

Human Krüppel-like Factor 8: a CACCC-box binding protein that associates with CtBP and represses transcription

Jane van Vliet, Jeremy Turner and Merlin Crossley*

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

Received December 20, 1999; Revised and Accepted March 20, 2000

DDBJ/EMBL/GenBank accession no. U28282

ABSTRACT

CACCC-boxes are recognised by transcription factors of the Sp/Krüppel-like Factor (Sp1/KLF) family. Here we describe one member of this family, KLF8/ZNF741/BKLF3 (KLF8). KLF8 contains a characteristic C-terminal DNA-binding domain comprised of three Krüppel-like zinc fingers, but also has limited homology to another family member, KLF3/Basic Krüppel-like Factor (KLF3/BKLF), in its N-terminus. Most significantly, it shares with KLF3/BKLF a Pro-Val-Asp-Leu-Ser/Thr motif. In KLF3/BKLF this motif mediates contact with the co-repressor protein C-terminal Binding Protein (CtBP). We demonstrate that the KLF8 Pro-Val-Asp-Leu-Ser motif also contacts CtBP. We show that the N-terminus of KLF8 functions as a repression domain and that its activity relies on the integrity of the CtBP recognition motif. We demonstrate that the zinc fingers of KLF8 recognize CACCC elements in DNA and that full-length KLF8 can repress a CACCC-dependent promoter. Finally we determine that KLF8 is broadly expressed in human tissues. These results establish KLF8 as a CACCC-box binding protein that associates with CtBP and represses transcription.

INTRODUCTION

CACCC (or GGGTG)-boxes and related GC-rich elements are present in the promoters and enhancers of a large number of mammalian and viral genes. Numerous different transcription factors have been identified that bind to these elements and either activate or repress gene expression. The first CACCC-box binding proteins to be described were Sp1 and its relatives Sp2–4, but later a number of additional factors including Krüppel-like Factor 1/Erythroid Krüppel-like Factor (KLF1/EKLF), Lung Krüppel-like Factor (KLF2/LKLF) and Basic Krüppel-like Factor (KLF3/BKLF) were reported (for reviews see 1,2). Sp/Krüppel-like Factor (Sp/KLF) proteins all contain a characteristic C-terminal DNA-binding domain comprising three Krüppel-like zinc fingers. The conservation of this domain is such that the proteins all recognise related CACCC- and GC-boxes, although some differences in binding preference have

been observed [for instance Sp1 binds with high affinity to certain GC-boxes, while KLF3/BKLF prefers CACCC elements (3)]. The mammalian Sp/KLF family now comprises 16 proteins (1,2). The Sp1 subfamily proteins are named Sp1–4, and it has been recommended by the Human Gene Nomenclature Committee that the remaining proteins be referred to as KLF1–12.

Several members of the family are broadly expressed in many different tissues throughout development and despite considerable work, the precise biological roles of these proteins have been difficult to define. In contrast, one family member, KLF1/EKLF (4), is expressed primarily in erythroid tissues and extensive experiments both *in vitro* (5) and in KLF1/EKLF knockout embryos (6,7) indicate that KLF1/EKLF binds to a CACCC-box in the β -globin promoter and activates globin transcription. The molecular mechanism by which KLF1/EKLF operates has been intensely investigated. Interestingly, in addition to binding the β -globin promoter, KLF1/EKLF also appears to function through a CACCC-box that lies in the globin Locus Control Region (LCR), >50 kb upstream of the β -globin gene (8,9). This site is within the core region of hypersensitive site 3 and the binding of KLF1/EKLF is required for the proper formation of this hypersensitive site and for β -globin gene expression. These results suggest that KLF1/EKLF may be involved in altering chromatin structure. Recent work has revealed that KLF1/EKLF recruits a SWI/SNF-related chromatin remodelling complex termed Erythroid Remodelling Complex 1 (E-RC1). It is, therefore, possible that E-RC1 activity may be required for the remodelling of chromatin in both the LCR and the globin promoter (10).

While KLF1/EKLF has been shown to activate transcription, other members of the family have been implicated in the repression of gene expression. In some instances, it appears that repressing and activating KLF family proteins co-exist within the same cell and it is possible that their competing activities may serve to balance transcriptional output (1,2 and references therein). The molecular mechanisms by which KLF proteins repress transcription have also been investigated and it has been found that KLF3/BKLF functions, at least in part, by recruiting transcriptional co-repressors of the C-terminal Binding Protein (CtBP) family (11). These proteins were first identified as cellular proteins of unknown function that associated with the Adenovirus oncoprotein E1A (12). It was shown that human CtBP specifically recognised a short Pro-Val-Asp-Leu-Ser

*To whom correspondence should be addressed. Tel: +61 2 9351 2233; Fax: +61 2 9351 4726; Email: m.crossley@biochem.usyd.edu.au

motif at the C-terminus of E1A and it has subsequently become apparent that CtBP proteins bind to related motifs found in DNA-binding proteins (11,13–16). For instance, the motif Pro-Val-Asp-Leu-Thr is present in the N-terminal repression domain of BKLf/KLF3 and the binding of CtBP is required for the activity of this domain (11). It is now known that at least two CtBP proteins (CtBP1 and -2) are present in mammals and related proteins are also found in *Caenorhabditis* and *Drosophila* (13,14). *In vivo* experiments in *Drosophila* have demonstrated that dCtBP associates with and contributes to the repression activity of a number of DNA-binding proteins including Krüppel, Snail and Hairy (13,14). The precise mechanism of action of CtBP is still under study but two reports have suggested that its action may involve histone deacetylase enzymes (15,17).

Here we report the characterisation of KLF8/ZNF741/BKLf3 (KLF8). We demonstrate that this protein is able to bind CACCC-boxes in DNA and can repress gene expression by associating with CtBP co-repressors.

MATERIALS AND METHODS

Plasmids and oligonucleotides

KLF8 was amplified from K562 cell cDNA using the Polymerase Chain Reaction (PCR), with the 5' primer CGGAATTCATGGTTCGATATGGATAAACTCATAAA and the 3' primer ACGGGTTCGACTCACATGGTGTTCATGGCGACGGC. The resulting fragment was digested with *EcoRI* and *SalI* and inserted between the *EcoRI* and *XhoI* sites of pcDNA3 (Invitrogen, Groningen, The Netherlands), to create pcDNA3/KLF8. Similar primers were used to create a related pcDNA3 construct (pcDNA3Flag/KLF8), which contained a Flag epitope at its N-terminus. The truncated version of *KLF8*, encoding the zinc finger region (residues 240–359), was generated by digesting the PCR fragment at a *BamHI* site which lies upstream of the finger encoding region and with *SalI* and inserting the resulting fragment between the *BamHI* and *XhoI* sites of pcDNA3Flag to form pcDNA3Flag/KLF8(240–359). The plasmid encoding the Gal4 DNA-binding domain-KLF8(1–262) fusion was generated by PCR as described above except that in this case a 3' primer with the sequence ACGGGTTCGACTCATCCCTGCATTTGGGCCATTG was used. The resulting fragment was first cloned into the yeast vector pGBT9 (Clontech, Palo Alto, CA) between the *EcoRI* and *SalI* sites downstream of the region encoding the Gal4 DNA-binding domain (residues 1–147) and then a *HindIII*–*SalI* fragment encoding the entire Gal4 DNA-binding domain-KLF8(1–262) fusion was inserted between the *HindIII* and *XhoI* sites of pcDNA3. The prey plasmid for the yeast two-hybrid experiment, pGAD10/KLF8(1–262) was generated by inserting the *EcoRI/SalI* fragment encoding residues 1–262 between the *EcoRI* and *SalI* sites of a pGAD10 derivative (Clontech).

KLF8 derivatives containing a two amino acid substitution Asp-Leu→Ala-Ser in the core of the Pro-Val-Asp-Leu-Ser motif were generated by overlap PCR, using the two complementary internal primers AGTGGAAACCAGTTGCCTCCTCCTTTCACAAGCC and GGCTTGTGAAAGGAGGAGGCAACTGGTTCCACT and the appropriate 5' and 3' primers described above.

The reporter plasmids used in the transfection experiments were (gal4)₅TK-CAT and (CACCC)₃TK-GH. The former plasmid contains five Gal4 recognition elements immediately upstream of the Herpes thymidine kinase promoter driving a bacterial chloramphenicol acetyl transferase (CAT) gene. The latter reporter contains three β-globin CACCC elements upstream of the Herpes thymidine kinase promoter driving a human growth hormone (GH) reporter gene. The plasmid pCMV-lacZ was used as a control to determine transfection efficiencies.

The oligonucleotides used in the gel retardation experiments were: β-globin CACCCTAGAGCCACACCCTGGTAAG and CTTACCAGGGTGTGGCTCTA; γ-globin CACCCAGCTC-TAAACTCCACCCATGGG and AGCTCCCATGGGTGGA-GTTTAG; GC-rich Sp1 site TAAATGACGCCCCGCCAGTCGAT and ATCGACTGGGGCGGGCGTCATTTA; and the mutant β-globin CACCC site AGCTACTAGGCATACCCTGTCC and AGCTGGACAGGGTATGCCTAGT.

Cell culture, transfections, nuclear extracts and gel retardation experiments

Cells were cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum, penicillin, streptomycin and glutamine. COS cells were transfected by the DEAE-mediated method and NIH 3T3 cells were transfected by the calcium phosphate method as previously described (3,11). Transfection efficiencies were normalised by reference to the co-transfected *lacZ* reporter gene. CAT activity was measured using the Quan-T-CAT assay system (Amersham Pharmacia Biotech, Uppsala, Sweden) and GH levels were assayed by the Allegro system (Nichols Institute) according to the manufacturers' instructions (Institute Diagnostics, San Juan Capistrano, CA). Nuclear extracts and gel retardation experiments were carried out as previously described (3).

Yeast two-hybrid and GST pulldown assays

The Clontech two-hybrid system was used according to the manufacturer's instructions. The bait plasmid pGBT9/mCtBP2, encoding a Gal4 DNA-binding domain mCtBP2 fusion (11), and the prey plasmid pGAD10/KLF8(1–262) encoding the N-terminus of KLF8 fused to the Gal4 activation domain, were co-transfected into the yeast strain HF7c. The resulting transformants were selected on minimal media deficient in tryptophan and leucine and the resulting colonies transferred to media that was also deficient in histidine. Growth was monitored and cells photographed after 3 days. A plasmid encoding a KLF8 derivative with the double amino acid substitution in the core of the Pro-Val-Asp-Leu-Ser motif was assayed in the same manner.

Recombinant GST-mCtBP2 was produced in bacteria as previously described (11) and immobilised on glutathione agarose beads. KLF8 radiolabelled with [³⁵S]methionine and a mutant version with the double amino acid substitution were prepared by *in vitro* transcription and translation from pcDNA3/KLF8 using T7 polymerase and the Promega TNT system (Promega, Madison, WI). The GST-pulldown procedure was carried out as previously described (11).

Northern blot analysis

A DNA fragment comprising nt 168–465 of *KLF8* was excised from pcDNA3/KLF8 by *EcoRI* and *BamHI* and labeled with ³²P using standard protocols. This fragment was then used to



Figure 1. Alignment of the primary amino acid sequences of KLF8, KLF3/BKLF and the related protein KLF12/Ap2-rep. KLF8 is identical to ZNF741, GenBank accession no. U28282, but contains a glutamate in place of a glycine at residue 263. Residues that are identical are marked with an asterisk below the sequence. The C-terminal zinc finger domain is overlined and the zinc coordinating cysteine and histidine residues are marked by arrows. Three additional stretches of homology are boxed and the CtBP contact motif Pro-Val-Asp-Leu-Ser/Thr is shaded.

probe a Clontech human multi-tissue northern blot and a Clontech cancer cell line blot. After hybridisation the blots were washed at high stringency and analysed using a Phosphorimager (Molecular Dynamics, Sunnyvale, CA). Similar results were obtained with a distinct fragment consisting of nt 465–718 of *KLF8*. The blots were then probed with β -actin to assess relative amounts of mRNA.

RESULTS

KLF8 encodes a CACCC-box binding protein

In order to identify additional KLF family proteins, we screened the human databases for proteins related to KLF3/BKLF and identified a number of expressed sequence tags and a longer unpublished sequence referred to in reviews as *ZNF741* (GenBank accession number U28282) or as BKLF3 (1,2). This gene encodes a protein that contains the key feature of the KLF family, that is, it contains three Krüppel-like zinc fingers at its C-terminus. We amplified this gene from K562 cell cDNA using PCR (as described in Materials and Methods) and confirmed its identity by sequencing. According to the convention of the Human Gene Nomenclature Committee we used the next available KLF number to name this protein (2), hence KLF8. We chose KLF8 rather than KLF13, which is numerically the next number in the series, since the name KLF8 had become vacant when it became apparent that the protein tentatively assigned the name KLF8 (but never published as such) was identical to a previously described KLF factor.

The full protein sequence is shown in Figure 1, together with the sequence of KLF3/BKLF and another related protein KLF12/Ap-2rep (18). As can be seen from the figure, the

proteins are highly related in their zinc finger region but contain very limited homology elsewhere. The similarity in the zinc finger region suggested that the zinc fingers of KLF8 might also bind CACCC sequences in DNA. In order to investigate this possibility, we cloned the region encoding the three fingers of KLF8 and the full-length *KLF8* sequence into the mammalian expression vector pcDNA3Flag and expressed these proteins in COS cells. We then used nuclear extracts from these cells in gel retardation experiments together with a radiolabelled probe comprising the CACCC-box from the human β -globin promoter. Nuclear extracts from control COS cells (transfected with pcDNA3 vector alone) and cells over-expressing BKLF/KLF3 were used for comparison. Figure 2a, lanes 1 and 5, shows the complexes produced by endogenous CACCC-binding proteins present in control COS cells. Lane 2 shows a prominent new complex produced by KLF3/BKLF (and faster migrating complexes formed by degradation products). In lane 3 there is evidence of a new complex generated by full-length KLF8 binding to the CACCC-box probe. This complex co-migrated with that formed by KLF3/BKLF. In lane 4 a faster migrating complex produced by the fingers of KLF8 can be seen. In lanes 6–8 antibodies specific to Sp1, Sp3 or both are included to identify the bands generated by these proteins. In summary, the results of this experiment indicated that the zinc fingers of KLF8 can bind to a typical CACCC-box element.

The structure of KLF family zinc fingers is such that each finger makes contact with three nucleotides and the general consensus for the recognition site is NCN CNC CCN (1,2,4). Nevertheless, there are some differences between the fingers of different family members and it has been observed that there is some variation in their preferred binding sites. For instance, it has been observed that KLF3/BKLF binds with higher affinity

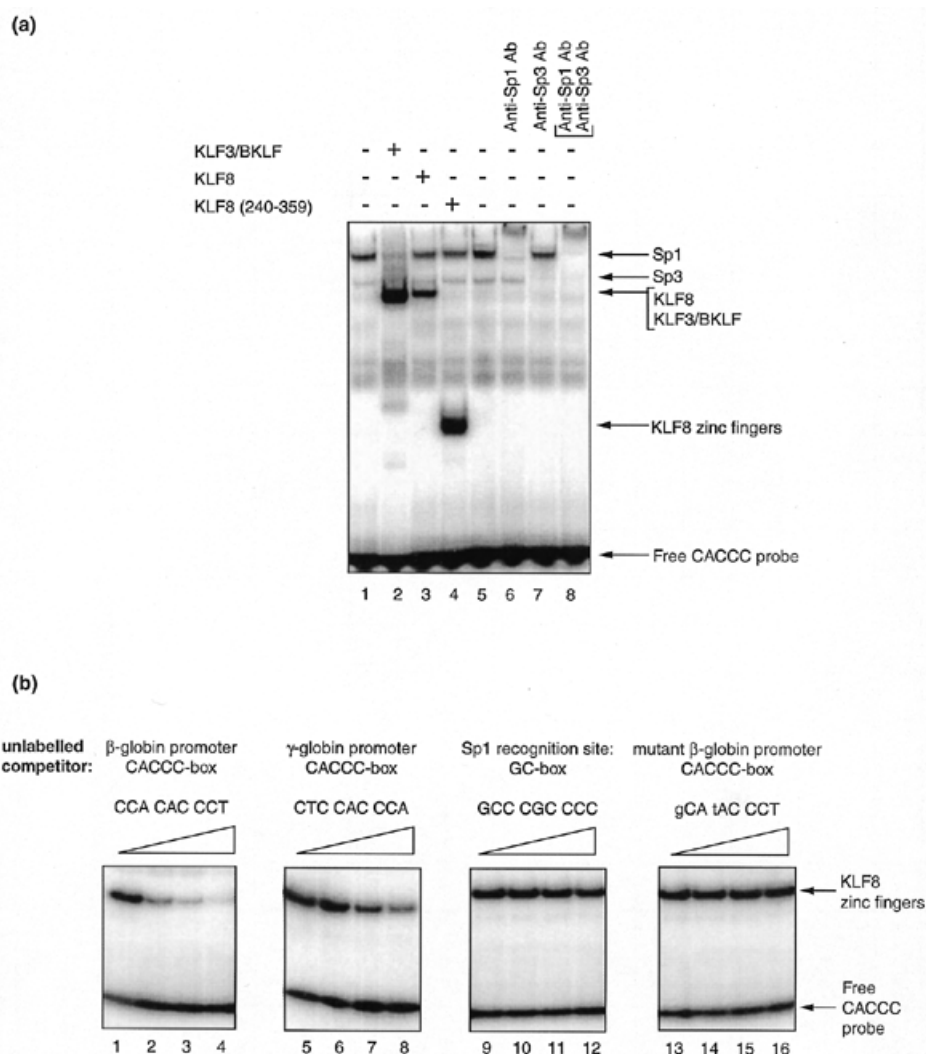


Figure 2. KLF8 recognises the CACCC-box from the β -globin promoter. **(a)** A gel retardation experiment showing retarded complexes generated by the binding of proteins present in COS cell nuclear extracts to a radiolabelled double-stranded oligonucleotide containing a CACCC element. Lanes 1–4 contain nuclear extracts from: COS cells transfected with pcDNA3 vector alone; cells transfected with pMT2/(KLF3/BKLF); cells transfected with pcDNA3/KLF8; cells transfected with pcDNA3/KLF8(240–359) encompassing the entire zinc finger domain of KLF8. Lanes 5–8 also contain nuclear extracts from COS cells transfected with pcDNA3 vector alone and in addition contain antibodies as follows: lane 6, anti-Sp1; lane 7, anti-Sp3; lane 8, anti-Sp1 and anti-Sp3. The positions of bands generated by endogenous Sp1 and Sp3 are shown to the right of the figure, together with the complexes generated by KLF3/BKLF, KLF8 and the zinc fingers of KLF8 and the position of the unbound probe. **(b)** Competition experiments demonstrate that the KLF8 zinc fingers bind most tightly to the β -globin CACCC box. All lanes contain nuclear extract from COS cells transfected with pcDNA3/KLF8(240–359) as in lane 4 above, but in addition unlabelled competitor double-stranded oligonucleotides have been added as follows: lanes 1–4 contain 0, 10, 30 and 100 ng of the β -globin CACCC site; lanes 5–8 contain the same amounts of the γ -globin CACCC site; lanes 9–12 contain the GC-rich Sp1 site; and lanes 13–16 contain the mutant β -globin CACCC site. See Materials and Methods for the precise sequence of all oligonucleotides.

to certain CACCC than to GC-rich sequences (3). To test specificity of binding by KLF8, competition experiments were performed using double-stranded oligonucleotides encompassing the β -globin CCA CAC CCT element, the γ -globin CTC CAC CCA element and a typical GC-element (known to be recognised by Sp1) GCC CGC CCC. A negative control oligonucleotide with two changes within the β -globin CACCC element (gCA tAC CCT) was also included. As shown in Figure 2b, the β -globin site competed more effectively than the γ -globin element, while the GC-rich sequence and mutant CACCC element competed very little. The pattern was the same as that previously reported for

KLF3/BKLF (3) and this result is consistent with the high level of conservation between KLF3/BKLF and KLF8 in the zinc finger region (Fig. 1).

KLF8 associates with the co-repressor protein CtBP

In many instances, comparisons between different members of the KLF family reveal no similarities outside the zinc finger region. Figure 1, however, shows that there are a number of additional short stretches of homology shared by KLF3/BKLF and KLF8. Most notably, the putative nuclear localisation sequence immediately upstream of the fingers (residues 269–272

in KLF8), a small hydrophobic domain rich in Val residues (residues 179–190) and finally a Pro-Val-Asp-Leu-Ser motif (residues 86–90). A related motif in KLF3/BKLF is known to be required for its physical association with the co-repressor protein CtBP (11).

We therefore tested whether KLF8 could also associate with CtBP and whether this association involved its Pro-Val-Asp-Leu-Ser motif. Specifically, we used the yeast two-hybrid assay and GST pull-down assays to investigate the association of CtBP with N-terminal portions of KLF8, containing intact or mutated Pro-Val-Asp-Leu-Ser motifs. Since mammalian CtBP1 and CtBP2 proteins are highly related and have indistinguishable binding characteristics, we confined our analysis to the murine CtBP2 protein (11). Yeasts were transfected with an expression vector encoding the Gal4 DNA-binding domain fused to CtBP2 and a vector encoding the Gal4 activation domain fused to amino acids 1–262 of KLF8 or a similar plasmid encoding an Asp-Leu→Ala-Ser double substitution in the putative CtBP recognition motif. The patterns of yeast growth shown in Figure 3a indicate that the intact KLF8 domain interacted with CtBP2, whereas the domain containing the mutated motif did not. Negative control yeast strains transfected with vectors encoding Gal4 domains without attached KLF8 and CtBP2 fusions also showed no growth (data not shown).

Further evidence for an interaction between KLF8 and CtBP2 was obtained using GST pull-down experiments (Fig. 3b). Here the ability of GST-CtBP2 to retain radiolabelled KLF8 and KLF8 containing a mutation in the Pro-Val-Asp-Leu-Ser motif was compared. As can be seen from the figure, GST-CtBP2 efficiently retained the intact KLF8 but retained very little of the mutant form. These results indicated that CtBP2 can physically associate with the N-terminal domain of KLF8.

KLF8 can repress a CACCC-box-dependent promoter

We next tested whether KLF8 acted as a repressor of transcription. We co-transfected NIH 3T3 cells with an expression vector encoding KLF8 and a reporter promoter containing three tandem β -globin CACCC elements upstream of the thymidine kinase promoter. We also tested a KLF8 derivative containing a mutation in the CtBP recognition motif. As shown in Figure 4a, KLF8 repressed transcription, as did the mutant protein. These results indicated that KLF8 can act as a transcriptional repressor but suggest that it is not solely reliant on its CtBP contact motif for this activity.

We have previously observed similar phenomena using KLF3/BKLF (11) and FOG (19). In both cases, these proteins contain discrete repression domains that are dependent on CtBP but appear to have additional repression domains that compensate when the CtBP-binding motif is mutated in the context of the full-length protein. We tested whether the N-terminal domain of KLF8, when fused to a heterologous DNA-binding domain, acted as a repression domain that was dependent on CtBP. Specifically, we tested intact and mutant N-terminal domains of KLF8 fused to the Gal4-DNA binding domain for their ability to repress a Gal4-dependent promoter in NIH 3T3 cells (Fig. 4b). As can be seen from the figure, the intact N-terminus of KLF8 functioned as a repression domain, whereas the mutant version did not. Taken together these results indicated that, like KLF3/BKLF, KLF8 does contain an N-terminal repression domain that requires CtBP for its activity, but that

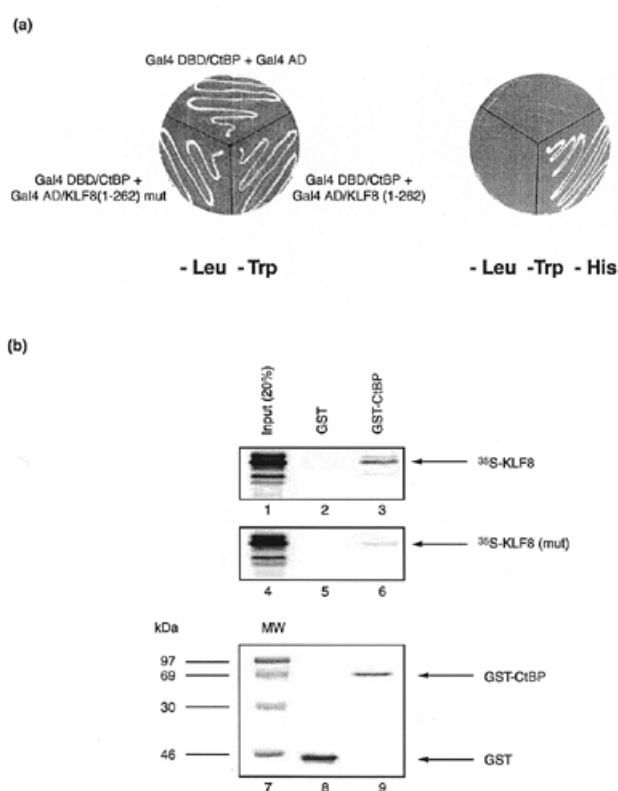


Figure 3. KLF8 physically interacts with CtBP and the Pro-Val-Asp-Leu-Ser motif is required for this interaction. (a) A yeast two-hybrid experiment showing the interaction between KLF8 and CtBP. On the left yeast strains harbouring the test plasmids indicated are shown growing on minimal media deficient in tryptophan and leucine. On the right the same yeast are shown on minimal media that also lacks histidine. As can be seen, neither the negative control strain, which harbours a plasmid encoding the Gal4 activation domain together with a plasmid encoding the Gal4 DNA-binding domain/mCtBP2 fusion, nor the strain which contains a plasmid encoding KLF8 with a Asp-Leu→Ala-Ser mutation in the core CtBP contact motif, grow on histidine deficient media. But in the remaining strain, the physical interaction of the Gal4 DNA-binding domain/mCtBP2 fusion and the intact Gal4 activation domain/KLF8(1–262) fusion allows the reconstitution of Gal4 activity, the activation of the *HIS3* reporter and growth on media lacking histidine. (b) A GST-pull-down experiment comparing the ability of GST-mCtBP2 to retain either radiolabelled KLF8 or a mutant version of KLF8 containing an Asp-Leu→Ala-Ser mutation within the Pro-Val-Asp-Leu-Ser motif. GST-mCtBP2 fusion protein and control GST were immobilised on agarose beads and mixed with radiolabelled KLF8. The beads were then washed repeatedly, boiled in loading buffer and subjected to electrophoresis. The resulting gel was examined by Phosphorimaging for the presence of KLF8 and mutant KLF8 which had been retained by the GST-mCtBP2 fusion or GST alone. Lanes 1 and 4 show 20% of the total input radiolabelled KLF8 and mutant KLF8, respectively. Lanes 2 and 5 show the background amounts of intact and mutant KLF8 retained by GST alone. Lanes 3 and 6 show the amounts of intact and mutant KLF8 retained by GST-mCtBP2, respectively. As can be seen by comparing lanes 3 and 6, a significant amount of intact KLF8 is retained by GST-mCtBP2, whereas very little of the mutant KLF8 is retained, indicating that an intact Pro-Val-Asp-Leu-Ser motif is involved in the contact of KLF8 and mCtBP2 *in vitro*. The bottom panel shows the gel stained with Coomassie blue, with molecular weight size markers in lane 7, GST protein in lane 8 and GST-mCtBP2 in lane 9.

the full-length protein has the ability to repress transcription even when the Pro-Val-Asp-Leu-Ser motif is mutated.

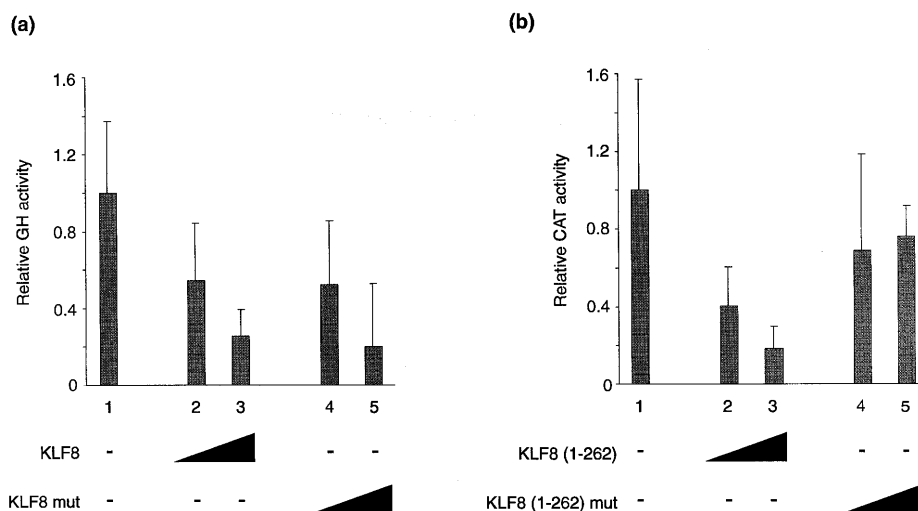


Figure 4. KLF8 represses transcription. (a) Full length KLF8 represses a CACCC-dependent promoter. The ability of full-length KLF8 to repress transcription was compared with a mutant version carrying an Asp-Leu→Ala-Ser mutation in the CtBP recognition motif. KLF8 expression plasmids were co-transfected together with a reporter containing three copies of the β -globin CACCC box upstream of the Herpes thymidine kinase promoter and a GH reporter gene. Reporter (0.5 μ g) was used in all lanes, together with 1 and 4 μ g of pcDNA3/KLF8 in lanes 2 and 3, or pcDNA3/KLF8 Asp-Leu→Ala-Ser in lanes 4 and 5. (b) The ability of the N-terminal domain of KLF8 to function as a repression domain was compared with that of a mutant version containing a Asp-Leu→Ala-Ser mutation in the core CtBP binding motif. Gal4 DNA-binding domain/KLF8(1-262) fusions were co-transfected into NIH 3T3 cells, together with a reporter containing five Gal4 recognition elements upstream of a Herpes thymidine kinase promoter and a CAT reporter gene. Reporter (2.5 μ g) was used in all lanes, together with 0.5 and 4 μ g of an expression plasmid encoding intact Gal4 DNA-binding domain-KLF8(1-262) in lanes 2 and 3, or the mutant version in lanes 4 and 5.

KLF8 is expressed widely in human tissues

To investigate the expression profile of *KLF8* we analysed mRNA from several human tissues by northern blot. To reduce the possibility of detecting cross-hybridising transcripts we did not use the region encoding the zinc finger domain as a probe but made two distinct probes from the non-finger region. The first probe consisted of nt 168–465 and the results are shown in Figure 5. Similar results were obtained with a probe containing nt 465–718 on this and an additional northern blot containing similar human RNA samples (data not shown). Four major transcripts were detected in most human tissues. The significance of these different transcripts is not known and it is possible that they arise from alternative promoter or polyadenylation usage or from alternative splicing. The relative levels of the four transcripts are similar in different tissues and the transcripts are most abundant in kidney, heart and placenta but are also detectable in other tissues. These findings are consistent with the fact that expressed sequence tags homologous to *KLF8* have been detected in libraries from several tissues, including kidney, heart, uterus, brain, skeletal muscle and testes. We also observed the same transcripts in K562 mRNA (data not shown).

DISCUSSION

KLF8 is a sequence-specific DNA-binding protein that can recognise CACCC-box elements and repress gene expression. It contacts DNA through its C-terminal zinc finger region and represses transcription, in part, by recruiting the co-repressor

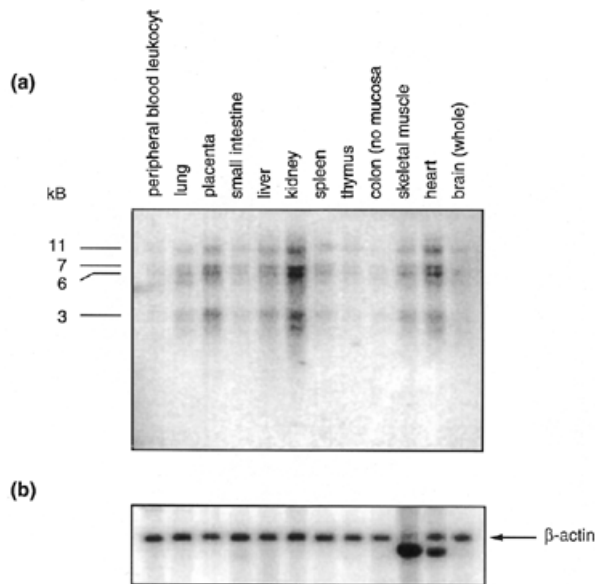


Figure 5. The expression of *KLF8* mRNA in human tissues analysed by northern blotting. (a) The tissues from which the mRNA was derived are shown above the figure. The major bands migrating at 3, 6, 7 and 11 kb are indicated. (b) The blot was then reprobbed with human β -actin in order to control for the loading and integrity of the mRNA samples.

CtBP through a Pro-Val-Asp-Leu-Ser motif in its N-terminus. *KLF8* is, therefore, very similar to *KLF3/BKLF*, which also

binds to CACCC sequences and also recruits CtBP. Both KLF3/BKLF and KLF8, however, may rely on additional repression mechanisms, since mutation of the CtBP contact motif in either protein does not completely abolish repression activity. The residual activity may in part involve the exclusion of endogenous activating CACCC-box proteins by competition or alternatively the proteins may recruit additional co-repressors. Interestingly, another protein that also binds CtBP, Drosophila Hairy, has a second repression domain that is known to recruit the co-repressor Groucho (13) and there is some evidence that KLF3/BKLF may also recruit a second co-repressor protein (11 and unpublished results).

Sp/KLF family members have been reported to bind to a large number of different CACCC-box elements *in vitro* and to regulate the activity of a range of promoters in transfection experiments. The *in vivo* targets of the different factors, however, have been difficult to define. The observation that the β -globin gene is under-expressed in KLF1/EKLF null embryos is strong evidence that it directly regulates this gene *in vivo* (6,7). The observation that in Sp1 deficient cells the thymidine kinase and methyl cytosine binding protein 2 genes are under-expressed suggests that Sp1 may regulate these genes (20). A number of other candidate target genes have been proposed for several other family members (1,2 and references therein) and as knockout work continues it will be interesting to see whether the expression of these putative targets is affected. In some instances, of course, redundancy among co-expressed family members may complicate the interpretation of these experiments.

We originally isolated KLF8 from cDNA derived from the erythroid cell line K562. A number of other CACCC-box binding proteins are co-expressed in erythroid cells, including KLF1/EKLF, KLF3/BKLF, KLF11/TIEG2/FKLF as well as Sp1, Sp3 and Sp4 (3,5,9,21,22 and unpublished results). Since functional CACCC-boxes are present in the promoters of the ϵ -, γ - and β -globin genes and in the β -globin LCR, it has been proposed that several of these proteins are involved in regulating globin gene expression (5–10,22). As previously mentioned, KLF1/EKLF has been shown to have a critical role in activating transcription of the adult β -globin gene *in vivo*, and it has been proposed that KLF11/TIEG2/FKLF plays a similar role in activating the γ -globin gene earlier in development (22). It is possible that the CACCC-box binding factors (such as KLF3/BKLF and KLF8) that are capable of associating with the co-repressor CtBP and repressing gene expression may have roles in the silencing of CACCC-dependent globin promoters during development. It has been anticipated that developmentally stage-specific transcription factors might mediate globin gene switching. The realisation, however, that high level transcription of the β -globin gene is specifically activated by KLF1/EKLF in adult cells, although KLF1/EKLF is present throughout development, raises the possibility that other more broadly expressed proteins may also facilitate specific developmental switches. If this is the case, it is to be expected that additional levels of control, such as protein phosphorylation, acetylation or the availability of specific cofactors, must operate to limit the activity of the broadly expressed protein, as they have been shown to influence the activity of KLF1/EKLF (1,2).

KLF8 shares some sequence similarity with KLF3/BKLF and both genes are co-expressed in many mammalian tissues.

Although we originally noted the similarity between human KLF8 and murine KLF3/BKLF proteins, the sequence of human KLF3/BKLF is also now available in the sequence databases (GenBank accession number AR001759) and it appears that the human and murine KLF3/BKLF protein (but not gene) sequences are identical. The murine *KLF8* gene has not yet been reported. The murine *KLF3/BKLF* gene has recently been inactivated by homologous recombination in murine embryonic stem cells and the resulting knockout mice appear to suffer from a myeloproliferative disorder (S.H.Orkin *et al.*, unpublished results). Since KLF3/BKLF and KLF8 seem to be co-expressed in certain tissues and both bind the protein partner CtBP and similar DNA sequences, it will be interesting to investigate possible overlap in the functions of the two proteins, or possible interplay in the expression of the two genes. As well as being related to KLF3/BKLF, KLF8 is also closely related to another family member KLF12/Ap2-rep (18). This protein was originally identified as a protein that recognized a CACCC-related element in the Ap-2 α promoter and repressed gene expression. Work is currently being undertaken to investigate its mechanism of action (R.Buettner *et al.*, personal communication and unpublished results). Unlike KLF3/BKLF and KLF8, which are expressed in a range of tissues, the expression of KLF12/Ap2-rep is confined to the kidney, with low levels also occurring in liver, lung and brain. The three proteins form a subfamily of KLF factors that share some common sequences. It should be noted, however, that outside the zinc finger region, the homology is confined to very small sequence motifs and the proteins also contain large domains where little similarity exists. Future work in mice and *in vitro* will undoubtedly shed further light on the similarities and differences in the biological roles and molecular mechanisms of action of these proteins.

ACKNOWLEDGEMENTS

We are grateful to members of the laboratory for their reading of the manuscript and helpful suggestions. J.v.V. and J.T. are supported by Australian Postgraduate Awards. This work is supported by a grant from the Australian NHMRC to M.C.

REFERENCES

- Philipsen,S.P. and Suske,G. (1999) *Nucleic Acids Res.*, **27**, 2991–3000.
- Turner,J. and Crossley,M. (1999) *Trends Biochem. Sci.*, **24**, 236–241.
- Crossley,M., Whitelaw,E., Perkins,A., Williams,G., Fujiwara,Y. and Orkin,S.H. (1996) *Mol. Cell. Biol.*, **16**, 1695–1705.
- Miller,I.J. and Bieker,J.J. (1993) *Mol. Cell. Biol.*, **13**, 2776–2786.
- Donze,D., Townes,T.M. and Bieker,J.J. (1995) *J. Biol. Chem.*, **270**, 1955–1959.
- Nuez,B., Michalovich,A., Bygrave,R., Ploemacher,R. and Grosveld,F. (1995) *Nature*, **375**, 316–318.
- Perkins,A.C., Sharpe,H. and Orkin,S.H. (1995) *Nature*, **375**, 318–322.
- Wijgerde,M., Gribnau,J., Trimbom,T., Nuez,B., Philipsen,S., Grosveld,F. and Fraser,P. (1996) *Genes Dev.*, **10**, 2894–2902.
- Gillemans,N., Tewari,R., Lindeboom,F., Rottier,R., de Wit,T., Wijgerde,M., Grosveld,F. and Philipsen,S. (1998) *Genes Dev.*, **12**, 2863–2873.
- Armstrong,J.A., Bieker,J.J. and Emerson,B.M. (1998) *Cell*, **95**, 93–104.
- Turner,J. and Crossley,M. (1998) *EMBO J.*, **17**, 5129–5140.
- Schaepfer,U., Boyd,J.M., Verma,S., Uhlmann,E., Subramanian,T. and Chinnadurai,G. (1995) *Proc. Natl Acad. Sci. USA*, **92**, 10467–10471.
- Poortinga,G., Watanabe,M. and Parkhurst,S.M. (1998) *EMBO J.*, **17**, 2076–2078.
- Nibu,Y., Zhang,H. and Levine,M. (1998) *Science*, **280**, 101–104.

15. Criqui-Filipe,P., Ducret,C., Maira,S.M. and Wasyluk,B. (1999) *EMBO J.*, **15**, 3392–3403.
16. Postigo,A.A. and Dean,D.C. (1999) *Proc. Natl Acad. Sci. USA*, **96**, 6683–6688.
17. Sundqvist,A., Sollerbrant,K. and Svensson,C. (1998) *FEBS Lett.*, **429**, 183–188.
18. Imhof,A., Schuierer,M., Werner,O., Moser,M., Roth,C., Bauer,R. and Buettner,R. (1999) *Mol. Cell. Biol.*, **19**, 194–204.
19. Fox,A.H., Liew,C., Holmes,M., Kowalski,K., Mackay,J. and Crossley,M. (1999) *EMBO J.*, **18**, 2812–2822.
20. Marin,M., Karis,A., Visser,P., Grosveld,F. and Philipsen,S. (1997) *Cell*, **89**, 619–628.
21. Cook,T., Gebelein,B., Belal,M., Mesa,K. and Urrutia,R. (1999) *J. Biol. Chem.*, **274**, 29500–29504.
22. Asano,H., Li,X.S. and Stamatoyannopoulos,G. (1999) *Mol. Cell. Biol.*, **19**, 3571–3579.