

A gene encoding an intestinal-enriched member of the Krüppel-like factor family expressed in intestinal epithelial cells

Michael D. Conkright, Maqsood A. Wani, Kathleen P. Anderson and Jerry B. Lingrel*

Department of Molecular Genetics, Biochemistry and Microbiology, University of Cincinnati College of Medicine, 231 Bethesda Avenue, Cincinnati, OH 45267-0524, USA

Received as resubmission December 21, 1998; Revised and Accepted January 14, 1999

DDBJ/EMBL/GenBank accession no. AF079852

ABSTRACT

The Krüppel-like factors make up a multigene family of transcription factors that have discrete patterns of expression, implying they play important biological roles in the tissues in which they are expressed. We have identified and characterized the cDNA for a novel murine transcription factor that is an additional member of the Krüppel-like family of transcription factors, named intestinal-enriched Krüppel-like factor (IKLF). This gene appears to be a homolog of the human BTEB-2 gene, although it exhibits a different pattern of tissue expression and the translated product is larger. IKLF is expressed in a limited number of tissues; the highest levels of IKLF expression are found in the digestive tract. IKLF shows temporal changes in expression during embryogenesis indicating that this gene is developmentally regulated. In addition, IKLF expression is limited to the epithelial lining of the intestine and is localized primarily to the base of the crypts in the adult intestine. The IKLF cDNA encodes for a 446 amino acid protein and is able to transactivate by binding specific DNA elements that are also recognized by other members of the Krüppel-like family. In addition, mutations in the activation domain attenuate the ability of this protein to function as a transcription factor. Collectively, these findings show that we have identified a transcription factor that is expressed predominantly in the epithelial crypt cells of the gastrointestinal tract and is a member of the Krüppel-like family of transcription factors.

INTRODUCTION

The epithelial lining of the intestine is a highly prolific tissue composed of two distinct compartments, the crypts of Lieberkühn and the villus. The adult mouse small intestine contains ~1 100 000 crypts with the stem cells located in each crypt producing ~300 cells/day (1). These daughter cells migrate from the crypt to the apex of the surrounding villus. Once the cells migrate from the crypt they lose their proliferative capacity and undergo terminal differentiation. The change in cellular phenotype during

migration from the crypt to the apex is mediated by the complement of transcription factors (both positive and negative) present in an epithelial cell at any given time. Thus, changes in the repertoire of gene expression create regional differences in gene expression along the crypt–villus axis (2,3). It has been suggested that the epithelial cells receive signals during their migration away from the crypts that instruct them to commit themselves to terminal differentiation (4). However, the source of such signals or their molecular mediators are currently unknown. The prolific nature of the intestine and continuous cellular differentiation in the crypt–villus axis makes the intestine an attractive model for studying the mediators of cellular commitment and differentiation.

Examination of factors that mediate intestinal-specific genes should be helpful in understanding the complexities of epithelial cell differentiation and intestinal-specific gene expression. To date, only a limited number of transcription factors have been implicated in regulating the expression of intestinal-specific genes. One such protein, gut-enriched Krüppel-like factor (GKLF/EZF), is expressed primarily in the digestive tract and epidermal layer of the skin with minimal expression in other tissues (5,6). Specifically, GKLF expression occurs in the middle to upper crypt region, which contains cells that have primarily undergone growth arrest (5,7). Expression of GKLF is down-regulated during intestinal tumorigenesis and in proliferating cells (5,8). Similarly, overexpression of GKLF in cell culture inhibits DNA synthesis (5). These studies suggest that GKLF/EZF has an important role in regulating growth of intestinal epithelial cells migrating to the apex of the villus. In addition GKLF potentially functions in the terminal differentiation of these cells.

This study describes a novel murine transcription factor that is an additional member of a multigene family of transcription factors called the Krüppel-like factors. Similar to GKLF, this novel murine transcription factor is expressed in a tissue-selective manner, with the highest expression occurring in the gastrointestinal tract. *In situ* hybridization shows that IKLF is expressed in intestinal epithelial cells. The overall expression pattern as well as the lack of homology in the activation domain, however, clearly distinguish GKLF and IKLF as separate gene products. We have named this protein intestinal-enriched Krüppel-like factor to denote the relatedness to other Krüppel-like factors and underscore the predominant expression in the intestinal tract.

*To whom correspondence should be addressed. Tel: +1 513 558 5324; Fax: +1 513 558 1190; Email: jerry.lingrel@uc.edu

MATERIALS AND METHODS

Isolation of IKLF cDNA

A sequence comparison of the zinc finger region of LKLF was performed using the BLAST algorithm provided by the National Center for Biotechnology Information (Rockville, MD) to search the expressed sequence tag (dbEST) nucleotide database. Four overlapping IMAGE Consortium (LLNL) cDNA clones (accession nos AA623934, AA538230, AA791786 and W62866) were obtained from Genome Systems Inc. (St Louis, MO). The cDNA clone was subjected to double-stranded nucleotide sequence analysis with a model 373 DNA sequencer and the *Taq* Dye deoxy sequencing protocol (Applied Biosystems Inc., Foster City, CA). The DDBJ/EMBL/GenBank accession no. for IKLF is AF079852.

Plasmid constructs

The full-length IKLF cDNA fragment was subcloned into the expression vector pBK-CMV (Stratagene, La Jolla, CA) which utilizes the cytomegalovirus immediate-early promoter. Five IKLF mutants with internal deletions in the activation domain fused to the DNA-binding domain were constructed. An additional five IKLF mutants were constructed that contained the first 231 amino acid residues in addition to the mutations mentioned above. Each deletion mutant was constructed by amplifying a portion of the activation domain and the DNA-binding domain by PCR. The two PCR products were ligated at a common *Xho*I restriction site introduced by a silent mutation. In addition, restriction sites were added to the distal primers to allow cloning into the expression vector pBK-CMV in *Bam*HI and *Hind*III sites. The 5'-end of the cDNA was then added by use of *Pvu*II and *Eco*RI restriction sites. The oligos used to amplify the DNA-binding domain were 5'-CCCAAGCTTGGATGGCAAAGCGGACTG-3' and 5'-GGG-AAGCTTCTATGCATAGTCCGGGACGTCATAGGGATAGTT-CTGGTGGCGCTTCATG-3'. The 3' DNA-binding domain primer also coded for a hemagglutinin (HA) epitope, thereby adding an epitope tag on the C-terminus of the protein for detection by western blotting. The five activation domain mutants were constructed by a common 5' oligo, 5'-CGGGATCCCAGTATCTTCAGTCC-3', and an oligo unique for each mutant: IKLF Δ 345-357, 5'-GACAGGCTCGAGAGTTGGCG-3'; IKLF Δ 327-357, 5'-GG-TTGTGAATCTCGAGTTTGG-3'; IKLF Δ 311-357, 5'-GAGT-GGCTCGAGATTCTG-3'; IKLF Δ 283-357, 5'-CCTGCTGCT-CGAGAAACTGAC-3'; IKLF Δ 252-357, 5'-GGGTGTGGCA-GGCCTGC-3'. All PCR amplified fragments were verified by nucleotide sequence analysis.

The reporter construct used in the initial transactivation assay is HS2- β CAT, containing an HS2 fragment from the locus control region and the human β -globin promoter region from -265 to the translation start site as previously described (9). Restriction sites were added to the primers such that this fragment could be cloned with *Sal*I and *Hind*III. By using these sites, the fragment was subcloned into a Bluescript vector (Stratagene) containing the chloramphenicol acetyltransferase (CAT) gene in the *Hind*III and *Bam*HI sites. The HS2 fragment was subsequently added. The two consensus CACCC elements in the β -globin promoter were destroyed by introduction of mutations as previously described elsewhere (9).

Transfections and CAT assays

Mouse NIH 3T3 fibroblasts were plated at a density of 5×10^5 cells/100 mm tissue culture dish. The following day, cells were transfected by CaPO₄ precipitation (10) with 10 μ g of reporter (HS2- β CAT) plasmid, 5 μ g of pSV2LUC control plasmid and 10 μ g of IKLF or IKLF deletion mutant. The medium was changed 24 h after the CaPO₄ treatment and the cells were harvested 48 h after transfection. The cell pellet was divided into two fractions; one fraction was used to assay CAT activity and the second was used for western blot analysis. The luciferase and CAT activities were determined by disrupting the cells by three cycles of freeze-thaw lysis in 0.25 M Tris (pH 7.5) to make crude protein extract. An aliquot of protein extract was used for analysis of luciferase activity (Promega, Madison, WI). The remaining extract was heat inactivated at 65°C for 10 min. Extract amounts were normalized for transfection efficiencies and CAT assays were performed at 37°C for 1 h. The thin-layer chromatography plates were exposed to a PhosphorImager plate (Molecular Dynamics, Sunnyvale, CA) for quantitation. Normalized values for the CAT activity are based on percent conversion of [¹⁴C]chloramphenicol substrate to the acetylated forms and corrected for transfection efficiency with luciferase activity.

Western blot analysis of IKLF and IKLF mutants

IKLF or IKLF deletion mutants were transfected into NIH 3T3 cells for 48 h as described above. Cells were lysed with Laemmli buffer (0.0625 M Tris, pH 6.8, 2% SDS, 6 M urea, 0.150 M DTT, 0.005% BPB) and equal amounts of lysate based on luciferase activity of each sample were separated by electrophoresis on a 9% SDS-polyacrylamide gel. Following electrophoresis, the proteins were transferred to Hybond-P (Amersham) PVDF transfer membrane. After blocking, 0.1 μ g/ml rabbit polyclonal antibody raised against influenza HA protein (Santa Cruz Biotechnology) was incubated with the membrane. The secondary antibody, horseradish peroxidase-conjugated goat anti-rabbit (Amersham) was used at a 1:10 000 dilution. Protein-antibody interaction was visualized by chemiluminescence detection with the ECL western blotting analysis system (Amersham).

Northern blot analysis of IKLF

Total RNA was isolated from the uterus, placenta, testis, skin, kidney, brain, adipose, heart, spleen, liver, lung, skeletal muscle, stomach, small intestine, large intestine, proximal colon, distal colon and cecum of mouse. A total of 15 μ g of RNA from each tissue was used for northern blot analysis (10). The developmental blot containing 2 μ g of poly(A) RNA from mouse embryos at days 7, 11, 15 and 17 of gestation was purchased from Clontech Laboratories Inc. (Palo Alto, CA). The samples labeled on days 7 and 11 also included placenta and yolk sac. All blots were probed with the entire 1.4 kb IKLF cDNA and with an actin control probe supplied with the commercially prepared blot. No cross-hybridization with other Krüppel-like factors was noted under the conditions used in these analyses.

In situ hybridization

For *in situ* hybridization a 369 bp *Eco*RI-*Acc*I fragment of IKLF was subcloned into pBluescript and the recombinant plasmid was used to generate ³⁵S-labeled riboprobes. The plasmid DNA was linearized either with *Eco*RI and transcribed with T7 polymerase

to synthesize antisense RNA probe or linearized with *AccI* and transcribed with T3 polymerase to yield a sense RNA probe. 17.5 day mouse embryos were obtained from a timed pregnant mouse. The embryos were fixed in 4% paraformaldehyde overnight at 4°C and were embedded in paraffin. The adult tissues were dissected from mouse after whole animal perfusion and were similarly fixed and embedded. The embedded tissues were sectioned at 5 µm and the sections were layered onto RNase-free microscopic slides. The tissue sections were processed as described (11) for *in situ* hybridization at high stringency with either an antisense or a sense riboprobe. After hybridization the slides were washed and dipped in Kodak NTB-2 liquid emulsion. The emulsified slides were placed in lightproof boxes and exposed at 4°C for 2–3 weeks. The exposed slides were developed and viewed under dark-field microscopy.

RESULTS

Identification of mouse Krüppel-like cDNAs and their deduced amino acid sequence

To identify novel members of the Krüppel-like family of transcription factors a Blast search of the expressed sequence tag

(dbEST) database (12) was conducted with the zinc finger of LKLF and four overlapping IMAGE consortium cDNA clones obtained from Genome Systems Inc. which are homologous to the Krüppel-like family of transcription factors. The complete sequence of the cDNA obtained from multiple clones is 1591 bp in length and contains a 1338 bp open reading frame (ORF) along with a 166 bp 5'-untranslated sequence and an 84 bp 3'-untranslated sequence (Fig 1). An in-frame stop codon exists 129 nt (43 codons) upstream from the first ATG. The start of translation was confirmed by creating constructs that contained different potential start sites and size comparison by western blot analysis (data not shown). The ORF contains three zinc finger motifs, common to the Cys₂His₂ zinc finger class of DNA-binding proteins.

IKLF is a member of the Krüppel-like family of transcription factors

The mRNA corresponding to the new Krüppel-like factor cDNA is expressed largely in the intestine, as shown latter. We have named the coded protein intestinal-enriched Krüppel-like factor (IKLF) to underscore the predominant expression in the intestinal tract and denote the relationship to the other Krüppel-like factors. The nucleotide and amino acid sequence comparison between

ccgagcccaggagcccgatctccgtgcccgccttcg tg agcgtctggtgcccggcccagggtccccgcggcggccccc	82
gccgagtccgcccgtcccgtgcccagcccagcagaggtgggatcgcgatcgcctccgtgtcccgcctcccgtaatccccagaccgctcc	166
ATGCCACGCGGGTGCTGACCATGAGCGCCCGCCTGGGACCACTGCCCCAGCCGCGCGCCGCGCAGGCCGCGTGTTCGCG	250
M P T R V L T M S A R L G P L P Q P P A A Q D E P V F A	28
CAGCTCAAGCCGGTGTGGCGCTGCGAACCCGCGCCGCGAGCGGGCGCTCTCTCCGGAGACGATCTGAAACACGCGCACCAC	334
Q L K P V L G A A N P A R D A A L F S G D D L K H A H H	56
CACCCGCTCGCCCGCCGACCGCGTGGCCCGGACTGCCCTCGGAGGAGTGGTCCAGACAAGATGTAAATGGAGAAGTAT	418
H P P A P P P A A G P R L P S E E L V Q T R C E M E K Y	84
CTGACCCCTCAGTCCCTCCAGTTCCCGATAATTCAGAGCATAAAAAGTATAGACGAGACAGTGCCTCAGTGGTAGACAGTTC	502
L T P Q L P P V P I I S E H K K Y R R D S A S V V D Q F	112
TTCACTGACACTGAAGGCATACCTTACAGCATCAACATGAACGCTCTCTCCCTGACATCACTCACCTGAGAAGTGGCCTCTAC	586
F T D T E G I P Y S I N M N V F L P D I T H L R T G L Y	140
AAATCCCAGAGACCATGCGTAACACAGATCAAGACAGAACCTGTTACCATTTCAGCCACAGAGCGAGTCGACGGCCCTCCT	670
K S Q R P C V T Q I K T E P V T I F S H Q S E S T A P P	168
CCTCCTCGGCCCCACCAGGCTCTCCCGAGTTCACCTAGTATCTTCAGCTCCCACAGACACAGCGCCACCACAGGAGGTG	754
P P P A P T Q A L P E F T S I F S S H Q T T A P P Q E V	196
AACAATATCTTCATCAACAAGAAGTTCCTATACAGATCTTCATCTCTCTGCTCCCTCCAGCAGGCCACCTGTACCAGCTG	838
N N I F I K Q E L P I P D L H L S V P S Q Q G H L Y Q L	224
TTGAATACACCGGATCTAGACATGCCAGTTCGACAAACAGACGGCAGTAATGGACACCCCTAATGTTTCTATGGCAGGCCCT	922
L N T P D L D M P S S T N Q T A V M D T L N V S M A G L	252
AACCCACACCCCTCTGCTGTTCCACAGAGCTCAATGAAACAGTTCAGGGCATGCCCCCTTGCAAGTACACCATGCCAAGTACAG	1006
N P H P S A V P Q T S M K Q F Q G M P P C T Y T M P S Q	280
TTTCTCCACAGCAGGCCACTTATTTTCCCGGTCACCCAAAGCTCAGAGCCTGGAAGTCCCAGATAGACAAGCTGAGATGCTG	1090
F L P Q A T Y F P P S P P S S E P G S P D R Q A E M L	308
CAGAATCTCACCCACCTCCGCTATGCCGCTACAATGCTTCCAAACTGGCGATTACAAACCAATTTACCTGCCACTCTG	1174
Q N L T P P P S Y A A T I A S K L A I H N P N L P A T L	336
CCAGTAAATCGCCAATCTCCACCTGTGATACACAGAGGAGTAACCGGATCTGGAGAAGCGACGTATCCACTTCTGC	1258
P V N S P T L P P V R Y N R R S N P D L E K R R I H F C	364
GATTATAATGGTTGCACRAAAGTTTATACAAAGTCGTCTCACTTAAAGCTCACCTGAGGACTCATACGGGCGAGAAGCCCTAC	1342
D Y N G C T K V Y T K S S H L Y A H L R T H T G E K P Y	392
AAGTGCACTGGGAGGGCTGCGACTGGAGGTTTGCCCGGTCCGGATGAGCTGACCCGCCACTACAGGAAGCACAGGGCCGCAAG	1426
K C T W E G C D W R F A R S D E L T R H Y R K H T G A K	420
CGTTCCAGTGCATGGTGTGCCAACGAGCTTCTCCCGCTCCGACCACTCGCGTGCACATGAAGCGCCACCAGAAGTGAAGcg	1510
P F Q C M V C Q R S F S R S D H L A L H M K R H Q N *	446
agcgaacgctgcccaccgcccctgacgccttgagctccgctttgccatcctttaaacgcagacctaacttcataaaaaag	1591

Figure 1. Nucleotide and deduced amino acid sequence of the IKLF cDNA. The 5'- and 3'- untranslated nucleotide regions are in lower case letters and an in-frame stop codon upstream of the first ATG is shown in bold; the coding region is denoted in upper case. Amino acids are identified by their one-letter code and proline residues are presented in bold type. The three zinc fingers are underlined.

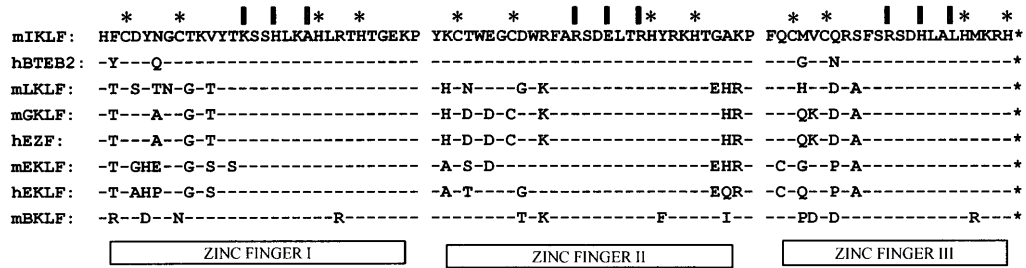


Figure 2. Alignment of the zinc fingers of IKLF with other members of the Krüppel-like family of transcription factors. Black squares highlight the amino acids that contact the DNA and asterisks indicate the cysteine and histidine residues of three zinc fingers.

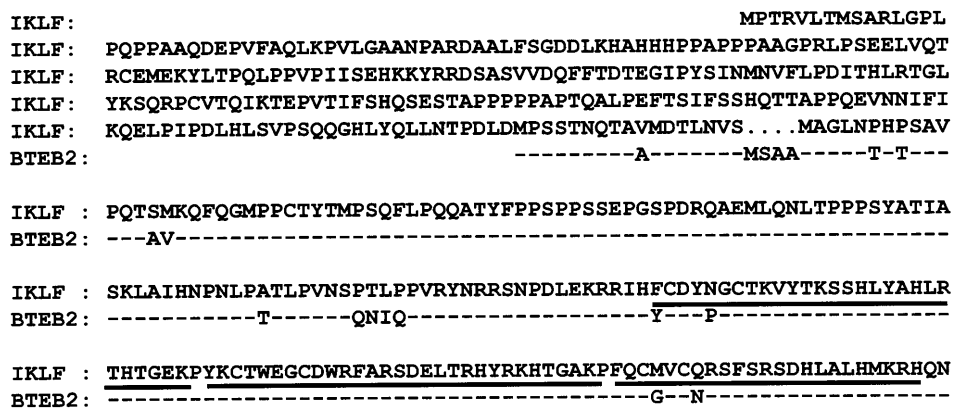


Figure 3. Comparison of murine IKLF and human BTEB-2 amino acids. Dashes represent identity to IKLF and dots represent gaps. The first methionine of BTEB-2 (15) corresponds to amino acid residue 232 of IKLF. The three zinc fingers are underlined.

Krüppel-like family members is highest in the DNA-binding domain but diverges significantly in the rest of the molecule, analogous to the GATA family of transcription factors. Figure 2 illustrates the similarity between the IKLF binding domain and the binding domain of other members of the Krüppel-like family, LKLF (9), GKLF/EZF (5,6), hEZF (13), BKLF (14), BTEB-2 (15) and murine and human EKLF (16,17). Clearly members of this family are conserved in the zinc finger region. In addition, the putative amino acids that contact the DNA major groove are identical, suggesting that members of this family bind to similar response elements.

Despite the lack of sequence homology, the activation domains of various members of Krüppel-like family members contain between 14 and 23% proline residues (9,16). IKLF contains 16% proline residues compared with a marginal 6% proline content in BTEB and SP1 (18). The significance of the proline-rich activation domain is currently unknown, however, the proline residues have been implicated in protein-protein interactions (19). The striking similarity of IKLF with other members of the Krüppel-like family suggests that this protein is an additional member of this family of transcription factors.

Additional sequence analysis reveals that IKLF is possibly a murine homolog of the human transcription factor BTEB-2 (15; Fig. 3). An amino acid comparison between BTEB-2 and IKLF shows 92% homology for the regions that overlap. However, BTEB-2 is only 219 amino acids (15) compared with IKLF which is 446 amino acids. Moreover, IKLF is expressed predominantly

in the gastrointestinal tract whereas BTEB-2 expression is less restrictive with expression seen in heart, placenta, lung, skeletal muscle, pancreas, bladder, prostate, small intestine, colon and stomach (data not shown).

IKLF DNA-binding sites

A hallmark feature of proteins belonging to the Krüppel-like family is the ability to bind the regulatory element CACCC. This is due to the high degree of sequence conservation among the zinc finger region of such proteins. Because the putative DNA contact sites of the zinc fingers are very well conserved between IKLF and the other members of the Krüppel-like family (Fig. 2), it is likely that IKLF can also bind the *cis* element CACCC. To test the ability of IKLF protein to mediate transcriptional activation by binding this element, the β -globin promoter, which contains two CACCC sites, driving the expression of CAT and an expression vector containing full-length IKLF cDNA or vector alone were co-transfected into mouse NIH 3T3 fibroblasts. In addition, the same promoter with the CACCC sites mutated was transfected in the absence or presence of IKLF. In these experiments IKLF is able to function as a transcription factor by transactivating the reporter gene by a 5-fold increase above vector alone. However, when the CACCC sites are mutated, IKLF is no longer able to transactivate. These results demonstrate that IKLF can transactivate and the ability to do so is dependent on an intact CACCC element (Fig. 4).

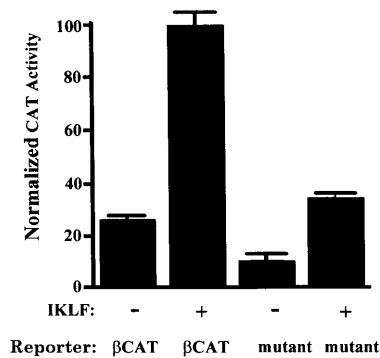


Figure 4. IKLF or pBKCMV (vector DNA) were co-transfected with either the β -globin promoter-CAT reporter construct or the β -globin promoter-CAT reporter construct with mutations in the two CACCC elements. Normalized CAT activity is represented in the bar graph. The activity obtained with the β cat construct was arbitrarily set at 100% for comparison of the mutant constructs (mean \pm SEM, $n = 3$ experiments).

IKLF activation domain

The activation domain of several of the Krüppel-like factors can be divided into repressor and activator subdomains (13,20). In order to further characterize the transactivation domain of IKLF, a series of deletions in the putative activation domain of this protein were produced. Five deletion mutants, shown in Figure 5A, IKLF- Δ 252-357, IKLF Δ 283-357, IKLF Δ 311-357, IKLF Δ 327-357 and IKLF Δ 345-357, were constructed and assayed for their ability to transactivate a reporter gene as described above. Deletions in the activation domain attenuate the ability of IKLF to transactivate (Fig. 5). To further characterize the activation domain, the first 231 amino acid residues were truncated from each mutant described above, creating double deletion mutants (IKLF Δ 1-231 + Δ 252-357, IKLF Δ 1-231 + Δ 283-357, IKLF Δ 1-231 + Δ 311-357, IKLF Δ 1-231 + Δ 327-357 and IKLF Δ 1-231 + Δ 345-357) (Fig. 5C). This additional deletion in all the mutants further suppressed the ability of the transcription factor to transactivate. There does not appear to be a particular region or subdomain that activates transcription, rather the entire activation domain is necessary for transactivation. In addition, the transactivating efficiency dramatically decreases with the increase in deletions in the activation domain. The attenuation in transactivation by the IKLF deletion mutants is unlikely due to instability of the expressed proteins, since an equivalent or greater amount of protein, as compared with wild-type, is detected from cell lysates by western blot analysis (Fig. 5). The use of a C-terminal epitope tag for western blot analysis also demonstrates that all mutants are being translated in the correct reading frame. In contrast to other members of this family, IKLF does not appear to contain an inhibitory subdomain (13,20). These results suggest that the ability of IKLF to transactivate is not dependent on a specific region of the transactivation domain and unlike other Krüppel-like factors it does not consist of separable repressor and activator subdomains.

Expression pattern of the IKLF gene

Transcription factors belonging to class II zinc fingers are known to be developmentally regulated and important during embryogenesis as well (21). To determine if IKLF is also developmentally regulated, IKLF expression during development was determined

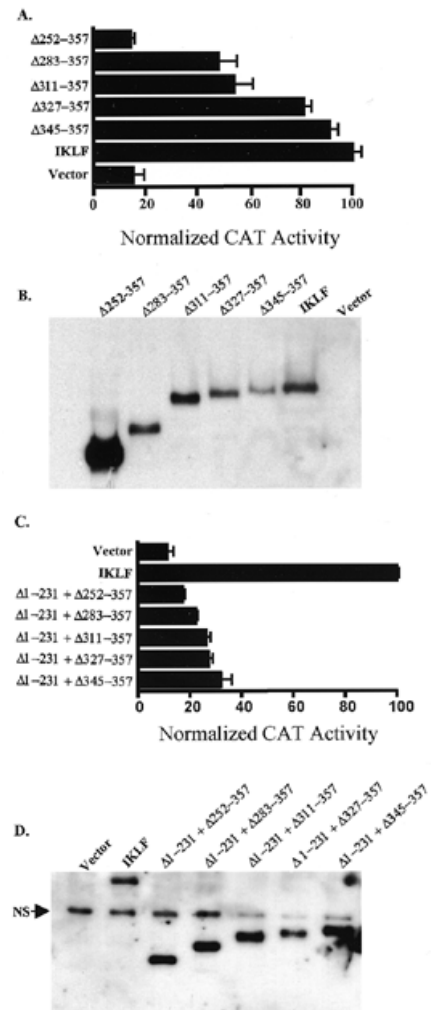


Figure 5. Deletion mutations in the activation domain of IKLF. IKLF or IKLF with deletions in the activation domain were co-transfected with HS2 β CAT reporter construct into NIH 3T3 cells. (A) CAT activity of IKLF and IKLF deletion mutants normalized to wild-type IKLF activity is represented in the bar graph (mean \pm SEM, $n = 3$ experiments). (B) Western blot of IKLF and IKLF deletion mutants tagged with the HA epitope. (C) CAT activity of IKLF and IKLF double deletion mutants normalized to wild-type IKLF activity is represented in the bar graph (mean \pm SEM, $n = 3$ experiments). (D) Western blot of IKLF and IKLF double deletion mutants tagged with the HA epitope. A non-specific band (NS) is seen in each lane.

by northern blot analysis. A northern blot containing poly(A) RNA from mouse embryos at 7, 11, 15 and 17 days of gestation was hybridized with a probe specific for IKLF (Fig. 6A). IKLF is abundantly expressed at day 7 of gestation but is completely absent at day 11. Low levels of expression return at days 15 and 17 of gestation. Similar to IKLF, other members of the Krüppel-like family such as EKLF, LKLF and GKLF also show changes in expression patterns during development, each having a unique temporal pattern (6,9). These temporal changes in expression during embryogenesis suggest that this gene is developmentally regulated.

To determine the tissues expressing IKLF, northern analysis with total RNA from 15 different tissues and three regions of the large intestine of adult mouse were conducted. IKLF is expressed

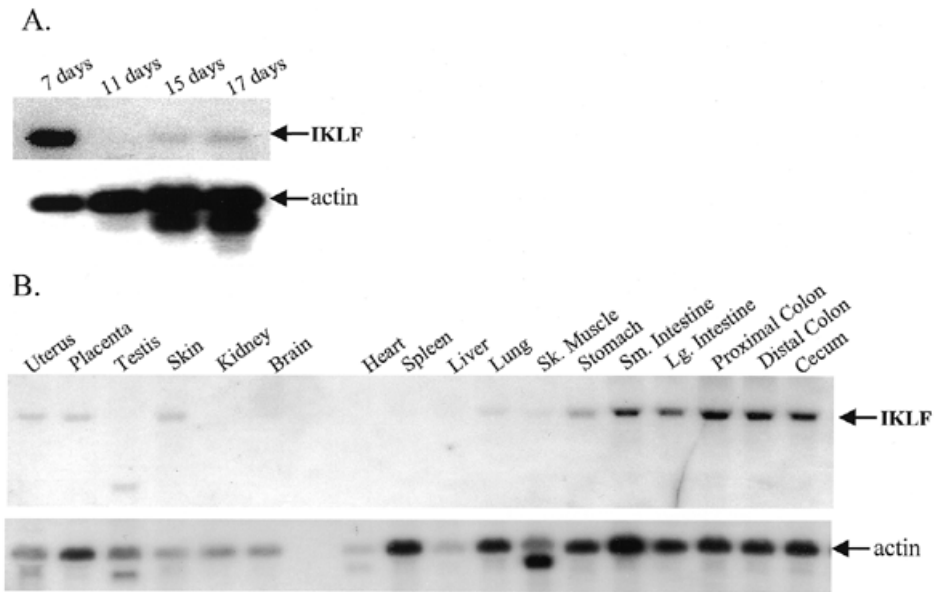


Figure 6. Northern blot of IKLF expression in adult mouse and embryo tissues. (A) Northern blot containing 15 μ g of RNA/lane was hybridized with a 1442 bp cDNA fragment of IKLF. (B) Northern blot of IKLF expression in mouse embryo. A northern blot containing 2 μ g of poly(A) RNA/lane was obtained from a commercial supplier and was hybridized with a 1442 bp cDNA fragment of IKLF. Both blots were stripped and hybridized with β -actin control probe.

predominantly in the gastrointestinal tract with high levels of expression of a 4.8 kb transcript observed in the stomach, small intestine and large intestine (Fig. 6B). Further subdivision of the large intestine into the proximal colon, distal colon and cecum revealed that IKLF is expressed equally in these different regions of the large intestine. In addition, lower levels of the IKLF transcript are present in the uterus, placenta, skin, lung and testis. The transcript present in the testis is smaller in size, similar to the northern analysis of EKLF (16) and LKLF (9), and may represent an alternatively spliced transcript. An identical pattern of tissue expression and level of expression was obtained when a rat poly(A) northern blot was hybridized with the same DNA probe, although placenta and skin were not included on the rat blot (data not shown). Like other members of the Krüppel-like family, IKLF is expressed predominantly in a tissue-restricted manner with some lower levels of expression in other tissues.

Whilst IKLF expression is predominant in the gastrointestinal tract, the stomach, small intestine and large intestine are complex organs with multiple tissue layers and cell types. To better localize the cell types expressing IKLF, *in situ* hybridization was conducted on sections of these organs. IKLF expression was observed only in the epithelial layer of the stomach, small intestine and colon. In contrast to GKLF, which is primarily found in the middle to upper region of the intestinal crypts (5), IKLF is enriched in intestinal epithelial cells located in the lower to middle crypt region (Fig. 7). In addition to examining adult organs, *in situ* hybridization on day 17 whole embryos were performed to examine expression of IKLF in multiple organs during development (Fig. 8). Similar to the adult tissues, IKLF is expressed in embryos in the epithelial layer of the gastrointestinal tract (Fig. 8). These results show that IKLF is expressed in the gastrointestinal tract and is further localized to the epithelial layer. Moreover, in the intestine IKLF expression is localized in the base of the crypts where active cell division is occurring.

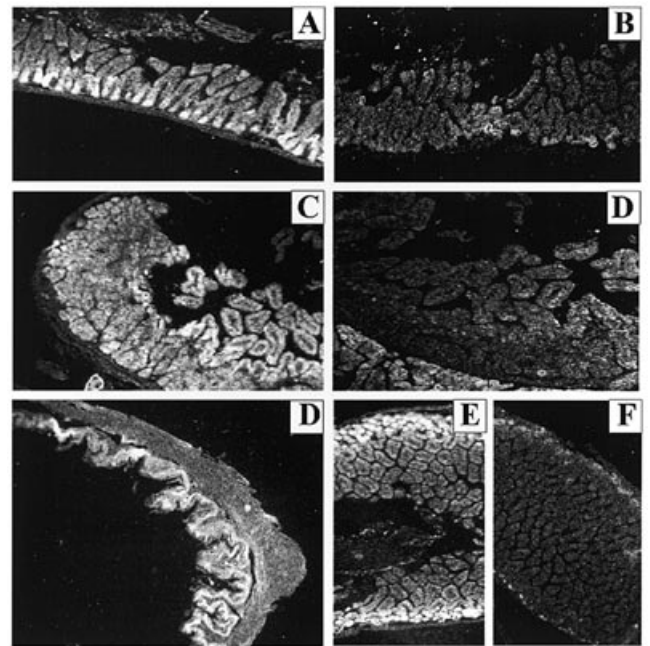


Figure 7. Expression of IKLF in adult mouse tissue as detected by *in situ* hybridization. Dark-field microscopy of (A) colon, (C) duodenum, (E) stomach and (F) small intestine, hybridized with antisense riboprobe or colon (B), duodenum (D) and (G) small intestine hybridized with sense riboprobe.

DISCUSSION

The cDNA sequence for a novel intestinal-enriched Krüppel-like factor was obtained by utilizing the expressed sequence tag database. Members of the Krüppel-like transcription factor possess structural features that classify them as part of a common

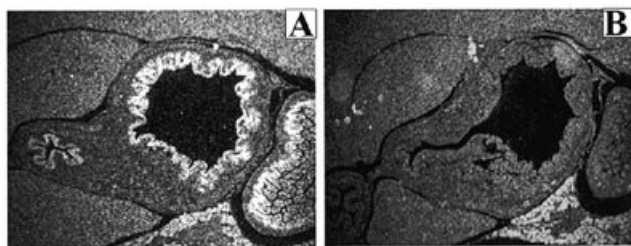


Figure 8. Expression of IKLF in mouse E17.5 embryo as detected by *in situ* hybridization. Dark-field microscopy of mouse E17.5 embryo hybridized with (A) antisense or (B) sense riboprobe.

family. One of the defining features is the similarity between the Cys₂His₂ zinc finger DNA-binding domain (9). The putative DNA contact sites are identical in each member of this family with variations occurring in the zinc finger in the regions that are outside of the DNA contact sites. It is therefore not surprising that members of this family bind a similar core element 5'-CNCCC-3' (5'-GGGNG-3') although preferences for variations of this sequence have been reported (9,14,15). Moreover, by utilizing PCR casting, the preferred binding site for GKLf has been determined to be 5'-G/AG/AGGC/TGC/T-3' (5'-CACCC-3') which is not only in agreement with the prediction based on the structural modeling of other Cys₂Hys₂ zinc fingers but also in agreement with the functional studies mentioned above (22–24).

IKLF is possibly the murine homolog of the human transcription factor BTEB-2 (15). The IKLF cDNA codes for a 446 amino acid protein while BTEB-2 codes for a 219 amino acid protein (15). Despite these differences in size, BTEB-2 is 92% homologous to IKLF. In addition, the nucleotide homology is also quite high and extends far beyond the coding region of BTEB-2. The difference in size of IKLF compared with BTEB-2 may be the result of alternative splicing. IKLF and BTEB-2 are expressed in the gastrointestinal tract, however, BTEB-2 expression is less restrictive, with expression occurring in multiple tissues. Despite the difference in size and expression pattern between IKLF and BTEB-2, IKLF appears to be a homolog of the human BTEB-2 because of the high homology of the activation domain.

BTEB-2 was named based on its ability to bind the basal transcription element (BTE) of the rat cytochrome P450 1A1 (CYP1A1) gene (17). The cDNA encoding the zinc finger domain of BTEB was used as hybridization probe in the isolation of BTEB-2 cDNA (24), hence the name assignment. However, upon further sequence analysis, it became evident that BTEB-2 bears more resemblance to the Krüppel-like family of transcription factor than it does to BTEB (9). IKLF appears to be a homolog of the human BTEB-2 gene although the translated product is larger and exhibits a different pattern of expression. The murine gene has been named intestinal-enriched Krüppel-like factor (IKLF) to denote the relatedness to members of the Krüppel-like family and restricted expression pattern.

Transcription factors are often composed of two separate domains, a DNA-binding domain and a transactivation domain. The activation domain of at least two members of the Krüppel-like family comprise subdomains that mediate protein–protein interactions (20), resulting in activation of transcription and repression of transcription (13,20). In both EKLF (20) and GKLf (13), the N-terminal half of the activation domain contains a transactivating

subdomain while the portion proximal to the zinc finger consists of a repressor subdomain. Despite the lack of similarity in the amino acid sequence, the subdomains of EKLF and GKLf are localized to the same region of the activation domain. In contrast, IKLF does not appear to contain a repressor domain and may function only as an activator. The activation domain of IKLF, like its other family members, does not contain amino acid sequences which represent a FAX (finger-associated box) (25) motif or KRAB (Krüppel-associated box) (26) motif. Such motifs are present in class I zinc finger transcription factors, which are expressed in various adult tissues, as well as during embryogenesis (21). IKLF belongs to class II zinc finger proteins that are expressed in a restricted number of tissues as well as during embryogenesis (21).

The Krüppel-like factors exhibit highly restricted patterns of expression in embryogenesis, (6,9) and their expression seems to be regulated during development. High expression levels of IKLF are present at day 7 of gestation, followed by a complete lack of expression at day 11 with a slight return of expression on days 15 and 17. These temporal changes in expression during embryogenesis suggest that this gene is also developmentally regulated. High expression of IKLF at day 7 of gestation correlates with the beginning of foregut formation, although one is only able to speculate on the role of this transcription factor during development. The Krüppel-like factors, like most transcriptional regulators of embryogenesis, share a DNA-binding domain that is highly homologous and bind similar regulatory elements (21), further suggesting a potential role of IKLF in development.

IKLF expression occurs predominantly in the gastrointestinal tract with minimal amounts of expression in other tissues. Closer examination by *in situ* hybridization reveals IKLF expression is limited to the epithelial layer of the gastrointestinal tract. No transcript of IKLF is detected in the muscular layer of any of the organs. Interestingly, in contrast to GKLf/EZF which is expressed in the middle to upper region of the crypt epithelium (5), IKLF expression occurs in the middle to lower region of the crypt epithelium. Thus, IKLF expression is abundant in the base of the crypt, which is composed of undifferentiated cells.

IKLF and GKLf are able to bind similar DNA regulatory elements, potentially targeting the same genes. However, they appear to be expressed at different stages of cellular differentiation, suggesting that each have a unique role during different states of differentiation. IKLF and GKLf/EZF may target the same gene with opposing effects on transcription. For example, during the undifferentiated state of the crypt epithelium cell, IKLF may activate transcription of specific target genes. As the cell develops and differentiates, IKLF expression is decreased while GKLf/EZF expression is increased. Once a change in threshold concentration is reached, GKLf/EZF displaces IKLF and attenuates transcription. Indeed, GKLf/EZF not only contains a repression subdomain but also functions as a repressor of transcription (13). Moreover, GKLf/EZF expression appears to correlate with the onset of differentiation and decreases during cellular proliferation (5,6), further supporting the idea of opposing roles during differentiation of gut epithelial cells.

To date only a limited number of gut-specific transactivating DNA-binding proteins have been identified. Intestinal-enriched Krüppel-like factor is a novel murine transcription factor that is expressed predominantly in the gastrointestinal tract and its expression is highly restricted during embryogenesis. Collectively, these results suggest IKLF has an important role during

development and tissue-specific gene expression. Future experiments will determine the function of IKLF in the context of intestinal epithelial differentiation and morphogenesis.

ACKNOWLEDGEMENTS

This research was supported in part by National Institutes of Health grant HL 57281 (J.B.L.) and NIH training grant 5-T32 HL07752 (M.D.C.).

REFERENCES

- 1 Potten, C.S. and Loeffler, M. (1990) *Development*, **110**, 1101–1020.
- 2 Cohn, S.M., Roth, K.A., Birkenmeier, E.H. and Gordon, J.I. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 1034–1038.
- 3 Hermiston, M.L. and Gordon, J.I. (1993) *Semin. Dev. Biol.*, **4**, 275–291.
- 4 Hermiston, M.L., Wong, M.H. and Gordon, J.I. (1996) *Genes Dev.*, **10**, 985–996.
- 5 Shields, J.M., Christy, R.J. and Yang, V.W. (1996) *J. Biol. Chem.*, **271**, 20009–20017.
- 6 Garrett-Sinha, L.A., Eberspaecher, H., Seldin, M.F. and de Crombrughe, B. (1996) *J. Biol. Chem.*, **271**, 31384–31390.
- 7 Gordon, J.I., Schmidt, G.H. and Roth, K.A. (1992) *FASEB J.*, **6**, 3039–3050.
- 8 Ton-That, H., Kaester, K.H., Shields, J.M., Mahatanakoon, C.S. and Yang, V.W. (1997) *FEBS Lett.*, **419**, 239–243.
- 9 Anderson, K.P., Kern, C.B., Crable, S.C. and Lingrel, J.B. (1995) *Mol. Cell. Biol.*, **15**, 5957–5965.
- 10 Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- 11 Wert, S.E., Glasser, S.W., Korfhagen, T.R. and Whitsett, J.A. (1993) *Dev. Biol.*, **156**, 426–443.
- 12 Lennon, G.G., Auffray, C., Polymeropoulos, M. and Soars, M.B. (1996) *Genomics*, **333**, 151–152.
- 13 Yet, S.-F., McA'Nulty, M.M., Folta, S.C., Yen, H.-W., Yoshizumi, M., Hsieh, C.-M., Layne, M.D., Chin, M.T., Wang, H., Perrella, M.A., Jain, M.K. and Lee, M.-E. (1998) *J. Biol. Chem.*, **273**, 1026–1031.
- 14 Crossley, M., Whitelaw, E., Perkins, A., Williams, G., Fujiwara, Y. and Orkin, S.H. (1996) *Mol. Cell. Biol.*, **16**, 1695–1705.
- 15 Sogawa, K., Imataka, H., Yamasaki, Y., Kusume, H., Abe, H. and Fujii-Kuriyama, Y. (1993) *Nucleic Acids Res.*, **21**, 1527–1532.
- 16 Miller, I.J. and Bieker, J.J. (1993) *Mol. Cell. Biol.*, **13**, 2776–2786.
- 17 Bieker, J.J. (1996) *DNA Cell Biol.*, **15**, 347–352.
- 18 Imataka, H., Sogawa, K., Yasumoto, K., Kikuchi, Sasano, K., Lobayashi, A., Hayami, M. and Fujii-Kuriyama, Y. (1992) *EMBO J.*, **11**, 3663–3671.
- 19 Williamson, M. (1994) *Biochem. J.*, **297**, 249–260.
- 20 Chen, X. and Bieker, J.J. (1996) *EMBO J.*, **15**, 5888–5896.
- 21 Pieler, T. and Bellefroid, E. (1994) *Mol. Biol. Rep.*, **20**, 1–8.
- 22 Shields, J.M. and Yang, V.W. (1998) *Nucleic Acids Res.*, **26**, 796–802.
- 23 Klevit, R.E. (1991) *Science*, **253**, 1367–1393.
- 24 Pavletich, N.P. and Pabo, C.O. (1991) *Science*, **252**, 809–817.
- 25 Knöchel, W., Pötting, A., Köster, M., El-Baradi, T., Nietfeld, W., Bouwmeester, T. and Pieler, T. (1989) *Proc. Natl Acad. Sci. USA*, **86**, 6097–6100.
- 26 Bellefroid, E.J., Poncelet, D.A., Lecocq, P.J., Revelant, O. and Martial, J.A. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 3608–3612.