The tissue-dependent Keratin 19 gene transcription is regulated by GKLF/KLF4 and Sp1

Felix H. Brembeck and Anil K. Rustgi

Gastroenterology Division, Cancer Center, Department of Genetics, University of Pennsylvania

Running title: Tissue specific K19 gene regulation by GKLF/KLF4 and Sp1

*Corresponding author:

Anil K. Rustgi, M.D.
Gastroenterology Division and Cancer Center
600A CRB
University of Pennsylvania
415 Curie Blvd.
Philadelphia, PA 19014
Phone 215-898-0154
Fax 215-573-5412
E-mail anil2@mail.med.upenn.edu
Summary

Keratins play critical roles in cellular differentiation and cytoskeletal organization. Keratin 19 (K19) is unique because it has been implicated as a marker of stem cells in some tissues such as the hair follicle in the skin. It is also associated with malignant transformation in esophageal and pancreatic cancers. Here we show that the K19 promoter is active in a subset of gastrointestinal cancer cells derived from esophageal and pancreas, but inactive in other contexts. This activity was mapped to a short region containing an overlapping binding site for gut-enriched Krüppel-like factor (GKLF/KLF4) and Sp1. GKLF has a higher binding affinity, and is the predominant binding factor in cells with low Sp-1 protein levels. Pancreatic acinar cells normally do not express K19, but overexpression of GKLF and Sp1 in these cells leads to aberrant expression similar to what is observed in pancreatic cancer. These results demonstrate the functional interaction of ubiquitous and tissue-restricted transcription factors in determining tissue-and neoplasm-specific patterns of gene expression.
Introduction

Cytokeratins belong to the family of intermediate filaments (IFs) and are associated with cellular differentiation and cytoskeleton organization. Keratin proteins can be further subdivided in two different types, the acidic type I (K9-20) and basic type II (K1-8) keratins (1-3). There is at least one pair of acidic and basic keratins expressed as heterodimers in any given epithelial tissue throughout development and differentiation. For example, in stratified epithelia, basal cells are characterized by the expression of K5 with K14/K15. Upon differentiation, stratified keratinizing epithelia express K1 and K10, whereas non-keratinizing stratified epithelia (e.g. the human esophagus) express K4 and K13. Simple epithelia express K8 and K18 (4). Regulation of keratin expression in the differentiating stratified squamous epithelium is orchestrated by a complex interplay of both ubiquitous and tissue-specific transcription factors (5,6). Analysis of the molecular mechanisms underlying provides insights into the normal program of differentiation as well dedifferentiation during malignant transformation. Moreover, possible progenitor stem cells can be potentially identified or defined by studying keratin gene expression (7) as exemplified by the hair follicle.

The most striking exception to the keratin-pair rule is the smallest known acidic keratin, namely keratin 19 (K19). K19 has no known basic type II keratin partner, although its expression is often found in cells which express K8 (2,8). K19 has the highly conserved alpha-helical central domain which is essential for filament formation. However, this keratin lacks the C-terminal nonhelical “tail domain” present in other acidic keratins. K19 is fully competent in filament formation with a basic keratin partner in vitro, although these complexes are unstable (9). Indeed, in some tissues, K19 can substitute for K18, the usual partner of K8 (10). In K18 knockout mice, K19 compensates for the loss of K18 expression (11). Stasiak et al. (1989) proposed that K19 might act as
“neutral” keratin in terms of differentiation. By dimerization with any early-synthesized basic type II keratin and with resulting delayed expression of the usual type I keratin partner, a cell could be held in a “flexible state of differentiation” (12).

The human K19 gene was sequenced by Bader et al. (1986) (13) and Stasiak and Lane (1987) (14) and the gene localizes to chromosome 17q21-q22 (15,16). The murine K19 gene shows high homology to its human counterpart (17) and the gene is located in the acidic keratin cluster on mouse chromosome 11 (17,18). Precursor cells in different tissues display high K19 levels and upon differentiation, K19 expression is epithelial cell-specific. For example, in the developing pancreas, duct-like precursor cells harbor high K19 expression. After these precursor cells transdifferentiate into endo- and exocrine compartments, K19 expression resides predominantly in ductal epithelial cells, whereas islet cells show markedly decreased K19 expression and acinar cells are K19 negative (19). Likewise, the fetal liver shows high K19 expression. Upon differentiation, hepatocytes lose expression of K19, whereas its expression is retained in bile ducts (20). Sites of K19 expression in the human adult include the esophagus, stomach, pancreas, small intestine and colon. In the human skin, K19 expression is limited to a small portion of cells in the outer root sheath of the hair follicle, but no expression is evident in the interfollicular suprabasal epidermis (12). These K19 positive cells in the hair follicle demonstrate features of slow-cycling cells and may represent stem cells (21). Many premalignant and malignant tissues display K19 expression, such as dysplasia and carcinoma of squamous epithelia, adenocarcinoma of the lung, breast, pancreas, stomach and colon (22).

Little is known about the regulation of K19 expression. We hypothesized that the transcriptional regulation of K19 is important in understanding differentiation programs in different cell types, namely squamous epithelial and pancreatic ductal cells, and that differences may arise from the manner in which K19 is regulated by different transcriptional factors. In this study we describe the tissue specific activity and regulation
of a 5' 2.0 kB fragment flanking the K19 gene. We have identified a critical element that is an important determinant of basal transcription of this gene and is involved in its tissue specific expression. The interplay between gut-enriched Krüppel-like factor (GKLF/KLF4) and Sp1 modulates K19 promoter activity through this cis-regulatory element and contributes to tissue specificity of K19 expression. In particular, GKLF regulates K19 expression in pancreatic ductal, but not acinar, cells, whereas GKLF and Sp1 cooperate to modulate K19 expression in esophageal squamous epithelial cells.
Experimental procedures

Cell culture and transient transfections studies

T.T, TE-12, Hs766T, Panc-1, Capan-2, ARIP, HeLa and NIH 3T3 fibroblasts (ATCC) were grown in 5% CO₂ at 37°C as subconfluent monolayers in Dulbecco’s Modified Eagle’s Medium (DMEM) supplemented with 10% v/v fetal calf serum (FCS), 100 U/ml penicillin, 100 mg/ml streptomycin and 1% L-glutamine (all reagents obtained from Sigma). AR42J cells (ATCC) were grown in Ham’s F12K medium (Gibco) supplemented with 20% FCS. Experiments were carried out in the log phase of growth after the cells were seeded for 24 hours. For luciferase reporter gene assays, transient transfection of cells was performed using the calcium phosphate precipitation technique (5′-3′, Inc.). Cells were plated at a density of 1 x 10⁶ cells/35-mm well and transfected 24 h later. Per 35 mm well, 0.5 -2 µg of the respective K19 promoter-luciferase reporter constructs were transfected. The transfectant mixture consisted of a 250 µl solution of 125 mM CaCl₂, 25 mM Hepes (pH 7.05), 0.75 mM Na₂HPO₄, 5 mM KCl, 140 mM NaCl and 6 mM glucose. After 12 hrs. incubation, cells were washed twice with phosphate-buffered saline (PBS) followed by exchange with fresh DMEM supplemented as above. AR42J and ARIP cells were transfected using the liposomal DNA transfer technique. After an attachment period of 12 hours, cells were incubated with 800 µl of the transfection mixture containing 0.5-2.0 µg plasmid DNA and 20 µl Lipofectamine reagents in Opti-MEM medium (Life Technologies). Cells were incubated for 5 hrs., followed by exchange with fresh media supplemented with FCS. For cotransfection studies, 1 µg of either Sp1 or GKLF expression vector were used. 48 hrs after transfection start, cells were washed twice with PBS and lysed in 250 µl of 1x cell culture lysis reagent (Promega). Luciferase activity of 100 µl cell lysate was determined using 100 µl luciferase reagent (Promega) by measurement of relative luciferase units (RLU) in a luminometer (MLX Microtiter Plate Luminometer, Dynex Tech.). All controls were performed using the same amount of
DNA of the respective empty vector. Transfection efficiency was routinely monitored by transfection with 1 µg of RSV-Luc in each set of experiments. Luciferase activity was determined and results of luciferase activity for each subset of experiment were corrected accordingly. Transfections were performed in triplicate, and results of at least three independent experiments were calculated as the mean ± SEM values for luciferase activity.

**DNA sequencing**

A 2.0 kB BamHI fragment containing the 5’ regulatory region of the mouse K19 gene (17) was subcloned into pcDNA 3.1 (Invitrogen) for sequencing. Two different clones were independently sequenced in sense and antisense direction using vector specific primers. In addition, four additional internal primers were designed to obtain the complete sequence (Table 1). Results were aligned and compared for both clones and the complete −1970 to +46 K19 DNA sequence established. Sequencing was performed by the dye terminator cycle sequencing method (ABI) in the automated DNA sequencing facility of the University of Pennsylvania.

**DNA plasmid constructs**

*K19 promoter-luciferase reporter and serial deletion constructs*

The −1970+46 K19 luciferase reporter vector, designated K19-1970, was constructed by PCR amplification of the DNA using the K19-pcDNA3.1 plasmid as template. For directional subcloning of the PCR product into the luciferase reporter vector pGL3 basic (Promega), KpnI and SacI restriction enzyme sites were generated using the K19−1970 5’-primer and K19+46 3’-primer, respectively (Table 1). Serial deletions, designated K19−960, K19−654, K19−288, K19−181, K19−102, and K19−58 were accordingly generated by combining the respective 5’-primer (Table 1) and the K19+46 3’-primer. The PCR reaction was performed using 20 ng plasmid DNA, 200 µM
of each dNTP, 0.5 µM of each primer and 2.5 U Pfu Turbo DNA polymerase (Stratagene) in a 1x-reaction buffer provided by the manufacturer. After an initial 60 sec. denaturation step at 94°C, 30 cycles of amplification were done as follows: denaturation 60 sec/94°C, annealing 60 sec/60°C, extension 2 min/72°C and a final extension at 72°C for 10 minutes. For all constructs, a single PCR product of the expected size was generated. The PCR products and pGL3 basic vector were digested with KpnI and SacI restriction enzymes (New England Biolabs), agarose gel-purified (Qiagen) and ligated with T4 DNA Ligase (New England Biolabs). All constructs were subjected to confirmatory DNA sequencing. The plasmids were purified by a modified alkaline lysis method (Qiagen).

**Mutant K19 promoter constructs**

Mutations of the Sp1 sites at –198, –62 and –33 and GKLF site mutants (MT1 and MT2) at position –58 in the full-length K19 were generated by site-directed mutagenesis using overlap extension PCR (23). Complementary oligonucleotide primers (Table 1) containing the mutated nucleotide sequence were designed to generate two PCR fragments having overlapping ends in a primary amplification reaction. In this first step, wild-type K19–1970+46 cDNA was used as template. The first PCR amplification step of each mutant required the outer K19–1970 5’ primer and the respective internal mutant 3’ primer to generate one PCR fragment. The second PCR fragment was generated using the respective internal mutant 5’ primer and the outer K19+46 3’ primer. The PCR products for each mutant in this first amplification step were gel purified (Qiagen) and combined in a subsequent "fusion" PCR reaction. The overlapping ends of these fragments anneal and serve as template and as primer for the second amplification reaction. Using the outer K19–1970 5’ primer and the K19+46 3’ primer, the different mutated full-length K19 promoter fragments were generated. PCR reaction components and amplification conditions for first and second steps were the same as
described above. The mutated full-length K19 promoter fragments were then subcloned into the pGL3 basic vector as outlined above. All constructs were subjected to confirmatory sequencing.

**GKLF and Sp1 expression vectors**

Construction of the expression vectors containing human GKLF and Sp1 cDNAs was described previously. Briefly, the cDNA of human GKLF was subcloned into pcDNA 3.1 HIS (Invitrogen) (24). A 4.0 kB *Xba*I fragment containing the cDNA of human Sp1 was obtained from the pEVR2-Sp1 vector (25) and subcloned into pRc/CMV (Invitrogen).

**Western Blot analysis**

K19, Sp1 and GKLF protein expression was determined using lysates from subconfluent cell cultures. For the preparation of total cell lysates, cells were washed twice with cold PBS, detached from culture dishes with a plastic cell scraper and resuspended in a lysis buffer containing 50 mM HEPES (N-[2-Hydroxyethyl]Piperazine-N'-[2-ethane sulfonic acid] pH 7.0, 250 mM sodium chloride, 5 mM EDTA and 0.1 % NP-40 (Gibco BRL) supplemented with 1 mM phenylmethylsulfonyl fluoride, 1 mM dithiothreitol and 0.1 % aprotinin (Sigma). 10 µg of total protein from each sample was separated on a 10-15% SDS-polyacrylamide gel. Following electrophoresis, the protein was transferred to an Immobilon membrane (Millipore Corp.). After transfer, membranes were stained with 0.1 % Ponceau-S Solution (Sigma) to confirm equal loading of proteins. Membranes were blocked with 5% milk, 10 mM Tris-HCl, pH 7.4, 150 mM NaCl, 0.2% Tween 20 for 1 h. The mouse monoclonal anti-Keratin 19 antibody (Sigma) and the rabbit GKLF anti-serum (provided by V. Yang) were used at a 1:2000 dilution in 10 mM Tris-HCl, pH 7.4, 150 mM NaCl, 0.2% Tween 20, 1% bovine serum albumine overnight at 4°C. A polyclonal rabbit anti-Sp1 antibody (SantaCruz) was used in a
1:1000 dilution in the same solution. Secondary antibody, either horseradish peroxidase-conjugated sheep anti-mouse or goat anti-rabbit antibody respectively (Amersham Pharmacia Biotech), were used at a 1:4000 dilution for 1 hour at room temperature. Horseradish peroxidase activity was detected with a chemiluminescence system (ECL system, Amersham Pharmacia Biotech).

**Electrophoretic Mobility Shift Assays (EMSAs)**

Nuclear extracts from different cell lines were prepared as described by Schreiber et al. (26) except the buffers were supplemented with a mixture of 0.5 µg/ml protease inhibitors (aprotinin, chymostatin, pepstatin) (Boehringer Mannheim). 5 µg nuclear protein extracts were used as determined by a colorimetric protein method (Bio-Rad protein assay). Oligonucleotide DNA probes were synthesized by the phosphoramidite procedure (Integrated DNA Technologies), purified by gel electrophoresis and annealed. 5 pmol of double-stranded oligonucleotides were radioactive labeled with $[^{32}\text{P}]$. Radiolabeling was done by a Klenow fill-in reaction in a buffer consisting of 10 mM Tris-HCl, pH 7.5, 5 mM MgCl$_2$, 7.5 mM dithiothreitol, 33 µM dATP, 33 µM dGTP, 33 µM dTTP, 0.33 µM $[^{32}\text{P}]$dCTP (NEN Life Science Products), 1 unit of DNA polymerase I Klenow fragment (Amersham Pharmacia Biotech). EMSAs were carried out by incubating 5 µg of nuclear extract with 5 fmol of the $[^{32}\text{P}]$-labeled oligonucleotide DNA probe (20,000 cpm) in a 20-µl binding reaction containing 10 mM Tris-HCl, 50 mM NaCl, 1 mM dithiothreitol, 1 mM EDTA, 10% glycerol, and 1.0 µg of poly(dI-dC) (Amersham Pharmacia Biotech). After incubation at room temperature for 15 min, the samples were loaded on a 6% polyacrylamide, 0.25x Tris borate gel and electrophoresed at 10 V/cm for 2 h. The gels were dried and exposed to x-ray film (Kodak) at 80 °C for 1 to 12 hrs. For competition experiments, the nuclear extract was co-incubated with 100-fold excess of unlabeled double-stranded oligonucleotides. Sequences of wild-type and mutated oligonucleotides are summarized in Table 2. Immune supershift assays were performed.
using the Sp1 and GKLF antibodies. The antibody was preincubated with the nuclear extract at room temperature for 15 min. prior to the addition of the $[^{32}\text{P}]$-labeled oligonucleotide DNA probe. Generation and purification of recombinant GKLF protein was described previously (24). Recombinant Sp1 protein was obtained from Promega. Other conditions for EMSAs are described above.
Results

The 5' regulatory region of the mouse K19 gene contains several putative cis-regulatory elements

A 2.0 kB fragment of the 5' regulatory region of the mouse K19 gene (Lussier, 1990) was completely sequenced. Using overlapping internal primers, the sequence was confirmed for two independent clones containing this fragment (Genbank submission number AF237661). The proximal 350 base pairs correspond to the published sequence containing the ATA box and putative transcription start site (Genbank accession number M316120) (17). Analysis of the promoter region (27) reveals several potential cis-regulatory elements, including several AP-1 and ets factor binding sites, previously shown to regulate expression for other keratins. Of particular interest, there are three canonical Sp1 sites at −198, −62 and −33 bp, two of which (−198, −62) are also conserved in the human K19 gene (15,17).

The 5' regulatory region of the K19 gene is variably active in different cell types

The full-length K19 (~1970+46) fragment was subcloned into the pGL3 luciferase reporter vector. After transient transfection of a panel of different cell lines, K19 promoter activity was determined and a dose-response curve was established for each cell line (Figure 1A). K19 promoter activity also correlated with expression of the endogenous K19 protein (Figure 1B).

As a model for simple epithelia we focused on a panel of pancreatic cancer lines of different origin with a well characterized keratin expression profile (28,29). Cell lines with a ductal origin (Panc-1, Capan-2) demonstrated the highest K19 promoter activity, consistent with the observation of high K19 protein expression in these (Figures 1A, 1B). Hs766T, representing an undifferentiated pancreatic cancer cell line, demonstrated low, but dose-dependent K19 promoter activity. In contrast, pancreatic cancers cells of acinar
origin (AR42J, ARIP) had little, if any, reporter gene activity which corresponded to the absence of endogenous K19 protein expression (Figures 1A, 1B). These results in aggregate support the notion that K19 plays a more important role in the regulation of pancreatic ductal epithelial cells than pancreatic acinar cells.

As a model for the stratified squamous epithelium, we examined two esophageal cancer cell lines, T.T and TE-12. Both cell lines displayed a dose-dependent K19 promoter reporter gene activity (Figure 1A) with evidence of endogenous K19 protein (Figure 1B). This result is consistent with the expression of K19 in normal and malignant human esophageal cells. 3T3 fibroblasts served as a negative control since they do not harbor K19 promoter activity or protein expression.

Serial deletions of the K19 promoter reveal important cis-regulatory elements within -288 bp

In order to determine the localization of important cis-regulatory elements within the K19 promoter, we constructed serial deletions of the K19 promoter at –960, –654, –288, –181, –102 and –58 bp which were designed to not disrupt putative cis-regulatory elements based upon computer analysis. These constructs were generated by PCR using the full-length promoter as template and subcloned into the luciferase reporter system. The activity of these deletion constructs was tested in TE-12 and Panc-1 cells as models for stratified and simple epithelia, respectively (Figure 2). Both cell lines displayed a similar pattern of promoter activity. High level K19 promoter activity was observed using the –960, –654 and –288 deletion constructs. This was followed by a significant reduction of approximately 50% in TE-12 and 70% in Panc-1 cells at –181. At –102, activity was further decreased about 50% in both cell lines. At –58 bp, almost no activity was found. Thus, the region critical for K19 expression resides between -58 and -288 bp. Of note, two canonical Sp1 DNA binding sites are located in within this region,
one in the plus strand at $-198$, and another in the minus strand at $-62$; of note, another Sp1 DNA binding site is located at $-33$ bp.

Point-mutations of canonical Sp1 sites in the full-length K19 promoter reveal a critical element at $-62$ bp

We next analyzed if these Sp1 DNA binding motifs could participate in the regulation of basal K19 promoter activity, especially since Sp1 is known to modulate the expression of other keratins (6,30-32). In order to examine the role of these three Sp1 sites, we mutated the GGCGGGG motif to GTTCGGGG in the context of the full-length promoter. This approach preserves the complete promoter sequence and focuses on the role of a putative binding site by mutation of only two nucleotides. Using a PCR-based site-directed mutagenesis approach (23), we generated three K19 promoter reporter gene constructs containing the K19 promoter ($-1970+46$) with the mutated Sp1 element at $-198$, $-62$ and $-33$ respectively. K19 promoter activity was tested after transient transfection of TE-12 (Figure 3) and Panc-1 cells and compared to the wild-type full-length K19 promoter.

As Figure 3 reveals, point mutation of the Sp1 site at $-198$ leads to a reduction of K19 promoter activity of approximately 50% in TE-12 cells. When the Sp1 site at $-62$ is mutated, we observed a nearly complete loss of K19 promoter activity in TE-12 cells. In contrast, mutation of the Sp1 site at $-33$ resulted in partial reduction of activity in TE-12 cells. Similar, if not identical, results were noted in Panc-1 cells (data not shown). These results indicated that the full-length K19 promoter harbors a critical cis-regulatory element at $-62$ bp that contributes significantly to basal activity.
**Binding of Sp1 and GKLF to the –67 to –56 element**

To characterize DNA-protein interactions between transcriptional factors and cis-regulatory elements that reside within the K19 promoter sequence at –62 bp, we performed electromobility shift assays (EMSAs) using nuclear extracts from TE-12 and Panc-1 cells. A radioactive labeled double-stranded oligonucleotide probe spanning from –81 to –48 bp was incubated with 5 µg of nuclear extracts (Figures 4 A and B). Sequences of wild-type and mutated double stranded oligonucleotides are indicated in Table 2. Computer-based analysis revealed that this region contains putative binding sites not only for Sp1, but also for GKLF/KLF4 which is important in epithelial cell differentiation and also GATA of which certain isoforms are expressed in the gastrointestinal tract.

Using TE-12 nuclear extracts, we detected two specific bands (designated as I, II) binding to the wild-type oligonucleotide probe (Figure 4 A, lane 1). In contrast, only band II was present in Panc-1 nuclear extracts (lane 7). Both bands were competed with 100x excess of unlabeled wild-type probe (lane 2 and 8). Band I was not competed by 100x excess of unlabeled mutant Sp1 oligonucleotide (lane 3). Using a specific Sp1 antibody, band I was competed away (lane 4) and, as detected with long exposure of the gel, also supershifted (not shown). Moreover, band I did not bind to the radioactive labeled Sp1 mutant probe (lane 6). These results indicate that band I represents Sp1.

Band II is not competed by 100x excess of unlabeled mutant Sp1 competitor (lane 3 and 9) and still binds the radioactive labeled Sp1 mutant probe (lane 6 and 12). Incubation of the nuclear extracts with the Sp1 antibody did not shift band II (lane 4 and 10), but using GKLF antibody, we detected a supershift of band II (lane 5 and 11). This indicates that band II represents GKLF. Incubation of the nuclear extracts with an unrelated antibody did not shift either bands I or II (data not shown). In addition, we detected very faint bands that migrated above Sp1, the identity of which are not clear.
since they were partially competed by wild-type and mutated oligonucleotide probes. However, none of the bands observed were super-shifted by a Sp3 antibody (data not shown). Of note, there was a nonspecific band that migrated between bands I and II in both TE-12 and Panc-1.

To further elucidate which nucleotides in the –81 to –48 bp element bind Sp1 and GKLF, we performed EMSAs using different unlabeled mutated oligonucleotides as competitors (Figure 4B). In addition to the Sp1 MT oligonucleotide (GGGC\texttt{T}GGAAGT), we used two other mutant sequences, the GKLF MT1 (GGGCGGT\texttt{T}AAGT) and GKLF MT2 (GGGCGGGG\texttt{GG}GT), as competitors in EMSAs. Sp1 was not competed by GKLF MT1 (lane 5), whereas the GKLF MT2 competed away Sp1 binding (lane 6). This indicates that Sp1 mediated DNA binding requires the presence of 4 guanosine nucleotides in GGGC\texttt{GG}GAAGT. GKLF was competed away by wild-type (lane 2 and 8), GATA MT (lane 3 and 9) or Sp1 MT (lane 4 and 10) oligonucleotides. GKLF MT1 failed to compete the GKLF band (lane 5 and 11), whereas GKLF MT2 was able to do so (lane 6 and 12), indicating that GKLF mediated DNA binding requires the presence of two guanosine nucleotides in GGGCGGGG\texttt{GG}AAGT. It should be noted that for both the Sp1 and GKLF bands similar results were obtained when the respective mutated oligonucleotides were radioactively labeled and used as probes in EMSAs (data not shown). Since the –81 to –48 probe contains a putative binding site for GATA transcription factors at –71, we mutated the 5’-\texttt{GATA} motif to 5’-\texttt{CTTA}. Using this mutated oligonucleotide (Figure 4B, lanes 3 and 9), we obtained in both cell lines the same competition of the specific bands as with the wild-type competitor (lane 2 and 8), indicating that none of the proteins represent a GATA related protein. This was further corroborated by immune supershift assays using specific antibodies against the GATA-4, -5 and -6 isoforms in which no supershift of the bands was observed (data not shown). Taken together, the –67 to –56 bp element in the K19 promoter contains in the
3’ strand an overlapping binding site for Sp1 and GKLF which is specifically recognized by these two transcription factors.

Recombinant protein assays show predominant binding of GKLF

To independently confirm the results from the EMSAs with nuclear extracts, we performed EMSAs with purified recombinant proteins for GKLF (24) and Sp1 (Promega). Purified human GKLF protein was obtained from cells transfected with a HIS-tagged GKLF expression vector and purified over a Nickel column as previously described (24). The purified protein was separated on a SDS-PAGE and probed with the GKLF antiserum. A band of the expected size was detected and the protein was not degraded. (data not shown).

Figure 5A depicts DNA binding activity of recombinant GKLF. Approximately 150 ng purified protein was used (lane 1) and increased to 300 (lanes 2) and 600 ng (lanes 3 to 9), respectively. Specific binding of GKLF to this nucleotide sequence is observed. With increasing amounts of GKLF protein, a second higher migrating band was observed, which might represent dimerization. Indeed, the lower band migrates identically to what is found with nuclear extracts, suggesting the higher band is detected under conditions with high GKLF protein content. The GKLF band in this assay were competed away by 100x excess of unlabeled wild-type (lane 4), Sp1 MT (lane 5) and GKLF MT2 (lane 7) oligonucleotide, but not by GKLF MT1 (lane 6) oligonucleotide. Incubation with the Sp1 antibody did not alternate the DNA binding pattern of recombinant GKLF (lane 8). In contrast, we observed a supershift using the GKLF antibody (lane 9).

Figure 5B shows binding of recombinant Sp1 protein to the wild-type oligonucleotide. Increasing amounts of recombinant Sp1 protein (150 to 600 ng) demonstrate specific binding (lanes 1 to 3). Of note, using recombinant Sp1 we observed a different migration pattern of the Sp1 band compared to endogenous Sp1
found in nuclear extracts which may be attributable to the high salt purification method used by the manufacturer (Promega). Sp1 binding was at least partially competed away by 100x excess of unlabeled wild-type (lane 4) and GKLF MT2 (lane 7) oligonucleotides, but not by Sp1 MT (lane 5) and GKLF MT1 (lane 6) oligonucleotides, reminiscent of what was observed for Sp1 in TE-12 nuclear extracts. Titration of recombinant GKLF protein (75 and 150 ng) demonstrated that increasing amounts of GKLF protein abrogated the Sp1 binding pattern (lanes 8 and 9). Approximately 150 ng of purified GKLF was sufficient to abrogate the binding of 600 ng Sp1 fusion protein (lane 9), suggesting that GKLF likely has higher binding activity to this sequence than Sp1. In addition, we performed EMSAs using nuclear extracts from TE-12 cells (Figure 5C). After addition of 150 ng purified recombinant GKLF protein into the assay, we observed complete loss of Sp1 binding (band I; lane 2). Titration of recombinant Sp1 protein (300 and 600 ng) to the nuclear extracts did not abolish GKLF binding (band II; lane 3 and 4). The same titration experiments were performed using nuclear extracts from Panc-1 cells with GKLF as the predominant DNA binding protein to the wild-type sequence (data not shown).

**Disruption of the Sp1 and GKLF binding site leads to loss of K19 promoter activity**

The EMSAs indicate that four guanosines within the GGGC\textbf{GAGT} element of the K19 promoter are especially important for binding of Sp1 and GKLF. We next elucidated the functional consequences on K19 promoter activity with the introduction of mutations in these nucleotides. In addition to the Sp1–62 mutant already mentioned, we generated two GKLF–58 mutants (MT1 and MT2) by a PCR-based site-directed mutagenesis approach in the context of the full length (~1970+46) K19 promoter. Promoter activity was determined after transient transfection of TE-12 (Figure 6) and Panc-1 cells and compared to the wild-type full-length K19 promoter.
Mutation of the Sp1 binding site (GGGCGGGAAGT to GGGCTTGAAGT) resulted in a significant loss of K19 promoter activity in TE-12 (Figure 6). Mutation of the adjacent two guanosines (GGGCGGGGAAGT to GGGCGGTTAAGT, designated as GKLF MT1), which resulted in disruption of both Sp1 and GKLF binding, lead to dramatic loss of promoter activity in TE-12 (Figure 6). Of note, nearly identical results were observed in Panc-1 cells (data not shown). GKLF MT2 (GGGCGGGGAAGT to GGGCGGGGGAAGT), which preserves the DNA binding of Sp1 and GKLF, accordingly did not alter activity of the full-length K19 promoter (Figure 6).

Endogenous Sp1 and GKLF level correlate with K19 promoter activity

Based upon the results in the EMSAs and protein competition experiments, we next tested the hypothesis that endogenous Sp1 and GKLF protein levels influence K19 promoter activity. While Sp1 was variably expressed in the cell lines, its level is reduced in Panc-1 cells (Figure 7). This is supportive of the notion that we did not observe Sp1 binding in EMSAs with Panc-1 nuclear extracts. GKLF expression also varied but is found to be at high levels in both esophageal squamous and pancreatic ductal cancer cells (Figure 7). However, GKLF expression was not detected in pancreatic acinar cancer cells (AR42J, ARIP). Thus, it appears that cells with low expression levels of both Sp1 and GKLF have little, if any, K19 promoter activity and K19 protein expression as shown in figure 1.

Overexpression of Sp1 and GKLF restores K19 promoter activity in cells lacking K19 promoter activity

If endogenous expression levels of Sp1 and GKLF contribute differentially to K19 promoter activity and expression of its gene product in different cell types, it is reasonable to postulate that overexpression of these transcription factors would lead to increase of K19 promoter activity. We therefore cotransfected different cell lines with the
full-length K19 promoter reporter and either empty control vector, Sp1 or GKLF expression vectors and combination of both (Figure 8). Overexpression of Sp1 or GKLF in TE-12 and Panc-1 cells leads only to a slight change in the K19 promoter activity, since in these cells the promoter is already maximally stimulated (data not shown). However, we were able to reconstitute K19 promoter activity in the two pancreatic acinar cancer cell lines, AR42J and ARIP. Both cell lines have only minimal promoter activity. Sp1 overexpression alone leads to a 3-fold increase in K19 promoter activity. When GKLF alone was overexpressed, we also obtained a similar stimulation of K19 promoter activity. Overexpression of both Sp1 and GKLF demonstrated additive stimulation of K19 promoter activity. This supports the notion that Sp1 and GKLF both bind and transactivate the K19 promoter through an additive mechanism.
Discussion

Analysis of the transcriptional control of keratin genes provides important insights into cellular differentiation and malignant transformation. K19 is expressed during early embryonic development (33,34) and has a restricted epithelial pattern in the adult (22). K19 expression has been implicated as a marker for stem cells in the skin (21) and perhaps other stratified squamous epithelial systems as well in duct-like precursor cells in the pancreas (35). In both contexts, knowledge of how pluripotential stem cells undergo transdifferentiation into different lineages would be extremely helpful. Stem cells are believed to reside in the outer root sheath of the hair follicle which ultimately give rise to the various components of the epidermis (21). In contrast, putative stem cells are believed to reside in the proliferating basal cell compartment of other stratified squamous epithelia such as the esophagus, which rapidly become differentiated and ultimately undergo senescence followed by continuous renewal (36). Pancreatic ductal precursor or putative stem cells, undergo transdifferentiation into ductal, acinar and islet cells. In addition, since K19 is implicated in the malignant transformation in these cell types, it is important to define how K19 may play a role in the squamous epithelial and pancreatic ductal cancers. In particular, the preponderance of pancreatic cancers arise from ductal cells, not acinar or islet cells.

To elucidate how K19 is regulated at the transcriptional level in certain cell types, we focused our attention upon esophageal squamous and pancreatic ductal cell lineages. We observed high K19 promoter activity in certain cell lineages, namely in esophageal squamous and well-differentiated pancreatic ductal cancer cells. Strikingly, activity was essentially negative in pancreatic cell lines of acinar origin. Recognizing that the K19 promoter has multiple positive and negative \textit{cis}-regulatory elements, serial deletions fused to the luciferase reporter gene were generated and tested in esophageal squamous and pancreatic ductal cancer cell lines. While there was high K19 promoter
activity proximal to −288 bp, a remarkable reduction in activity was found when sequences containing one of the three canonical Sp1 binding sites at -62 was deleted. Mutation of the Sp1 core motif in the context of the full-length K19 promoter revealed a critical role for the apparent Sp1 site at −62 for basal promoter activity. When mutated, K19 promoter activity dropped significantly in both cell lines tested. Of particular interest, this Sp1 element is conserved in the mouse, rat and human K19 promoter at approximately the same position (Table 3), underscoring its functional importance.

Biochemical analysis of the K19 promoter sequence centered at −62 bp revealed recognition and binding by Sp1 and GKLF using both nuclear extracts and recombinant fusion proteins. Moreover, recombinant protein assays demonstrated that GKLF mediated DNA binding activity is higher since it abrogates binding of both recombinant Sp1 and endogenous Sp1. Sequence specific mutations within this 3′-GGGCGGGGAAGT element revealed, that binding of Sp1 requires the presence of 4 guanosine nucleotides (GGGC GGGAAGT) in the K19 promoter, whereas binding of GKLF requires the presence of two guanosine nucleotides (GGGCGGGGGAAGT). As a functional consequence, when these two guanosine nucleotides are mutated in the full-length K19 promoter (disrupting both Sp1 and GKLF binding to this site), we observe the most prominent loss of basal promoter activity both in TE-12 and Panc-1 cells.

Our biochemical and functional studies suggest that GKLF contributes to specific activation of the K19 promoter in pancreatic ductal epithelial cells, but not pancreatic acinar cells. This may in part be attributable to differences in GKLF protein levels with abundant expression in pancreatic ductal epithelial cells but virtually none in acinar cells, but perhaps even more so, promotes the notion that K19 indeed might be a marker for pancreatic ductal cells which to date has escaped such a specific marker. Furthermore, GKLF expression correlates with K19 promoter activity, but appears to be only minimally dependent upon Sp1 in these cell types, whereas both are required in esophageal
squamous epithelial cells. Accordingly, when Sp1 and GKLF are overexpressed in cell lines with low or absent K19 promoter activity, promoter activation can be restored.

The Sp1/KLF family comprises more than 16 different mammalian members of zinc finger transcription factors that bind GC/GT rich DNA elements (37,38). These transcription factors share within the C-terminal DNA binding domain three zinc-fingers, similar to the *Drosophila melanogaster* zinc finger protein Krüppel. However, they have different functions in regulating gene transcription. Among them, the most and best-studied member is Sp1, which is ubiquitously expressed and has been involved in cellular control of cell-cycle regulation and differentiation (39-41). Involvement of Sp1 in the regulation of keratins has been previously described. The promoter of the corneal-specific K3 gene is regulated by an overlapping Sp1/AP-2 site. Interestingly, endogenous levels of Sp1 and AP-2 define K3 gene transcription in differentiating corneal epithelial cells (32,42). K5, a marker of basal epithelial cells in the epidermis, is regulated within a cluster of transcription sites, including Sp1 binding (43). Moreover, Sp1, among other transcription factors, regulates expression of both K16 and K17, keratins which are induced under pathologic conditions of the skin (30,31). K18, overexpressed in a tumorigenic clone of a human colon carcinoma cell line, is regulated by Sp1 within a minimal promoter element (44,45). Our previous studies indicate that interaction between cyclin D1 and Sp1 is important in the regulation of the K4 promoter in primary versus malignant squamous epithelial cells (46). The induction of keratin gene expression is a complex interplay of different transcription factors acting via different cis-regulatory elements, including Sp1, AP1, AP2 and ets factors (6,47).

The KLF family of transcription factors is phylogenetic linked to the Sp1 family (37). One subgroup among KLFs, sharing structural and functional similarities, comprises the closely related gut- (GKLF, also named EZF or according to a newly proposed nomenclature, KLF4), erythroid cells- (EKLF or KLF1) and lung- (LKLF or KLF2) enriched Krüppel-like factors. Representing relatively tissue-restricted
transcription factors, an important role in cell proliferation and differentiation has been proposed by several studies (48-51). GKLF is found in epithelial cells of the gastrointestinal tract in the intestine and esophagus (50), pancreas (Brembeck and Rustgi, unpublished observations), lung and skin (51). Induction of GKLF mRNA levels is observed in the mouse embryo at day 15 (51) and is important for intestinal development (52). Additionally, GKLF is down regulated in late stages of tumorigenesis of the min mouse model (52). Expression of GKLF is high in growth-arrested fibroblasts and almost absent in exponentially growing (50). In addition, GKLF expression is found in multilayered epithelia, where there switch from proliferation to differentiation (50,51).

GKLF binds DNA through a core DNA motif of 5'- (G/A)(G/A)GG(C/T)G(C/T)-3' (53). Only few gene targets for GKLF transactivation have been previously characterized. Our own work identified the esophageal squamous epithelium specific K4 and Epstein-Barr virus ED-L2 promoters as targets for GKLF (24). Both promoters harbor a CACCC-like motif, which has been described as the core motif for EKLF (48). In the K4 and ED-L2 promoter, this motif is recognized by GKLF resulting in activation of transcription (24). Other studies described the GKLF core motif in the basic transcription element of the cytochrome P-450IA1 promoter CYP1A1. Both Sp1 and GKLF bind and recognize this element. However, GKLF inhibits Sp1-mediated activation of the CYP1A1 promoter by competing for this binding site and by protein-protein interactions with Sp1 (54). In contrast, our studies reveal that both Sp1 and GKLF act positively on an overlapping Sp1/GKLF site in the K19 promoter with preferential binding by GKLF. However, in the context of the K19 promoter, GKLF and Sp1 do not appear to interact directly, since the immune supershift assays are specific for either Sp1 or GKLF. We postulate that the endogenous protein levels of Sp1 and GKLF are important determinants in binding to this element. The GKLF binding motif represented in the 3'-GGGCCGGGGAAGT element of the K19 promoter slightly differs from the postulated GKLF minimal core motif (G/A)(G/A)GG(C/T)G(C/T). However, in the same study by
Shields and Yang, analysis of DNA sequences recognized by GKLF includes the 3’-GGGAAG motif (53).

Our novel finding that GKLF contributes to the tissue-specific transcriptional regulation of K19 expression provides a new paradigm for the role of this Krüppel-like factor in the regulation of gene expression. GKLF appears to play an important role in cell-specific differentiation by activating K19 expression which is influenced by Sp1 in stratified squamous epithelial cells but is relatively independent of Sp1 in pancreatic ductal epithelial cells. Furthermore, pancreatic acinar cells, in contrast to pancreatic ductal cells, do not harbor K19 promoter activity or protein expression which may help to explain in part to explain lineage differences between acinar and ductal cells. It is possible that Sp1 and GKLF modulate K19 gene transcription differently, where Sp1 is important developmentally (41) and GKLF directs cell fate decisions during differentiation.
Acknowledgements

This work was supported by NIH grant R01-DK53377 (AKR), The Leonard and Madlyn Abramson Family Cancer Research Institute (AKR), and Deutsche Forschungsgemeinschaft Br 1806/1-1 (FHB). We thank Dr. Mitch Lazar for critical reading of this manuscript. We also thank Dr. D. Chung for providing the pEVR2-Sp1 vector, Dr. A. M. Mes-Mason for the K19 promoter fragment, and Dr. V. Yang for the GKLF antiserum. The authors appreciate the support of Prof. Dr. S. Rosewicz and Prof. Dr. E. O. Riecken.

Footnotes

The nucleotide sequence reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number AF237661.

The abbreviations used are: K, keratin; GKLF, gut-enriched Krüppel-like factor; EKLF, erythroid Krüppel-like factor; LKLF, lung Krüppel-like factor; bp, base pair(s); kb, kilobase pair(s); PCR, polymerase chain reaction; EMSA, electrophoretic mobility shift assay.
References

1. Jorcano, J. L., Rieger, M., Franz, J. K., Schiller, D. L., Moll, R., and Franke, W. W. (1984) J Mol Biol 179(2), 257-81

2. Quinlan, R. A., Schiller, D. L., Hatzfeld, M., Achtstatter, T., Moll, R., Jorcano, J. L., Magin, T. M., and Franke, W. W. (1985) Ann N Y Acad Sci 455, 282-306

3. Pendleton, J. W., Violette, S. M., Hunihan, L. W., Greene, L. A., and Ruddle, F. H. (1991) Genomics 9(2), 369-72

4. Nadeau, J. H., Berger, F. G., Cox, D. R., Crosby, J. L., Davisson, M. T., Ferrara, D., Fuchs, E., Hart, C., Hunihan, L., Lalley, P. A., and et al. (1989) Genomics 5(3), 454-62

5. Byrne, C., Tainsky, M., and Fuchs, E. (1994) Development 120(9), 2369-83

6. Eckert, R. L., and Welter, J. F. (1996) Mol Biol Rep 23(1), 59-70

7. Jones, P. H., Harper, S., and Watt, F. M. (1995) Cell 80(1), 83-93

8. Sun, T. T., Tseng, S. C., Huang, A. J., Cooper, D., Schermer, A., Lynch, M. H., Weiss, R., and Eichner, R. (1985) Ann N Y Acad Sci 455, 307-29

9. Fradette, J., Germain, L., Sesaiah, P., and Coulombe, P. A. (1998) J Biol Chem 273(52), 35176-84

10. Baribault, H., and Oshima, R. G. (1991) J Cell Biol 115(6), 1675-84

11. Magin, T. M., Schroder, R., Leitgeb, S., Wanninger, F., Zatloukal, K., Grund, C., and Melton, D. W. (1998) J Cell Biol 140(6), 1441-51

12. Stasiak, P. C., Purkis, P. E., Leigh, I. M., and Lane, E. B. (1989) J Invest Dermatol 92(5), 707-16

13. Bader, B. L., Magin, T. M., Hatzfeld, M., and Franke, W. W. (1986) Embo J 5(8), 1865-75

14. Stasiak, P. C., and Lane, E. B. (1987) Nucleic Acids Res 15(23), 10058

15. Bader, B. L., Jahn, L., and Franke, W. W. (1988) Eur J Cell Biol 47(2), 300-19
16. Ceratto, N., Dobkin, C., Carter, M., Jenkins, E., Yao, X. L., Cassiman, J. J., Aly, M. S., Bosco, P., Leube, R., Langbein, L., Feo, S., and Romano, V. (1997) *Cytogenet Cell Genet* **77**(3-4), 169-74

17. Lussier, M., Filion, M., Compton, J. G., Nadeau, J. H., Lapointe, L., and Royal, A. (1990) *Gene* **95**(2), 203-13

18. Filion, M., Sarafian, V., Lussier, M., Belanger, C., Lapointe, L., and Royal, A. (1994) *Genomics* **24**(2), 303-10

19. Bouwens, L., Lu, W. G., and De Krijger, R. (1997) *Diabetologia* **40**(4), 398-404

20. Stosiek, P., Kasper, M., and Karsten, U. (1990) *Liver* **10**(1), 59-63

21. Michel, M., Torok, N., Godbout, M. J., Lussier, M., Gaudreau, P., Royal, A., and Germain, L. (1996) *J Cell Sci* **109**(Pt 5), 1017-28

22. Bartek, J., Bartkova, J., Taylor-Papadimitriou, J., Rejthar, A., Kovarik, J., Lukas, Z., and Vojtesek, B. (1986) *Histochem J* **18**(10), 565-75

23. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pease, L. R. (1989) *Gene* **77**(1), 51-9

24. Jenkins, T. D., Opitz, O. G., Okano, J., and Rustgi, A. K. (1998) *J Biol Chem* **273**(17), 10747-54

25. Hagen, G., Muller, S., Beato, M., and Suske, G. (1994) *Embo J* **13**(16), 3843-51

26. Schreiber, E., Matthias, P., Muller, M. M., and Schaffner, W. (1989) *Nucleic Acids Res* **17**(15), 6419

27. Quandt, K., Frech, K., Karas, H., Wingender, E., and Werner, T. (1995) *Nucleic Acids Res* **23**(23), 4878-84

28. Rosewicz, S., Stier, U., Brembeck, F., Kaiser, A., Papadimitriou, C. A., Berdel, W. E., Wiedenmann, B., and Riecken, E. O. (1995) *Gastroenterology* **109**(5), 1646-60

29. Brembeck, F. H., Kaiser, A., Detjen, K., Hotz, H., Foitzik, T., Buhr, H. J., Riecken, E. O., and Rosewicz, S. (1998) *Br J Cancer* **78**(10), 1288-95
30. Magnaldo, T., Bernerd, F., Freedberg, I. M., Ohtsuki, M., and Blumenberg, M. (1993) *DNA Cell Biol* **12**(10), 911-23

31. Milisavljevic, V., Freedberg, I. M., and Blumenberg, M. (1996) *DNA Cell Biol* **15**(1), 65-74

32. Chen, T. T., Wu, R. L., Castro-Munozledo, F., and Sun, T. T. (1997) *Mol Cell Biol* **17**(6), 3056-64

33. Jackson, B. W., Grund, C., Schmid, E., Burki, K., Franke, W. W., and Illmensee, K. (1980) *Differentiation* **17**(3), 161-79

34. Lussier, M., Ouellet, T., Lampron, C., Lapointe, L., and Royal, A. (1989) *Gene* **85**(2), 435-44

35. Bouwens, L. (1998) *J Pathol* **184**(3), 234-9

36. Karam, S. M. (1999) *Front Biosci* **4**, D286-98

37. Philipsen, S., and Suske, G. (1999) *Nucleic Acids Res* **27**(15), 2991-3000

38. Turner, J., and Crossley, M. (1999) *Trends Biochem Sci* **24**(6), 236-40

39. Birnbaum, M. J., van Wijnen, A. J., Ogden, P. R., Last, T. J., Suske, G., Stein, G. S., and Stein, J. L. (1995) *Biochemistry* **34**(50), 16503-8

40. Karlseder, J., Rotheneder, H., and Wintersberger, E. (1996) *Mol Cell Biol* **16**(4), 1659-67

41. Marin, M., Karis, A., Visser, P., Grosveld, F., and Philipsen, S. (1997) *Cell* **89**(4), 619-28

42. Wu, R. L., Chen, T. T., and Sun, T. T. (1994) *J Biol Chem* **269**(45), 28450-9

43. Ohtsuki, M., Flanagan, S., Freedberg, I. M., and Blumenberg, M. (1993) *Gene Expr* **3**(2), 201-13

44. Gunther, M., Frebourg, T., Laithier, M., Fossar, N., Bouziane-Ouartini, M., Lavialle, C., and Brison, O. (1995) *Mol Cell Biol* **15**(5), 2490-9

45. Prochasson, P., Gunther, M., Laithier, M., Fossar, N., Lavialle, C., and Brison, O. (1999) *Exp Cell Res* **248**(1), 243-59

46. Opitz, O. G., and Rustgi, A. K. (2000) *Cancer Res* in press
47. Brembeck, F. H., Opitz, O. G., Libermann, T. A., and Rustgi, A. K. (2000) *Oncogene* 19(15), 1941-9

48. Miller, I. J., and Bieker, J. J. (1993) *Mol Cell Biol* 13(5), 2776-86

49. Anderson, K. P., Kern, C. B., Crable, S. C., and Lingrel, J. B. (1995) *Mol Cell Biol* 15(11), 5957-65

50. Shields, J. M., Christy, R. J., and Yang, V. W. (1996) *J Biol Chem* 271(33), 20009-17

51. Garrett-Sinha, L. A., Eberspaecher, H., Seldin, M. F., and de Crombrugghe, B. (1996) *J Biol Chem* 271(49), 31384-90

52. Ton-That, H., Kaestner, K. H., Shields, J. M., Mahatanankoon, C. S., and Yang, V. W. (1997) *FEBS Lett* 419(2-3), 239-43

53. Shields, J. M., and Yang, V. W. (1998) *Nucleic Acids Res* 26(3), 796-802

54. Zhang, W., Shields, J. M., Sogawa, K., Fujii-Kuriyama, Y., and Yang, V. W. (1998) *J Biol Chem* 273(28), 17917-25
Figure Legends

Figure 1: Activity of the 5’ regulatory region of the K19 gene in different cell types correlates with endogenous K19 protein expression

A. Dose-response curve of the –1970+46 K19 promoter in different cell lines. Cells were transiently transfected with the indicated amount of the K19-1970+46 luciferase reporter construct. Hela cells were used as positive control and NIH 3T3 fibroblasts served as negative control. Luciferase activity is expressed as x-fold activity over control (pGL3 basic vector) ± SEM. Transfections were performed in triplicate and confirmed through three independent experiments.

B. Expression of the 40kDa K19 protein as determined by Western Blot Analysis. Total cell lysates were separated by SDS page and protein expression determined using a specific K19 antibody.

Figure 2: Serial deletion analysis of the K19 promoter

Serial deletions of the K19-1970+46 promoter at -960, -654, -288, -181, -102 and –58 bp were generated and subcloned into the luciferase reporter gene vector. Activity was tested in TE-12 and Panc-1 after transient transfection with 1 µg of the respective construct. Relative Luciferase Units (RLU) were determined and results of luciferase activity for each subset of experiment were corrected to RSV-Luc control. Results of at least three independent experiments are shown.

Figure 3: Point-mutations of canonical Sp1 sites in the full-length K19 promoter

The canonical Sp1 sites at –198, -62 and –33 in the K19 promoter were mutated by PCR based site-directed mutagenesis (GGGGCGGGG → GGTTCGGGG) and the
mutated K19 full-length promoter subcloned into the luciferase reporter gene vector. 1 µg of each construct was transiently transfected into TE-12 cells. Activity of wild-type and mutated full-length K19 promoter is shown as relative luciferase units (RLU) corrected to RSV-Luc control. Results of at least three independent experiments ± SEM are indicated.

Figure 4: Electrophoretic Mobility Shift Assays (EMSAs) show binding of Sp1 and GKLF to the –67 to –56 element in the K19 promoter

A. 5 µg nuclear extracts from TE-12 and Panc-1 cells were incubated with radioactive labeled double-stranded oligonucleotide probe spanning from –81 to –48 bp. WT= wild-type, MT = mutant oligonucleotide. Competition and immunosupershift experiments were performed as indicated. Band I correspondents to the Sp1-DNA complex and band II corresponds to the GKLF-DNA complex.

B. Competition experiments with 100x unlabeled oligonucleotides in TE-12 and Panc-1 cells. Different mutated oligonucleotides (Table 2) were used as competitors with the –81 to –48 wild-type probe.

Figure 5: Recombinant protein assays show predominant binding of GKLF to the –67 to –56 element of the K19 promoter

A. DNA binding activity of recombinant GKLF protein. Approximately 150 ng purified recombinant GKLF protein was used and increased to 300 and 600 ng. DNA binding was competed with 100x of unlabeled wild-type and mutant oligonucleotides. Immunosupershift was performed using Sp1 and GKLF antibodies.

B. DNA binding activity of recombinant Sp1 protein. 150 ng purified protein was used and increased to 300 and 600 ng. Competition was performed using 100x of
unlabeled wild-type and mutant oligonucleotides. In addition, Sp1 binding was assessed by titration of 75 and 150 ng of purified GKLF protein.

C. Recombinant protein assay using 5 µg nuclear extracts from TE-12 cells. 150 ng purified recombinant GKLF protein and 300 to 600 ng purified recombinant Sp1 protein were added into the reaction(s).

Figure 6: Biological consequences of point-mutations of the Sp1 and GKLF sites at –62/–58 in the full-length K19 promoter

One mutant of the Sp1 sites at –62 and two mutants of the GKLF binding site at –58 were generated by PCR based site-directed mutagenesis in the context of the K19 full-length promoter and subcloned into the luciferase reporter vector. 1 µg of each construct was transiently transfected into TE-12 cells. Activity of wild-type and mutated full-length K19 promoter is shown as relative luciferase units (RLU) corrected to RSV-Luc control. Results of at least three independent experiments ± SEM are indicated.

Figure 7: Protein expression levels of endogenous Sp1 and GKLF in different cell types

Expression of the Sp1 and GKLF protein in the indicated cell lines was determined by Western Blot Analysis. Equal amounts of total cell lysates (10 µg) were separated by SDS-PAGE (10-15%) and protein expression determined using a specific Sp1 or GKLF antibody, respectively.

Figure 8: Overexpression of Sp1 and GKLF restores K19 promoter activity in cells lacking K19 promoter activity

Acinar pancreatic cancer cells AR42J and ARIP cells were cotransfected with 1µg K19 full-length luciferase reporter vector and equal amounts of either empty control
vector or the Sp1 or GKLF expression vectors. Luciferase activity is expressed as x-fold activity over control (pGL3 basic vector) ± SEM.
**Table 1:** Sequences of K19 promoter specific primers. Nucleotides in bold represent the mutated nucleotides (MT) compared to the wild-type sequence (WT). Note that oligonucleotides for the K19-Sp1-62 MT, K19-GKLF-58 MT1 and MT2 have been used for EMSAs and for PCR-generation of the respective luciferase reporter gene constructs (*see Methods*).

| Primer               | Nucleotide sequence (5' to 3')                                                                 |
|----------------------|------------------------------------------------------------------------------------------------|
| **Sequencing primers** |                                                                                              |
| K19 INT1  5'         | GAG TCC TTC TAC CAC TGA TTC                                                                    |
| K19 INT2  3'         | CTT AGA AAT CAC ACC CCT CTG                                                                    |
| K19 INT3  5'         | TGA AGA CAA GAG GAC TGA CTT                                                                    |
| K19 INT4  3'         | CAG AGT GGG ACA TCT TCA TAA                                                                    |
| **K19 promoter deletion primers** |                                                                                             |
| K19-1970  5'         | CGG GGT ACC AAG TTC CTT TCT AAG ACC CA                                                        |
| K19-960  5'          | CGG GGT ACC GAA TTG TTG CAC TGC AGG CT                                                         |
| K19-654  5'          | CGG GGT ACC AGG GAA GGG TGG AGG TGT CTT                                                        |
| K19-288  5'          | CGG GGT ACC GAA GTT AGA GTT GTG TAA GC                                                         |
| K19-181  5'          | CGG GGT ACC AAG GAC TGA GAC CTC TGG CT                                                         |
| K19-102  5'          | CGG GGT ACC ATA TTT GCA CTC TGG GAG CTT                                                        |
| K19-58   5'          | CGG GGT ACC TCA ACA TCT CCA TCC CCC TT                                                         |
| K19+46   3'          | AGG GAG CTC GGA AAA AGG GAC GCA GGT CT                                                         |
| **Sp1 binding site mutant primers and oligonucleotides** |                                                                                             |
| K19-Sp1-198 MT  5'   | AGT GGG CTG G**TT** CGG GGC AGC TCT GGG AA                                                     |
| K19-Sp1-198 MT  3'   | AGA GCT GCC CCG A**AC** CAG CCC ACT CCA AAG                                                    |
| K19-Sp1-62 MT  5'    | GGA AAT TTC TGA TAC CCG A**AC** CTT CAA CAT                                                     |
| Oligonucleotide | Type | Sequence 5' | Sequence 3' |
|-----------------|------|-------------|-------------|
| K19-Sp1-62 MT 3' | GGA GAT GTT GAA GGT **T**CG GGT ATC AGA AAT |
| K19-Sp1-33 MT 5' | ATC CCC CTT CCC **G**AA CCG GGC ATA AAA AG |
| K19-Sp1-33 MT 3' | TTT TAT GCC CGG **T**TC GGG AAG GGG GAT GG |
| **K19-81 to –48 wildtype oligonucleotides** |
| K19-81 WT 5' | GGA AAT TTC TGA TAC CCG CCC CTT CAA CAT |
| K19-48 WT 3' | GGA GAT GTT GAA GGG GCG GGT ATC AGA AAT |
| **GATA binding site mutant oligonucleotides** |
| K19-GATA-71 MT 5' | GGA AAT TTC TCT TAC CCG CCC CTT CAA CAT |
| K19-GATA-71 MT 3' | GGA GAT GTT GAA GGG GCG GGT **A**AG AGA AAT |
| **GKLF binding site mutant primers and oligonucleotides** |
| K19-GKLF-58 MT1 5' | GGA AAT TTC TGA TAC CCG CCA ATT CAA CAT |
| K19-GKLF-58 MT1 3' | GGA GAT GTT GAA **T**TG GCG GGT ATC AGA AAT |
| K19-GKLF-58 MT2 5' | GGA AAT TTC TGA TAC CCG CCC C**G**G CAA CAT |
| K19-GKLF-58 MT2 3' | GGA GAT GTT G**C** GGG GCG GGT ATC AGA AAT |
Table 2: Sequences of double-stranded oligonucleotides used in EMSAs. Sequences of wild-type (WT) and mutated (MT) oligonucleotides corresponding to position –81 to –48 in the K19 promoter. The potential binding sites for GATA, Sp1 and GKLF are indicated.

|                | Sp1 | GATA | GKLF |
|----------------|-----|------|------|
| **K19-81/-48 WT** | 5’ GGA AAT TTC TGA_TAC CCG CCC CTT CAA CAT CTC C | 3’ CCT TTA AAG ACT ATG GGC GGG GAA GTT GTA GAG G |
| **K19-GATA-71 MT** | 5’ GGA AAT TTC TCT_TAC CCG CCC CTT CAA CAT CTC C | 3’ CCT TTA AAG AGA ATG GGC GGG GAA GTT GTA GAG G |
| **K19-Sp1-62 MT** | 5’ GGA AAT TTC TGA_TAC CCG AAC CTT CAA CAT CTC C | 3’ CCT TTA AAG ACT ATG GGC TTG GAA GTT GTA GAG G |
| **K19-GKLF-58 MT1** | 5’ GGA AAT TTC TGA_TAC CCG CA_TTT CAA CAT CTC C | 3’ CCT TTA AAG ACT ATG GGC GTT TAA GTT GTA GAG G |
| **K19-GKLF-58 MT2** | 5’ GGA AAT TTC TGA_TAC CCG CCC CCG CAA CAT CTC C | 3’ CCT TTA AAG ACT ATG GGC GGG GCC GTT GTA GAG G |
Table 3: Comparison of mouse, rat and human K19 promoter sequences. At approximately the same position relative to the transcription start site, the Sp1 and GKLF binding site is conserved in all three species. (n.a., not published or no GenBank entry available; straight line, Sp1 DNA binding site; dashed line, GKLF DNA binding site)

| Species | Position | Sequence | Reference | GenBank       |
|---------|----------|----------|-----------|---------------|
| mouse   | -71      | 5’GATACCCGCCCCTTCAACATCTCCATCCC<br>3’CTATGGCGGGGAAAGTTGAGGCTAGG<br>................. | this paper, (17) | AF237661, M36120 |
| rat     | -68      | 5’GATACCGCCCTTCAACATCTCCACTCC<br>3’CTATGGCGGGGAAAGTTGAGGCTAGG<br>................. | n.a. | AF089865          |
| human   | -83      | 5’GATATCGCCCCTGACACCATTCCCTCCTT<br>3’CTATAGGGGACTGTTGTAAGGGGAA<br>............... | (15) | n.a.              |
B

**TE-12**

- WT probe
- WT
- GATA MT
- Sp1 MT
- GKLF MT1
- GKLF MT2

| 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| free probe |

**Panc-1**

- WT probe
- WT
- GATA MT
- Sp1 MT
- GKLF MT1
- GKLF MT2

| 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|----|----|----|
| free probe |
|            | esophageal | pancreatic ductal | pancreatic acinar |
|------------|------------|-------------------|------------------|
| T.T        |            |                   |                  |
| TE-12      |            |                   |                  |
| Hs766T     |            |                   |                  |
| Panc-1     |            |                   |                  |
| Capan-2    |            |                   |                  |
| AR42J      |            |                   |                  |
| ARIP       |            |                   |                  |
| Hela       |            |                   |                  |
| 3T3        |            |                   |                  |

**Sp1**

- 97 kDa

**GKLF**

- 48 kDa
The tissue-dependent Keratin 19 gene transcription is regulated by GKLF/KLF4 and Sp1
Felix H. Brembeck and Anil K. Rustgi

J. Biol. Chem. published online June 19, 2000

Access the most updated version of this article at doi: 10.1074/jbc.M004013200

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts