Identification and Characterization of a Gene Encoding a Gut-enriched Krüppel-like Factor Expressed during Growth Arrest*  

(Received for publication, December 29, 1995, and in revised form, April 16, 1996)

Janiel M. Shields§, Robert J. Christy¶, and Vincent W. Yang**

From the Departments of Medicine and Biological Chemistry, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205 and the Center for Molecular Medicine, Institute of Biotechnology, University of Texas Health Science Center, San Antonio, Texas 78250

A cDNA clone, named gut-enriched Krüppel-like factor (GKLF), was isolated from an NIH 3T3 library using a probe encoding the zinc finger region of the immediate-early transcription factor zif/268. The deduced GKLF amino acid sequence contains three tandem zinc fingers that are related to members of the Krüppel family of transcription factors. By indirect immunofluorescence, GKLF is localized to the cell nucleus. In cultured fibroblasts, GKLF mRNA is found in high levels in growth-arrested cells and is nearly undetectable in cells that are in the exponential phase of proliferation. The growth-arresting nature of GKLF is demonstrated by an inhibition of DNA synthesis in cells transfected with a GKLF-expressing plasmid construct. In the mouse, GKLF mRNA is present in select tissues and is most abundant in the colon, followed by the testis, lung, and small intestine. In situ hybridization experiments indicate that GKLF mRNA is enriched in epithelial cells located in the middle to upper crypt region of the colonic mucosa. Taken together, these results suggest that GKLF is potentially a negative regulator of cell growth in tissues such as the gut mucosa, where cell proliferation is intimately coupled to growth arrest and differentiation.

Eukaryotic transcription factors are classified according to the structural motif that contacts the DNA. The zinc finger motif is one such example, in which a zinc atom is tetrahedrally coordinated by 4 amino acid residues (usually cysteine or histidine) within a 30-amino acid sequence to form the DNA-binding domain. The biological importance of this structure is reflected by estimates that the human genome has between 300 and 700 genes containing the zinc finger motif (Klug and 1996). They are characterized by multiple zinc fingers that are related to members of the Kru¨ppel family of transcription factors. By indirect immunofluorescence, GKLF is localized to the cell nucleus. In cultured fibroblasts, GKLF mRNA is found in high levels in growth-arrested cells and is nearly undetectable in cells that are in the exponential phase of proliferation. The growth-arresting nature of GKLF is demonstrated by an inhibition of DNA synthesis in cells transfected with a GKLF-expressing plasmid construct. In the mouse, GKLF mRNA is present in select tissues and is most abundant in the colon, followed by the testis, lung, and small intestine. In situ hybridization experiments indicate that GKLF mRNA is enriched in epithelial cells located in the middle to upper crypt region of the colonic mucosa. Taken together, these results suggest that GKLF is potentially a negative regulator of cell growth in tissues such as the gut mucosa, where cell proliferation is intimately coupled to growth arrest and differentiation.

*This work was supported in part by grants from the National Institutes of Health (to V. W. Y. and R. J. C.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Supported by a National Research Service Award from the National Institutes of Health.

¶ To whom correspondence should be addressed: Dept. of Medicine, Ross 918, The Johns Hopkins University School of Medicine, 720 Rutland Ave., Baltimore, MD 21205. Tel.: 410-955-9691; Fax: 410-955-9677; E-mail: vyang@welchlink.welch.jhu.edu.

** To whom correspondence should be addressed: Dept. of Medicine, Ross 918, The Johns Hopkins University School of Medicine, 720 Rutland Ave., Baltimore, MD 21205. Tel.: 410-955-9691; Fax: 410-955-9677; E-mail: vyang@welchlink.welch.jhu.edu.

1 The abbreviations used are: EKLF, erythroid Krüppel-like factor; GKLF, gut-enriched Krüppel-like factor; LKLF, lung Krüppel-like factor; bp, base pair(s); ORF, open reading frame; PBS, phosphate-buffered saline; BrdUrd, bromodeoxyuridine; DMEM, Dulbecco's modified Eagle's medium; FCS, fetal calf serum; BTEB2, basic transcription element-binding protein 2.
Gut-enriched Krüppel-like Factor

Haley et al., 1995. These findings suggest that depending on the signaling event, the mammalian cell cycle is regulated by both positive and negative control mechanisms.

In this study, we describe the identification and characterization of GKLF, a zinc finger-containing nuclear protein that is a member of the Krüppel family of transcription factors. Expression of GKLF in cultured fibroblasts is highest in growth-arrested cells and lowest in cells in the exponential phase of proliferation. Moreover, constitutive expression of GKLF in transfected cells results in the inhibition of DNA synthesis. In vivo, GKLF mRNA is very abundant in the crypt epithelium of the mouse colon. These findings suggest that GKLF may act as a negative regulator of proliferation of intestinal epithelial cells.

**EXPERIMENTAL PROCEDURES**

**Materials—** Restriction endonucleases and modifying enzymes were purchased from New England Biolabs Inc. (Beverly, MA). Sequences were purchased from U. S. Biochemical Corp. Radiolabels were purchased from DuPont NEN. Media and serum were purchased from Life Technologies, Inc. and Hyclone Laboratories (Logan, UT), respectively. Fluorescein isothiocyanate, Texas Red, and horseradish peroxidase-conjugated secondary antibodies were purchased from Amersham Corp. The prokaryotic expression vectors pET3d and pET16b (Studier et al., 1990) were purchased from Novagen (Madison, WI). The constitutive eukaryotic expression vector PMT3 (Swick et al., 1992), which yields with 1 the simian virus 40 enhancer and the adenovirus major late promoter largely in the form of inclusion bodies, was solubilized by the method of DEAE-dextran technique (Lopata et al., 1992). Overlapping restriction endonuclease fragments of the cDNA with sizes of 2800 bp in length were amplified, bidirectionally by the dideoxy chain termination method using the Sequenase kit (U. S. Biochemical Corp.). Sequence comparison was performed bidirectionally by the dideoxy chain termination method using the Sequenase kit (U. S. Biochemical Corp.). Sequence comparison was performed using the BLAST algorithm provided by the National Center for Biotechnology Information (Rockville, MD).

Production of GKLF in Bacteria and Generation of Anti-GKLF Serum—Complementary DNA containing the open reading frame (ORF) of GKLF was subcloned into pET3d and pET16b to generate pET3d-GKLF and pET16b-GKLF, respectively. The pET3d-GKLF construct was introduced into the BL21(DE3) strain of Escherichia coli cells, which were subsequently induced by the addition of 0.4 mM isopropyl-β-D-thiogalactopyranoside to produce GKLF. The protein, present largely in the form of inclusion bodies, was solubilized by the method of Feder et al. (1993) and separated by SDS-polyacrylamide gel electrophoresis. GKLF was electroeluted from the gel and used to raise a rabbit polyclonal antiserum by HRP Inc. (Denver, PA).

Cell Transfection—Transient transfections were performed by the DEAE-dextran technique (Lopata et al., 1984) or by lipofection (Felgner et al., 1987) using the Lipofectin reagent as recommended by the manufacturer (Life Technologies, Inc.). The ORF of the GKLF cDNA was subcloned into the constitutive mammalian expression vector PMT3 to generate PMT3-GKLF. Transfections of the monkey kidney-derived cell line COS-1 (American Type Culture Collection, Rockville, MD) were accomplished with 1 μg/mL PMT3-GKLF DNA in 3.5- or 10-cm culture dishes. Two days following transfection, cells were examined for GKLF production by Western blot or immunocytochemical analysis (see below).

**Library Screening and DNA Sequencing—** A 575-bp cDNA probe. To control for loading of RNA samples, all blots were stripped and reprobed with a radiolabeled DNA fragment encoding the 18S ribosomal RNA gene and, in some cases, with a cDNA fragment encoding the constitutively expressed glyceraldehyde-3-phosphate dehydrogenase (CLONTECH, Palo Alto, CA). When RNA from fractionated colonic tissue was used, the inner surface of an everted piece of freshly obtained mouse colon was gently scraped with a razor blade, and RNA was isolated as described above. This method yields a highly enriched population of cells of epithelial origin as noted before (Oliva et al., 1993).

**In Situ Hybridization—** Slices, and layered onto microscope slides. Slides were processed for in situ hybridization as described previously (Tijssen et al., 1994a, 1994b) and probed with either the antisense or sense control sense probe at high stringency. After washing, slides were submerged in Kodak NTB-2 liquid emulsion and stored in light-proof boxes. After development, slides were counterstained with hematoxylin and eosin and viewed under dark-field microscopy.

**20010** Gut-enriched Krüppel-like Factor

Western Blot Analysis—SDS-polyacrylamide gel electrophoresis was performed according to Laemmli (1970) with the following modifications. The acrylamide concentrations of the stacking and running gels were 5% and 10%, respectively. Protein samples were dissolved in loading buffer (60 mM Tris-HCl, pH 6.8, 2% SDS, 100 mM dithiothreitol, and 0.01% bromphenol blue) heated to 100°C for 3 min; and loaded onto the gel in electrophoresis buffer containing 25 mM Tris-HCl, pH 8.3, 250 mM glycine, and 0.1% SDS. At the completion of electrophoresis, proteins were transferred to nitrocellulose membranes according to the method of Towbin et al. (1979) and immunoblotted with rabbit anti-GKLF serum (1:1000 dilution) or preimmune serum. Following incubation with the secondary antibody conjugated to horseradish peroxidase (conjugated donkey anti-rabbit IgG), GKLF was visualized with enhanced chemiluminescence (Amersham Corp.).

Immunocytochemistry—Immunocytochemical studies of transiently transfected COS-1 cells grown on plastic coverslips were performed 2 days following transfection. Coverslips were washed with PBS and fixed in 4% paraformaldehyde (PFA) for 10 min. Cells were washed with 0.1% Nonidet P-40 in PBS for 10 min, washed with PBS, blocked with 10% fetal calf serum at 37°C for 15 min, and incubated with rabbit anti-GKLF serum (1:500 dilution) in PBS at room temperature for 1 h. After washing with PBS, the coverslips were incubated with fluorescein isothiocyanate-conjugated goat anti-rabbit IgG (1:200 dilution) at room temperature for 1 h, washed with PBS followed by water. Coverslips were then treated with 2% HCl at room temperature for 30 min to denature the DNA, neutralized with 0.1 M sodium borate at room temperature for 5 min, and washed with PBS. A mouse monoclonal antibody raised against BrdUrd (Sigma) for an additional 4 h. Coverslips bearing transfected cells were washed with PBS, fixed with 3% paraformaldehyde in PBS, permeabilized with 0.25% Triton X-100 in PBS for 10 min, and washed with PBS followed by water. Coverslips were then treated with 2% HCl at room temperature for 15 min, after which they were washed with PBS. Coverslips were subsequently incubated with a mixture of fluorescein isothiocyanate-conjugated goat anti-rabbit IgG (1:200) and Texas Red-conjugated sheep anti-mouse IgG (1:20) at 37°C for 1 h, mounted, and visualized with a fluorescence microscope as described above.

RNA Isolation and Northern Blot Analysis—RNA was isolated from NIH 3T3 cells and from various mouse tissues by the guanidinium thiocyanate method (Chirgwin et al., 1979). Twenty μg of total RNA were size-fractionated in 1.2% agarose gels containing 2.4 M formaldehyde (Ausubel et al., 1991) and transblotted onto nylon membranes (Hybond-N, Amersham Corp.). Hybridizations and washings were performed under high stringency conditions as described by Oliva et al. (1993) using a radiolabeled labeled GKLF cDNA probe. To control for loading of RNA samples, all blots were stripped and reprobed with a radiolabeled DNA fragment encoding the 18S r ribosomal RNA gene and, in some cases, with a cDNA fragment encoding the constitutively expressed glyceraldehyde-3-phosphate dehydrogenase gene (CLONTECH, Palo Alto, CA). When RNA from fractionated colonic tissue was used, the inner surface of an everted piece of freshly obtained mouse colon was gently scraped with a razor blade, and RNA was isolated as described above. This method yields a highly enriched population of cells of epithelial origin as noted before (Oliva et al., 1993).

In Situ Hybridization—Slices, and layered onto microscope slides. Slides were processed for in situ hybridization as described previously (Tijssen et al., 1994a, 1994b) and probed with either the antisense or sense control sense probe at high stringency. After washing, slides were submerged in Kodak NTB-2 liquid emulsion and stored in light-proof boxes. After development, slides were counterstained with hematoxylin and eosin and viewed under dark-field microscopy.
Gut-enriched Krüppel-like Factor

20011

RESULTS

Cloning of Mouse GKL F cDNA—To identify transcription factors that are involved in growth regulation, a cDNA library generated with mRNA from NIH 3T3 cells that were rendered quiescent and stimulated with serum for 3 h (Lanahan et al., 1993) was screened under reduced stringency conditions using a DNA probe containing the zinc finger region of the immediate-early transcription factor zif268 (Christy et al., 1992). A search in the GenBank™ nucleic acid database, two additional cDNA clones with significant homology were identified. These findings indicate that GKL F is a newly identified transcription factor.

Analysis of GKL F cDNA and Deduced Amino Acid Sequences—The GKL F cDNA contains a 311-bp 5'-untranslated region, a single ORF of 1449 bp, and a 977-bp 3'-untranslated region that is trailed by a poly(A) tail. The ORF potentially encodes a polypeptide of 483 amino acids with a predicted molecular mass of 53 kDa. Three potential methionine initiation codons are present in frame near the amino terminus (amino acids 1, 10, and 51), although nucleotide sequences surrounding the second methionine codon conform more closely to Kozak's rule for translation initiation (Kozak, 1987). A consensus sequence for polyadenylation, AATAAA (Proudfoot and Brownlee, 1976), is found 24 bp upstream of the poly(A) tail.

The deduced GKL F amino acid sequence contains three tandem zinc finger motifs at the carboxyl terminus. The three zinc fingers are closely related to the consensus sequence CX_{2,4}CT_{x}CT_{x}CT_{x} for zinc finger-containing transcription factors and are separated from each other by a 7-amino acid inter-finger spacer similar to the H/C link consensus sequence, TGEK P/Y/FX. These features classify GKL F as a member of the Krüppel family of proteins (Chowdhury et al., 1986; Morris et al., 1994; Schuh et al., 1986). GKL F is rich in proline and serine residues, which constitute 12.8 and 11.6% of the total amino acids, respectively. The clustering of these 2 amino acids is reminiscent of the transactivation domains of previously established transcription factors (Nakamura et al., 1993). In addition, a potential nuclear localization signal (Boulikas, 1992) that is rich in arginine and lysine residues is found between amino acid residues 384 and 390. A "PEST" sequence, which is found in proteins with intracellular half-lives of <2 h (Rogers et al., 1986), is present between amino acid residues 113 and 152. Finally, a 7-amino acid sequence, PPLPGRP, present between amino acid residues 55 and 61, is highly related to the consensus sequence to which proteins containing the SH3 domain bind (Alexandropoulos et al., 1995; Ren et al., 1993). Taken together, these features suggest that GKL F is a nuclear-localized transcription factor.

A search in the GenBank™ protein database for amino acid sequences related to that of GKL F revealed many sequences similar to its zinc finger-containing region. In particular, the closest alignment was found with three recently identified eu-karyotic transcription factors: 1) the lung Krüppel-like factor (LKL F) (Anderson et al., 1995); 2) EKL F (Miller and Bieker, 1993); and 3) basic transcription element-binding protein 2 (BTEB2) (Sogawa et al., 1993), expressed in the placenta and testis. The degree of sequence identity in the B2-amino acid zinc finger region of GKL F to these proteins is 92, 84, and 82%, respectively (Fig. 2). Other sequences that are related (but less conserved) to the zinc finger region of GKL F include WT1 (Call et al., 1990) and Sp1 (Kadonaga et al., 1987), showing 59 and 52% sequence identity, respectively. In contrast, the amino acid sequence outside the zinc finger region of GKL F bears no significant homology to any previously identified proteins with one notable exception: the 20 amino acids immediately preceding the first cysteine residue of the first zinc finger are 90% identical between GKL F and KLF (Anderson et al., 1995). These findings indicate that GKL F is a newly identified polypeptide.

When the nucleotide sequence of GKL F was compared with those stored in the GenBank™ nucleic acid data base, two additional cDNA clones with significant homology were identified. Both sequences correspond to the extreme 3'-end of the 3'untranslated region of the GKL F cDNA. The first (GenBank™ accession number D25944) was obtained during a ran-
Gut-enriched Krüppel-like Factor

**Fig. 2. Amino acid sequence alignment between GKLF, LKLF, EKLF, and BTEB2.** The sequences presented are those from the zinc finger regions of the four proteins. The species from which the sequences are derived are as follows: GKLF, mouse (this study); LKLF, mouse (Anderson et al., 1995); EKLF, mouse (Miller and Bieker, 1993); and BTEB2, human (Sogawa et al., 1995). Identical sequences are boxed. Numbers on the right are the amino acid positions of the GKLF sequence. Asterisks indicate those amino acid residues involved in the coordination of the zinc atom, and the two inter-finger spacer regions that are highly conserved in the Krüppel family of transcription factors are bracketed (>).

| Protein | Amino Acid Sequence |
|---------|---------------------|
| GKLF    | TCVDYGGCKTYYTKSSHLKAHLRHTGKPY |
| LKLF    | TCVDYGGCKTYYTKSSHLKAHLRHTGKPY |
| EKLF    | TCVDYGGCKTYYTKSSHLKAHLRHTGKPY |
| BTEB2   | TCVDYGGCKTYYTKSSHLKAHLRHTGKPY |

GKLF is 93% identical to the sequence of LKLF, 87% nucleotide identity to GKLF. The second cDNA, named clone 59 (GenBank™ accession number L26292), was obtained during a differential screening of a cDNA library made from rat Sertoli cells that had been stimulated with follicle-stimulating hormone (Hamill and Hall, 1994). This cDNA is 800 bp long and is 93% identical to the GKLF sequence. The high degree of sequence identity in the 3'-untranslated region between GKLF and these two partial cDNA clones suggests that the latter two potentially represent the human and rat homologues of GKLF, respectively.

Characterization of the GKLF Protein—To characterize the protein encoded by GKLF, a cDNA fragment containing the ORF was subcloned into two prokaryotic expression vectors, pET3d and pET16b (Studier et al., 1990), the latter creating a fusion protein, which contains an additional 3 kDa of bacterial sequence when expressed. The BL21(DE3) strain of host E. coli cells transformed with pET3d-GKLF and induced with isopropyl-β-d-thiogalactopyranoside to produce GKLF. The protein was resolved by SDS-polyacrylamide gel electrophoresis, purified, and used to generate a rabbit polyclonal antiserum. The specificity of the antiserum was confirmed by the ability of purified, bacterially produced GKLF to block the immunoreactivity of the 53-kDa protein band in PMT3-GKLF-transfected COS-1 cells (data not shown).

Expression of GKLF in Response to Serum Stimulation in Cultured Fibroblasts—GKLF was isolated from an NIH 3T3 cDNA library generated with RNA from cells that had been rendered quiescent and then stimulated for 3 h with serum. The serum responsiveness of GKLF in NIH 3T3 cells was therefore examined by Northern blot analysis. The mRNA content of GKLF from growth-arrested quiescent NIH 3T3 cells (which had been maintained in 0.5% FCS for 5 days) was compared with that from cells induced to enter the cell cycle by the addition of medium containing 15% FCS for various lengths of time between 0 and 24 h. In addition, RNA isolated from proliferating cells in the exponential phase of growth was analyzed.

Expression of GKLF in Response to Serum Stimulation in Cultured Fibroblasts—GKLF was isolated from an NIH 3T3 cDNA library generated with RNA from cells that had been rendered quiescent and then stimulated for 3 h with serum. The serum responsiveness of GKLF in NIH 3T3 cells was therefore examined by Northern blot analysis. The mRNA content of GKLF from growth-arrested quiescent NIH 3T3 cells (which had been maintained in 0.5% FCS for 5 days) was compared with that from cells induced to enter the cell cycle by the addition of medium containing 15% FCS for various lengths of time between 0 and 24 h. In addition, RNA isolated from proliferating cells in the exponential phase of growth was analyzed. As shown in Fig. 4, it is apparent that a significant level of GKLF mRNA was present in quiescent cells (lane 2), but was nearly absent in actively proliferating cells (lane 1). The addition of 15% FCS not only failed to increase the abundance of GKLF mRNA, but a decrease in the message content was detected beginning at 8 h after treatment. By 24 h after serum stimulation, a partial recovery of the GKLF mRNA level was observed. Fig. 5 summarizes the results of quantitative densitometric measurements of mRNA band intensities from Northern blot analyses of three independent experiments. The data confirmed that the level of GKLF transcript was much higher...
in growth-arrested cells compared with proliferating cells and that when cells were induced to enter the cell cycle, a reproducible decrease in the transcript level was observed beginning at a time period that precedes the commencement of DNA synthesis in NIH 3T3 cells (Quelle et al., 1993).

The growth arrest-specific nature of GKLF expression was confirmed by a reverse experiment to that of Fig. 4 in which NIH 3T3 cells in the exponential phase of proliferation were induced to enter quiescence by a reduction of the serum content in the medium from 10 to 0.5% for various lengths of time. Fig. 6 shows that the initial GKLF transcript was barely detectable, which only became apparent beginning 2 days after the cells were deprived of serum. A further increase in the GKLF mRNA level was seen between days 2 and 3, after which nearly equivalent levels of transcript were present up to 7 days. Similarly, when actively proliferating NIH 3T3 cells were left in 10% FCS without any additional feedings, levels of GKLF mRNA increased in a time-dependent manner (Fig. 7). This result was in clear contrast to the largely constant levels of the 18S ribosomal RNA and of the mRNA encoding glyceraldehyde-3-phosphate dehydrogenase (GAPDH).

Effect of Constitutive GKLF Expression on DNA Synthesis—The results of the preceding experiments suggest that expression of GKLF is temporally associated with growth arrest and that a down-regulation of GKLF occurs in cells that are stimulated to enter the cell cycle. To assess the effect of constitutive expression of GKLF on cell growth, COS-1 cells were transiently transfected with PMT3-GKLF or PMT3 for 24 h, at which time DNA synthesis was determined by [3H]thymidine incorporation. Table I shows the results of three independent experiments. In each experiment, the percentage of cells expressing GKLF in PMT3-GKLF-transfected cells was determined in parallel dishes by indirect immunofluorescence and found to vary between 25 and 30%. As shown, a statistically significant decrease in the incorporation of [3H]thymidine by

Fig. 4. Time course of GKLF expression in serum-starved NIH 3T3 cells stimulated with 15% FCS. NIH 3T3 fibroblasts were rendered quiescent by maintenance in DMEM containing 0.5% FCS for 5 days, at which time they were ~50% confluent (lane 2 or time 0). Cells were stimulated to enter the cell cycle by the addition of fresh DMEM containing 15% FCS for various periods of time between 1 and 24 h. Total RNA was isolated from cells at the times indicated, and 20 μg were examined by Northern blot analysis using a radioactively labeled GKLF cDNA probe. RNA from exponentially proliferating, nonsynchronized cells (lane 1 or P), maintained in DMEM supplemented with 10% FCS at ~30% confluency, was also examined for comparison. To control for RNA loading, the blot was subsequently stripped and reprobed with a DNA fragment encoding the 18S ribosomal RNA gene.

Fig. 5. Quantification of GKLF expression during serum stimulation of quiescent NIH 3T3 cells. Densitometric tracings of mRNA band intensities at different times following serum stimulation were performed, and the values were standardized to that of the mRNA level observed in quiescent (time 0) cells, taken as 100%. Shown are the mean values of three independent experiments. The vertical bars represent standard errors. P denotes proliferating cells.

Fig. 6. Time course of GKLF expression during serum starvation. Proliferating NIH 3T3 cells (lane 1 or P) were maintained in DMEM supplemented with 10% FCS for up to 6 days. The initial cell density was ~30%. By day 2, all the dishes were nearly 100% confluent. Northern blot analysis of 20 μg of total RNA was performed as described for Fig. 4. To control for RNA loading, the blot was stripped and sequentially reprobed with the DNA encoding the 18S ribosomal RNA gene, followed by the cDNA encoding glyceraldehyde-3-phosphate dehydrogenase (GAPDH).

Fig. 7. Effect of contact inhibition on GKLF expression. Proliferating NIH 3T3 cells (lane 1 or P) were maintained in DMEM containing 10% FCS without any additional feedings for up to 6 days. The initial cell density was ~30%. By day 2, all the dishes were nearly 100% confluent. Northern blot analysis of 20 μg of total RNA was performed as described for Fig. 4. To control for RNA loading, the blot was stripped and sequentially reprobed with the DNA encoding the 18S ribosomal RNA gene, followed by the cDNA encoding glyceraldehyde-3-phosphate dehydrogenase (GAPDH).
cells transfected with PMT3-GKLF as compared with cells transfected with vector alone was observed in all three experiments. The close correlation between the degree of inhibition of \[^{3}H\]thymidine incorporation into PMT3-GKLF-transfected cells and the percentage of cells expressing GKLF suggests that GKLF, when constitutively expressed, inhibits DNA synthesis.

To demonstrate directly that the cells expressing GKLF do not synthesize DNA, PMT3-GKLF-transfected COS-1 cells were first incubated with BrdUrd and then double-stained for GKLF and BrdUrd by indirect immunofluorescence. As shown in Fig. 8, two cells that expressed GKLF (arrowheads) failed to stain for BrdUrd. In contrast, four non-GKLF-expressing cells were positive for BrdUrd. After counting 100 consecutive GKLF-expressing cells, only five were noted to positively stain for BrdUrd, whereas 80% of the non-GKLF-expressing cells were BrdUrd-positive.

**Tissue Distribution of GKLF**—To determine the tissue distribution of GKLF, RNA was isolated from various adult mouse tissues and examined by Northern blot analysis. Of the tissues examined, the colon (both proximal and distal) contained the highest level of GKLF transcript (Fig. 9). Moderate levels of transcript were also noted in the distal small intestine (SI), testis, and lung. In addition, a small amount of GKLF transcript was present in the proximal small intestine. Smaller mRNA species were also noted in the intestinal tissues and testis and could represent other closely related sequences or products of alternate splicing. No appreciable amount of message was detected in the brain, kidney, liver, spleen, thymus, heart, muscle, and fat. These results indicate that the expression of GKLF is tissue-selective and is most abundant in the colon.

The mouse colon is composed of a heterogeneous population of cells, ranging from the epithelial cells lining the inner mucosal surface to various other non-epithelial cell types such as lymphocytes, fibroblasts, enteric neurons, and smooth muscle cells. To further localize the cellular origin of GKLF, Northern blot analysis was performed on RNA isolated from colonic mucosal scrapings (Fig. 10, lane 2), which represent an enrichment of the epithelial cell population as previously demonstrated (Oliva et al., 1993). As a comparison, Northern blot analysis was also performed on RNA isolated from the remaining colonic tissue after scraping (Fig. 10, lane 3). As shown, GKLF mRNA was highly enriched in the mucosal population of cells, suggesting that GKLF is mainly expressed in epithelial cells. The epithelium-specific expression of GKLF was confirmed by in situ hybridization. Fig. 11A shows that the antisense GKLF RNA probe hybridized primarily to the middle to

---

### Table I

| Construct | \[^{3}H\]Thymidine incorporation \(^a\) | \(n\) \(^b\) | \(p\) \(^c\) | Inhibition by GKLF \(\%\) |
|-----------|---------------------------------|-----|-----|------------------|
| Exp. 1    |                                 |     |     |                  |
| PMT3      | 588.3 ± 84.9                   | 6   | <0.01| 31.3             |
| PMT3-GKLF | 403.4 ± 66.4                   | 6   |       |                  |
| Exp. 2    |                                 |     |     |                  |
| PMT3      | 433.5 ± 90.3                   | 12  | <0.001| 26.3           |
| PMT3-GKLF | 319.7 ± 65.0                   | 12  |       |                  |
| Exp. 3    |                                 |     |     |                  |
| PMT3      | 319.3 ± 40.8                   | 12  | <0.0001| 28.2         |
| PMT3-GKLF | 229.0 ± 52.5                   | 12  |       |                  |

\(^a\) \[^{3}H\]thymidine incorporated was measured in triplicate over a 3-h period from one-tenth of the entire cell population in each dish. Data represent the mean ± S.D. for each construct.

\(^b\) Number of dishes assayed.

\(^c\) Statistical analysis was performed using the one-tailed t test.

---

**Fig. 8. Inhibition of DNA synthesis by constitutive expression of GKLF.** COS-1 cells were transiently transfected with PMT3-GKLF as described under "Experimental Procedures." Following transfection, cells were fed DMEM containing 0.5% FCS for 24 h, after which they were refed DMEM containing 10% FCS and 100 \(\mu\)M BrdUrd for an additional 24 h before being processed for immunocytochemical analysis. Cells were immunostained first with a mouse monoclonal BrdUrd antibody, followed by rabbit anti-GKLF serum. The secondary antibodies used were fluorescein isothiocyanate-conjugated goat anti-rabbit IgG (left panel) and Texas Red-conjugated sheep anti-mouse IgG (right panel). The arrowheads point to two GKLF-positive and BrdUrd-negative cells, whereas four other cells in the same field were GKLF-negative and BrdUrd-positive.

**Fig. 9. Tissue distribution of GKLF mRNA.** RNA was extracted from various mouse tissues, and 20 \(\mu\)g were analyzed by Northern blot hybridization as described for Fig. 4. SI, small intestine; prox., proximal.

**Fig. 10. Distribution of GKLF mRNA in a fractionated mouse colon.** The inner surface of a mouse colon was scraped with a razor blade to give rise to the mucosal fraction of cells from which RNA was extracted (lane 2). As a comparison, RNA was also extracted from the remaining tissue (lane 3). Lane 1 represents RNA obtained from an intact colon. Twenty \(\mu\)g of total RNA were analyzed by Northern blot hybridization.
upper region of the colonic crypt epithelium. In contrast, the control sense GKLF RNA probe produced a random background distribution of silver grains (Fig. 11B). These results indicate that GKLF is epithelium-specific and is expressed primarily in cells that are in the process of migrating from the base toward the top of the crypt.

**DISCUSSION**

The cDNA encoding GKLF was isolated by low stringency hybridization to a zinc finger-containing transcription factor, zif/268. Although not proven in this study, several features of GKLF strongly suggest that it is a transcription factor. First, the amino acid sequence at the carboxyl terminus containing the zinc fingers is similar to those of a number of proteins, many of which are proven transcription factors, such as LKLF, EKLF, and BTEB2 (Fig. 2). Second, GKLF contains a potential nuclear localization signal between amino acid residues 384 and 390 and is in fact localized to the cell nucleus in transfected cells (Fig. 3). Third, like many transcription factors with short half-lives (Chevaillier, 1993), GKLF contains a PEST sequence with a high PEST score of 5.8 (Rogers et al., 1990). Finally, GKLF contains abundant proline and serine residues, amino acid residues purportedly involved in the transactivation function of many transcription factors (Nakamura et al., 1993).

Furthermore, its proline-rich nature is shared by proteins closely related to GKLF, namely LKLF (Anderson et al., 1995), EKLF (Miller and Bieker, 1993), and BTEB2 (Sogawa et al., 1993). Based on the high degree of homology in the Krüppel region of the proteins and their proline-rich characteristics, the latter three transcription factors have recently been assigned to a new multigene family (Anderson et al., 1995). GKLF is thus the newest addition to this protein family.

Aside from the proline-rich domain of GKLF, amino acid sequence outside its zinc finger region further defines its phylogenetic origin. For example, absent from the GKLF sequence are the FAX (finger-associated box) (Knoechel et al., 1989) and KRAB (Krüppel-associated box) (Bellefroid et al., 1991) motifs, the latter estimated to be present in one-third of zinc finger-containing transcription factors. Genes encoding the FAX- and KRAB-containing zinc finger proteins (called class 1 zinc finger proteins by Pieler and Bellefroid (1994)) are generally organized in large clusters on certain chromosomes and are thought to have appeared late during evolution. Most of these genes are widely expressed in adult tissues and in most stages of embryogenesis. The Krüppel family of transcription factors (called class 2 zinc finger proteins by Pieler and Bellefroid (1994)), in contrast, includes fewer members that are highly conserved and that function in the context of cell differentiation and embryogenesis. The majority of class 2 zinc finger proteins exhibit highly restricted patterns of expression in tissues and during embryogenesis. The tissue-selective nature of GKLF and its close relationship with other Krüppel-like factors such as LKLF and EKLF indicate that it belongs to the class 2 zinc finger protein gene family according to the classification of Pieler and Bellefroid (1994).

Structural analysis of zinc finger-containing proteins indicates that each finger consists of a β-pleated sheet at the amino-terminal half and an α-helix at the carboxyl-terminal half and that the latter makes direct contact with DNA (Berg, 1990). A number of studies examined the relationship between the amino acid sequence in the helical portion of a finger and the DNA sequence to which the finger binds (Berg, 1992; Klug and Schwabe, 1995; Nardelli et al., 1991). These studies revealed a number of important conclusions that enable one to predict the DNA sequence to which a zinc finger protein may bind. Table II compares the amino acid sequences in the α-helical portions of each of the three zinc fingers of GKLF to those of several other Krüppel-like transcription factors along with the DNA sequences with which the fingers interact. It is ap-
parent that amino acid sequences in the helical region of each of the three zinc fingers of GKLF, BTEB2, and EKLF are identical. Moreover, finger 2 of GKLF is identical to finger 1 of zif/268 and nearly identical to finger 2 of Sp1 as well as finger 3 of zif/268. With the exception of EKLF, these amino acid sequences recognize the consensus trinucleotide GCG. In addition, finger 1 of GKLF is highly similar to finger 1 of Sp1, which binds the trinucleotide GGG. Finally, finger 3 of GKLF is very similar to finger 3 of Sp1 and finger 2 of zif/268, both of which bind to the sequence GGG. Based on these comparisons, one can predict a sequence of 5'-GGG GCG GGG-3' to which GKLF is potentially capable of binding. This sequence is identical to the binding sequences of BTEB2 and Sp1. Of note is that although EKLF recognizes a different DNA sequence despite an overall identity in the zinc finger sequences, it is capable of interacting with an Sp1-binding site (Hartzog and Myers, 1993).

The expression of GKLF in cultured fibroblasts is of interest. The cDNA library from which GKLF was initially derived was made with RNA from quiescent NIH 3T3 cells that had been stimulated with serum for 3 h. Because the number of positive clones obtained during a repeat screening of the same library with a partial GKL cDNA fragment was quite high (>50 positive plaques out of a total of 1 million), we initially thought that GKLF, like zif/268, would behave like an immediately-early gene. We were therefore surprised to find from the Northern blot experiment shown in Fig. 4 that GKLF behaved quite differently from zif/268 or other immediate-early genes. The differences include the following. 1) The steady-state GKLF transcript level was high in serum-starved quiescent NIH 3T3 cells. 2) The transcript level did not rise appreciably during the first few hours of serum induction. 3) The transcript level began to fall at 8 h of serum treatment. In addition, the GKLF transcript was nearly undetectable in RNA harvested from exponentially proliferating cells. Combining the results of serum stimulation (Figs. 4 and 5), serum starvation (Fig. 6), and contact inhibition (Fig. 7) experiments, it becomes apparent that expression of GKLF is associated with cessation of cell growth. It is especially of interest to note that in serum-stimulated cells, the GKLK transcript level first decreases at a time that immediately precedes the S phase of the cell cycle (Figs. 4 and 5) (Quelle et al., 1993). This observation suggests that GKLF may exhibit a negative effect on cell cycle progression, particularly at the G2/S transition phase. The diminished [3H]thymidine incorporation by cells in which GKLF was constitutively expressed (Table I) and the lack of BrdUrd uptake by cells expressing GKLF (Fig. 8) support this hypothesis.

The expression of GKLF in response to serum stimulation and deprivation in cultured fibroblasts is reminiscent of that of a group of genes exclusively expressed in the growth arrest state. These genes are divided into two categories: the gas (growth arrest-specific) genes (Ciccarelli et al., 1990; Gorski et al., 1993) and the gadd (growth arrest and DNA damage-inducible) genes (Fornace et al., 1989; Zhan et al., 1994). Like GKLF, expression of the gas genes is highest in quiescent cells and is down-regulated following mitogen stimulation. In particular, the time course of expression of GKLF following mitogen addition to quiescent cells (Figs. 4 and 5), serum deprivation (Fig. 6), and cell contact (Fig. 7) is remarkably similar to that for gas1 (Ciccarelli et al., 1990; Schneider et al., 1988) and gadd6 (Ciccarelli et al., 1990). Similar to the inhibitory effect of GKLF on DNA synthesis (Table I), when ectopically expressed, some of the gas genes cause growth arrest and are thought to be involved in a negative circuit that governs growth suppression (Del Sal et al., 1992). Nevertheless, despite an overall similarity in the pattern of expression of the gas gene family, they encode a diverse group of protein products. For example, gas1 and gas3 encode integral membrane proteins (Del Sal et al., 1992; Maniioletti et al., 1990); gas2 encodes a protein of the microfilament network system (Brancolini et al., 1992); and gas6 encodes a secreted, vitamin K-dependent protein that is a ligand for the Axl receptor tyrosine kinase (Varnum et al., 1995). The gas gene (growth arrest-specific homeobox) is a notable exception of the gas gene family in that it encodes a homeobox-containing transcription factor that is highly specific for vascular smooth muscle cells (Gorski et al., 1993). Finally, gadd153, the only gadd gene with an identified function, is a member of the C/EBP family of transcription factors and is the human homologue of the murine CHOP-10 gene (Ron and Habener, 1992).

In vivo, the expression of GKLF is highly tissue-selective and is enriched in regions of the intestinal tract, testis, and lung (Fig. 9). It is of interest to note that a partial cDNA fragment encoding the putative rat homologue of GKLF was identified in Sertoli cells that had been treated with follicle-stimulating hormone (clone 59) (Hamil and Hall, 1994). Other than its induction by follicle-stimulating hormone, little information regarding the expression of clone 59 is available. In situ hybridization experiments should help clarify the cellular origin of GKLF in the testis. It is clear from Figs. 10 and 11, however, that expression of GKLF in the colon is highly enriched in the crypt epithelium. This finding is supported by the cloning of a putative human homologue of GKLF from the colonic mucosa (see "Results"). Moreover, results of in situ hybridization indicate that the GKLF transcript is localized to a population of epithelial cells residing in the middle to upper crypt region. This portion of the colonic crypt is thought to consist of cells that have undergone growth arrest and that have begun to differentiate into mature colonocytes as they emerge from the proliferating compartment at the base of the crypt (Gordon et al., 1992). Thus, GKLF may play two potential physiological roles in this environment. It can act either as a growth-suppressing gene product that is involved in the growth arrest of epithelial cells as they exit the base of the crypt or as a differentiation-promoting gene product that is responsible for activating downstream genes that are required for the differentiated epithelial phenotype. These two functions are not mutually exclusive such that GKLF can potentially serve both, in a manner similar to MyoD, which promotes both myogenic differentiation and terminal withdrawal from the cell cycle (Haley et al., 1995). Clearly, the exact function of GKLF in the cell cycle and/or in terminal differentiation awaits further examination.

Acknowledgments—We thank Dr. Lanahan for providing the NIH 3T3 cDNA library and the Genetics Institute for the PMT3 plasmid. We also thank Dr. Corey Mjaatvedt for assisting in the in situ hybridization.

REFERENCES
Alexanderopoulos, K., Cheng, G., and Baltimore, D. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 3110–3114
Anderson, K. P., Kern, C. B., Crable, S. C., and Lingrel, J. B. (1995) Mol. Cell. Biol. 15, 5957–5965
Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1992) Current Protocols in Molecular Biology, 2nd Ed., John Wiley & Sons, Inc., New York
Beliefreet, E. J. P., Poncelet, D. A., Lecocq, P. J., Revelent, O., and Martial, J. A. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 3608–3612
Berg, J. M. (1990) Annu. Rev. Biophys. Biophys. Chem. 19, 405–421
Berg, J. M. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 11109–11110
Bouilloux, T. (1993) Crit. Rev. Eukaryotic Gene Expression 3, 193–227
Brancolini, C., Bottega, S., and Schneider, C. (1992) Cell Biol. 127, 1251–1261
Call, K. M., Glaser, T., Ito, C. Y., Buckler, A. J., Pelletier, J., Haber, D. A., Rose, E. A., Kral, A., Yeger, H., Lewis, W. H., Jones, C., and Housman, D. E. (1990) Cell 60, 509–520
Cao, X., Mahendran, R., Guy, G. R., and Tan, Y. H. (1993) J. Biol. Chem. 268, 16949–16957
Chevallier, P. (1993) Int. J. Biochem. 25, 479–482
