3 Results

The most complete and least biased method to analyze the expression control of a particular gene is to characterize the promoter of the gene in question and the regulatory regions positioned further upstream or downstream of the promoter. The promoter is defined as the sequence necessary for guiding the RNA polymerase to the transcription start site. In order to locate the promoter, the mRNA transcription start site has to be mapped first.

3.1 Transcription start site of the murine LI-cadherin gene

The transcription initiation site for the mLI-cadherin gene was mapped using 5′-RACE-PCR and primer extension.

3.1.1 5′-RACE-PCR

The 5′-RACE-PCR (rapid amplification of cDNA- ends) provides a method for the detection of messenger RNA from a source of poly(A)$^+$ RNA. First, single strand cDNA is synthesized from total RNA by reverse transcriptase (RT) using Superscript II and a 20-mer oligo(dT) primer (Overdier et al., 1997), then a cDNA double-strand is generated using DNA polymerase. An adapter-ligated DNA fragment with a known sequence is ligated to each blunt end of the cDNA double-strand. Now it is possible to amplify the unknown upstream (5′) region with a primer complementary to the adapter-ligated DNA fragment and a gene-specific reverse primer. When using a gene-specific forward primer, the downstream (3′) sequence can be amplified.

Murine intestinal mRNA was isolated from total RNA as described in section 2.2.12. 1 µg of the poly(A)$^+$ RNA was transcribed into cDNA, and a double stranded adapter was ligated to the ends of the cDNA. Then PCR was used to amplify the sequence between the 5′-oligonucleotide and a reverse primer specific for LI-cadherin (P400R) binding 400 bp downstream of the previously described 5′-end of the mRNA (Angres et al., 2001). A schematic set-up of the adapter-ligated cDNA and the positions of the used primers are shown in figure 3.1A. Figure 3.1B shows an ethidium bromide-stained agarose gel in which a 400 bp PCR product is observed when using the primer combination AP1 and P400R (lane P). A 300 bp PCR product was obtained using two LI-cadherin specific primers (in situ 1 and P400R, lane C). The 400 bp PCR product was cloned into the TA-vector pCR2.1 and amplified in *E.coli* TOP10F′. The lengths of the cloned LI-cadherin-RACE-PCR products were examined with colony-PCR (section 2.2.4, table 2.5). The primer Ex1R positioned 50 bp downstream of the known 5′-end and the adapter primer AP-1 (figure 3.1A) were used to investigate the plasmid content of about 100 colonies of the transformation. In figure 3.2, a collection of some of the PCR products is shown after separation in an agarose gel. The colonies not containing LI-cadherin cDNA, or shorter fragments than the hitherto known cDNA, did not yield any PCR products. Plasmids from the colonies containing the five longest PCR products were isolated and subsequently sequenced. By this process, the already known 23 bp long 5′-untranslated region of the LI-cadherin cDNA (Angres et al., 2001) was expanded to 127 bp.

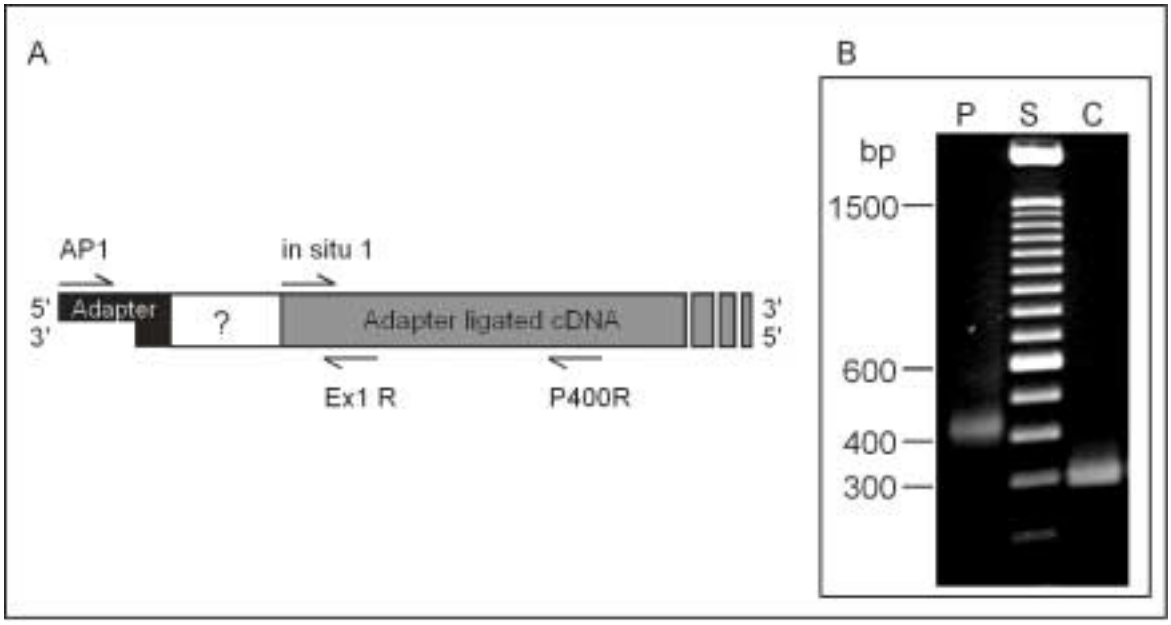


Figure 3.1: Adapter ligated cDNA and 5′-RACE-PCR. **(A)** Design of adapter-ligated cDNA including the binding positions of the 5′-RACE-primers (AP-1 and P400R), the secondary primer (Ex1R) and the control primers (in situ and P400R). **(B)** Ethidium bromide-stained agarose gel; P = 5 µl 5′-RACE-PCR product, C = 5 µl control-PCR product, and S = 100 bp DNA ladder. Positions of molecular weight standards (bp) are indicated on the left. The unknown 5′-region is indicated by a question mark.

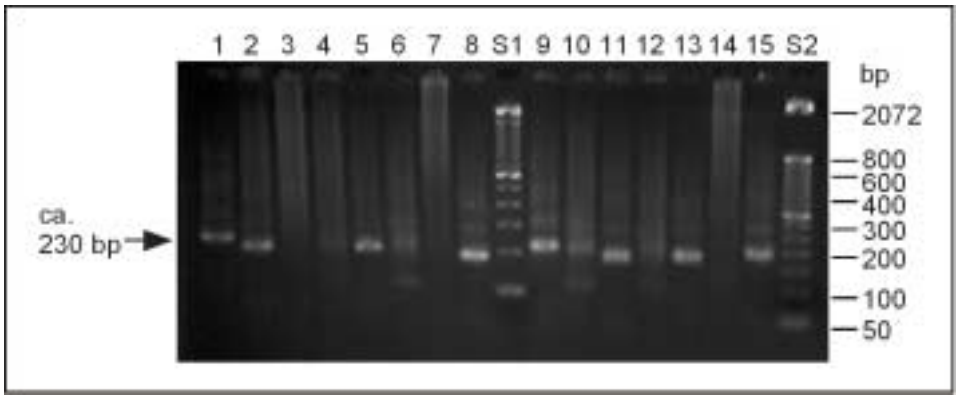


Figure 3.2: Colony-PCR of different LI-cadherin 5′-RACE-PCR clones. Ethidium bromide-stained agarose gel shows products of different colonies. S1 = 100 bp DNA ladder, S2 = 50 bp DNA ladder. The arrows indicate the insert length selected for sequencing.

3.1.2 Primer extension

The 5′-untranslated region of the murine LI-cadherin cDNA was further characterized by a primer extension reaction. With this method, the 5′-terminus of RNA can be mapped and the amount of RNA can be quantified. A primer is radioactively end-labeled, hybridized to the RNA, and extended by reverse transcriptase using unlabeled deoxynucleotides to form a single-stranded DNA complementary to the template RNA. The resulting DNA is analyzed on

a sequencing gel. The length of the extended primer reflects the maximal length of the mRNA and thus the first transcription start-site (Sambrook and Fritsch, 1989).

Two different oligonucleotides in different preparations with 90 µg total RNA each were used to generate extension products of the murine LI-cadherin. The reverse primers used differed in position by 31 bp, and were positioned close to the putative start site derived from the 5′-RACE-PCR. The results of the primer extension analysis are shown in figure 3.3. The reaction yielded a major product of 67 bp for primer 1 and 98 bp for primer 2. The difference between the two primer extension products was 31 bp, which corresponds to the distance of the two oligonucleotides on mRNA level. Two further experiments reproduced maximal extension products with equal sizes. A control reaction with yeast tRNA instead of mRNA did not produce any extension product.

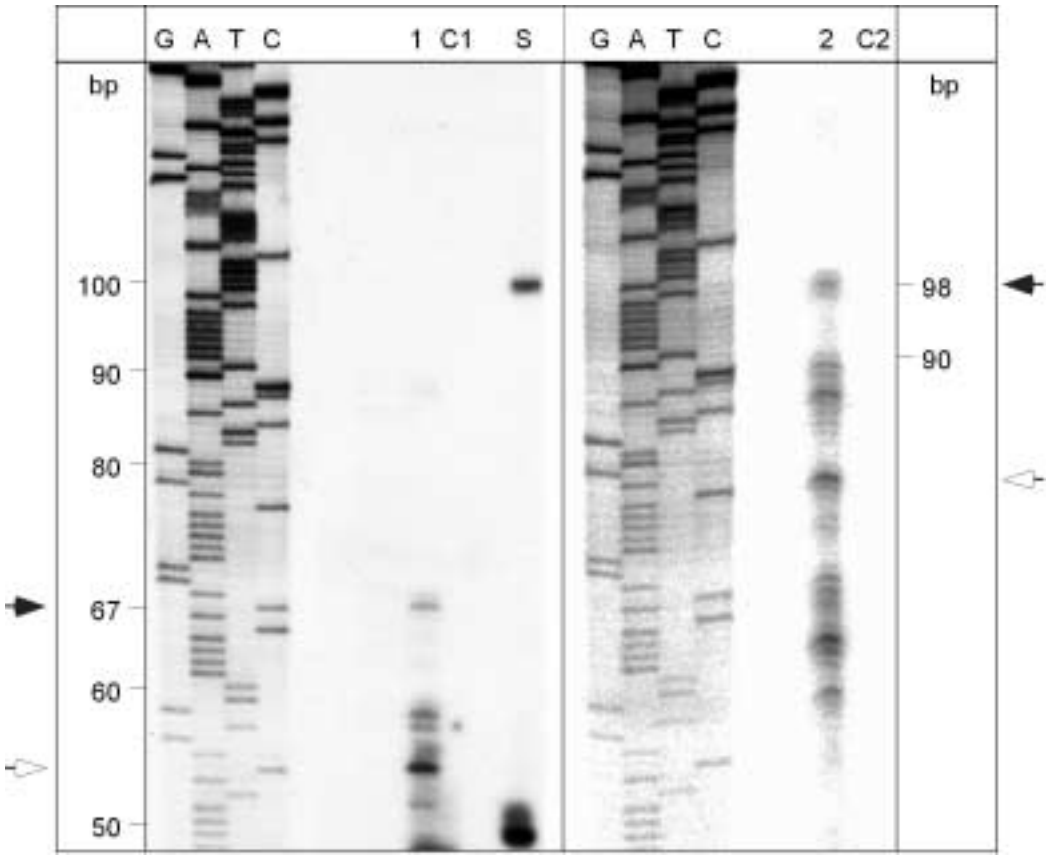


Figure 3.3: Autoradiogram of the primer extension after separation in a sequencing gel. Two murine LI-cadherin specific reverse DNA-primers differing by 31 bp in their target position were used to generate primer extension products with 90 µg of total murine intestinal RNA. They yielded products of 67 bp and 98 bp, respectively (lanes 1 and 2; filled arrows). A negative control reaction was performed with the same primers and 100 µg yeast tRNA (lane C1 and C2). Sequence reaction products (Lane G, A, T and C) from a third primer and a LI-cadherin-cDNA plasmid served as a 1 bp-ladder in addition to a P^{32}-labeled 50 bp DNA-ladder (Lane S). Open arrows = possible second start site.

Remarkably, many mRNA molecules seemed to have a start site beginning 18 bp further downstream (3′). Both primers showed increased signal intensity in this region. The question about a second transcription start site could not be fully answered due to the bad gel resolution in this region indicated by the bp ladder. However, shorter products in the primer extension reaction may define a second transcription start site. Alternatively, secondary structures of the isolated mRNA can be responsible for an earlier interruption of the reverse transcriptase-driven extension reaction. Heterogeneity of the 5′-terminus due to degradation can also explain the shorter extension products.

The specific nature of the primer extension was verified by control reactions carried out in parallel to the experiments. No products were obtained with either primer when yeast tRNA was substituted for intestine mRNA.

3.1.3 Sequence of the proximal 5′-flanking region of the LI-cadherin gene

The results from the 5′-RACE PCR, primer extension and the sequencing analysis of the respective genomic region are summarized in figure 3.4. The underlined part indicates the sequence obtained from the 5′-RACE PCR; the oligonucleotides 1 and 2 used by the primer extension are framed. The arrow depicts the transcription initiation site mapped by primer extension. It is apparent that the longest 5′-RACE PCR product is just 7 nucleotides shorter than the longest extension product.

The transcription initiation site mapped by primer extension analysis was located within the sequence AC<u>A</u>GGTT (underscore indicates nucleotide at +1), that is 134 bp upstream from the ATG translation initiation codon. A cytosine is located in front of the first transcribed adenine and the second transcribed nucleotide is guanine. This greatly resembles the transcriptional initiation site consensus sequence identified by an analysis of 502 vertebrate mRNAs (Bucher, 1990) and the consensus sequence pyrimidine-pyrimidine-adenine$_{+1}$-N-thymine/adenine-pyrimidine-pyrimidine found by an examination of TATA-less promoters (Javahery et al., 1994). In addition, two other sequences frequently found directly upstream of the initiation site were found: A variant CCAAT box was located at position –86, and a GC box was located at position –34. No TATA box was identified near the transcription start site.

3.2 Reporter gene systems for analyzing the LI-cadherin promoter

In order to verify that the proximal 5′-flanking region contains a functional promoter and to evaluate a possible cell specificity of the 5′-flanking region, reporter gene assays were performed in cultured cells. The reporter gene plasmid pEGFP-1 contains the cDNA of the “enhanced green fluorescent protein” (EGFP) without a promoter. The gene is transcribed and EGFP expressed only after the introduction of an activating DNA region upstream of the reporter gene. EGFP is a protein that shows fluorescence after excitation with light having a wave length of 488 nm. Hence, the expression of EGFP provides a basis for qualitative analysis of transcriptional activity which can be easily observed using a fluorescence microscope.

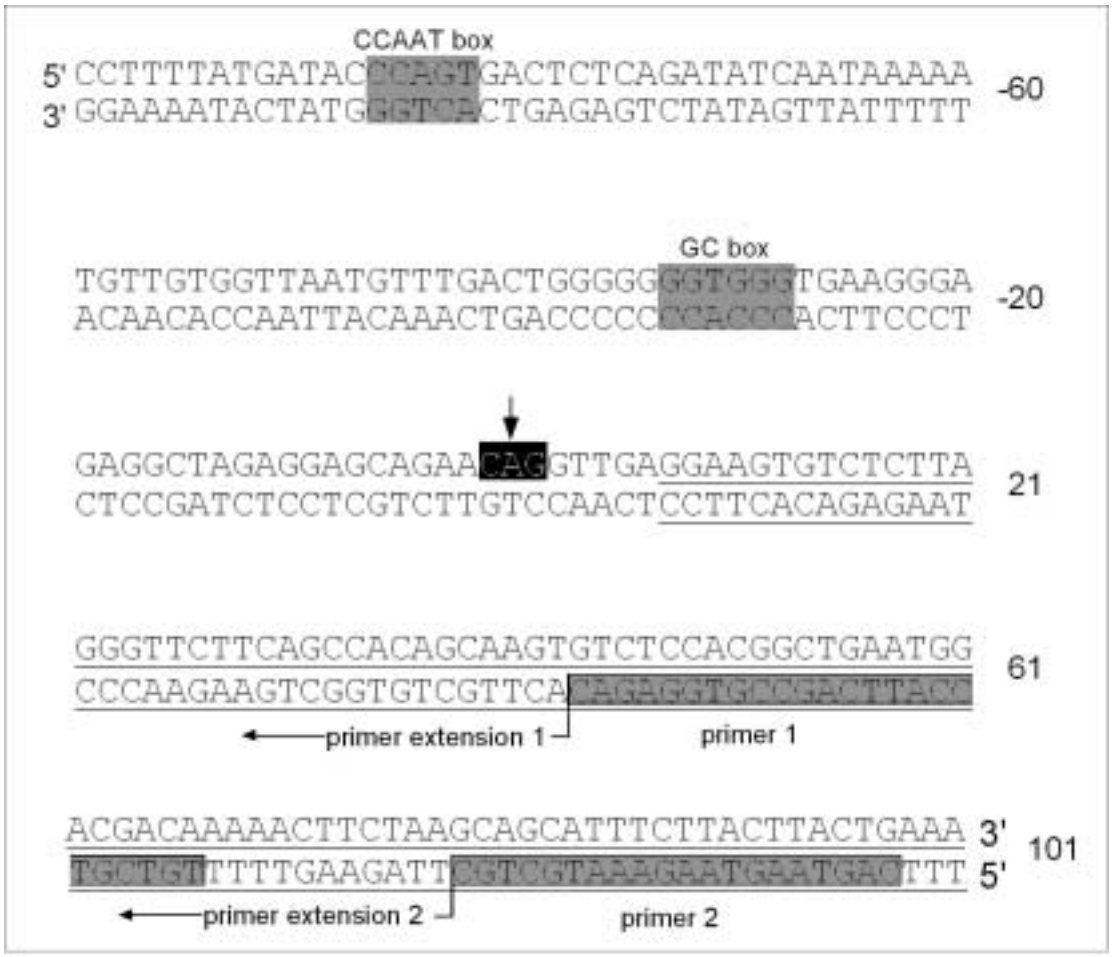


Figure 3.4: Localization of the transcriptional initiation site. The first nucleotide to be transcribed is marked by an arrow. The positions of the two oligonucleotides used in the primer extension reaction are shown in shadowed boxes. The sequence confirmed through 5′-RACE-PCR is underlined. The 5′-flanking region of the transcriptional start site reveals a potential GC-element located at position –33 and a CCAAT box located at position –86.

3.2.1 Cloning of the 5′-flanking region of the murine LI-cadherin gene

Published data of other intestine specific expressed proteins, like sucrose-isomaltase (Markowitz et al., 1993) and the fatty acid binding protein (Simon et al., 1993) indicate that most of the regulatory DNA elements important for a tissue-specific expression are localized within 3 kb upstream of the transcription initiation site. Based on these data, two different fragments of the 5′-flanking region of the murine LI-cadherin gene, 1.8 kb and 2.8 kb including the putative transcription start site and 67 bp of the first exon, were chosen for the initial study of the mouse LI-cadherin 5′-flanking region. Figure 3.5A shows a schematic set-up of the 5′-flanking region and the first untranslated exon of the LI-cadherin gene. The positions of primers used for PCR amplification of the 1.8 and 2.8 kb DNA fragments are shown. An isolated cosmid clone (cLI-1) containing the 5′-end of the LI-cadherin cDNA (Jung, 2000) was used as a template for the PCR-amplification of the 5′-flanking region of the LI-cadherin gene. Two fragments containing the transcription initiation site, 2774 bp or 1794 bp of the proximal 5′-flanking region and 70 bp of the first (non-coding) exon were amplified

from the cosmid clone cLI-1 using Taq-DNA-polymerase (figure 3.5A). The PCR products were ligated into the TA-vector pCR2.1 followed by transformation of *E.coli* TOP10' with the ligation products. The resulting plasmid clones were sequenced in order to exclude mutations due to amplification. After the ligation into the pCR2.1-vector, it was possible to excise the integrated amplification products with a wide range of different restriction enzymes. Next, clones containing the correct inserts were digested with EcoRI, followed by a separation of the inserts from the vector backbone by agarose gel electrophoresis (figure 3.5B). The two 5′- LI-cadherin gene fragments were subsequently excised and extracted from the gel. Finally, the EcoRI restriction fragments were inserted upstream to the EGFP reporter gene in the promoter-less vector pEGFP-1. After the transformation of *E.coli* DH5α with the reporter gene plasmids, clones containing the plasmid of the correct size and the insert in the right orientation were used to isolate a large quantity of highly purified plasmid-DNA, designated pLI(2.8F)-EGFP and pLI(1.8F)-EGFP. In figure 3.5C, an ethidium bromide-stained agarose gel shows the EcoRI-digested pEGFP-1 vector in lane P. The linearized pLI(2.8F)-EGFP and pLI(1.8F)-EGFP constructs are shown in the neighboring two lanes. A vector card of the pLI(2.8F)-EGFP construct is shown in figure 3.5D.

3.2.2 Sequence analysis of the putative promoter and the 5′-flanking region

The promoter is the most important element responsible for the expression of gene products, although elements regulating the cell specific expression pattern are not necessarily located in this region. The presence or absence of enhancer sequences and the interaction between multiple activator and inhibitor proteins are important regulatory elements of the transcriptional initiation. The number and type of regulatory elements to be found is different for every gene. Different combinations of transcription factors can also exert differential regulatory effects upon transcriptional initiation. The various cell types express characteristic combinations of transcription factors; this is the major mechanism for cell-type specificity in the regulation of mRNA gene expression.

The 5′-flanking region of the murine LI-cadherin gene was examined for the presence of typical eukaryotic gene regulatory elements using the transcription factor database TRANSFAC (http://transfac.gbf.de/TRANSFAC). The specific positions of the identified elements within the sequence are summarized in table 3.1. In addition to the detected CCAAT- and GC-box, several potential binding sites for transcription factors were identified. These transcription factors are involved in tissue-specific gene expression like the activator protein-1 (AP-1), the hepatocyte nuclear factor 1 (HNF-1), the homeodomain transcription factor Cdx2, and the GATA factors. Other conserved binding sites for transcription factors include sites for Oct1, RFX-1, and cMyb.

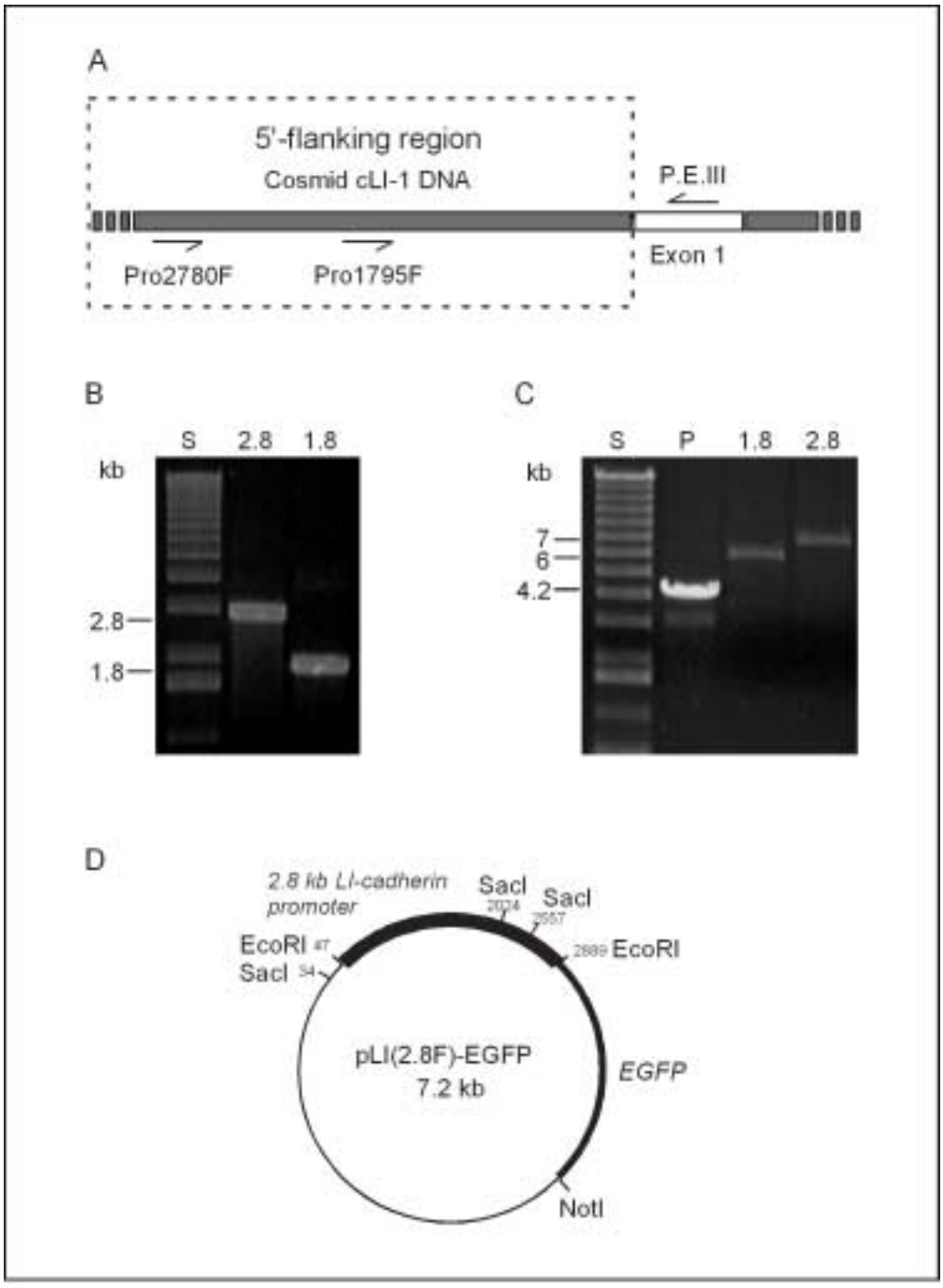


Figure 3.5: (A) Schematic view of the region chosen for the initial study of the LI-cadherin 5′-flanking region. The positions of the forward (Pro2780F, Pro1795F) and reverse primers (P.E.III) used for PCR amplification of the 5′-flanking region of the LI-cadherin gene are shown with the cosmid cLI-1 as template. 2774 bp and 1794 bp of the 5′-flanking region combined with 70 bp of exon 1 (downstream of the putative start site) were amplified in two different reactions. **(B)** Ethidium bromide-stained agarose gel showing the two different PCR-products, 2774 bp (2.8 kb) and 1794 bp (1.8 kb) fragments of the LI-cadherin 5′-flanking region. S = 1 kb DNA ladder. **(C)** Ligation of EcoRI-digested fragments of the 5′-flanking region into the vector pEGFP-1: Agarose gel electrophoresis of pEGFP-1 linearized with EcoRI (P) and pLI(1.8F)-EGFP and pLI(2.8)-EGFP linearized with NotI. S = 1 kb DNA ladder. **(D)** Vector card of the reporter gene construct pLI(2.8F)-EGFP.

Table 3.1: Putative regulatory elements in the LI-cadherin 5′-flanking region.

Name + reference	Consensus	Sequence	Location	Identity
Initiator (Javahery et al., 1994)	YYANWYY Y= C,T W= A,T N= A,C,G,T	aCAGgTT	-2 to +5	5/7
GC box (Briggs et al., 1986; Bucher, 1990)	GGRGGG R= C,T,A	GGTGGG	–33 to –28	6/6
AP-1 (Angel et al., 1991)	TGAGTCA	TGAcTgg TGAcTCt TtAtTCA	-43 to –37 -82 to –76 -215 to –209	4/7 5/7 5/7
HNF1 (Mendel et al., 1991)	GTTAATNATTAAC N= A,C,G,T	GTTAATGtTTgAC	–52 to –40	11/13
Cdx2 (Margalit et al., 1993)	ATAAAA TTTTAT	ATAAAA TTTTAT	–66 to –61 –97 to –91	6/6 7/7
GATA-1 GATA-2 GATA-3 (Ko and Engel, 1993)	WGATAV W= A,T V= A,G	AGATAt TGATAA TGATAc gGATAc AGATAG	–74 to –69 –92 to –87 –104 to –99 –259 to –254 –267 to –262	5/6 6/6 5/6 4/6 6/6
CCAAT box (Bucher, 1990)	CCAAT	CCAgT	–86 to –82	4/5
OCT1 (Kemler et al., 1989)	CTCATGA ATGCAAAT	GTaATGt tTCATat ATaCAcAc ATGCcAca	–109 to –103 –212 to –206 –257 to –249 –290 to –283	5/7 4/7 5/8 5/8
RFX1 (David et al., 1995)	CCCTAGCAACAG	aCtcAGCAACca	–177 to –166	7/12
cMyb (Howe and Watson, 1991)	YAACKG Y= C,T K= G,T	CAACTt CAACTa	–248 to –243 –264 to –259	5/6 5/6
IK2 (Molnar et al., 1996)	NNNYGGGAWNNN N= A,C,G,T Y= C,T W= A,T	GGGaGGGAGCT AGATGGGAAAG TGGgGGGATTT GAGTGGGAcTTA	–820 to –810 –969 to –959 –1504 to –1493 –642 to –631	10/11 11/11 10/11 10/11
STAT1 (Horvath et al., 1995)	TTCCCVKAA V= A,G K= G,T	TTCCaGGAA	–1039 to –1032	8/9
GFI1 (Zweidler-Mckay et al., 1996)	NNNAAATCASWGYNNN N= A,C,G,T S= G,C W= A,T Y= C,T	ACAtAATCcagGTCTG TGTAAATCAGAtaGGC CTGgAATCtaTtCATA CAGcAATCtaTaaATT AGTgAATCAtcagGCA AGTTAATCttAcTTCT ACAAAATCACTtgTCA	–1463 to –1448 –1433 to –1418 –1334 to –1319 –1230 to –1215 –1031 to –1016 –925 to –910 –448 to –432	12/16 15/16 12/16 12/16 11/16 13/16 14/16
HFH3 (Overdier et al., 1997)	RSRTVTTTVYRTW R= G,T,A S= G,T,C V= A,G Y= C,T W= A,T	TaATGTTTGacTg cTGTGTTTGTTcc	–50 to –38	9/13

3.2.3 Characterization of an *in vitro* cell-system

In order to analyze the reporter gene constructs containing the 5′-flanking region of the LI-cadherin *in vitro*, the following favorable characteristics were used to select suitable eukaryotic cell lines: The cell lines should express endogenous LI-cadherin, and should be efficiently transfectable with plasmid-DNA. Further, the cell lines should be morphologically similar to intestinal epithelial cells and should express typical intestinal marker genes. The murine intestinal cell line STC-1 (Rindi et al., 1990) and the murine colon carcinoma cell line CMT93 (ATTC nr. CCL-223) were tested for these criteria.

Cell lysates from each cell line were tested for endogenous expression of LI-cadherin by Western blotting. LI-cadherin was detected in the lysates from the STC-1 cells (figure 3.6A), whereas it was absent in the CMT93 cell lysate (not shown).

LI-cadherin detected in cell lysates from STC-1 cells is about the same size as LI-cadherin detected in murine intestinal lysates. The minimal size difference can be due to a different glycosylation state of the proteins caused by different cultivating conditions (Gawlitzek et al., 1995). The characteristic distribution of LI-cadherin along the cell membrane was verified for the STC-1 cells by immunofluorescent staining (figure 3.6B). LI-cadherin was detected at the cell-cell contact zones resembling the expression pattern of LI-cadherin in murine intestinal cells. It has also been shown that the hormone cholecystokinin and a functional $GABA_A$ receptor (Mangel et al., 1993; Snow et al., 1994) are secreted by STC-1. These proteins are characteristic for neuroendocrine cells in the gastrointestinal tract; their secretion confirmed that the STC-1 cell line has maintained at least some of its differentiation state. This assumption was supported by the detectable expression of E-cadherin and the proteins α- and β-catenin (figure 3.6C-E), which are important for the formation of the stable E-cadherin mediated cell-cell adhesion. In many carcinoma cell lines, the expression of E-cadherin is disturbed (Behrens et al., 1992).

Lysate from murine intestinal cells was used as a positive control. Based on these data, the STC-1 cell line was a useful candidate for the LI-cadherin promoter test system.

3.2.4 Qualitative reporter gene assay: GFP-readout

The DEAE-dextran method was used to transfect STC-1 cells with the EGFP reporter gene constructs pLI(1.8F)-EGFP and pLI(2.8F)-EGFP in order to analyze the putative activity of the LI-cadherin 5′-flanking region. As control plasmids, a plasmid containing the CMV promoter driving the expression of the EGFP gene (pEGFP-C1) and the promoter-less plasmid pEGFP-1 were used. $2x10^6$ cells were transfected with one of the four reporter genes in three parallel reactions. One representative experiment out of three independent experiments is shown in figure 3.7. Both EGFP constructs containing the LI-cadherin 5′-flanking region were able to activate the transcription of EGFP. After 48 hours, EGFP was expressed in STC-1 cells. The cells transfected with the pEGFP-1 plasmid without a promoter did not show a specific fluorescence, whereas the expression level in the cells transfected with pEGFP-C1 was much higher than the expression driven by the LI-cadherin 5′-flanking region.

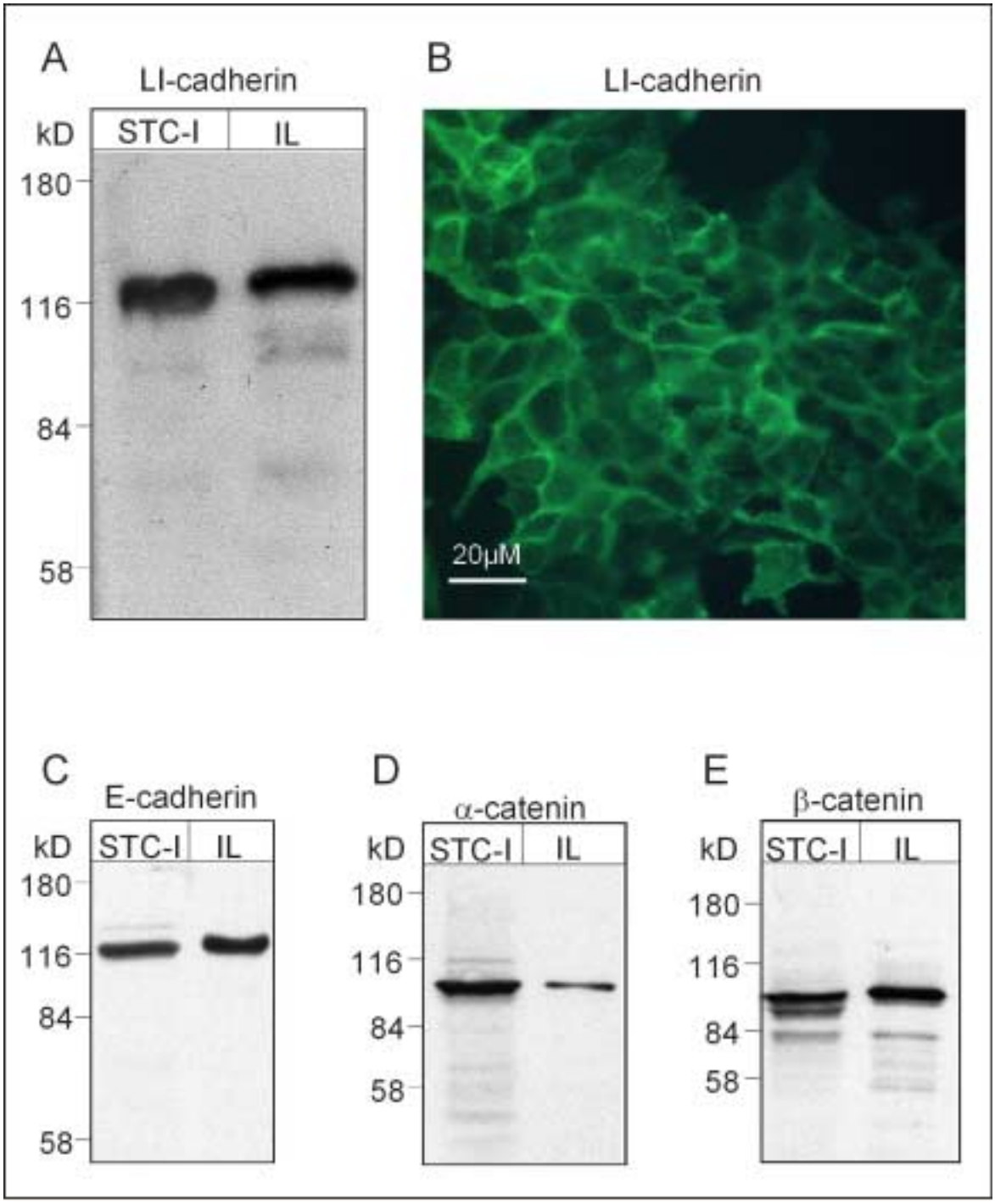


Figure 3.6: Expression of LI-cadherin and components of the cadherin-catenin complex in STC-1 cells. **(A)** Western blot of STC-1 cell lysate and lysate from murine intestine (IL) with anti-mLI-cadherin-pAb. **(B)** Immunofluorescence of STC-1 cells with anti-mouse LI-cadherin-pAb and Alexa 488 anti-rabbit-IgG-secondary antibody. Bar = 20 µm. Western blot of STC-1 cells and lysate from murine intestine (IL) with **(C)** anti-E-cadherin-mAb, **(D)** anti-α-catenin-mAb, and **(E)** anti-β-catenin-pAb. The primary antibodies in (A), (C), (D), and (E) were detected with a secondary antibody conjugated with HRP using a luminescence reaction.

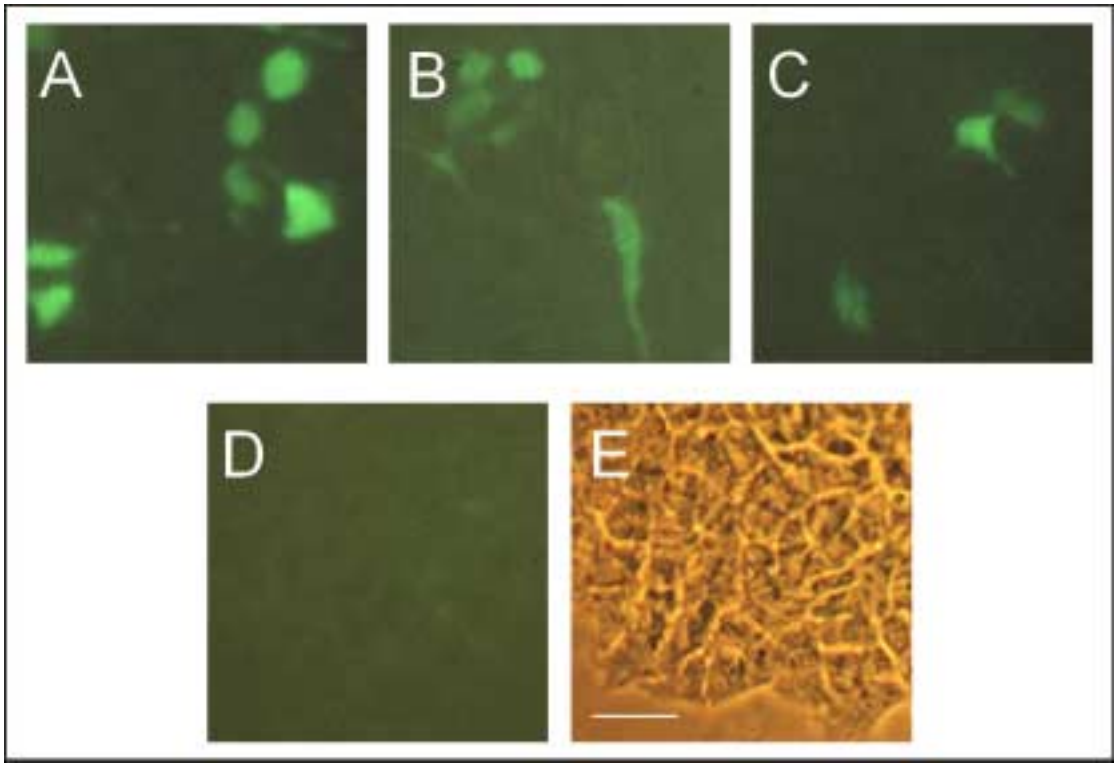


Figure 3.7: Promoter assay with EGFP as reporter gene. STC-1 cells were transfected with **(A)** pEGFP-C1, **(B)** pLI(1.8F)-EGFP, **(C)** pLI(2.8)-EGFP, and **(D)** pEGFP-1, incubated for 48 hours at 37°C and 5% CO_2 and visualized by fluorescence microscopy. **(E)** Phase contrast image of (D). Bar = 20 µm.

3.2.5 Quantitative reporter gene assay: Luciferase readout

The use of reporter genes is essential for the analysis of gene expression and regulation. While measurement of EGFP can give a qualitative evaluation of the transfection efficiency and the expression level, it is often necessary to have a system in which differences between varying regulatory regions can be measured with greater accuracy.

The reporter gene plasmid pGL3-basic provides a basis for the quantitative analysis of cis-acting factors such as promoters or enhancers that regulate mammalian gene expression. It contains the coding region for a specific luciferase, which is isolated from the North American firefly *Photinus pyralis* (DeLuca and McElroy, 1978). Firefly luciferase is one of the most extensively studied enzymes that catalyzes light production in bioluminescent organisms. It requires luciferin, ATP, and O_2 as substrates. The reactions catalyzed by firefly luciferase are:

$$\text{luciferase} + \text{luciferin} + \text{ATP} \xleftrightarrow{Mg^{2+}} \text{luciferase} \cdot \text{luciferyl-AMP} + PP_I$$

$$\text{luciferase} \cdot \text{luciferyl-AMP} + O_2 \rightarrow \text{luciferase} + \text{oxyluciferin} + \text{AMP} + CO_2 + h\nu$$

The vector pGL3-basic contains a firefly luciferase cDNA that has been optimized for monitoring transcriptional activity in transfected eukaryotic cells. Since it does not contain

any promoter sequences, putative regulatory sequences can be inserted upstream of the firefly luciferase coding region in order to study their influence on the transcriptional activity.

Two reporter gene constructs were generated using the same cloning strategy as in the construction of the reporter gene constructs pLI(1.8)-EGFP and pLI(2.8)-EGFP. Since pGL3-basic does not contain an EcoRI site in its multiple cloning region, a linker including the EcoRI recognition sequence was first introduced into the SmaI-restriction site. The 1.8 kb and 2.8 kb EcoRI-restriction fragments obtained from the pCR2.1 plasmid (section 2.1.1) were subsequently ligated into the EcoRI site in the pGL3-basic vector. *E.coli* DH5α were transformed with the reporter gene plasmids, and clones containing plasmid DNA with the correct size and insert in the right orientation were used to isolate a large quantity of highly purified reporter gene constructs designated pLI(1.8F)-luc and pLI(2.8F)-luc. An additional construct containing the 2.8 kb fragment in antisense orientation was also isolated and designated pLI(2.8R)-luc.

Differences in cell viability and transfection efficiency can cause experimental variability. To eliminate inherent assay-to-assay variability, the Dual-Luciferase Reporter Assay System was used. This system is based on a simultaneous expression and measurement of the firefly luciferase (experimental reporter) and the *Renilla* luciferase (internal control) within the same sample. Firefly and *Renilla* luciferases have different enzyme structures and substrate requirements which make it possible to selectively discriminate between their respective bioluminescent reactions. The luminescence from the firefly luciferase reaction may be quenched while simultaneously activating the luminescent reaction of *Renilla* luciferase. The uptake of the two different luciferase-containing vectors is supposed to be proportional in all transfections. Normalizing the activity of firefly luciferase to the activity of *Renilla* luciferase minimizes experimental variability caused by differences in cell viability or transfection efficiency. Other sources of variability, such as differences in pipetting volumes, cell lysis efficiency and assay efficiency, are also effectively minimized (Sherf et al., 1996).

3.2.6 Functional analysis of the LI-cadherin promoter-region

The promoter constructs pLI(1.8F)-luc, pLI(2.8F)-luc, and the positive control plasmid pGL3-control were co-transfected with the vector pRL-SV40 encoding the cDNA for *Renilla* luciferase into STC-1 cells with DEAE-dextran. Since promoter activity should be orientation-dependent, cells were also transfected with pGL3-basic containing the 2.8 kb LI-cadherin 5′-flanking region cloned in the reverse complementary orientation (pLI(2.8R)-luc) as a negative control. 48 hours after transfection, the cells were rinsed with PBS and lysed with passive lysis buffer (PLB). Firefly and *Renilla* luciferase activity was assayed in 20 µl of cell extract by adding the appropriate substrates according to the manufacturer's directions. Light emissions were integrated for the initial 10 sec of emission using a luminometer.

The activity of firefly and *Renilla* luciferase was examined for each plasmid in ten parallel transfection reactions (figure 3.8A). A plot of the luminescence of firefly against the luminescence of *Renilla* showed a constant correlation between the two enzymes.

Interestingly, cells transfected with pLI(1.8F)-luc showed 30% higher luciferase activity than pLI(2.8F)-luc transfected cells, whereas the pLI(2.8R)-luc construct was not able to activate

the transcription of luciferase (figure 3.8B). The SV40 promoter driving the transcription of luciferase in pGL3-control showed a very low activity in STC-1 cells as compared to the LI-cadherin promoter. Based on these data, the dual-luciferase reporter assay system was conformed as a useful system for normalizing the activity of the experimental reporter firefly luciferases.

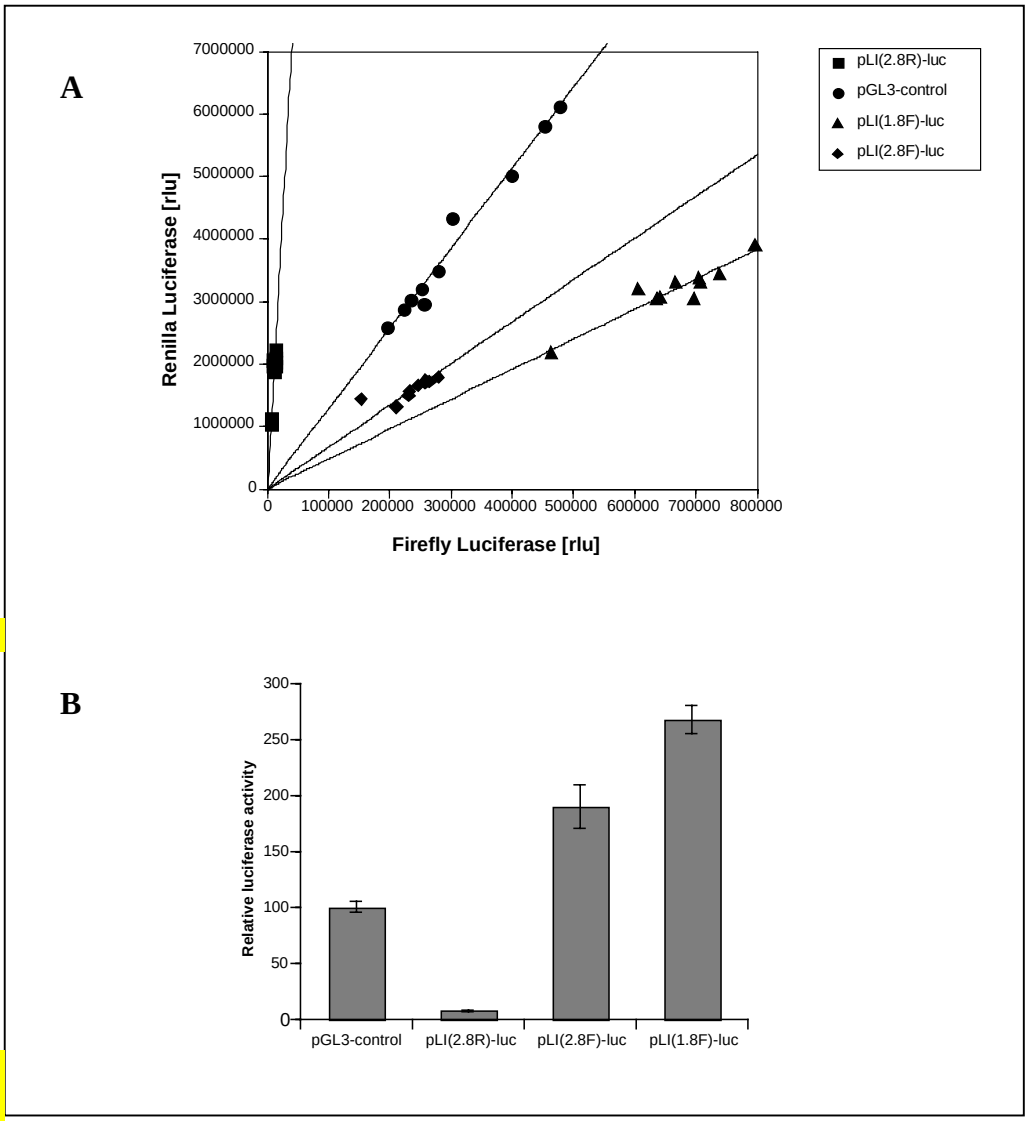


Figure 3.8: Correlation of firefly luciferase activity and *Renilla* luciferase activity. **(A)** Plot of the luminescence of firefly luciferase against the luminescence of *Renilla* luciferase. Each data point represents an independent transfection experiment. In each experiment, STC-1 cells were transfected with $1x10^{11}$ molecules of the respective firefly luciferase reporter gene construct including pGL3-control, pLI(1.8F)-luc, pLI(2.8F)-luc, and pLI(2.8R)-luc, and $1x10^{10}$ molecules of the control plasmid pRL-SV40 containing *Renilla* luciferase as reporter gene. The DEAE dextran method was used (rlu = relative light units). **(B)** The data in (A) plotted as relative luciferase activity of each plasmid (the mean relative activities ± SD). In (A), a constant correlation between firefly and *Renilla* can be seen. In (B), pLI(1.8F)-luc and pLI(2.8F)-luc show a 34- and 24-fold luciferase activity respectively, compared to pLI(2.8R)-luc. The activity of pLI(1.8F)-luc was significantly higher than the activity of pLI(2.8F) and pGL3-control.

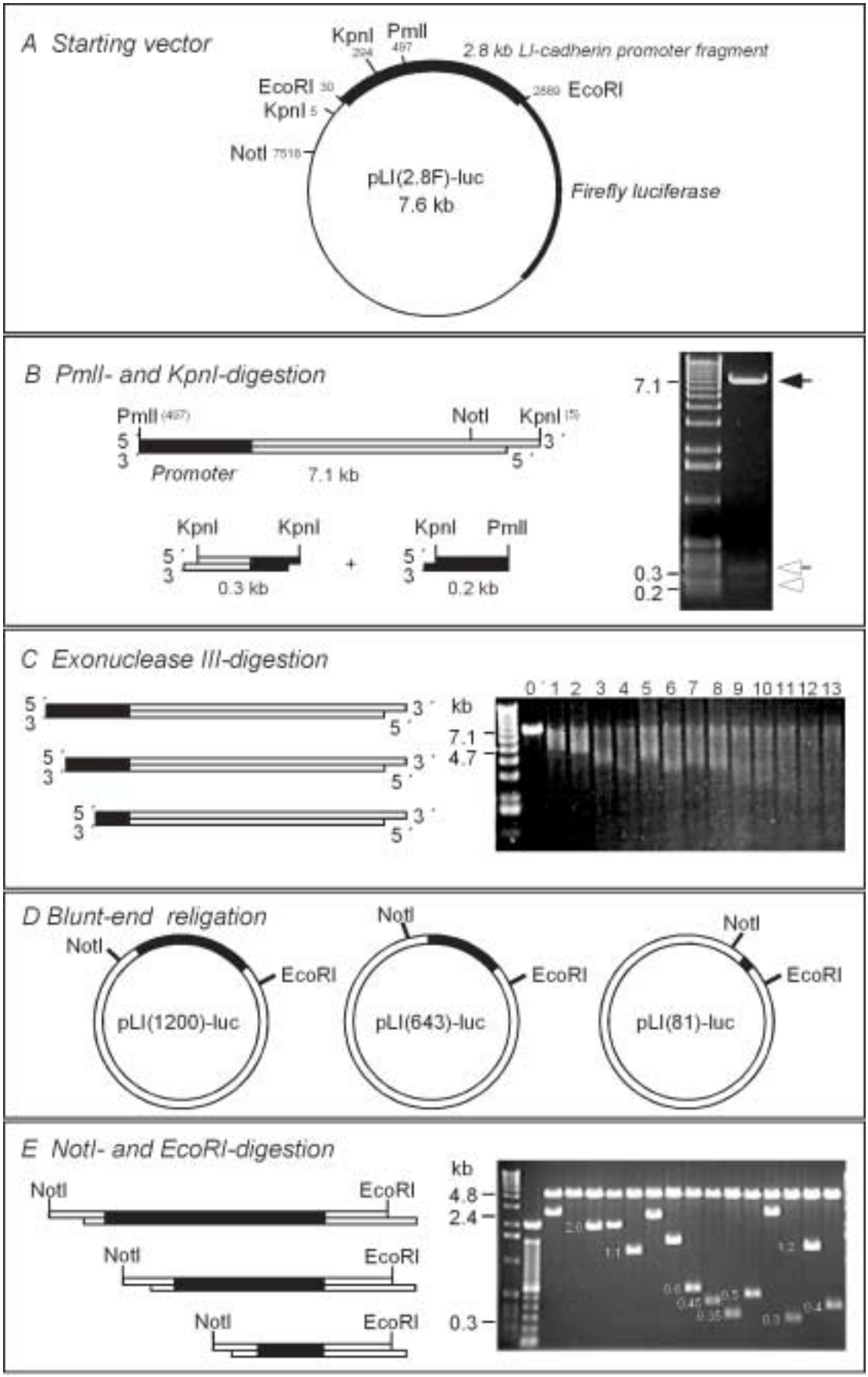
3.3 Serial 5′-deletion analysis of the LI-cadherin promoter

In order to identify transcriptional active regulatory regions upstream of the transcription start site, a large set of deletion mutants lacking progressively more nucleotides from the 5′-end was generated.

The pLI(2.8F)-luc plasmid served as starting DNA material for a digestion with exonuclease III (figure 3.9A). Exonuclease III catalyzes the stepwise removal of 5′-mononucleotides from recessed or blunt 3′-hydroxyl termini of double-stranded DNA (Weiss, 1976). For the creation of unidirectional deletion mutants (Henikoff, 1987), the double stranded DNA is digested with two restriction enzymes. The enzyme that cleaves closer to the target sequences generates a blunt end or a recessed 3′-terminus while the other enzyme generates a 3- or 4-nucleotide protruding 3′-terminus. Digestion with exonuclease III proceeds unidirectionally away from the site of cleavage into the target DNA-sequences. The exposed single strands are removed by digestion with nuclease S1 and by the 3′- $\rightarrow$ 5′-exonuclease activity of the Klenow fragment of *E.coli* DNA polymerase I. The linear DNA is then recirculated and sequenced with a universal primer complementary to a part of the vector backbone.

First, the pLI(2.8F)-luc construct was cleaved with the restriction endonucleases KpnI and PmlI which generated a 5′-blunt end fragment of 7.1 kb that served as target sequence for exonuclease III (figure 3.9B). The two faster-migrating bands in the gel are the KpnI/KpnI-fragment (0.3 kb) and the KpnI/PmlI-fragment (0.2 kb) resulting from the digestion. At 1 min intervals, aliquots of the exonuclease III-digestion mix were removed and analyzed on an agarose gel (figure 3.9C). Only samples containing DNA fragments between 4.7 kb and 7.1 kb were further processed, since shorter fragments lack the transcription start site and the 5′-flanking sequence. The 1 and 2 min samples contained DNA-fragments of various sizes, and the DNA in these samples was re-ligated with T4 ligase (figure 3.9D). *E.coli* DH5α cells were subsequently transformed with the ligation products. Plasmid DNA prepared from colonies of the two different transformations was cleaved with EcoRI and NotI and analyzed by agarose gel electrophoresis (figure 3.9E). A variety of clones from the 2 min sample is shown; the 4.8 kb fragment present in all lanes is the vector backbone, the other fragments are various 5′-deletions of the LI-cadherin regulatory region.

Figure 3.9: Generation of deletion mutants. **(A)** The vector pLI(2.8F)-luc and the appropriate restriction sites. **(B)** Left: Schematic view of the 7.1 kb blunted 5′-terminus target sequence for Exonuclease III. The black part represents the LI-cadherin 5′-flanking region. Right: An ethidium bromide-stained agarose gel showing the products of the KpnI- and PmlI-digestion of pLI(2.8F)-luc; filled arrow represents the 7.1 kb target fragment, open arrows represent the two smaller fragments. **(C)** Left: Schematic view of unidirectional deletions from the 5′-termini. Right: Ethidium bromide-stained agarose gel showing the 7.1 kb KpnI/PmlI-DNA fragment digested with Exonuclease III. 0′ = undigested DNA, 1-13 = 1-13 min digestion with Exonuclease III. **(D)** Religation of DNA with T4 ligase. **(E)** Ethidium bromide-stained agarose gel showing the products of the digestion of re-ligated deletion mutants with NotI and EcoRI. The 4.8 kb fragment present in all lanes corresponds to the promoter-less vector. The smaller fragment in each lane represents various 5′-deletions of the LI-cadherin promoter.


A Starting vector
KpnI
PmlI
2.8 kb LI-cadherin promoter fragment
EcoRI
KpnI
NotI
pLI(2.8F)-luc
7.6 kb
Firefly luciferase
B PmlI- and KpnI-digestion
PmlI
NotI
KpnI
Promoter
7.1 kb
KpnI
KpnI
0.3 kb
KpnI
PmlI
0.2 kb
7.1
0.3
0.2
C Exonuclease III-digestion
kb
7.1
4.7
0 1 2 3 4 5 6 7 8 9 10 11 12 13
D Blunt-end religation
NotI
EcoRI
pLI(1200)-luc
pLI(643)-luc
pLI(81)-luc
E NotI- and EcoRI-digestion
NotI
EcoRI
kb
4.8
2.4
0.3

Clones containing 0.1 - 2.4 kb fragments were chosen for sequencing using the primer pGL3-basicFor (5′-CCA GTG CAA GTG CAG GTG CC-3′) which binds to the vector backbone upstream (5′) of the KpnI digestion site. The sequencing analysis showed that the vector was correctly re-ligated and that the truncated promoter region had the expected length. Apart from the deletion, the various clones did not show any mutations. The resulting deletion mutants are shown in table 3.2. The deletion mutants are named according to the number of bp the LI-cadherin 5′-flanking region contains upstream of the transcription start site (indicated as starting position in table 3.2).

Table 3.2: Deletion mutants of the 5′-flanking region of the LI-cadherin gene.

Deletion mutant 1-12	**Starting position (relative to the transcription start site)**	**Deletion mutant 13-24**	**Starting position, nucleotide**
pLI(2269F)-luc	–2269	pLI(412)-luc	–412
pLI(1972F)-luc	–1972	pLI(380)-luc	–380
pLI(1664)-luc	–1664	pLI(275)-luc	–275
pLI(1477)-luc	–1477	pLI(259)-luc	–259
pLI(1200)-luc	–1200	pLI(245)-luc	–245
pLI(1128)-luc	–1128	pLI(197)-luc	–197
pLI(1027)-luc	–1027	pLI(184)-luc	–184
pLI(898)-luc	–898	pLI(161)-luc	–161
pLI(834)-luc	–834	pLI(106)-luc	–106
pLI(643)-luc	–643	pLI(91)-luc	–91
pLI(617)-luc	–617	pLI(81)-luc	–81
pLI(528)-luc	–528	pLI(68)-luc	–68

In addition, three specific truncations were created by PCR: Gene-specific forward primers (mLIPro55F, mLIPro33F, and mLIPro22F) were used to amplify 55, 33, and 22 bp upstream of the transcriptional start site. The reverse primer was the same as the one used for the amplification of the 2.8 kb fragment; the downstream region remained unchanged in all reporter gene constructs. The PCR products were analyzed by agarose gel electrophoresis (figure 3.10) and subsequently ligated into the TA-vector pCR2.1. After DNA preparation and EcoRI-digestion of the plasmids, clones with inserts in the correct orientation were verified by sequencing.

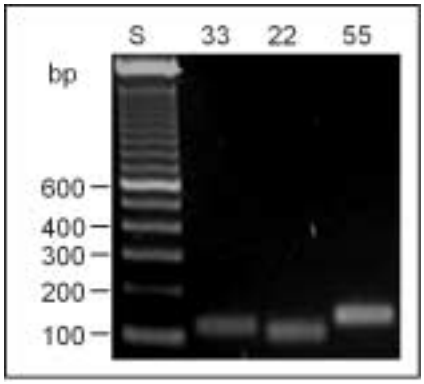


Figure 3.10: Generation of three promoter fragments by PCR. Ethidium bromide-stained agarose gel showing the amplified 33 bp, 22 bp, and 55 bp LI-cadherin promoter fragments. S = 100 bp DNA ladder.

Functional analysis of the deletion mutants

The effect of the different truncations of the LI-cadherin 5′-flanking region was examined with the dual-luciferase reporter assay system. STC-1 cell were transfected with the different deletion mutants using DEAE-dextran and luciferase activity was measured after 48 hours. The STC-1 cells were also transfected with pGL3-basic and pGL3-control as negative and positive controls. As shown in figure 3.11, truncation of the 5′-flanking region from –2774 bp to –55 bp did not significantly decrease luciferase activity. However, further truncation to nucleotide –33 resulted in 80% reduction in luciferase activity, indicating a crucial function of this region.

Interestingly, some of the reporter gene constructs were able to increase the luciferase activity significantly in comparison with the basic activity of the 2.8 kb promoter. This effect can be due to positive and negative regulatory elements that control the transcriptional activity. The most noticeable regions that possess potential regulatory properties are summarized in table 3.3. Positive regulatory elements controlling the activity within the 5′-flanking region of the LI-cadherin gene in transfected STC-1 cells were likely to be located within the region from –1200 to –1128, from –617 to –528, and from –161 to –81. Regulatory elements repressing the transcription were detected between –834 to –643, –259 to –184, and –81 to –55. A further analysis of the regulatory region concentrated on the proximal region from –55 to –33, since this fragment seemed to contain all elements necessary for stronger basic promoter activity. In addition, binding sites for several known transcription factors were found within the first 100 bp upstream of the transcription start site.

Table 3.3: Regions that possess potential regulatory properties.

Location	Regulatory element
–1200 to –1128	+
–834 to –643	–
–617 to –528	+
–259 to –184	–
–161 to –81	+
–81 to –55	–
–55 to +1	+++

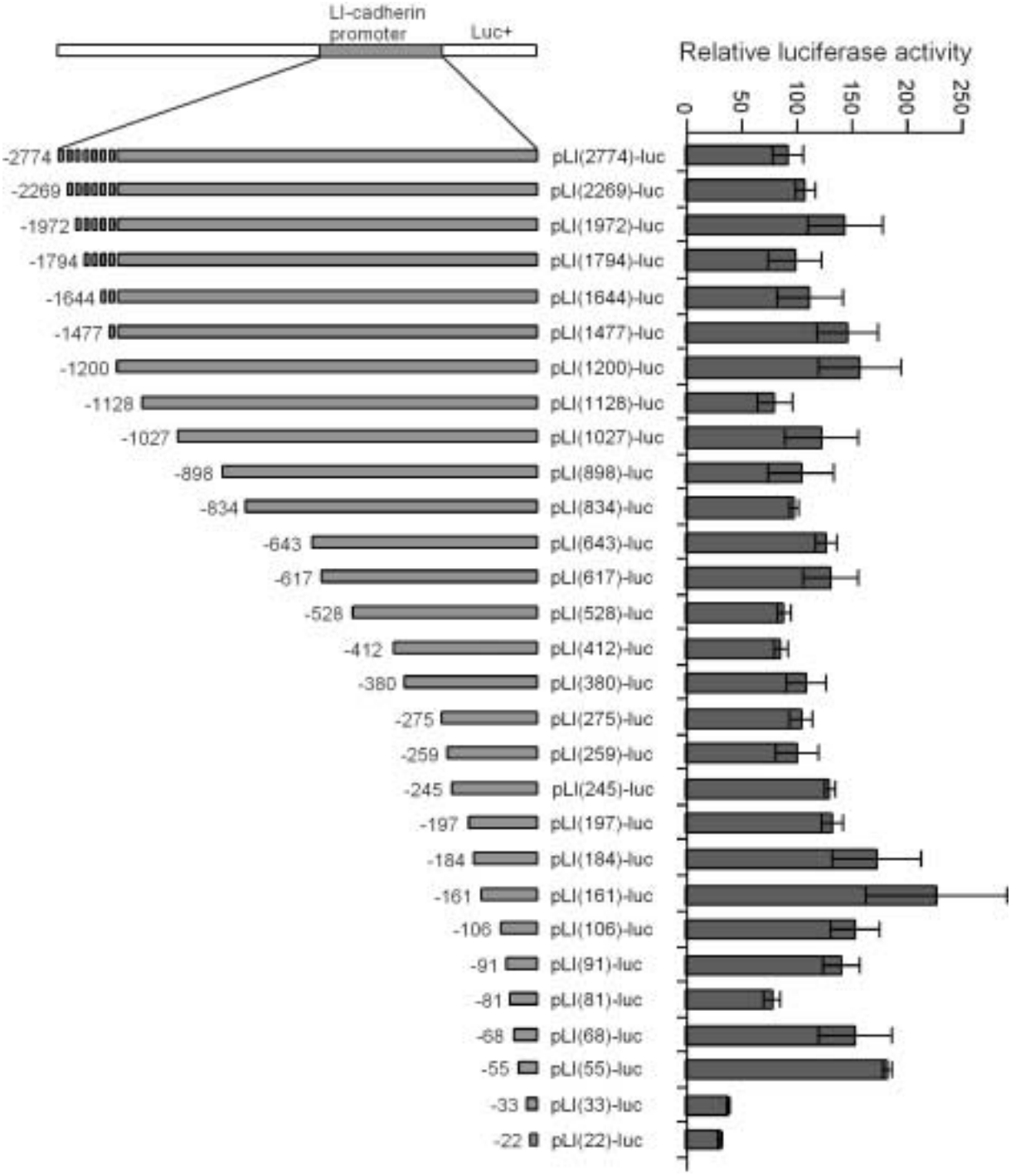


Figure 3.11: Deletion analysis of the proximal 5′-flanking region of the LI-cadherin gene using the dual-luciferase reporter assay system. Reporter gene plasmids ($1x10^{11}$ molecules) containing the indicated lengths of the 5′-flanking region were transiently transfected into STC-1 cells. Luciferase activity was measured after 48 hours. The ratios between firefly luciferase activity and *Renilla* luciferase activity are shown relative to the control plasmid pGL3-control. Data are means of at least 3 separate transfections (the mean relative activities +/-SE). All of the reporter gene plasmids had an activity significantly greater than pGL3-basic ($P<0.05$), except for pLI(33)-luc and pLI(22)-luc which had considerably less activity.

3.4 Binding of nuclear proteins to the LI-cadherin promoter

The computer analysis of the murine and human 5′-flanking region of the LI-cadherin gene regarding potential transcription factor binding sites showed several conserved consensus sequences within the first 100 bp upstream of the transcription start site. The conserved binding sites have now been tested for their ability to bind specific factors in intestinal cells. Electrophoretic mobility shift assays (EMSA) were performed in order to determine whether intestinal cells contain nuclear proteins that bind to the murine LI-cadherin. Double-stranded 30 bp oligonucleotides derived from the region between –100 and –21 were end-labeled with ^{32}P, incubated with nuclear extract, and the DNA-protein complexes were separated from unbound oligonucleotide by non-denaturing gel electrophoresis after an incubation at 4°C for 30 min. All binding reactions were performed in the presence of 2 µg poly(dI-dC) to inhibit non-specific binding. In some reactions, a 100-fold molar excess of unlabeled oligonucleotide was included as a specific competitor. Super shift analysis was also performed with antibodies directed against some of the potential binding candidates. Nuclear proteins isolated from cultured STC-1 cells or mouse intestinal enterocytes were used for all binding assays.

Figure 3.12A shows the binding of nuclear extracts from cultured STC-1 cells and from intestinal enterocytes to a 30-mer oligonucleotide containing the nucleotides –50 to –21 of the 5′-flanking region. In lane 2, the arrows indicate three complexes (1, 2 and 3) that were found with nuclear extracts from the STC-1 cells but were absent in reactions containing no nuclear extract (lane 1). Binding was specific, since a 100-fold molar excess of unlabeled oligonucleotide inhibited the formation of the DNA-protein complexes (lane 3). Lane 5 shows that similar complexes of comparable mobility were also detected in nuclear extracts from enterocytes, and that formation of these complexes was also specific (lane 6). Two additionally retarded bands were present in lane 5 that migrated somewhat faster than complex 1 and 2 (lane 5; arrows). The complex formations with both nuclear extracts were not competed by a 100-fold excess of unlabeled oligonucleotide (ex) containing a completely different sequence (30 nucleotides from exon 3 of the LI-cadherin gene) (lane 4 and 7), indicating that the interaction is sequence specific.

The region between nucleotide –50 and –21 contains two conserved DNA-binding motifs: An AP-1 binding site and a GC-box known to potentially bind SP-factors. In order to examine whether one of these proteins is binding to the LI-cadherin promoter, nuclear extract from STC-1 cells was pre-incubated with anti Sp1 and anti AP-1 (*c-Jun*) antibodies in two reactions. The respective EMSA with the Pro-50/-21 probe is shown in figure 3.12B. Complex 1 (lane 2) was partially supershifted by anti-Sp1 antibody but no changes in complexes 2 and 3 were observed (lane 6). None of the complexes supershifted after addition of anti-AP-1 (lane 5).

Figure 3.12C (lanes 1-9) shows the binding of nuclear extracts from cultured STC-1 cells to a 30-mer oligonucleotide containing the nucleotides –70 to –41 of the promoter. This region of the murine LI-cadherin gene contains binding sites for the intestine specific homeodomain protein Cdx2 and hepatocyte nuclear factor HNF-1. Several retarded bands were found (lane 2) that were absent in a reaction containing no nuclear extract (lane 1). The formation of all the complexes were inhibited by a 100-fold excess of unlabeled wild-type oligonucleotide

(lane 3). Since the complex formations were not competed by a 100-fold excess of unlabeled control oligonucleotide containing 30 nucleotides from exon 3 of the LI-cadherin gene (lane 4), the interactions must be specific. Complexes 1 and 2 were absent when the nuclear extract was incubated with a 100-fold molar excess of an unlabeled oligonucleotide containing the HNF-1 consensus sequence (lane 5). When an antibody directed against HNF-1 was included in the reaction, the intensity of complexes 1 and 2 corresponding to the putative HNF-1-DNA complexes was markedly diminished and a proportion of complex 1 was super-shifted when the antibody was diluted 1:100 (lane 7). Addition of undiluted HNF-1 antibody resulted in complete absence of complex 1 and a supershift (lane 6). 100-fold molar excess of an unlabeled Cdx2 consensus sequence (Cdx2-a) did not compete for complex formations (lane 8), indicating that Cdx2 does not interact with the region from –70 to –41. This was supported by the absence of a supershift after addition of an antibody directed against Cdx2 (lane 9).

A 30-mer oligonucleotide containing the nucleotides –100 to –71 of the promoter was further analyzed by EMSA for its capacity of binding proteins present in nuclear extracts from STC-1 cells (figure 3.12D). This region contains a second binding site for Cdx2 (Cdx2-b), the activator protein AP-1, and a CCAAT-element. Two complexes (1 and 2) were formed (lane 2) that were absent in the reaction containing no nuclear extract (lane 1). The formation of complex 1 was specific, since the binding was inhibited by a 100-fold excess of unlabeled wild-type oligonucleotide (lane 3). The complex formations were not competed by a 100-fold excess of unlabeled oligonucleotide containing 30 nucleotides from exon 3 of the LI-cadherin gene (lane 4), indicating that the interaction is sequence specific. Pre-incubation of STC-1 nuclear extract with anti Cdx2 and anti AP-1 (*c-Jun*) antibodies in two reactions resulted in a supershift (SS) only with the Cdx2 antibody (lane 5 and 6). The intensity of the lower part of complex 1 seemed to decrease, indicating that this complex may consist of two or more similarly migrating complexes. This was supported by competition of complex formation with a 100-fold excess of unlabeled Cdx2 consensus sequence (lane 8); the lower part of complex 1 disappeared whereas the upper part was not competed by the consensus sequence. Anti HNF-1 antibody was added as a control for specificity of the Cdx2 and AP-1 antibody; none of the complexes supershifted after addition of anti-HNF-1 (lane 7).

The transcription factor Sp1 has been described to be important for the regulation of the expression of E-cadherin and Ksp-cadherin. Sp1 is binding to a GC-box present in the promoter region of each gene. LI-cadherin and Ksp-cadherin are both co-expressed with E-cadherin, and it was thus of interest to see if the same factor is binding to all three cadherin promoters. Mobility shift analysis was performed with the GC-rich region of the E-cadherin and Ksp-cadherin promoter under the same conditions used with the LI-cadherin –50 to –21 region. The results obtained by the use of the E-cadherin GC-region is presented in figure 3.13. The retardation pattern of the E-cadherin GC-rich region achieved with the STC-1 nuclear extract was the same as achieved with the LI-cadherin –50 to –21 region (lane 2); three complexes were formed that were comparable with the LI-cadherin sequence, although complex 1 was weaker. The Ksp-cadherin GC-region was involved in a similar complex formation (not shown). 100-fold excess of the unlabeled GC-region of the E-cadherin and LI-cadherin gene were able to compete with the labeled probe, demonstrating a specific DNA-protein complex formation (lane 3 and 6).

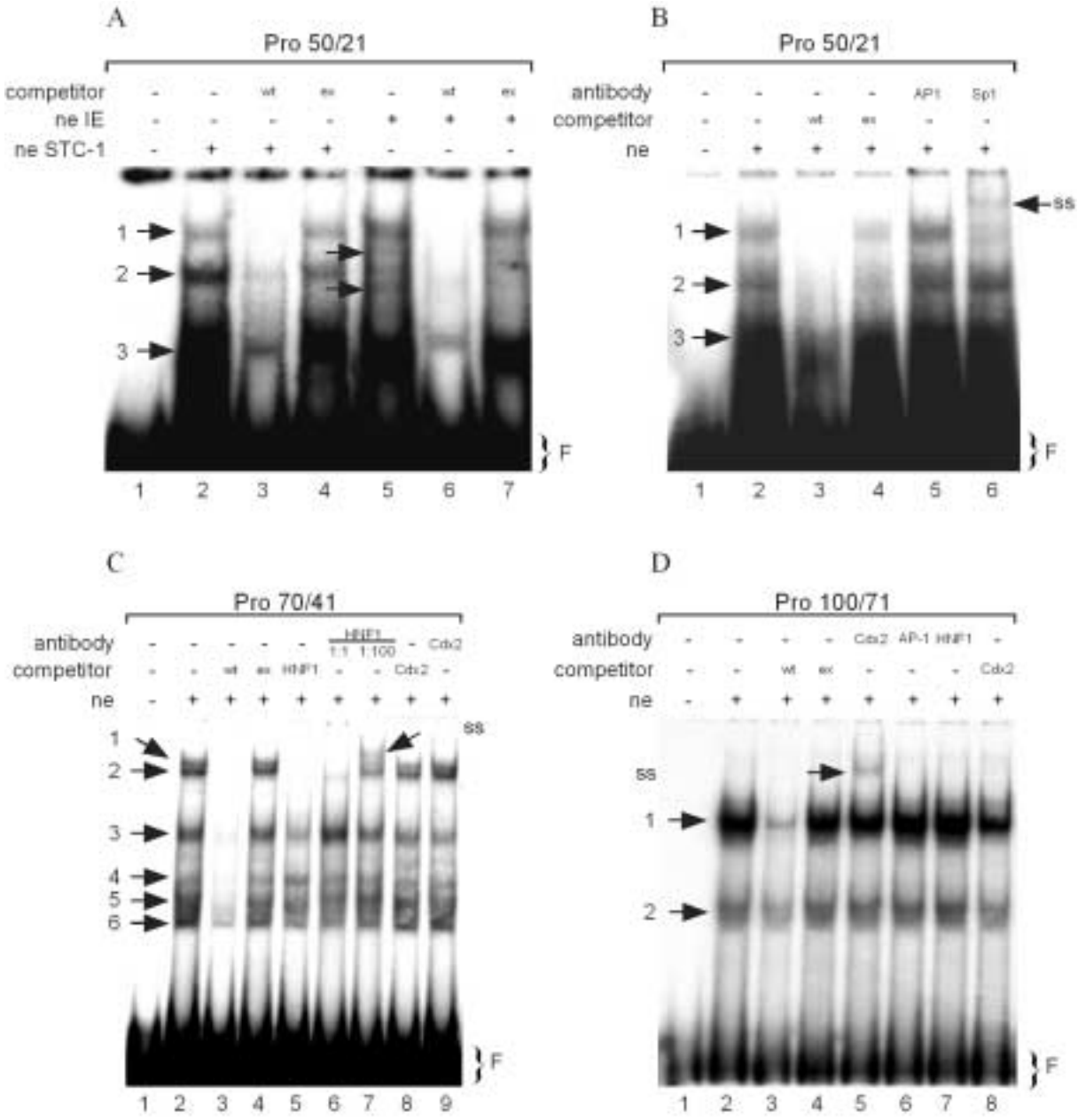


Figure 3.12: Binding of nuclear proteins to the 5′-flanking region of the murine LI-cadherin gene. Electrophoretic mobility shift assays (EMSA) were performed using radioactively labeled oligonucleotides containing the nucleotides –50 to –21 **(A and B)**, -70 to –41 **(C)**, and –100 to –71 **(D)**. "-" indicates no addition of nuclear extract (ne), competitor (unlabeled oligonucleotide), or antibody. "+" indicates addition of nuclear extract, competitor (100-fold molar excess of unlabeled oligonucleotide), or antibody. F = free probe.

The complex formation was not competed by a 100-fold excess of unlabeled oligonucleotide containing 30 nucleotides from exon 3 of the LI-cadherin gene (lane 4). The involvement of a Sp1 factor in the complex formation of the GC-rich region of both the E- and Ksp-cadherin promoter was confirmed by supershift analysis (Ksp-cadherin promoter is not shown here). The anti-Sp1 antibody supershifted complex 1 generated by the GC-rich region of the E-cadherin promoter (lane 5). The supershifted band was associated with a similar decrease in the faster-mobility complex 1 of the GC-rich region. The same results were obtained with the Ksp-cadherin promoter (not shown).

Taken together, these results demonstrated that STC-1 cells and enterocytes contained nuclear proteins that specifically bind to the regions of the promoter from –100 to –71, -70 to –41, and –50 to –21. Supershift analysis indicated that Sp1 is one of the factors binding to the GC-rich region in the LI-cadherin promoter, as well as to the GC-rich region in the E- and Ksp-cadherin promoter. HNF-1 is one of the factors binding to the –70 to –41 region whereas Cdx2 is binding to the consensus binding sites (Cdx2-b) within the –100 to –71 region.

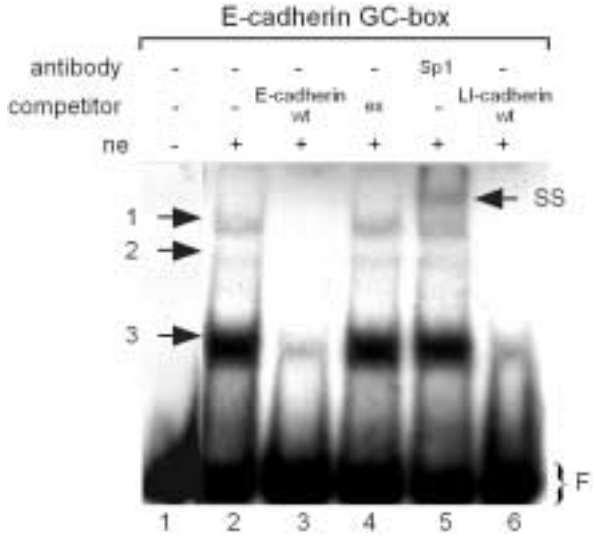


Figure 3.13: Binding of nuclear proteins to the GC-rich region in the E-cadherin promoter. Electrophoretic mobility shift assays (EMSA) were performed using a radioactively labeled oligonucleotide containing the GC-rich region of the E-cadherin promoter (5′-TCAGCGGCGCCGGGGGCGGTGCCTGCGGGC). "-" indicates no addition of nuclear extract, competitor (unlabeled oligonucleotide), or antibody. "+" indicates addition of nuclear extract (ne), competitor (100-fold molar excess of unlabeled oligonucleotide), or antibody. F = free probe.

3.5 Mutational analysis

3.5.1 Effects of mutations on binding of nuclear extract to the LI-cadherin promoter

Because the oligonucleotides used by EMSA contain more than one potential binding site, every one of them was separately mutated. Figure 3.14 shows the results of EMSA comparing the abilities of the mutated oligonucleotides to compete with the wild-type sequence for binding to nuclear extract of STC-1 cells. All the mutations are shown in figure 3.14A.

Mutations within the GC-box, the HNF-1 and the AP-1 binding site were sequentially introduced into the –50 to –21 oligonucleotide, and the ability of the mutated oligonucleotides to compete with the wild-type sequence for binding to nuclear proteins was compared. In lane 1 of figure 3.14B, the arrows show that nuclear extracts from the STC-1 cells contain proteins that bind to the region from –50 to –21, similar to the results shown in figure 3.12A. Binding was specific, since the formation of the retarded bands was inhibited by a 100-fold molar excess (lane 2) of unlabeled wild-type oligonucleotide; complex 1 was completely inhibited, while formation of complexes 2 and 3 were only partially inhibited. 100-fold molar excess of unlabeled oligonucleotides containing the mutations AP-1 (lane 4) and HNF-1 (lane 6) were able to inhibit the formation of the complexes similar as unlabeled wild-type oligonucleotide. In contrast, 100-fold molar excess of the oligonucleotide containing a GC-mutation (mGC1) had no effect on formation of complex 1 (lane 5), indicating that the mutated sequence was unable to compete for binding. Therefore, complex 1 seems to contain a GC-binding protein.

The putative binding sites for a Cdx2 (Cdx2-a) and HNF-1 were mutated one at a time in the –70 to –41 oligonucleotide and analyzed by EMSA (figure 3.14C). The results confirmed the super shift analysis; the mutated HNF-1 binding sites could not compete with the wild-type sequence for binding to nuclear extract (lane 4), indicating that the mutated nucleotides are involved in formation of complex 1 and 2. Unlabeled mutated Cdx2-a binding site inhibited complex formation (lane 5), which was consistent with the lack of a supershift after incubation with an antibody against Cdx2 (figure 3.12C, lane 9).

Within the region from –100 to –71, the putative binding sites for Cdx2-b, AP-1, and a CCAAT-binding protein were mutated one at a time in three different oligonucleotides. A 100-fold molar excess of the unlabeled wild-type sequence (lane 3) inhibited the DNA-protein complex formation (lane 2). The competition with mutated CCAAT-box (lane 4) and AP-1 binding site (lane 5) resulted in a similar inhibition. In contrast, 100-fold molar excess of the mutated oligonucleotide containing a mutated Cdx2-b binding site competed only for the upper (a) but not the lower part (b) of complex 1 (lane 7), indicating that this complex likely consists of two complexes as suggested when competing with a consensus Cdx2-b binding site or a Cdx2 antibody (figure 3.12D).

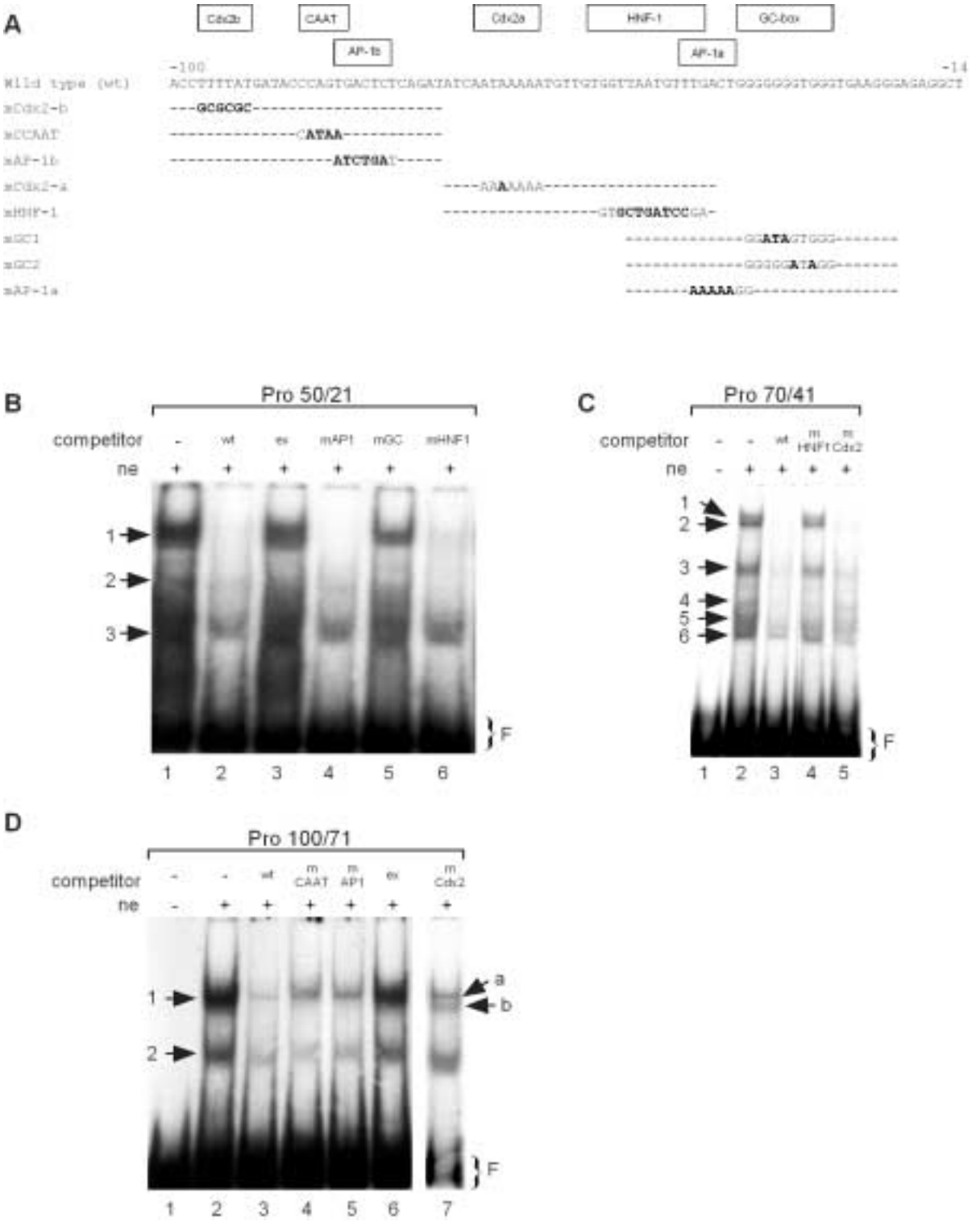


Figure 3.14: Effects of mutations of single binding consensus sequences within the LI-cadherin promoter on protein binding. **(A)** Sequence of the wild-type (wt) proximal region and mutated oligonucleotides. **(B)-(D)** EMSAs were performed using nuclear extracts from STC-1 cells and radioactively labeled wild-type oligonucleotides containing in **(B)** nucleotides –50 to –21, in **(C)** –70 to -41 and in **(D)** nucleotides –100 to –71. Binding was analyzed in absence or presence of a 100-fold molar excess of the respective unlabeled wild-type or mutated oligonucleotides. F = free probe.

Taken together, these results indicate that an exchange of three nucleotides within the GC-box in the region from –50 to –21 interfered with protein binding. This was consistent with the supershift of complex 1 obtained with the anti-Sp1 antibody. A mutation of the AP-1 binding site within the –50 to –21 region did not affect the ability to inhibit any complex formation, as already suggested by lack of a supershifted complex in the presence of an anti-AP-1 antibody.

Within the –70 to –41 region, only a mutation of the HNF-1 binding site resulted in loss of competition; the mutated Cdx2-a oligonucleotide could compete for interaction with nuclear extract. Mutating the Cdx2 binding site (Cdx2-b) within the region from –100 to –71 showed a weaker effect on formation of complex 1. The lower part of complex 1 was not affected by the mutated oligonucleotide, indicating that this sequence is involved in complex formation.

3.5.2 Effects of mutations on LI-cadherin promoter activity

In order to examine whether mutations of the GC-box, the two Cdx2-binding sites, and the HNF-1 binding site would affect the LI-cadherin promoter function, reporter gene constructs based on the 1.8 kb promoter construct were generated by site-directed mutagenesis each carrying different mutations, as indicated in figure 3.14A. In addition, the two mutations of the GC-box were also introduced into the 55 bp LI-cadherin promoter reporter gene construct pLI(55)-luc. Their effects on transcriptional activity were analyzed using the dual-luciferase reporter assay system. Transfection of STC-1 cells with the 1.8 kb wild-type promoter produced a 4.5-fold stimulation of luciferase activity compared to the promoter less control plasmid (figure 3.15A, right). Transfection with a reporter plasmid containing mutation mGC1 almost completely abolished the activity, whereas the mutation mGC2 resulted in a significant 50% decrease in activity compared to the wild type. When mGC1 and mGC2 were introduced into pLI(55)-luc, the luciferase activity decreased only about 25% and 15% respectively compared to the wild type 55 bp promoter fragment (figure 3.15A, left). Mutation of the distal Cdx2 (Cdx2-b) binding site in the 1.8F bp-promoter resulted in a similar reduction of luciferase activity as mutation mGC2 in the long fragment, whereas the more proximal site (Cdx2-a) gave rise to an almost 100% increase of luciferase activity. Mutation of both Cdx2 binding sites in one reporter gene construct reduced the activity similar to the Cdx2-b mutation. The mutated HNF-1 binding site showed the strongest effect; the luciferase activity decreased to less than 15% of the wild-type promoter fragment. The activity was further reduced after mutating both the HNF-1 and Cdx2-b binding site in one reporter gene construct.

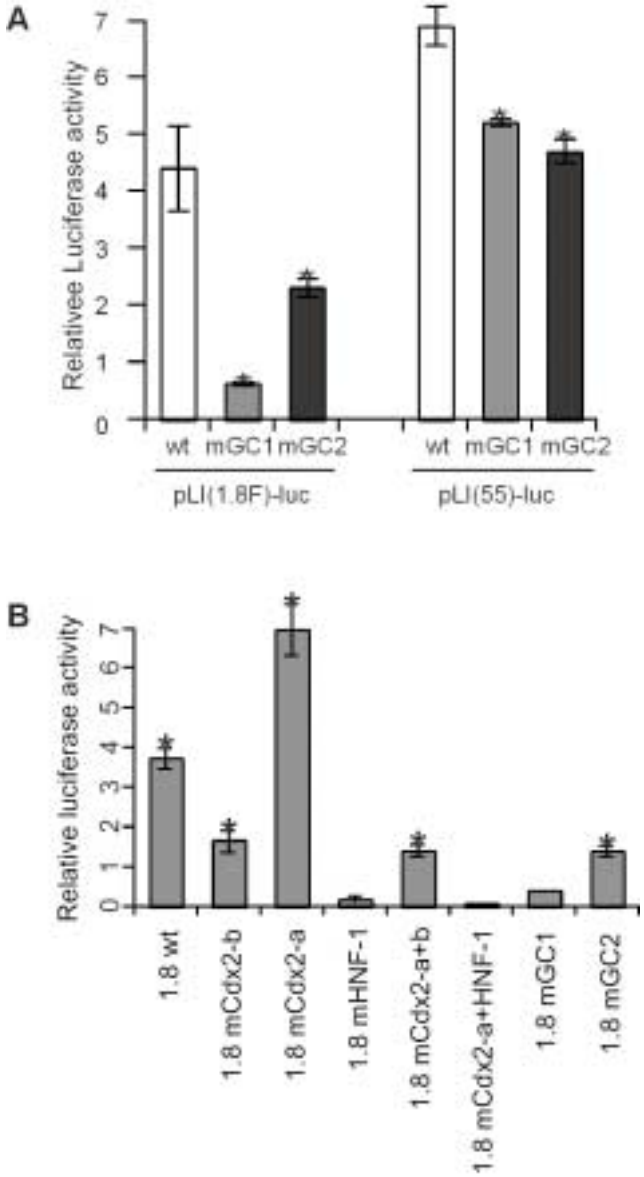


Figure 3.15: (A) Effect of mutation GC1 and GC2 in pLI(55)-luc compared to pLI(1.8F)-luc; reporter gene construct containing either wild-type (wt) or mutated sequences were transfected into STC-1 cells. Luciferase activity was measured after 48 h and normalized to transfection efficiency using the dual-luciferase reporter assay system. Data are means of three separate transfections **(B)** The reporter gene construct pLI(1.8F)-luc ($1x10^{11}$ molecules) containing either wild-type (wt) promoter sequences or mutated sequences (mCdx2-a, mCdx2-b, mHNF-1, mGC1, mGC2) were transfected into STC-1 cells, and luciferase activity was measured after 48 h. Luciferase activity was normalized to transfection efficiency using the dual-luciferase reporter assay system, and is shown relative to pGL3-basic. *Significantly less than wild type ($P< 0.05$).

3.6 Evaluation of tissue- and species-dependent activity of the LI-cadherin promoter

In order to evaluate whether the activity of the LI-cadherin promoter depends on cell specificity and endogenous expression of LI-cadherin, the reporter gene construct pLI(1.8F)-luc was transfected into fourteen different cell lines from a total of eight tissues and four different species (table 3.5). In addition, all cell lines were tested for endogenous expression of LI-cadherin by Western blotting. Human LI-cadherin was detected at different levels in the lysates from the cell lines BON, 23132/87, AGS, A549, and HuH7 (figure 3.16B) whereas LI-cadherin was absent in lysates from the cell lines HT29, HepG2, and 293. Murine LI-cadherin was either detected in the murine cell lines B16F0 or CMT93 (figure 3.16A). Detection of human and murine LI-cadherin in Caco2 and STC-1 cells served as positive controls. The transfection method used was dependent on the cell line (see *material and methods section 2.3.2*).

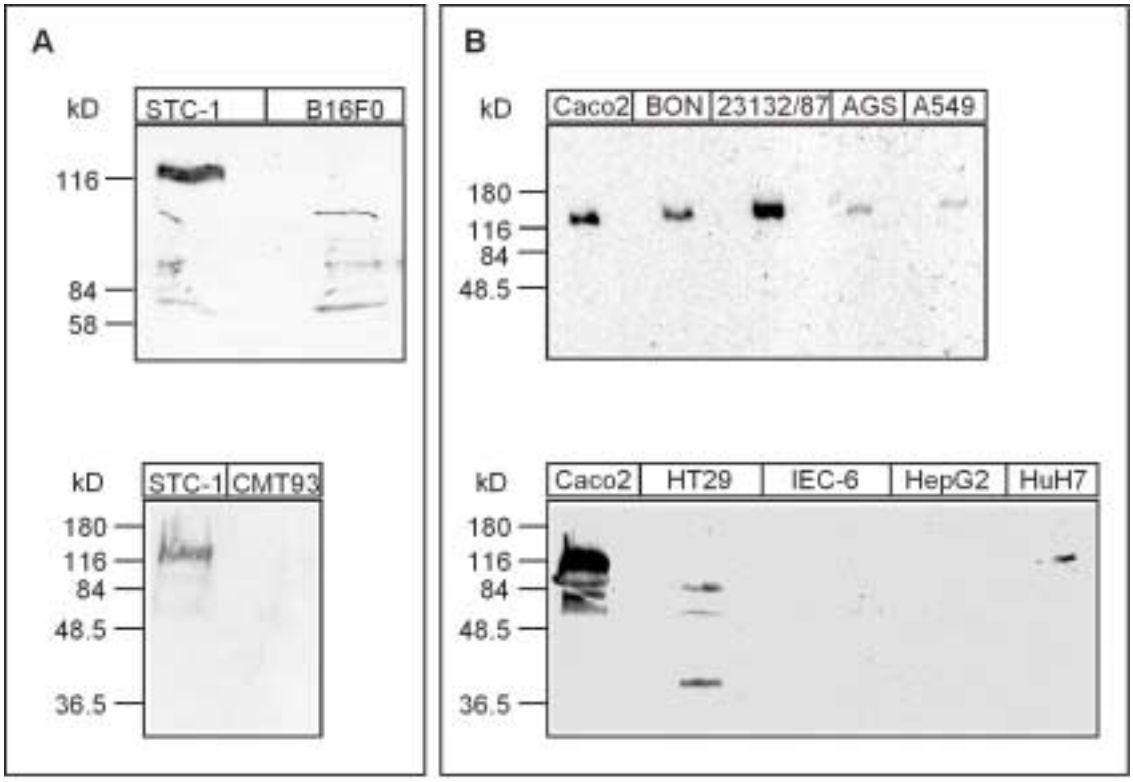


Figure 3.16: Analysis of LI-cadherin expression in cultured cell lines by Western blotting. 10 µg of total protein from each cell lysate were separated per lane. **(A)** Detection of mouse LI-cadherin in cell lysates from STC-1 cells, B16F0 cells and CMT93 cells with anti-mLI-cadherin-pAb. **(B)** Detection of human LI-cadherin in cell lysates from Caco2, BON, 23132/87, AGS, A549, HT29, IEC-6, HepG2, and HuH7 with anti-hLI-cadherin-pAb.

All these cell lines were subsequently cotransfected with pLI(1.8F)-luc and pRL-SV40. Luciferase activity was compared with that induced by the control vector pGL3-basic. Activity of the control vector in each cell line was defined as 1. Transfection of pLI(1.8F)-luc resulted in different stimulation of luciferase activity varying from cell line to cell line (figure 3.17). In transfected STC-1 and Caco2 cells, the pLI(1.8F)-luc vector induced a 6- and 10-fold higher luciferase activity than the control vector, whereas transfection of CMT93 and HuH7 resulted in two-fold or less stimulation. These increases in luciferase activity were statistically significant ($P < 0.001$). In addition, transfection of HepG2, BON, 23132/87, and A549 cells resulted in 10-, 15-, 16-, and 10-fold stimulation respectively. There was no significant stimulation of luciferase activity following transfection in IEC-6, 293, CHO, and B16F0 cells. The results from the Western blot analysis and the luciferase assay are summarized in table 3.5

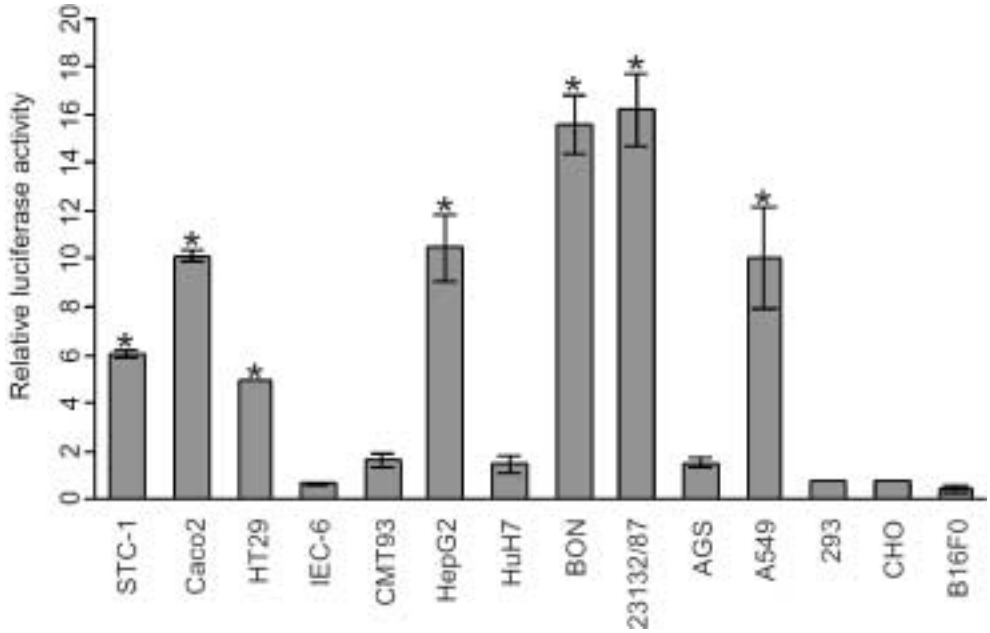


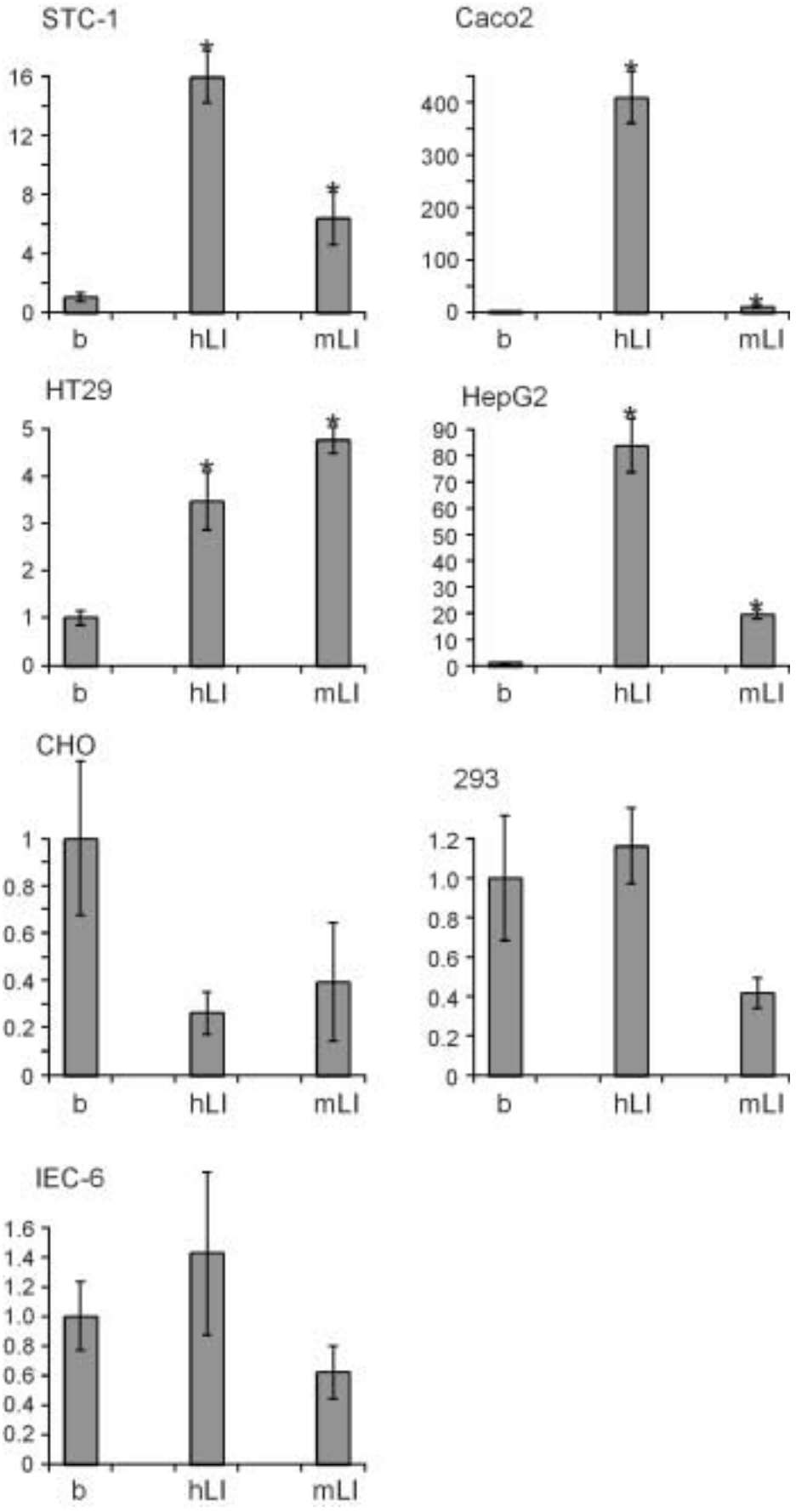
Figure 3.17: Activity of the murine LI-cadherin promoter in different cell lines. 1-5 x 10^5 cells were transfected with $1x10^{11}$ copies of the pLI(1.8F)-luc reporter gene constructs using different methods (*material and methods section 2.3.2*). Luciferase activity was measured after 48 hours and normalized to transfection efficiency using the dual-luciferase reporter assay system. Normalized luciferase activity is shown relative to pGL3-basic. Data are means of at least three separate transfections. *Significantly higher activity than pGL3-basic ($P<0.05$).

Table 3.5: Cell lines transfected with pLI(1.8F)-luc. *Significantly higher expression of the reporter gene than with pGL3-basic (P=0.05).

Cell line	Species	Tissue	Fold activation (pGL3-basic = 1)	Endogenous LI-cadherin
STC-1	Mouse	Intestine	6*	+
Caco2	Human		10*	+
HT29	Human		5*	-
IEC-6	Rat		<1	-
CMT93	Mouse		1.7*	-
HepG2	Human	Liver	10*	-
HuH7	Human		1.5	+
BON	Human	Pancreas	15*	+
23132/87	Human	Stomach	16*	+
AGS	Human		1.5	+
A549	Human	Lung	10*	+
293	Human	Kidney	<1	- (not shown)
CHO	Hamster	Ovary	<1	- (not shown)
B16F0	Mouse	Melanoma	<1	-

Moreover, cells were transfected with the reporter gene construct containing 1.1 kb of the human 5'-flanking region, phLI(1.1)-luc (Gimpel, 2001), in parallel with the murine construct pLI(1128)-luc to analyze whether the promoter activity was species-dependent. Figure 3.18 shows the normalized luciferase activity in each cell line compared to that induced by the control vector pGL3-basic. In STC-1 cells, the luciferase activity was stimulated 16-fold by the human promoter and 6-fold by the murine promoter compared to pGL3-basic. In the human Caco2 and HepG2 cells, transfection resulted in a larger difference in luciferase activity by the two promoters; the activity of luciferase in Caco2 cells was 40-fold higher with the human promoter compared to the murine promoter. In HepG2 cells, the human promoter was more than eight times as active. Transfection of HT29 resulted in 3.5- to 5-fold stimulation with both constructs, whereas transfected CHO, 293, and IEC-6 cells show a luciferase expression below the level of the control experiment, confirming the results above.

Figure 3.18: Comparison of the murine and human LI-cadherin promoter-activity in cultured cells. $1x10^{11}$ copies of the pLI(1128)-luc and phLI(1.1)-luc constructs were transfected into 1-5 x 10^5 cells using different methods (*material and methods section 2.3.2*), and luciferase activity was measured after 48 hours. Normalized luciferase activity is shown relative to pGL3-basic. Data are means of at least three separate transfections. *Significantly greater activity than pGL3-basic (P<0.05). b = pGL3-basic, hLI = phLI(1.1)-luc, and mLI = pLI(1128)-luc.


STC-1
16
12
8
4
0
b
hLI
mLI
Caco2
400
300
200
100
0
b
hLI
mLI
HT29
5
4
3
2
1
0
b
hLI
mLI
HepG2
90
80
70
60
50
40
30
20
10
0
b
hLI
mLI
CHO
1
0.8
0.6
0.4
0.2
0
b
hLI
mLI
293
1.2
1.0
0.8
0.6
0.4
0.2
0
b
hLI
mLI
IEC-6
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0
b
hLI
mLI