).

Specific target genes regulated by other members of the family have not been as clearly defined. LKLF was identified originally as a CACCC-box protein that was expressed abundantly in the lung, as well as in a number of other tissues (Anderson *et al.*, 1995). Targeted mutation of the LKLF gene in mice has revealed important roles for this protein in T cell activation (Kuo *et al.*, 1997b) and in vascular endothelial cells (Kuo *et al.*, 1997a). GKLF/EZF is also expressed in vascular endothelial cells (Yet *et al.*, 1998), as well as in the gut (Shields *et al.*, 1996) and in the epidermal layer of skin (Garrett-Sinha *et al.*, 1996). It is believed to be instrumental in controlling the proliferation of epithelial tissues, and there is some evidence that it can act as either an activator or a repressor of transcription (Yet *et al.*, 1998).

Like EKLF, BKLF is expressed at high levels in erythroid cells, but it is found additionally in other tissues, most notably the brain (Crossley *et al.*, 1996). *In vitro* binding studies have demonstrated that like EKLF, it binds to the CACCC motifs found in the β -globin promoter as well as to CACCC-boxes in the globin locus control regions, and several of the heme biosynthetic genes. The exact target genes regulated by BKLF, however, have not been defined. BKLF knockout mice have been generated, and initial analysis indicates that they suffer from a myeloproliferative disorder, and thus it seems likely that BKLF plays a role in hematopoiesis (Perkins *et al.*, 1997). It has been demonstrated previously that BKLF can activate transcription from a minimal promoter containing a single BKLF-binding site, although activation was significantly weaker than achieved by other Krüppel family proteins and was only observed with very high levels of BKLF (Crossley *et al.*, 1996). In this study, we have extended our analysis of the functional activity of BKLF and demonstrate that it can also act as a potent repressor. We show that the minimal repression domain maps to a 74 amino acid region within the N-terminus, and that this domain retains its activity when fused to a heterologous DNA-binding domain. In an effort to understand the mechanism by which BKLF represses transcription, we carried out a yeast two-hybrid screen against the repression domain of BKLF, and identified a novel cofactor protein which we refer to as murine C-terminal-binding protein 2 (mCtBP2).

mCtBP2 is a new member of the CtBP family of related proteins. The first member of the family, human CtBP (now hCtBP1), was identified initially as a cellular protein that bound to the C-terminal region of the adenoviral E1a oncoprotein (Boyd *et al.*, 1993; Schaeper *et al.*, 1995). It was shown that deletion of this region dramatically enhanced the tumorigenicity of E1a, suggesting that the binding of hCtBP1 significantly modulated the effect of

E1a *in vivo*. Subsequently, it was demonstrated that an equivalent deletion to a gal4–E1a fusion protein dramatically enhanced its ability to activate transcription (Sollerbrant *et al.*, 1996). This result suggested that hCtBP1 might be involved in the negative regulation of transcription. Recently, a *Drosophila* homologue of human CtBP1 has been identified as a partner protein of the transcriptional repressors Hairy, Snail, Knirps and Enhancer of split [E(spl)] mδ (Nibu *et al.*, 1998; Poortinga *et al.*, 1998). It has also been suggested that dCtBP binds the repressor Krüppel (Nibu *et al.*, 1998). The *in vivo* studies in *Drosophila* indicate that dCtBP is a genuine co-repressor protein that plays an important role in repressing gene expression during development.

Human CtBP1 recognizes signature Pro-X-Asp-Leu-Ser motifs in the C-terminus of E1a (Boyd *et al.*, 1993; Schaeper *et al.*, 1995; Sollerbrant *et al.*, 1996). *Drosophila* CtBP has been shown to bind similar motifs found within the C-terminus of Hairy and E(spl)mδ (Poortinga *et al.*, 1998), and within the body of the Snail and Knirps proteins (Nibu *et al.*, 1998). The mammalian BKLF protein contains a similar motif within its repression domain. We show that mCtBP2 associates with BKLF through this element and that it is instrumental in mediating the repressor activity of BKLF.

We searched the sequences of previously reported mammalian repressor proteins and found that potential CtBP recognition motifs occurred in several of these proteins, most notably in the zinc finger proteins Evi-1 (Morishita *et al.*, 1988), AREB6 (Watanabe *et al.*, 1993) and ZEB (Genetta *et al.*, 1994). The signature motif also occurs in Friend of GATA (FOG) (Tsang *et al.*, 1997), and a related *Drosophila* repressor protein, U-shaped (Haenlin *et al.*, 1997). We used the two-hybrid system to demonstrate that mCtBP2 is able to associate specifically with the mammalian proteins AREB6, ZEB, Evi-1 and FOG. Our results indicate that mCtBP2 is a co-repressor protein that can associate with a number of mammalian transcription factors.

Results

BKLF can act as a potent repressor of transcription

In order to assess the function of BKLF, we carried out co-transfection assays in *Drosophila* Schneider line 2 (SL2) cells. These cells are used conventionally for the study of CACCC-box factors since they are devoid of ubiquitous CACCC proteins (such as Sp1) which complicate interpretations of experiments in mammalian cell lines. It previously has been reported that CACCC-box factors can synergize with the glucocorticoid receptor (GR) on appropriately configured composite CACCC-glucocorticoid response elements (GREs) (Schüle *et al.*, 1988; Strähle *et al.*, 1988). With a view to assessing whether BKLF would co-operate with or quench the activity of GRs, we prepared a test promoter containing three copies of a composite CACCC-GRE site [which had a spacing that previously had been reported to give maximal synergy (Strähle *et al.*, 1988)] (Figure 1). First we tested the activity of BKLF alone, and in the absence of the glucocorticoid dexamethasone, and found it reduced the expression of this reporter gene slightly (Figure 1A, columns 1–4). In contrast, EKLF potently activated transcription (columns 5–7). This result suggested that BKLF

was capable of repressing transcription. We next co-expressed increasing amounts of BKLF together with a constant amount of EKLF and found that BKLF could silence the activation mediated by EKLF (Figure 1B). We noted that this inhibition occurred with relatively low levels of BKLF (gel retardation data not shown), suggesting that the effect was not due entirely to competition between BKLF and EKLF for the three CACCC-boxes and that it might be due to an active repression domain within BKLF. We therefore tested a deletion derivative of BKLF, in which the entire N-terminus of the protein had been removed, leaving only the zinc finger DNA-binding domain (Figure 1C). This truncated protein was able to compete with and inhibit EKLF-mediated activation to some extent, but was significantly less effective than full-length BKLF. These results raised the possibility that BKLF contained a repression domain, capable of repressing activated transcription.

We then proceeded to assess whether BKLF could repress activation mediated by the GR in the presence of dexamethasone. We found that BKLF potently repressed GR-activated transcription (Figure 1D, columns 1–4). In contrast, EKLF (and Sp1) (Figure 1D, columns 5–7, and data not shown) efficiently synergized with the GR on our promoter. We also tested the deletion derivative of BKLF, that only contains the zinc fingers, and found that it synergized weakly with the GR (Figure 1E). We conclude that BKLF contains a repression domain in its N-terminus that can silence both EKLF- and GR-activated transcription. In order to map this repression domain more precisely, we prepared a number of additional deletion constructs. We first used gel shift assays to verify that these mutants were all expressed at equivalent levels (data not shown), and then tested their ability to silence GR-activated transcription (Figure 1E). Deletion of the first 74 amino acids dramatically reduced repression activity, suggesting that the relevant domain included this region.

The N-terminal domain of BKLF interacts with the co-repressor mCtBP2

In order to identify possible cofactors which might bind to this domain and mediate BKLF's ability to repress transcription, we carried out a yeast two-hybrid screen using a murine erythroleukaemia (MEL) cell library (Tsang *et al.*, 1997) and a gal4 DNA-binding domain (DBD)–BKLF(1–268) fusion bait protein. This bait protein contains the repression domain of BKLF and adjoining sequences, but does not contain the BKLF zinc fingers. The analysis of $\sim 1.2 \times 10^6$ colonies led to the isolation of 110 that were His⁺, and 35 that were both His⁺ and β -galactosidase positive. Sequencing of these isolates revealed a full-length clone that encodes a member of the CtBP family of proteins. This protein is most highly related to a sequence generated through the assembly of database expressed sequence tags (ESTs) termed mCtBP2 (Katsanis and Fisher, 1998); however, it differs at four amino acids, and contains an additional 25 amino acids at the N-terminus. The sequences of recognized CtBP family proteins are shown in Figure 2.

Finally we showed that as well as interacting with gal4DBD–BKLF(1–268), mCtBP2 interacted with gal4DBD–full-length BKLF (Figure 3), but did not interact with three control baits, p53(72–390) (Iwabuchi *et al.*,

Fig. 1. BKLF represses transcription of a (CACCC-GRE)₃-CAT reporter gene in SL2 cells, whilst EKLF activates transcription. **(A)** BKLF represses basal expression in a dose-dependent manner, whereas EKLF activates. In addition to 500 ng of reporter plasmid in all columns, 0, 60 and 250 ng, and 1 µg of pPac-BKLF and 60 ng, 250 ng and 4 µg of pPac-EKLF were used in columns 1–4 and 5–7, respectively. **(B)** Increasing amounts of BKLF compete with and inhibit activation by EKLF. pPac-EKLF (500 ng) was used alone in column 1 and together with 30, 60 and 250 ng, and 1 µg of pPac-BKLF in columns 2–5, respectively. **(C)** Increasing amounts of BKLF zinc finger DNA-binding domain alone are less effective than full-length BKLF but can compete with and inhibit activation by EKLF at high concentrations. pPac-EKLF (500 ng) was used alone in column 1 and together with 30, 60 and 250 ng, and 1 µg of pPac-BKLF-finger domain in columns 2–5, respectively. **(D)** Increasing amounts of BKLF repress transcription activated by the GR, whereas EKLF synergistically activates transcription. pPac-GR (500 ng) was used alone in column 1 and together with 10, 30 and 60 ng of pPac-BKLF in columns 2–4, and with 60 ng, 250 ng and 1 µg of pPac-EKLF in columns 5–7, respectively. **(E)** Deletion analysis demonstrates that the first 74 amino acids of BKLF are required for repression of GR-activated transcription. pPac-BKLF deletion constructs (60 ng), as shown, were used together with 500 ng of pPac-GR.

1993), the N-finger of GATA-1 (Tsang *et al.*, 1997) or human lamin C(66–230) (Bartel *et al.*, 1993). We also exchanged the bait and prey, cloning BKLF(1–268) in-frame with the gal4 activation domain (AD) and mCtBP2 in-frame with the gal4DBD. A strong interaction was also observed in this experiment (Figure 3).

The repression domain and mCtBP2-binding domain of BKLF co-localize

We next used the two-hybrid assay to determine whether mCtBP2 bound to the region of BKLF previously implicated in transcriptional repression. We constructed a panel of gal4DBD–BKLF deletion constructs and used the two-

mCtBP2	MALVDKHKVKRQLDRICEGIRPQ	IMNGPLHPRPLVALLDGRDCTVEMPILKDLATVAFCDQAQSTQEIHEKVLN	74
hCtBP2	MALVDKHKVKRQLDRICEGIRPQ	IMNGPLHPRPLVALLDGRDCTVEMPILKDLATVAFCDQAQSTQEIHEKVLN	74
hCtBP1	MGSSHLNKG...LPLGVRPP	IMNGPLHPRPLVALLDGRDCTVEMPILKDVATVAFCDQAQSTQEIHEKVLN	68
rCtBP1	MGSSHLNKG...LPLGVRPP	IMNGPLHPRPLVALLDGRDCTVEMPILKDVATVAFCDQAQSTQEIHEKVLN	68
dCtBP	MDKNLMPKRSR...IDVKG	FANGPLQARPLVALLDGRDCSIEMPILKDVATVAFCDQAQSTQEIHEKVLN	68
cCtBP	MAGGRDEDEEYMGNGKANLNPASPHTPHPVLPDPQALNDILSMTST	RMNGPSSSRPLVALLDGRDCSVEMPILKDVATVAFCDQAQSTQEIHEKVLN	100
	***	*****	
mCtBP2	EAVGAMMYHTITLTREDLEKFKALRVIVRIGSGYDNVDIKAAGELGIAVC	NIPSAVEETADSTVCHILNLYRRNTWLYQALREGTRVQSVEQIREVASG	174
hCtBP2	EAVGAMMYHTITLTREDLEKFKALRVIVRIGSGYDNVDIKAAGELGIAVC	NIPSAVEETADSTVCHILNLYRRNTWLYQALREGTRVQSVEQIREVASG	174
hCtBP1	EAVGALMYHTITLTREDLEKFKALRIIVRIGSGFDNIDIKSAGDLGIAVC	NVPAASVEETADSTLCHILNLYRRAT...GCTRRCGRAHESRASSR	161
rCtBP1	EAVGALMYHTITLTREDLEKFKALRIIVRIGSGFDNIDIKSAGDLGIAVC	NVPAASVEETADSTLCHILNLYRRAT...GCTRRCGRAHESRASSR	161
dCtBP	EAVGALMWHTITLTREDLEKFKALRIIVRIGSGTDNIDVKAAGELGIAVC	NVPGYGVVEEADTMTCLILNLYRRTYWLANMVREGKFTGPEQVREAAHG	168
cCtBP	EAVAALMYHSIKLEKEDLEKFKVLKVVFIRIGYGIDNIDVKAATELGIAVC	HAPGDYVEDVADSTLSLILDLFRTRYWHAKSYSERKTIGADQVRENAVG	200
	*****	*****	
mCtBP2	AARIRG...ETLGLIGFGRT...GQAVAVRAKAFGFSVIFYDPYLQDQIE	RS...LGVQRVYTLQDLLYQSDCVSLHCLNNEHNHHLINDFTIKQMRQG	264
hCtBP2	AARIRG...ETLGLIGFGRT...GQAVAVRAKAFGFSVIFYDPYLQDQIE	RS...LGVQRVYTLQDLLYQSDCVSLHCLNNEHNHHLINDFTIKQMRQG	264
hCtBP1	SARWRPRCQDPRGDLGHHRTWSRGAGSGAAGQVRVGFNVLFDYDYLSDGVE	RA...LGLQRVSTLQDLLFHSDCVTLHCLNNEHNHHLINDFTVVKQMRQG	257
rCtBP1	SARWRPRCQDPRGDLGHHRTWSRGAGSGAAGQVRVGFNVLFDYDYLSDGVE	RA...LGLQRVSTLQDLLFHSDCVTLHCLNNEHNHHLINDFTVVKQMRQG	257
dCtBP	CARIRG...DTLGLVGLGRI...GSAVALRAKAFGFSVIFYDPYLDPGID	KS...LGLTRVYTLQDLLFQSDCVSLHCLNNEHNHHLINFTIKQMRPG	258
cCtBP	SKKVRG...SVLGILGCGRV...GTAVGLRARAFGLHIIIFYDPFVREGHD	KASVTDLLELFRVYTMDEFMSRSDCISLHCLNCLGDETRGIIINADSLRQCKSG	300
	*****	*****	
mCtBP2	AFLVNAARGGLVDEKALAQAALKEGRIRGAALDVHESEPFSS.FAQGPLKDA	PNLICTPHTAWYSEQASLEMRERAAATEIRRAITGRIPESLRNCVNKEFFV	363
hCtBP2	AFLVNAARGGLVDEKALAQAALKEGRIRGAALDVHESEPFSS.FAQGPLKDA	PNLICTPHTAWYSEQASLEMRERAAATEIRRAITGRIPESLRNCVNKEFFV	363
hCtBP1	AFLVNTARGGLVDEKALAQAALKEGRIRGAALDVHESEPFSS.FSQGPLKDA	PNLICTPHAAWYSEQASLEMRERAAATEIRRAITGRIPDSLKNCVNKDHLT	356
rCtBP1	AFLVNTARGGLVDEKALAQAALKEGRIRGAALDVHESEPFSS.FSQGPLKDA	PNLICTPHAAWYSEQASLEMRERAAATEIRRAITGRIPDSLKNCVNKDHLT	356
dCtBP	AFLVNTARGGLVDDTLALALKQGRIRAAALDVHENEPYN.VFGALKDA	PNLICTPHAAFFSDASATELREMAATEIRRAIVGNIIPDVLNRCVNKEYFM	357
cCtBP	VYIVNTSHAGLINENDLAAALKNGHVKGAAALDVHDSVRFDPNCLNPLVGC	PNIINTPHSAWMTEASCKDLRINAAKEIRKAINGRCQDLTHCINKEAVM	400
	*****	*****	
mCtBP2	TSAPWSVIDQQAHPHPELN...GATYRYPGIVGVAPGG.	LPPAMEGIIPGGIPVTHNLPTVAHPSQAPSPNQPTKH	435
hCtBP2	TSAPWSVIDQQAHPHPELN...GATYRYPGIVGVAPGG.	LPPAMEGIIPGGIPVTHNLPTVAHPSQAPSPNQPTKH	435
hCtBP1	AATHWASMDPAVHVHPELN...GAAYRYPPGVGVAPGTG.	IPAAVEGIVPSAMSLSHGLPPVHPPHAPSPGQTVKP	428
rCtBP1	AATHWASMDPAVHVHPELN...GAAYRYPPGVGVAPGTG.	IPAAVEGIVPSAMSLSHGLPPVHPPHAPSPGQTVKP	428
dCtBP	RTPP.....AAAAGVAAASV.	YPEGKLMISNQEK	386
cCtBP	RNSNPINRRITSSAHLNLMGFPTLPNFPMSMSPHFPYPNLLAMGAMQG	ALNPFMNGALPFPNPAALSSSLAAQAANAQRGSPANRSSRSSPSPTNK	500
	*****	*****	
mCtBP2	GDNREHPNEQ		445
hCtBP2	GDNREHPNEQ		445
hCtBP1	EADRDHASDQL		439
rCtBP1	EADRDHASDQL		439
cCtBP	SSVSPGNNGHVKTEPSSPAKIEVDIAENDKHSMMTFLQRLIAPNGDSGA	STADSGIEGGDKVKVQSDGDNEMEDMEVIDAEKLKEELNIGQLEEPPEIS	600
	*****	*****	
cCtBP	VGLNNGNRINIDEQPLAT		618

Fig. 2. The amino acid sequences of existing members of the CtBP family. Conserved residues are indicated by asterisks. The putative active site histidine conserved in certain dehydrogenase enzymes is shown in bold, with an arrowhead. Accession numbers are as follows: mCtBP2, AF059735; hCtBP1, G1063638; hCtBP2, G2909777; rCtBP1, g1585432; dCtBP, G2950374/G2982720; cCtBP, Q20596.

hybrid system to test which of these interacted with mCtBP2 (Figure 3). Significantly, deletion of the N-terminal repression domain abolished the interaction with mCtBP2, but not with other BKLF-interacting proteins isolated from the library screen (data not shown). This result suggests that mCtBP2 binds within the BKLF repression domain.

Within this N-terminal repression domain of BKLF, we noted the motif Pro-Val-Asp-Leu-Thr, which is closely related to a recognized hCtBP1-binding site in E1a (Pro-Val-Asp-Leu-Ser) (Schaeper *et al.*, 1995). It previously has been reported that mutation of the core Asp-Leu residues to Ala-Ser abolished the interaction between E1a and hCtBP1 (Schaeper *et al.*, 1995). We therefore used the two-hybrid system to test whether the analogous mutation in either the original BKLF(1–268) bait or in the full-length BKLF bait affected their interaction with mCtBP2. In both cases, this mutation significantly impaired the interaction (Figure 3), suggesting that mCtBP2 directly targets this motif in BKLF. Furthermore, exchange of the bait and prey (as above) also demonstrated

that this mutation impaired the interaction between BKLF and mCtBP2 (Figure 3).

mCtBP2 binds BKLF *in vitro*

We next sought to determine whether the interaction between mCtBP2 and BKLF could be detected *in vitro*. We prepared GST–BKLF fusion proteins immobilized on agarose beads and tested their ability to retain *in vitro* translated ³⁵S-radiolabelled mCtBP2. GST–full-length BKLF and GST–BKLF(1–268) fusions were able to bind mCtBP2 efficiently, whereas proteins containing the core Ala–Ser mutation, the zinc fingers of BKLF alone or GST alone could not retain mCtBP2 (Figure 4A). Thus, the interaction of mCtBP2 with BKLF *in vitro* requires the N-terminal repression domain, and is dependent on the integrity of the Pro-Val-Asp-Leu-Thr motif.

BKLF, mCtBP2 and CACCC-box DNA form a ternary complex

We also investigated whether mCtBP2 could associate with full-length BKLF bound to a CACCC sequence in

Fig. 3. mCtBP2 recognizes a Pro-Val-Asp-Leu-Thr motif in the repression domain of BKLF. A deletion series of gal4DBD-BKLF fusions was assayed for binding to a gal4AD-mCtBP2 fusion using the yeast two-hybrid system. Mutation of the core CtBP-binding site from Asp-Leu to Ala-Ser impairs the interaction. An interaction was also observed when the bait and prey were exchanged. A gal4DBD-mCtBP2 fusion interacts with a gal4AD-BKLF(1-268) fusion, but not the mutant gal4AD-BKLF(1-268)mut, where the core CtBP-binding motif is mutated as above. The rate of yeast growth on His/Leu/Trp-deficient medium observed after 30 h incubation is shown on the right.

Fig. 4. mCtBP2 associates with BKLF *in vitro*. (A) GST pull-down experiments show that mCtBP2 recognizes the core motif in BKLF. The amounts of ³⁵S-radiolabelled mCtBP2 retained by GST (lane 2), GST-BKLF fingers (lane 3), GST-BKLF(1-268) (lane 4), GST-mutant BKLF(1-268) (containing the core Ala-Ser mutation) (lane 5) and GST-full-length BKLF (lane 6) are shown. The input lane (lane 1) contains 50% of the ³⁵S-radiolabelled mCtBP2 used in the binding assays. Equivalent amounts of GST fusion proteins were used in each lane. (B) Gel mobility shift experiments using BKLF and mutant BKLF (containing the core Ala-Ser mutation) and mCtBP2 co-expressed in SL2 cells show that BKLF and mCtBP2 can form a complex on a double-stranded CACCC-box oligonucleotide. Lanes 1-16 contain SL2 cell nuclear extracts prepared from cells transfected with the following expression vectors: lane 1, 1 µg of pPac alone; lane 2, 1 µg of pPac-BKLF; lanes 3-6, 1 µg of pPac-BKLF and 10 and 100 ng, 1 and 5 µg of pPac-mCtBP2, respectively; lane 7, 1 µg of pPac-mutant BKLF; lanes 8-11, pPac-mutant BKLF and pPac-mCtBP2 as for lanes 3-6; lane 12, 1 µg of pPac; lane 13, 1 µg of pPac-BKLF and 100 ng of pPac-mCtBP2; lane 14, as for 13 but with anti-BKLF sera; lane 15, 1 µg of pPac-BKLF and 1 µg of pPac-mCtBP2; lane 16, as for 15 but with anti-BKLF antisera. The arrows indicate the site of migration of BKLF-DNA and the BKLF-mCtBP2-DNA complexes as indicated.

DNA. We expressed full-length BKLF, or full-length BKLF containing the core Ala-Ser mutation, in SL2 cells and used the nuclear extracts in gel mobility shift experiments with the β-globin CACCC site as a probe

(Figure 4B). Retarded complexes corresponding to BKLF and mutant BKLF can readily be observed (Figure 4B, lanes 2 and 7). However, when increasing amounts of mCtBP2 are co-expressed, the intensity of the BKLF-

DNA complex diminishes, whereas the intensity of the mutant BKLF–DNA complex is unaffected (Figure 4B, compare lanes 2–6 and 7–11). The reduction in the observable BKLF–DNA complex suggests that either mCtBP2 is interfering with BKLF's ability to bind DNA, or that a higher molecular weight complex is formed but is obscured by another band. Additional experiments revealed the presence of a high molecular weight complex which co-migrates with a fainter endogenous band (Figure 4B, lanes 13 and 15). This complex is only observed when intact BKLF is co-expressed with mCtBP2, and addition of anti-BKLF antisera leads to the disruption of this complex and the formation of a new supershifted complex (Figure 4B, lanes 14 and 16). This latter complex is of only slightly lower mobility than the BKLF–mCtBP2–DNA complex, and co-migrates with the BKLF–Ab–DNA complex (data not shown), suggesting that the antibody displaces mCtBP2. These results indicate that BKLF and mCtBP2 can form a complex on a CACCC-box oligonucleotide.

mCtBP2 is a co-repressor

To determine whether BKLF was repressing transcription primarily by recruiting mCtBP2 (or a related family member) to the promoter, we tested whether the core Ala–Ser mutation [which reduces the binding of mCtBP2 to BKLF (Figures 3 and 4)] would impair BKLF's ability to repress transcription in the SL2 cell assay (Figure 5A). In this system, full-length BKLF repressed GR-mediated activation in a dose-dependent manner, to >35-fold (Figure 5A, columns 2–4). In contrast, the mutant version of BKLF (which contains the core Ala–Ser mutation) caused only a 5-fold repression, which was not dose dependent (Figure 5A, columns 5–7). This result strongly suggests that BKLF is repressing transcription by recruiting CtBP to the reporter gene promoter. In this case, we conclude that BKLF is associating directly with endogenous dCtBP.

We carried out an additional experiment to demonstrate that mCtBP2 itself was capable of directly silencing gene expression. We constructed a chimaeric gene in which the coding sequence of mCtBP2 was fused directly to the zinc finger region of BKLF (Figure 5B, columns 1–6). In this case, the entire BKLF repression domain is removed and replaced by mCtBP2. mCtBP2 can therefore be targeted to the CACCC-box promoter, not by piggy-backing onto BKLF, but directly by means of the linked zinc finger domain. We tested the ability of this mCtBP2–finger chimaeric protein to repress GR-mediated transcriptional activation in SL2 cells. As shown in Figure 5B (columns 1–6), it repressed transcription efficiently in a dose-dependent manner. This result indicates that mCtBP2 is capable of repressing transcription, at least in *Drosophila* SL2 cells.

Additional experiments in mammalian cells provide further evidence that BKLF and mCtBP2 associate to mediate repression. We constructed a gal4DBD–BKLF(1–268) fusion protein and tested its ability to repress a gal4-dependent promoter driving growth hormone expression in NIH 3T3 cells (Figure 5C). We chose these cells since they express both endogenous BKLF and mCtBP2 (unpublished results). As shown in Figure 5C (columns 1–4), this fusion protein efficiently silenced gene expres-

sion in a dose-dependent manner. This result is consistent with our previous conclusion that the N-terminus of BKLF contains a repression domain. In order to determine whether the abrogation of mCtBP2 binding influenced BKLF's ability to repress transcription, we tested whether a mutant gal4DBD–BKLF(1–268) fusion protein, containing the core Ala–Ser substitution, could also mediate repression (Figure 5C, columns 5–7). In contrast to the result in SL2 cells, in NIH 3T3 cells this mutation had no discernible effect on BKLF's repression activity.

Reasoning that BKLF might possess an additional repression domain that functions in NIH 3T3 cells and compensates for the loss of CtBP-mediated repression, we constructed a minimal gal4DBD–BKLF(1–75) fusion protein that contained only the previously defined minimal repression domain of BKLF (see Figure 1). This protein also repressed reporter gene expression efficiently in a dose-dependent manner (Figure 5D, columns 1–4). In the case of the minimal construct, however, we found that mutation of the core Asp–Leu residues, within the mCtBP2-binding site, severely impaired its ability to repress transcription (Figure 5D, columns 5–7). This result suggests that the gal4DBD–BKLF(1–75) fusion represses transcription by recruiting mCtBP2 (or another member of the CtBP family expressed in NIH 3T3 cells).

Finally, in order to determine whether mCtBP2 itself was capable of repressing transcription in mammalian cells, we constructed a gal4DBD–mCtBP2 fusion and tested its activity against the gal4 site-dependent promoter. As shown in Figure 5E (columns 1–6), the gal4DBD–mCtBP2 chimaeric protein efficiently repressed the expression of the reporter gene in a dose-dependent manner.

Taken together, these results demonstrate that mCtBP2 is a true co-repressor in that it is capable of repressing transcription when delivered to a target promoter, either by binding to BKLF or when provided with its own BKLF zinc finger or gal4DBD. While mutant versions of BKLF, unable to bind CtBP, are unable to repress transcription in SL2 cells, it is interesting that the long form of the mutant gal4DBD–BKLF fusion could still repress gene expression in NIH 3T3 cells. This result suggests that additional cofactors may exist in mammalian cells that compensate for the loss of direct CtBP binding. Thus, our results indicate that while BKLF and CtBP co-operate to repress transcription, additional proteins are also likely to be involved (see Discussion).

The putative dehydrogenase activity of mCtBP2 is not required for transcriptional repression

It has been noted that CtBP family proteins have significant homology to the family of D-isomer-specific 2-hydroxy acid dehydrogenases (Schaeper *et al.*, 1995), and it has been suggested that dCtBP may repress transcription by means of this dehydrogenase activity (Nibu *et al.*, 1998). Attempts to demonstrate dehydrogenase activity of hCtBP1, however, have been unsuccessful, and an alternative hypothesis that the homology indicates structural similarity only and may reflect a conserved dimerization domain has been proposed (Schaeper *et al.*, 1995; Poortinga *et al.*, 1998). The dehydrogenases most similar to CtBP family proteins function as homodimers *in vivo* (Goldberg *et al.*, 1994), and it has also been shown that dCtBP can dimerize (Poortinga *et al.*, 1998). We tested

Fig. 5. BKLf and CtBP associate to repress transcription (A) Mutation of the CtBP-binding motif in BKLf impairs BKLf's ability to repress gene expression in SL2 cells. Increasing amounts of pPac-BKLf (columns 2–4) or pPac-mutant BKLf (containing the core Asp–Leu to Ala–Ser mutation) (columns 5–7) were co-transfected with 500 ng of pPac-GR and 500 ng of the (CACCC-GRE)₃-CAT reporter vector. In addition to reporter, 500 ng of pPac-GR was used alone in column 1 and together with 10, 30 and 60 ng of pPac-BKLf in columns 2–4, and 10, 30 and 60 ng of pPac-mutant BKLf in columns 5–7. (B) When tethered to DNA as a BKLf zinc finger fusion, mCtBP2 can repress directly GR-activated transcription in SL2 cells. pPac-GR (500 ng) alone is used in column 1 and together with 1, 5, 10, 20 and 40 ng of pPac-mCtBP2–BKLf fingers in columns 2–6. Columns 7–11 contain 1, 5, 10, 20 and 40 ng of a plasmid encoding a protein which is identical except that the putative active site histidine, H321 in mCtBP2, is replaced by alanine. (C) The N-terminus of BKLf, when fused to the gal4DBD, functions as a repressor in NIH 3T3 cells but does not appear to require the CtBP-binding motif. Columns 1–4 contain 0, 5, 20 and 80 ng of pcDNA3–gal4DBD–BKLf(1–268). Columns 5–7 contain 5, 20 and 80 ng of pcDNA3–gal4DBD–BKLf(1–268) containing the core Ala–Ser mutation. (D) The minimal repression domain of BKLf requires CtBP to repress transcription. Columns 2–4 contain 10, 20 and 80 ng of pcDNA3–gal4DBD–BKLf(1–75). Columns 5–7 contain 10, 20 and 80 ng of pcDNA3–gal4DBD–BKLf(1–75) containing the core Ala–Ser mutation. (E) mCtBP2 can repress gene expression in NIH 3T3 cells when fused to the gal4DBD. Columns 1–6 contain 0, 20, 80 and 250 ng, 1 and 4 μ g of pcDNA3–gal4DBD–mCtBP2; columns 7–11 contain 20, 80 and 250 ng, 1 and 4 μ g of an equivalent plasmid, but in this case the putative catalytic site histidine 321 is altered to alanine.

mCtBP2's ability to dimerize in both the yeast two-hybrid assay and in *in vitro* GST pull-down experiments, and found that mCtBP2 was also able to homodimerize effi-

ciently (data not shown). Thus it is clear that like the dehydrogenase enzymes, CtBP family proteins can dimerize, and it is possible that the similarity to dehydro-

genes arises primarily from conservation of the dimerization surfaces.

Nevertheless, it is noteworthy that all reported CtBP family members contain a conserved histidine residue that appears to correspond to the active site histidine in the enzyme D-lactate dehydrogenase (H296 in D-lactate dehydrogenase, H321 in mCtBP2 and H314 in hCtBP1) (Taguchi and Ohta, 1993; Schaeper *et al.*, 1995) (Figure 2). Thus it remains possible that CtBP proteins possess undetected dehydrogenase activity. It has been shown that mutation of H296 in D-lactate dehydrogenase severely impaired its enzymatic activity (Taguchi and Ohta, 1993). We therefore constructed a mutation of H321 in mCtBP2 and carried out experiments to determine whether this mutation influenced the ability of mCtBP2 to repress gene expression. We tested both mutant mCtBP2-BKLF zinc finger and gal4DBD-mCtBP2 fusion proteins, each containing an H321 to A substitution (Figure 5B and E respectively). The substitution had no significant effect on the repression activity of either the mCtBP2-BKLF finger chimera in SL2 cells (Figure 5B, columns 7–11) or the gal4DBD-mCtBP2 fusion in NIH 3T3 cells (Figure 5E, columns 7–11). These results suggest that even if mCtBP2 does possess undetected dehydrogenase activity, this activity is not essential for its role in the silencing of gene expression.

mCtBP2 binds a number of mammalian transcription factors

Noting that CtBP family members recognize the consensus motif Pro-X-Asp/Asn-Leu-Ser/Thr, we carried out database searches to determine whether this motif was found in other mammalian repressors. As shown in Table I, this motif (or closely related sequences) occurs in a number of proteins and most notably within the previously defined small repression domains of the zinc finger proteins Evi-1 and AREB6 (Bartholomew *et al.*, 1997; Ikeda *et al.*, 1998) (Figure 6). The motif also occurs in FOG (Tsang *et al.*, 1997) and in the AREB6-related protein ZEB (Genetta *et al.*, 1994). Evi-1 is an oncogene, whose expression is associated with myeloid leukaemias in humans and mice (Nucifora, 1997). Although Evi-1 appears to be capable of activating transcription, there is now good evidence that it can also act as a transcriptional repressor (Bartholomew *et al.*, 1997). AREB6 is a large zinc finger homeodomain protein, which has also been implicated in both the repression and activation of transcription (Watanabe *et al.*, 1993).

We used the yeast two-hybrid system to test whether a gal4DBD-mCtBP2 bait could interact with gal4AD hybrids containing portions of these mammalian transcription factors (Figure 6). We first tested the repression domain of Evi-1 (amino acids 514–726; Bartholomew *et al.*, 1997). This region contains the sequences Pro-Phe-Asp-Leu-Thr and Pro-Leu-Asp-Leu-Ser, and was strongly positive for interaction with mCtBP2. We also tested a 32 amino acid region within the repression domain of AREB6, encompassing a Pro-Leu-Asn-Leu-Thr motif. This domain also interacted efficiently with mCtBP2. We then studied a slightly larger domain from the protein ZEB (amino acids 697–795). ZEB is highly related to AREB6 but differs in a small region between amino acids 784 and 802, and therefore lacks the Pro-Leu-Asn-Leu-Thr motif

Table I. Numerous mammalian and *Drosophila* repressors and viral proteins contain CtBP recognition motifs

BKLF	⁵⁸ QVE	PVDLT	VNKR ⁶⁹
Evi-1	⁵⁵⁰ SES	PFDLT	TKRK ⁵⁶¹
	⁵⁸¹ QDQ	PLDLS	MGSR ⁵⁹²
AREB6	⁷⁶⁴ QEE	PLNLS	CAKK ⁷⁷⁵
ZEB	⁷³⁸ SPS	PLNLS	SSRN ⁷⁴⁹
	⁷⁶⁵ QVE	PLDLS	LKQQ ⁷⁷⁷
FOG	⁸⁰⁸ PDG	PIDLS	KRPR ⁸¹⁹
CtIP	³⁸⁹ MDK	PLDLS	DRFS ⁴⁰⁰
Krüppel	⁴¹⁰ EEA	PLDLS	EDGA ⁴²²
	⁴⁶¹ QTE	PEDLS	MHSP ⁴⁷²
Zfh-1	⁷⁸⁹ EDQ	PLDLS	VKRD ⁸⁰⁰
U-shaped	⁵³⁷ ESA	PLDLS	LRRS ⁵⁴⁸
Knirps	³²⁸ QEG	PMDLS	MKTS ³³⁹
Snail	³⁸ QDQ	PQDLS	LKRG ⁴⁹
	¹⁸³ GSE	PEDLS	VRND ¹⁹⁴
Hairy	³¹⁵ EQQ	PLSLV	IKKQ ³²⁶
E(spl)mδ	¹⁴³ AEE	PVNLA	DQKR ¹⁵⁴
Ad2 243R	²³⁰ PGQ	PLDLS	CKRP ²⁴¹
Ad12 235R	²²¹ QTV	PVDLS	VKRP ²³²
Consensus	qdq	PLDLS	Xkkk
	eee		rrr

The consensus recognition sequence is shown below the table, with highly conserved residues shown in upper case and less well conserved residues shown in lower case; X denotes any amino acid. Note: AREB6 is a repressor highly related to ZEB, and contains the two motifs contained in ZEB as well as its own unique motif (which is shown). Proteins shown in bold have been demonstrated to bind CtBP family members; proteins shown in plain text contain putative CtBP-binding sites, but have not been tested experimentally. Accession numbers are as follows: BKLF, U36340; Evi-1, G50874; AREB6, G558522; ZEB, G644840; Zfh-1, G158821; FOG, G2252814; U-shaped, G2597859; Knirps, G8154; Snail, G8630; E(spl)mδ, G158497; Krüppel, G1517920; CtIP, G1730321.

of AREB6. However, N-terminal to this region two additional Pro-X-Asp-Leu-Thr-like motifs occur (Table I). We found that mCtBP2 interacted strongly with this subdomain of ZEB. Finally, mCtBP2 was found to bind to a region within FOG (amino acids 563–859) which contains the motif Pro-Ile-Asp-Leu-Ser. Our results therefore suggest that mCtBP2 can interact with recognition motifs found in several different mammalian transcription factors.

Discussion

Sequence-specific DNA-binding proteins play a primary role in the regulation of gene expression. The DNA recognition domains of these proteins tend to form distinct structures, such as zinc fingers or homeodomains, and are usually easily recognizable. The activation and repression domains of transcription factors have been much more difficult to define. Activation domains have been classified in general terms as acidic, glutamine or proline rich (Mitchell and Tjian, 1989), while repression domains are often basic, alanine or proline rich (Hanna-Rose and Hansen, 1996). It now appears that many of these domains function by recruiting accessory proteins that mediate the activation or repression of transcription, and that in some cases the crucial protein–protein interactions involve only very small sequence motifs within the larger functional domains (Fisher *et al.*, 1996; Heery *et al.*, 1997). For example, it recently has been reported that the interaction between certain steroid receptor proteins and their co-activators occurs through short Leu-X-X-Leu-Leu motifs

Fig. 6. Various transcription factors implicated in transcriptional repression contain one or more CtBP recognition motifs in their repression domains; these motifs are shown as black boxes. Overlining indicates the regions of BKLF, Evi-1, AREB6, FOG and Krüppel that have been shown here to bind mCtBP2 in the two-hybrid assay (numbers indicate amino acid positions). Mapped repression domains are indicated by grey boxes. Zinc fingers are shown as striped boxes. The asterisk above AREB6 indicates the Asn→Gln substitution (within the CtBP recognition motif), which abolishes its repression activity. For accession numbers, see Table I footnotes.

in the co-activators (Heery *et al.*, 1997). We have shown here that Pro-X-Asp/Asn-Leu-Ser/Thr motifs within the repression domains of mammalian DNA-binding proteins mediate their interaction with the co-repressor mCtBP2, while others have demonstrated the necessity of this motif for interaction of dCtBP with the repression domains of *Drosophila* transcription factors (Nibu *et al.*, 1998; Poortinga *et al.*, 1998).

Recent evidence indicates that *Drosophila* CtBP interacts with the helix-loop-helix repressors Hairy and E(spl)mδ (Poorting *et al.*, 1998), the zinc finger protein Snail and the nuclear receptor protein Knirps (Nibu *et al.*, 1998). Here we show that the related murine protein, mCtBP2, interacts with the zinc finger transcription factors BKLF, AREB6, ZEB, Evi-1 and FOG. As we have demonstrated, BKLF can act as a potent repressor of transcription, though it has also been shown to activate transcription, albeit weakly (Crossley *et al.*, 1996). It has been proposed that BKLF may be involved in the activation of the SHP-1 phosphatase gene (Perkins *et al.*, 1997), and it is possible that BKLF acts as either an activator or repressor of transcription, depending on promoter or cellular contexts. Whilst ZEB is well recognized as a transcriptional repressor, AREB6 and Evi-1 have been reported to be involved in either the repression or activation of transcription, depending on contexts (Watanabe *et al.*, 1993; Genetta *et al.*, 1994; Morishita *et al.*, 1995; Tanaka *et al.*, 1995; Bartholomew *et al.*, 1997; Ikeda *et al.*, 1998). Likewise, the GATA-1-binding protein FOG appears to be able to activate or repress transcription depending

on context (Tsang *et al.*, 1997, and unpublished data). Interestingly, a CtBP recognition motif also occurs in the *Drosophila* repressor protein U-shaped which is similar to FOG (Figure 6) in that it also contains nine zinc fingers and specifically binds to a GATA factor, the *Drosophila* factor Pannier (Haenlin *et al.*, 1997). It will be interesting to determine whether the motif in U-shaped is required for the protein's repression activity during *Drosophila* development. The motif also occurs in the apparent *Drosophila* counterpart of AREB6, Zfh1, a transcription factor required for germ cell migration and gonadal mesoderm development (Broihier *et al.*, 1998). It has also been suggested that dCtBP may bind the founding member of the Krüppel family, the *Drosophila* Krüppel protein (Nibu *et al.*, 1998). We have demonstrated that mCtBP2 interacts with a small region of Krüppel (amino acids 356–502), encompassing the previously defined C-terminal repression domain (Hanna-Rose and Hansen, 1996) (Figure 6) and two potential CtBP recognition motifs (Table I) (unpublished results).

We have presented evidence that mCtBP2 is involved directly in the repression activity of BKLF. mCtBP2 binds to the minimal repression domain of BKLF, and mutations which abrogate its binding significantly impair BKLF's ability to repress transcription. Our results also suggest that CtBP proteins play a direct role in mediating repression by Evi-1 and AREB6. In the case of Evi-1, it is significant that two CtBP-binding motifs map within the repression domain of this protein. In the case of AREB6, there is even stronger evidence that CtBP is involved. We have

demonstrated that mCtBP2 binds a region of AREB6 that contains a single CtBP recognition motif (Pro-Leu-Asn-Leu-Thr). Moreover, it has been reported previously that an Asn to Gln substitution within this motif abolished repression, though the explanation for this loss of activity was not elucidated (Ikeda *et al.*, 1998). Our results strongly suggest that CtBP is involved in facilitating repression by AREB6. This example, however, is complicated by the fact that AREB6 also contains two additional potential CtBP-binding motifs slightly N-terminal to the critical motif. These additional motifs are conserved in the related protein ZEB, and we have demonstrated that these motifs can also bind mCtBP2 in two-hybrid assays. It may be that AREB6 contains three CtBP-binding sites and, since the inactivation of one of these compromises repression, it is likely that all three are required for full repressor activity. Interestingly, Snail and Krüppel also appear to have more than one CtBP-binding site (Nibu *et al.*, 1998), as does Evi-1 (Figure 6 and Table I). Thus, it seems that like the Leu-X-X-Leu-Leu motifs that are repeated in the co-activators of nuclear receptor proteins (Heery *et al.*, 1997), Pro-X-Asp-Leu-Ser motifs also occur in clusters within a single protein.

It is notable that several of the proteins targeted by mCtBP2 can act as either activators or repressors. It will be important to determine whether the activity of CtBP family members in particular cell lines correlates with the ability of these factors to repress transcription. Existing evidence from Northern blot analysis and EST studies suggests that hCtBP1, mCtBP2 and dCtBP mRNAs are widely expressed (Katsanis and Fisher, 1998; Poortinga *et al.*, 1998) but, clearly, the presence of CtBP mRNA may not indicate that the protein is active. It has been reported that hCtBP1 is modified differentially by phosphorylation during the cell cycle (Boyd *et al.*, 1993) and, moreover, repression by CtBP is likely to require additional cofactors (Nibu *et al.*, 1998; Poortinga *et al.*, 1998), so the availability of these proteins must also be taken into account.

Although the four mammalian proteins we have studied are all zinc finger proteins that have established roles in haematopoiesis, we believe that mammalian CtBP proteins will target a much wider group of factors. The results from *Drosophila* indicate that dCtBP interacts with a variety of different DNA-binding proteins and, likewise, our analysis of databases suggests that CtBP recognition motifs occur in a number of other transcriptional repressors and, interestingly, also in co-repressor proteins (data not shown). The observation that we and others have made, that both mCtBP2 and dCtBP (Poortinga *et al.*, 1998) can dimerize, raises the possibility that CtBP dimers may function to bring together repressors and co-repressors containing Pro-X-Asp-Leu-Ser motifs. CtBP may thus be an adaptor protein that facilitates the assembly of a larger repressor complex. Such a role is in keeping with the suggestion that dCtBP acts as a member of a large complex (Nibu *et al.*, 1998; Poortinga *et al.*, 1998), and with our result that dehydrogenase activity is not required for its ability to repress gene expression.

Although our results fit with the suggestion from work on dCtBP that CtBP acts as part of a repressor complex, the mechanism by which this complex represses gene expression is not clear. It may be significant that dCtBP

is believed to target short-range repressors (Nibu *et al.*, 1998). These repressors are classified as proteins that quench the action of activators bound within 100 bp or directly inhibit the core transcriptional machinery. Here we show that mCtBP2 quenches the activity of GR bound to an adjacent element, but we cannot exclude the possibility that it is working primarily on TATA-box-binding protein (TBP) or an associated factor critical to the activity of the basal transcriptional machinery. The previously described co-repressors NCoR, SMRT and Rb have been shown to mediate repression by recruiting histone deacetylase proteins (Nagy *et al.*, 1997; Brehm *et al.*, 1998; Lin *et al.*, 1998; Magnaghi-Jaulin *et al.*, 1998). It may be that mCtBP2 also recruits a histone deacetylase, but other mechanisms are also possible, and currently are being investigated.

Although CtBP proteins associate with DNA-binding proteins, it should be remembered that hCtBP1 was first isolated as a protein that interacted with the adenovirus protein E1a (Boyd *et al.*, 1993). E1a is generally regarded as a virally encoded transcriptional cofactor, and it is interesting that as well as hCtBP1, other transcriptional regulatory proteins (such as p300, Rb, TBP and certain TBP-associated factors) (Lee *et al.*, 1991; Zamanian and La Thangue, 1992; Geisberg *et al.*, 1995; Mymryk, 1996; Yang *et al.*, 1996) are also targeted by E1a. Unlike these proteins, hCtBP1 binds the C-terminus of E1a. This region is not required for oncogenic transformation but it has been shown to play a significant role in modulating the transforming activity of E1a (Boyd *et al.*, 1993). Thorough *in vivo* studies have demonstrated that the deletion of the CtBP-binding site significantly increases E1a's transforming and tumorigenic activities, while reducing its ability to inhibit metastases (Boyd *et al.*, 1993). Although the molecular mechanisms underlying these activities are not known, it seems likely that by binding CtBP, E1a alters normal patterns of gene expression and in this way alters the proliferative capacity of cells. It is possible that E1a is involved in delivering CtBP to additional target genes or, alternatively, by sequestering CtBP, it may deprive other transcription factors of this protein. It will be particularly interesting to determine whether the sequestering of CtBP alters the expression of genes regulated by oncogenic transcription factors such as Evi-1.

If mCtBP2 is proven to be exclusively a short-range repressor, it may be useful as a tool for specifically silencing particular genes, without affecting the activity of closely linked neighbouring genes. It appears that CtBP normally is delivered to target genes by piggy-backing onto DNA-bound transcription factors. In our study, we fused the zinc fingers from BKLF, which recognize CACCC-box sequences, to mCtBP2 and showed that the resulting protein could silence a CACCC-dependent promoter. Thus we demonstrated that if CtBP is provided with its own zinc finger DNA-binding domain it can silence gene expression directly. It is likely that zinc fingers with different specificities could also be used and, in this way, mCtBP2 could be targeted to other promoters. Indeed much progress has been made recently in engineering novel zinc fingers with a wide range of desired sequence specificities (Choo *et al.*, 1994, 1997; Kim and Pabo, 1997, 1998; Kim *et al.*, 1997). It thus seems possible that artificial zinc finger-mCtBP2 fusions with the required

sequence specificity could be directed against various chosen target promoters. It should be remembered, however, that our experiments only demonstrated silencing of reporter plasmids in transient transfection assays. Nevertheless, there is reason to believe that the silencing of endogenous genes may be successful. Firstly, it should be noted that the aim would be to target CtBP to fully active promoters, likely to be contained in accessible chromatin. Secondly, results from *Drosophila* have already indicated that a gal4DBD-dCtBP fusion is able to repress the transcription of an integrated gal4-dependent promoter *in vivo* (Nibu *et al.*, 1998). In this case, however, repression was weaker than expected, suggesting the requirement for additional cofactors. We have evidence of a novel mCtBP2-interacting protein, which also associates with BKLf (unpublished results). This latter protein binds to a distinct domain within BKLf and may account for the residual repressor activity observed in assays with the mutant BKLf proteins that are unable to bind CtBP directly (Figure 5A and C). We currently are investigating this protein, in the expectation that it is another component of a larger repressor complex. Ultimately, it may be possible to deliver the entire complex to chosen promoters, and in this way to control their activity artificially.

Materials and methods

Plasmids

The EKLF expression vector pPac-EKLF (Merika and Orkin, 1995) was a gift from M. Merika. pPac-BKLf was constructed by inserting a blunted *EcoRI* fragment containing the entire BKLf coding sequence (Crossley *et al.*, 1996) into the blunted *BamHI* site of pPac. pPac-GR was constructed by similarly cloning an *Asp718-DraI* fragment containing the entire coding region of the human GR (Hollenberg *et al.*, 1985) into the blunted *BamHI* site of pPac. The gal4DBD-mCtBP2 fusion consisted of the mCtBP2 coding sequence fused downstream of the gal4DBD(1–147) all within the mammalian expression vector pcDNA3 (Invitrogen). The mCtBP2-BKLf fingers fusion consisted of mCtBP2 fused upstream of the region encoding the three BKLf zinc fingers (residues 246–344) all cloned into the *BamHI* site of pPac. SL2 cell reporter plasmids contained three copies of a composite CACCC-GRE site (tgctAGAACAAtccGTGACAgagagCCACACCAAtctg; Strähle *et al.*, 1988) upstream of the minimal *Adh* promoter and the CAT gene in p1970 (Fan and Maniatis, 1989). NIH 3T3 cell reporters contained five gal4-binding sites upstream of a minimal promoter in pOGH (Nichols Institute).

Transactivation and transrepression assays

SL2 cells and NIH 3T3 cells were cultured in Shields and Sang medium and Dulbecco's modified Eagle's medium (DMEM) respectively, both supplemented with 10% fetal calf serum, penicillin, streptomycin and glutamine. Transfections were carried out by the calcium phosphate precipitation method (Chen and Okayama, 1987; Sambrook *et al.*, 1989) with the amounts of plasmid DNA indicated. A 0.5 µg aliquot of CAT reporter plasmid was used in all SL2 experiments, and 5 µg of growth hormone reporter was used in all NIH 3T3 cell experiments. All data shown represent the sum of several sets of experiments with duplicate preparations of plasmid DNA. Transfections were normalized by reference to co-transfected *lacZ* reporter plasmids and/or gel shifts were carried out to verify expression levels of DNA-binding proteins. In experiments using the human GR, dexamethasone was added to a final concentration of 10^{-8} M 20 h after the transfection. Cells or media were harvested 48 h after transfection. CAT assays and growth hormone assays were carried out using Amersham Quanti-Cat and Nichols Inst. Allegro GH assay kits according to the manufacturers' instructions.

Two-hybrid screening and assays

The Clontech two-hybrid system was used according to the manufacturer's instructions. A murine erythroleukaemia cell cDNA library in the gal4AD fusion vector, pGAD10, was transfected into the yeast strain HF7c harbouring the gal4DBD-BKLf(1–268) fusion protein expressed

from pGBT9. Two full-length mCtBP2 clones that differed only in the length of their 5'-untranslated regions were isolated. Surprisingly, intervening stop codons between the gal4AD and the mCtBP2 translational initiation codon suggested that neither clone was being expressed as an in-frame gal4AD fusion. Interestingly, dCtBP was cloned using a similar strategy and was also present as an out-of-frame fusion (Nibu *et al.*, 1998). Interactions between mCtBP2 and Evi-1, AREB6, ZEB and FOG subdomains were carried out by co-transfecting a gal4DBD-mCtBP2 bait plasmid together with gal4AD-prey proteins into HF7c. Transformants were selected on Trp/Leu-deficient plates and patched onto Trp/Leu/His-deficient media. Growth was scored after 30 h. Controls containing bait and prey alone were also conducted and were negative for growth up to >72 h.

In vitro binding assays

Full-length mCtBP2 was cloned into the vector pcDNA3 (Invitrogen), and 35 S-labelled *in vitro* translated protein was generated using T7 polymerase and the TNT system from Promega. GST fusion proteins and binding assays were carried out as previously described (Merika and Orkin, 1995). Gel retardation assays were performed using a 32 P-labelled double-stranded oligonucleotide (TAGAGCCACACCTGTGTAAG and CTTACCAGGGTGTGGCTCTA) (Crossley *et al.*, 1996).

Acknowledgements

We are grateful to Gerd Blobel, Archa Fox, Melissa Holmes and Margot Kearns for critical reading of the manuscript, to Louise Rafty for preliminary work on this project, to T. Genetta and T. Kadesch for ZEB plasmids, and to G. Chinnadurai for communicating unpublished results. J.T. is supported by an Australian Postgraduate Award. This work was supported by a grant from the Australian NHMRC to M.C.

References

- Anderson, K.P., Kern, C.B., Crable, S.C. and Lingrel, J.B. (1995) Isolation of a gene encoding a functional zinc finger protein homologous to erythroid Krüppel-like factor: identification of a new multigene family. *Mol. Cell. Biol.*, **15**, 5957–5965.
- Bartel, P.L., Chien, C.-T., Sternglanz, R. and Fields, S. (1993) Using the two-hybrid system to detect protein-protein interactions. In Hartley, D.A. (ed.), *Cellular Interactions in Development: A Practical Approach*. Oxford University Press, Oxford, pp. 153–179.
- Bartholomew, C., Kilbey, A., Clark, A.M. and Walker, M. (1997) The Evi-1 proto-oncogene encodes a transcriptional repressor activity associated with transformation. *Oncogene*, **14**, 569–577.
- Boyd, J.M., Subramanian, T., Schaeper, U., La Regina, M., Bayley, S. and Chinnadurai, G. (1993) A region in the C-terminus of adenovirus 2/5 E1a protein is required for association with a cellular phosphoprotein and important for the negative modulation of T24-ras mediated transformation, tumorigenesis and metastasis. *EMBO J.*, **12**, 469–478.
- Brehm, A., Miska, E.A., McCance, D.J., Reid, J.L., Bannister, A.J. and Kouzarides, T. (1998) Retinoblastoma protein recruits histone deacetylase to repress transcription. *Nature*, **391**, 597–601.
- Broihier, H.T., Moore, L.A., Van Doren, M., Newman, S. and Lehmann, R. (1998) *zfh-1* is required for germ cell migration and gonadal mesoderm development in *Drosophila*. *Development*, **125**, 655–666.
- Chen, C. and Okayama, H. (1987) High-efficiency transformation of mammalian cells by plasmid DNA. *Mol. Cell. Biol.*, **7**, 2745–2752.
- Choo, Y., Sanchez-Garcia, I. and Klug, A. (1994) *In vivo* repression by a site-specific DNA-binding protein designed against an oncogenic sequence. *Nature*, **372**, 642–645.
- Choo, Y., Castellanos, A., Garcia-Hernandez, B., Sanchez-Garcia, I. and Klug, A. (1997) Promoter-specific activation of gene expression directed by bacteriophage-selected zinc fingers. *J. Mol. Biol.*, **273**, 525–532.
- Crossley, M., Whitelaw, E., Perkins, A., Williams, G., Fujiwara, Y. and Orkin, S.H. (1996) Isolation and characterization of the cDNA encoding BKLf/TEF-2, a major CACCC-box-binding protein in erythroid cells and selected other cells. *Mol. Cell. Biol.*, **16**, 1695–1705.
- Fan, C.M. and Maniatis, T. (1989) Two different virus-inducible elements are required for human beta-interferon gene regulation. *EMBO J.*, **8**, 101–110.
- Feng, W.C., Southwood, C.M. and Bieker, J.J. (1994) Analyses of β -thalassaemia mutant DNA interactions with erythroid Krüppel-like factor (EKLF), an erythroid cell-specific transcription factor. *J. Biol. Chem.*, **269**, 1493–1500.

- Fisher, A.L., Ohsako, S. and Caudy, M. (1996) The WRPW motif of the hairy-related basic helix-loop-helix repressor proteins acts as a 4-amino-acid transcription repression and protein-protein interaction domain. *Mol. Cell. Biol.*, **16**, 2670–2677.
- Garrett-Sinha, L.A., Eberspaecher, H., Seldin, M.F. and de Crombrughe, B. (1996) A gene for a novel zinc-finger protein expressed in differentiated epithelial cells and transiently in certain mesenchymal cells. *J. Biol. Chem.*, **271**, 31384–31390.
- Geisberg, J.V., Chen, J.L. and Ricciardi, R.P. (1995) Subregions of the adenovirus E1A transactivation domain target multiple components of the TFIID complex. *Mol. Cell. Biol.*, **15**, 6283–6290.
- Genetta, T., Ruezinsky, D. and Kadesch, T. (1994) Displacement of an E-box-binding repressor by basic helix-loop-helix proteins: implications for B-cell specificity of the immunoglobulin heavy-chain enhancer. *Mol. Cell. Biol.*, **14**, 6153–6163.
- Goldberg, J.D., Yoshida, T. and Brick, P. (1994) Crystal structure of a NAD-dependent D-glycerate dehydrogenase at 2.4 Å resolution. *J. Mol. Biol.*, **236**, 1123–1140.
- Haenlin, M., Cubadda, Y., Blondeau, F., Heitzler, P., Lutz, Y., Simpson, P. and Romain, P. (1997) Transcriptional activity of Pannier is regulated negatively by heterodimerization of the GATA DNA-binding domain with a cofactor encoded by the u-shaped gene of *Drosophila*. *Genes Dev.*, **11**, 3096–3108.
- Hanna-Rose, W. and Hansen, U. (1996) Active repression mechanisms of eukaryotic transcription repressors. *Trends Genet.*, **12**, 229–234.
- Heery, D.M., Kalkhoven, E., Hoare, S. and Parker, M.G. (1997) A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature*, **387**, 733–736.
- Hollenberg, S.M., Weinberger, C., Ong, E.S., Cerelli, G., Oro, A., Lebo, R., Thompson, E.B., Rosenfeld, M.G. and Evans, R.M. (1985) Primary structure and expression of a functional human glucocorticoid receptor cDNA. *Nature*, **318**, 635–641.
- Ikeda, K., Halle, J.P., Stelzer, G., Meisterernst, M. and Kawakami, K. (1998) Involvement of negative cofactor NC2 in active repression by zinc finger-homeodomain transcription factor AREB6. *Mol. Cell. Biol.*, **18**, 10–18.
- Iwabuchi, K., Li, B., Bartel, P. and Fields, S. (1993) Use of the two-hybrid system to identify the domain of p53 involved in oligomerization. *Oncogene*, **8**, 1693–1696.
- Katsanis, K. and Fisher, E.M.C. (1998) A novel C-terminal binding protein (CtBP2) is closely related to CtBP1, an adenovirus E1A-binding protein and maps to human chromosome 21q21.3. *Genomics*, **47**, 294–299.
- Kim, J.S. and Pabo, C.O. (1997) Transcriptional repression by zinc finger peptides. Exploring the potential for applications in gene therapy. *J. Biol. Chem.*, **272**, 29795–29800.
- Kim, J.S. and Pabo, C.O. (1998) Getting a handhold on DNA: design of poly-zinc finger proteins with femtomolar dissociation constants. *Proc. Natl Acad. Sci. USA*, **95**, 2812–2817.
- Kim, J.S., Kim, J., Cepek, K.L., Sharp, P.A. and Pabo, C.O. (1997) Design of TATA box-binding protein/zinc finger fusions for targeted regulation of gene expression. *Proc. Natl Acad. Sci. USA*, **94**, 3616–3620.
- Kuo, C.T., Veselits, M.L., Barton, K.P., Lu, M.M., Clendenin, C. and Leiden, J.M. (1997a) The LKLF transcription factor is required for normal tunica media formation and blood vessel stabilization during murine embryogenesis. *Genes Dev.*, **11**, 2996–3006.
- Kuo, C.T., Veselits, M.L. and Leiden, J.M. (1997b) LKLF: a transcriptional regulator of single-positive T cell quiescence and survival. *Science*, **277**, 1986–1990.
- Lee, W.S., Kao, C.C., Bryant, G.O., Liu, X. and Berk, A.J. (1991) Adenovirus E1A activation domain binds the basic repeat in the TATA box transcription factor. *Cell*, **67**, 365–376.
- Lin, R.J., Nagy, L., Inoue, S., Shao, W., Miller, W.H., Jr and Evans, R.M. (1998) Role of the histone deacetylase complex in acute promyelocytic leukaemia. *Nature*, **391**, 811–814.
- Magnaghi-Jaulin, L., Groisman, R., Naguibneva, I., Robin, P., Lorain, S., Le Villain, J.P., Trouche, D. and Harel-Bellan, A. (1998) Retinoblastoma protein represses transcription by recruiting a histone deacetylase. *Nature*, **391**, 601–605.
- Merika, M. and Orkin, S.H. (1995) Functional synergy and physical interactions of the erythroid transcription factor GATA-1 with the Krüppel family proteins Sp1 and EKLF. *Mol. Cell. Biol.*, **15**, 2437–2447.
- Miller, I.J. and Bieker, J.J. (1993) A novel, erythroid cell-specific murine transcription factor that binds to the CACCC element and is related to the Krüppel family of nuclear proteins. *Mol. Cell. Biol.*, **13**, 2776–2786.
- Mitchell, P.J. and Tjian, R. (1989) Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science*, **245**, 371–378.
- Morishita, K., Parker, D.S., Mucenski, M.L., Jenkins, N.A., Copeland, N.G. and Ihle, J.N. (1988) Retroviral activation of a novel gene encoding a zinc finger protein in IL-3-dependent myeloid leukemia cell lines. *Cell*, **54**, 831–840.
- Morishita, K., Suzukawa, K., Taki, T., Ihle, J.N. and Yokota, J. (1995) EVI-1 zinc finger protein works as a transcriptional activator via binding to a consensus sequence of GACAAGATAAGATAAN1–28 CTCATCTTC. *Oncogene*, **10**, 1961–1967.
- Mymryk, J.S. (1996) Tumour suppressive properties of the adenovirus 5 E1A oncogene. *Oncogene*, **13**, 1581–1589.
- Nagy, L., Kao, H.Y., Chakravarti, D., Lin, R.J., Hassig, C.A., Ayer, D.E., Schreiber, S.L. and Evans, R.M. (1997) Nuclear receptor repression mediated by a complex containing SMRT, mSin3A and histone deacetylase. *Cell*, **89**, 373–380.
- Nibu, Y., Zhang, H. and Levine, M. (1998) Interaction of short-range repressors with *Drosophila* CtBP in the embryo. *Science*, **280**, 101–104.
- Nucifora, G. (1997) The EVI1 gene in myeloid leukemia. *Leukemia*, **11**, 2022–2031.
- Nuez, B., Michalovich, D., Bygrave, A., Ploemacher, R. and Grosveld, F. (1995) Defective haematopoiesis in fetal liver resulting from inactivation of the EKLF gene. *Nature*, **375**, 316–318.
- Perkins, A.C., Sharpe, A.H. and Orkin, S.H. (1995) Lethal β -thalassaemia in mice lacking the erythroid CACCC-transcription factor EKLF. *Nature*, **375**, 318–322.
- Perkins, A.C., Yang, H., Crossley, P., Fujiwara, Y. and Orkin, S.H. (1997) Deficiency of the CACCC-element binding protein, BKLF, leads to a progressive myeloproliferative disease and impaired expression of SHP-1. *Blood*, **90**, Suppl. 1, 575.
- Poortinga, G., Watanabe, M. and Parkhurst, S.M. (1998) *Drosophila* CtBP: a Hairy-interacting protein required for embryonic segmentation and Hairy-mediated transcriptional repression. *EMBO J.*, **17**, 2067–2078.
- Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Schaeper, U., Boyd, J.M., Verma, S., Uhlmann, E., Subramanian, T. and Chinnadurai, G. (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. USA*, **92**, 10467–10471.
- Schüle, R., Müller, M., Otsuka-Murakami, H. and Renkawitz, R. (1988) Cooperativity of the glucocorticoid receptor and the CACCC-box binding factor. *Nature*, **332**, 87–90.
- Shields, J.M., Christy, R.J. and Yang, V.W. (1996) Identification and characterization of a gene encoding a gut-enriched Krüppel-like factor expressed during growth arrest. *J. Biol. Chem.*, **271**, 20009–20017.
- Sollerbrant, K., Chinnadurai, G. and Svensson, C. (1996) The CtBP binding domain in the adenovirus E1A protein controls CR1-dependent transactivation. *Nucleic Acids Res.*, **24**, 2578–2584.
- Strähle, U., Schmid, W. and Schütz, G. (1988) Synergistic action of the glucocorticoid receptor with transcription factors. *EMBO J.*, **7**, 3389–3395.
- Taguchi, H. and Ohta, T. (1993) Histidine 296 is essential for the catalysis in *Lactobacillus plantarum* D-lactate dehydrogenase. *J. Biol. Chem.*, **268**, 18030–18034.
- Tanaka, T., Mitani, K., Kurokawa, M., Ogawa, S., Tanaka, K., Nishida, J., Yazaki, Y., Shibata, Y. and Hirai, H. (1995) Dual functions of the AML1/Evi-1 chimeric protein in the mechanism of leukemogenesis in t (3;21) leukemias. *Mol. Cell. Biol.*, **15**, 2383–2392.
- Tsang, A.P., Visvader, J.E., Turner, C.A., Fujiwara, Y., Yu, C., Weiss, M.J., Crossley, M. and Orkin, S.H. (1997) FOG, a multitype zinc finger protein, acts as a cofactor for transcription factor GATA-1 in erythroid and megakaryocytic differentiation. *Cell*, **90**, 109–119.
- Watanabe, Y., Kawakami, K., Hirayama, Y. and Nagano, K. (1993) Transcription factors positively and negatively regulating the Na,K-ATPase alpha 1 subunit gene. *J. Biochem.*, **114**, 849–855.
- Yang, X.J., Ogryzko, V.V., Nishikawa, J., Howard, B.H. and Nakatani, Y. (1996) A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A. *Nature*, **382**, 319–324.
- Yet, S.F. et al. (1998) Human EZF, a Krüppel-like zinc finger protein, is expressed in vascular endothelial cells and contains transcriptional activation and repression domains. *J. Biol. Chem.*, **273**, 1026–1031.
- Zamanian, M. and La Thangue, N.B. (1992) Adenovirus E1A prevents the retinoblastoma gene product from repressing the activity of a cellular transcription factor. *EMBO J.*, **11**, 2603–2610.

Received June 1, 1998; revised and accepted July 6, 1998