Transcriptional regulation of the human *CD97* promoter by Sp1/Sp3 in smooth muscle cells

Key words: EGF-TM7 receptor, GC-rich region, transcription factor

Manja Wobus[1,2*], Elke Wandel[1], Sonja Prohaska[3], Sven Findeiß[4], Katrin Tschöp[5,6], Gabriela Aust[1]

[1] *Department of Surgery, Research Laboratories, University of Leipzig, Leipzig, Germany*

[2] *Medical Clinic and Polyclinic I, University Hospital Carl Gustav Carus, Dresden, Germany*

[3] *Biomedical Informatics, Department of Computer Science and Engineering, Arizona State University, Tempe, AZ, USA*

[4] *Bioinformatics Group, Department of Computer Science, University of Leipzig, Leipzig, Germany*

[5] *Department of Internal Medicine II, University of Leipzig, Leipzig, Germany*

[6] *Department of Otolaryngology, University of Leipzig, Leipzig, Germany*

Abbreviations: EGF-TM7, Epidermal growth factor-like- seven transmembrane; GPCR, G-protein coupled receptor; Sp, specificity protein; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; HbSMC, human bronchial smooth muscle cell; wt, wildtype; mut, mutated; EMSA, electrophoretic mobility shift assay; ChIP, chromatin immunoprecipitation; UTR, untranslated region.

* Corresponding author. Medical Clinic and Polyclinic I, University Hospital Carl Gustav Carus, Fetscherstr. 74, D-1307 Dresden, Germany, Tel.: +49-351-458 4765; Fax: +49-351-458 5362. *E-mail address*: manja.wobus@uniklinikum-dresden.de

Abstract

The EGF-TM7 receptor CD97 shows different features of expression and function in muscle cells compared to hematopoetic and tumor cells. Since the molecular function and regulation of CD97 is poorly understood, this study aimed at defining its basal transcriptional regulation in smooth muscle cells (SMCs).

The computational analysis of the *CD97* 5'-flanking region revealed that the TATA box-lacking promoter possesses several GC-rich regions as putative Sp1/Sp3 binding sites. Transfection studies with serially deleted promoter constructs demonstrated that the minimal promoter fragment resided in the –218/+45 region containing one out of five identified GC-boxes in the leiomyosarcoma cell line SK-LMS-1 and human bronchial smooth muscle cells (HbSMCs). Mutation of the most proximal GC-site in *CD97* reporter gene constructs caused a significant decrease in promoter activity. Gel shift assays and chromatin immunoprecipitation revealed that Sp1 and Sp3 bound specifically to the most proximal GC-site. Furthermore, we showed that Sp1 and Sp3 over-expression activates *CD97* promoter activity in HEK 293 cells. Our data characterize for the first time the activity of the human *CD97* promoter which is controlled by Sp1/Sp3 transcription factors in SMCs.

1. Introduction

CD97 belongs to a family of G-protein coupled receptors (GPCRs) with an unusual large N-terminal extracellular domain (LNB-TM7) (Kwakkenbos et al., 2004; Stacey et al., 2000) and is the founding member of a small subfamily, the EGF-seven-transmembrane (EGF-TM7) receptors (Eichler et al., 1994; Hamann et al., 1995). It consists of a variable number of N-terminally located EGF-domains, a long extracellular stalk, the TM7 and a short intracellular region. Alternative splicing results in three isoforms containing three, four or five EGF-domains.

CD97 shows interesting features in the expression pattern in different cell types. We described CD97 expression in cells of the hematopoetic system and detected CD97 over-expression in tumor cells of different origin (Aust et al., 2002; Steinert et al., 2002). Moreover, CD97 can be proven in smooth muscle cells of the lung, gastrointestinal tract, urinary bladder and uterus (Aust et al., 2006). Surprisingly, vessels are heterogeneous: venous smooth muscle cells strongly express CD97, whereas elastic arteries are nearly CD97 negative. Furthermore, CD97 heterogeneity was observed during smooth muscle cell transformation. All leiomyomas and 9 out of 21 leiomyosarcomas were CD97 positive (Aust et al., 2006). The disability of muscular CD97 to bind the ligand CD55 implicates another function in this cell type (Wobus et al., 2004) as the suggested role as an adhesion molecule (Hamann et al., 1996) in tumor cells where CD97 promotes migration capacity (Galle et al., 2006; Steinert et al. 2002). Based on these findings, we are eminently interested in CD97 regulatory mechanisms in SMCs.

At present, no information regarding the basal or tissue-specific transcriptional regulation of any human EGF-TM7 receptor is available. Detailed sequence analysis of the *CD97* 5'-flanking region revealed the presence of several GC-rich boxes, putative binding sites for the Sp family of transcription factors. Mithramycin interferes with the binding of Sp1 transcription factors to GC-rich promoter regions in-vitro (Blume et al., 1991). We observed

decreased *CD97* promoter activities as well as mRNA expression levels after mithramycin treatment of SMC cultures (unpublished data) which might be an indication for a *CD97* gene regulation by Sp transcription factors. The Sp transcription factor family belongs to the conserved zinc finger DNA-binding domain proteins that recognize the DNA-binding motifs GC-box (GGGCGGG) and GT-box (GGTGTGGGG) (Suske, 1999). They are important for the expression of many different housekeeping genes as well as tissue-specific genes (Bouwman and Philipsen, 2002) that generally do not contain TATA- or CAAT-boxes in their proximal promoters. Several Sp proteins (Sp1-Sp8) have been identified (Bouwman and Philipsen, 2002; Suske et al., 2005). Whereas most Sp proteins show tissue-restricted expression patterns, Sp1 and Sp3 are ubiquitously expressed and compete for common target sequences (Shin et al., 2005). Sp1 is well known as a transcriptional activator, whereas Sp3 can be either a transcriptional activator or repressor of Sp1-mediated transcription, depending on the promoter context and cell type (Majello et al., 1997).

Here, we provide evidence that one of the identified GC-rich sites within the *CD97* promoter is functional active by binding Sp1 as well as Sp3 in SMCs. This is the first study providing comprehensive information on the structure and features of the human *CD97* promoter as well as a general molecular mechanisms by which the human *CD97* gene expression is regulated.

2. Materials and Methods

2.1. Construction of plasmids for promoter analysis

Progressive deletion constructs of the human *CD97* promoter region were generated from a 3116 bp PCR fragment of the human *CD97* 5'-flanking region (Wobus et al., 2006) and cloned into the *Kpn*I and *Nhe*I site of the promoterless firefly luciferase reporter vector pGL-3 basic (Promega GmbH, Mannheim, Germany). All nucleotide numberings are relative to the translation initiation codon (A=0). Primers used for amplification are listed in table 1. For amplification, we used high fidelity PCR enzyme mix (Fermentas GmbH, St. Leon-Rot,

Germany) and PCR conditions for long fragment amplification. Each construct was sequenced from both ends to confirm the correct fidelity. The expression vectors pN3/Sp1 and pN3/Sp3 were kindly provided by G. Suske (Philipps University, Marburg, Germany).

2.2. 5' RACE

To determine the transcription start of the *CD97* gene, 5' RACE (Roche Diagnostics GmbH, Mannheim, Germany) was used according to the instruction manual. Briefly, mRNA from SK-LMS-1 cells was transcribed into cDNA with a *CD97* specific revers primer HCD97r34 (Tab. 1). After adding a 3' poly (A) tail, amplification was carried out using an Oligo dt-anchor primer and a nested *CD97* specific primer HCD97r38 (Tab. 1). The PCR product was cloned into pGEM-T vector (Promega) and subsequently sequenced.

2.3. Computational identification of transcription factor binding sites

Computational binding site detection was performed to select a likely candidate for *CD97* regulation. In an initial pattern matching, using patch and TRANSFAC 6.0, several GC-rich boxes were identified. To strengthen the prediction we applied alternative methods and used alternative data resources. Weight matrix matching was performed using the match web interface at www.biobase.de and our implementation of the same algorithm (Kel et al., 2003) called pwmatch. Sp1 and Sp3 binding profiles were taken from TRANSFAC 6.0 and JASPAR 2005. The high scoring hits selected for site mutation studies are given in table 2.

2.4. Site-directed mutagenesis

Site-directed mutagenesis of potential Sp1/Sp3 binding sites in the *CD97* promoter region (Tab. 1, 2) was carried out using the QuickChange site-directed mutagenesis kit (Stratagene Europe, Amsterdam, The Netherlands).

2.5. Cell culture, transient transfection and reporter gene assay

The leiomyosarcoma cell line SK-LMS-1 and the embryonic kidney cell line HEK293 were obtained from American Type Culture Collection (ATCC, Rockville, MD). Human bronchial smooth muscle cells (HbSMC) were purchased from PromoCell GmbH (Heidelberg, Germany).

Transfection experiments were performed in 24-well plates using Lipofectamine 2000 (Invitrogen, GmbH, Karlsruhe, Germany). 200 ng of the various *CD97*-pGL-3 constructs were co-transfected with 10 ng of the *Renilla* reference plasmid pRL-null (Promega) to normalize the variations in transfection efficiency. Each plasmid was transfected in duplicate in at least four independent experiments. Fresh media was replaced 24 h later.

Cells were lysed 48 h post-transfection in passive lysis buffer (Promega). The luciferase activity was determined using 40 µl of cell extract with the dual luciferase assay system (Promega). The luciferase activity was normalized to the *Renilla* activity and expressed relative to the activity of the empty pGL-3 plasmid.

To determine the significance of data differences, an unpaired *t* test was used and a *p* value of less than 0.05 is considered to be significant.

2.6. Electrophoretic Mobility Shift Assay (EMSA)

Oligonucleotide probes of the five predicted GC-boxes (Tab. 1) were radiolabeled using T4 polynucleotide kinase in the presence of [γ-^{32}P]ATP and purified using the Nucleotide Removal Kit (Qiagen GmbH, Hilden, Germany). Nuclear extracts from SMCs were prepared according to the method of Schreiber et al. (1989). For the binding reaction nuclear extracts with ~5 µg protein were incubated for 15 min at room temperature in 10 mM Tris pH 7.5, 0.5 mM EDTA, 1 mM $MgCl_2$, 50 mM NaCl, 0.5 mM DTT, 6% glycerol and 0.75 µg of polydeoxyinosinic-deoxycitidylic acid (poly dI-dC; Sigma-Aldrich Chemie GmbH, Taufkirchen, Germany) in a total volume of 10 µl. For supershifts, Sp1 and Sp3 specific

antibodies (sc-420, sc-644, Santa Cruz Biotechnology, Heidelberg, Germany) were incubated with protein mixtures prior to probe addition. Specificity of binding was controlled by the addition of 50 fold excess of cold competitor DNA to either wt or mutant Sp1/Sp3 sites of the *CD97* promoter (Tab. 1). After probe addition, reactions were incubated for 20 min at room temperature and than separated on a 5% native polyacrylamide gel at room temperature. Gels were visualized using a FLA-3000 Fujifilm phosphorimaging system (Fuji, Düsseldorf, Germany).

2.7. Chromatin immunoprecipitation assay (ChIP)

ChIP assays were performed using a kit from Upstate (Lake Placid, NY, USA). All buffers contained 1 mM PMSF as well as Complete Protease Inhibitor Cocktail (Roche). $1x10^6$ SMCs were fixed in a final concentration of 1% formaldehyde for 10 min at 37°C. Chromatin was sheared by sonication (Bandelin Sonoplus, Berlin, Germany) four times for 10 s, yielding DNA fragments of 200-800 bp. After preclearing with DNA- and albumin-blocked Protein A-agarose for 30 min at 4°C, the samples were incubated overnight with 2 µg Sp1, Sp3 or control IgG antibody (DakoCytomation GmbH**,** Hamburg, Germany) with rotation at 4°C and precipitated with salmon sperm DNA-Protein A agarose for 1 h at 4°C. To reverse cross-links, chromatin complexes including input samples were incubated in 5 M NaCl at 65°C for 4 h, resuspended in Tris-EDTA proteinase K buffer for 1 h at 45°C and purified using a PCR purification kit (Qiagen). PCR amplifications were conducted using the *CD97* promoter-specific primers as indicated in table 1. Samples from at least three independent immunoprecipitations were analyzed.

3. Results

3.1. Cloning and characterization of the human CD97 promoter

A comprehensive computational analysis of the human *CD97* 5'-flanking region from -3071 to +45 relative to the *CD97* translation start sequence did not show putative TATA or CCAAT boxes but revealed the presence of five GC-rich boxes, potential binding sites for Sp1 and Sp3 transcription factors (Fig. 1A; Tab. 2).

To precisely determine the transcription start site in the flanking region of the *CD97* gene, 5' RACE analysis was carried out using total RNA from SK-LMS-1 cells. Sequencing of the RACE products revealed that transcripts are initiated from a site 78 nt upstream of the ATG translation initiation codon.

A 5'-flanking region fragment spanning nucleotides from – 3071 to +45 relative to the *CD97* translation start was cloned recently and displayed promoter activity in SW480 and DLD-1 colorectal tumor cells (Wobus et al., 2006). To identify the sequence that is necessary for basal transcription of *CD97* in SMCs, we examined the ability of various 5' deletions of that construct to drive luciferase expression in the promoterless pGL-3 plasmid (Fig. 1B). Each construct drove luciferase transcription in HbSMCs and SK-LMS-1 cells. The pGL-3 empty plasmid served as a negative control. The longest promoter construct (–3071/+45) containing all five identified potential Sp1/Sp3 sites resulted in a moderate luciferase expression (Fig. 1B). The -2101/+45 construct showed a higher promoter activity, but further deletion of the 5' sequence to -1801/+45 and –1151/+45 resulted in decreased activities. Interestingly, the sequence –218/+45 carrying one Sp1/Sp3 site showed the highest level of expression, whereas the activity is almost completely abrogated at -111/+45 lacking GC-sequences (Fig. 1B). The ratio between the various 5' deletion constructs was similar in both cell lines with higher promoter activities in the SK-LMS-1 leiomyosarcoma cells.

Figure 1

3.2. Functional characterization of Sp1/Sp3 for the CD97 promoter activity

To investigate the role of the GC-rich sites in the *CD97* promoter in detail, site-specific mutagenesis for each site was performed. The luciferase activities in the mutated reporter constructs were compared with corresponding wildtype constructs which was the -2101/+45 construct for the sites 1-4 and the –3071/+45 construct which contains also the site 5 located at –2317, respectively (Fig. 1C). Mutation of the most proximal GC-site 1 at -122 caused an almost complete abrogation of promoter activity in both cell lines. Mutation of the other potential Sp1/Sp3 sites did not cause significant changes in the promoter activities (Fig. 1C).

Figure 1

3.3. Sp1/Sp3 transcription factors bind to the CD97 promoter in-vitro and in-vivo

To test specific binding of Sp1 and Sp3 to the identified GC-rich boxes in-vitro, we performed EMSAs using end-labeled double-stranded probes covering the potential binding sites within the *CD97* promoter as indicated in Tab. 1. Except for GC-site 4, shifted DNA-protein complexes were observed after incubation with nuclear extract from SMCs similar to those obtained with a GC-consensus-oligo as representatively shown for the GC-site 1 (Fig. 2A). The binding was specific as shown by inhibition with an excess of unlabeled probe whereas the mutated oligos did not alter the DNA-protein binding.

To determine whether these bands were DNA-Sp1 or DNA-Sp3 complexes, we used an interference assay with specific antibodies. The upper band was supershifted by the anti-Sp1 antibody and the lower band by the anti-Sp3 antibody (Fig. 2B). We concluded that the upper band was a DNA-Sp1 complex and the lower band was the DNA-Sp3 complex, respectively.

To examine whether Sp1 and Sp3 in fact bind to the CD97 promoter in-vivo, we performed ChIP with anti-Sp1 and anti-Sp3 antibodies in SMCs. A 230 bp DNA fragment covering the first GC-site at –122 was amplified by PCR immunoprecipitated with anti-Sp1 as well as anti-Sp3 antibodies (Fig. 2C). The same band was obtained in the input DNA, whereas the normal

IgG control immunoprecipitates did not show a signal. Amplification with primers covering the other potentially Sp1/Sp3 binding sites produced no specific bands (not shown).

These results indicate that Sp1 as well as Sp3 bind specifically to one of the identified GC-rich sites within the *CD97* proximal promoter in-vivo and that the other sites are obviously not functionally relevant.

Figure 2

3.4. Sp1 and Sp3 activate CD97 promoter activity and expression

After having determined that Sp1 and Sp3 can bind to the *CD97* promoter and that this interaction is essential for basal promoter activity in SMCs, we examined whether Sp1 and/or Sp3 can directly activate the *CD97* promoter in co-transfection experiments in HEK293 cells. The endogenous Sp1/Sp3 expression levels are low in HEK293 cells, therefore they can be used for co-transfection of Sp1/Sp3 with promoter reporter gene constructs (Banks et al., 1999).

Figure 3 presents the increase of low basal *CD97* promoter activities after co-transfection with a Sp1 or Sp3 expressing plasmid in comparison to the empty vector pN3 control, suggesting that both factors act as transcriptional activators on *CD97* promoter activities. Expression of Sp1 and Sp3 together did not result in an additive effect of activity increase (not shown). Co-transfection of Sp1 or Sp3 with the GC-site 1 mutated construct had no effect on *CD97* promoter activity (Fig. 3) which emphasizes the importance of this Sp1/Sp3 site for *CD97* promoter regulation.

Figure 3

4. Discussion

In the present study, we characterized the 5'-flanking region of the human *CD97* gene as an initial step in understanding the regulation of *CD97* in SMCs. We showed that Sp1 and Sp3 transcription factors bind directly to the *CD97* promoter and upregulate its activity. Of the five identified Sp1/Sp3 binding sites in a 3116 nt *CD97* promoter region, the most proximal site relative to the translation start codon is critical for maintaining basal promoter activity. In linking the expression of *CD97* to the activity of a specific family of transcription factors in human cells, we believe this is the first description of one mechanism controlling the expression of a human EGF-TM7 molecule.

A comprehensive computational analysis revealed the absence of canonical TATA and CCAAT boxes in the 5'-flanking region of the *CD97* gene. In fact, typical TATA and CCAAT boxes are also absent in numerous human G protein-coupled receptor genes, including the alpha1β-adrenergic receptor, glucagon receptor, glucagon-like peptide receptor, growth hormone-releasing hormone receptor gene and the secretin receptor genes (Pang et al., 2004).

However, we identified five GC-rich sites in the 3116 nt *CD97* promoter region. The presence of GC-rich sequences within the 5'-flanking region is a feature of TATA-less genes, and the expression of such genes is regulated by members of the Sp-family (Philipsen and Suske, 1999; Suske, 1999). It has been proposed that many TATA-less and GC-rich promoters bind one or more Sp1 molecules to recruit specific cofactors such as TATA-binding protein-associated factors, which subsequently interact with transcription factor IID (TFIID) comprising TATA box-binding protein (TBP) and a number of TBP-associated factors to initiate gene transcription (Pang et al., 2004; Papadodima et al., 2005).

To get an indication whether the GC-rich sites influence *CD97* transcriptional regulation, we treated cell cultures with mithramycin, a cell-permeable agent that binds to GC-rich DNA sequences and that is frequently used to explore the sequence dependency of DNA-binding

factors (Liu et al., 2006). Since *CD97* promoter activities as well as mRNA expression levels were decreased by mithramycin in SMCs, we studied the GC-sites in more detail using reporter gene assays as well as gelshift and ChIP analysis.

The *CD97* promoter reporter analysis using 5' deletion constructs did not reveal a gradual decrease in promoter activity dependent on the GC-rich sites. These observations imply that also other unidentified positive/negative factors are effective in the proximal promoter of the *CD97* gene (Shin et al., 2005). The construct -218/+45 carrying one GC-site showed the highest promoter activity whereas activity is almost abolished in the -111/+45 construct lacking GC-sequences. It is likely that further upstream located transcription factor bindings influence the *CD97* promoter activity also in an inhibiting manner and act contrary to Sp1/Sp3. Generally, one or more transcription factors recruit specific proteins to cooperatively mediate the regulation of a target gene. There may also be competitive binding when multiple transcription factors simultaneously interact with the promoter region of a target (Wang et al., 2007).

To investigate which of the predicted Sp1/Sp3 sites is functional relevant, we performed reporter gene assays with *CD97* promoter constructs each with a mutated binding site. Interestingly, mutation of the most proximal site at –122 caused a significant decrease in promoter activity in comparison to the -2021 nt wildtype construct.

Although Sp1 has been found to be expressed in a great variety of cell types, its function can be regulated through several different mechanisms and these differences could explain its involvement also in the tissue-specific promoter modulation. The Sp1/Sp3-dependent transcriptional regulation of *CD97* in SMCs but not in colorectal tumor cells (own unpublished data) is a further example. Differences in localisation, biochemical modification and ligand binding of CD97 in SMCs suggest a different function and regulation from that at the cell surface of leukocytes and tumor cells (Wobus et al., 2004). The definition of CD97 regulatory processes in SMCs is also important with respect to a potential involvement in

pathological processes as in vascular disorders. Interestingly, there is evidence that Sp1-dependent mechanisms may be responsible for activation of genes characteristic of phenotypically modulated SMCs, such as the constitutive expression of platelet-derived growth factor (PDGF)- β chain gene during mouse development or smooth muscle myosin heavy chain (MHC) in neointimal SMCs after ballon injury in the rat aorta (Owens et al., 2004). Sp1 was shown to have an anti-mitogenic, pro-apoptotic role in SMCs of the artery wall in atherosclerosis and restenosis (Kavurma and Khachigian, 2003).

To evaluate whether increased levels of Sp1 and/or Sp3 would increase *CD97* promoter activity, we co-transfected *CD97* promoter constructs with Sp1 or Sp3 expression vectors in HEK293 cells which express low endogenous levels of Sp proteins (Banks et al., 1999) and can therefore be used for such Sp1/Sp3 over-expression experiments. Since dose-response studies have indicated that the Myeloid elf-1-like factor (MEF) and involucrin gene promoter activity is directly related to the concentration of Sp1 expression plasmid (Banks et al., 1999; Koga et al., 2005), we accomplished co-transfections of either Sp1 or Sp3. We demonstrated that both Sp1 and Sp3 can activate *CD97* promoter activity in all constructs carrying at least the GC-box at the –122 site. Over-expression of Sp1 and Sp3 together did not show an additive effect which is in accordance with the assumption that both proteins bind to identical sequences (Suske, 1999). We showed that Sp3 acts as transcriptional activator, like Sp1, in the proximal region of the *CD97* promoter. This suggests that Sp1 and Sp3 cross-regulate the expression of each other which could be explained by the fact that the Sp3 protein is able to bind within the *Sp1* promoter (Shin et al., 2005). Although Sp1 and Sp3 can bind to the same GC-rich box, Sp3 was originally found to suppress Sp1-mediated activation by binding to the this sequence, thereby preventing Sp1-binding and activation. However, whether Sp3 acts as an activator or a suppressor of Sp1-mediated activation depends on the cellular conditions (Sato and Furukawa, 2004). The interaction between CD97 and Sp1 might affect other transcription factors which are important for basal and/or induced promoter activity (Shimada

et al., 2001). A large number of Sp1-interacting proteins have been identified. The interaction of Sp1 and MyoD is also important for the expression of various muscle-specific genes (Sato and Furukawa, 2004).

In summary, this is the first report showing the transcriptional regulation of *CD97* by Sp1/Sp3 transcription factors in SMCs. Further on, it will be important to analyze other transcription factors to clarify the mechanisms underlying the regulation of CD97 in a cell- and tissue-specific manner.

Acknowledgements

We thank Dr. G. Suske (Philipps University, Marburg, Germany) and Dr. K. Engeland (University of Leipzig, Germany) for providing plasmid DNA as well as B. Rosin, U. Schmiedek and D. Sittig for skillful technical assistance.

This study was supported by the Interdisziplinary Center of Clinical Research (IZKF) Leipzig at the Faculty of Medicine of the Universität Leipzig (project D6), and by the Roland Ernst Stiftung.

References

Aust, G., Steinert, M., Schutz, A., Boltze, C., Wahlbuhl, M., Hamann, J., Wobus, M., 2002. CD97, but not its closely related EGF-TM7 family member EMR2, is expressed on gastric, pancreatic, and esophageal carcinomas. Am. J. Clin. Pathol. 118, 699-707.

Aust, G., Wandel, E., Boltze, C., Sittig, D., Schutz, A., Horn, L.C., Wobus, M., 2006. Diversity of CD97 in smooth muscle cells. Cell Tissue Res. 324, 139-147.

Banks, E.B., Crish, J.F., Eckert, R.L., 1999. Transcription factor Sp1 activates involucrin promoter activity in non-epithelial cell types. Biochem. J. 337, 507-512.

Blume, S.W., Snyder, R.C., Ray, R., Thomas, S., Koller, C.A., Miller, D.M., 1991.

Mithramycin inhibits SP1 binding and selectively inhibits transcriptional activity of the dihydrofolate reductase gene in vitro and in vivo. J. Clin. Invest. 88, 1613-1621.

Bouwman, P., Philipsen, S., 2002. Regulation of the activity of Sp1-related transcription factors. Mol. Cell Endocrinol. 195, 27-38.

Eichler, W., Aust, G., Hamann, D., 1994. Characterization of an early activation-dependent antigen on lymphocytes defined by the monoclonal antibody BL-Ac(F2). Scand. J. Immunol. 39, 111-115.

Galle, J., Sittig, D., Hanisch, I., Wobus, M., Wandel, E., Loeffler, M., Aust, G., 2006. Individual Cell-Based Models of Tumor-Environment Interactions: Multiple Effects of CD97 on Tumor Invasion. Am. J. Pathol. 169, 1802-1811.

Hamann, J., Eichler, W., Hamann, D., Kerstens, H.M., Poddighe, P.J., Hoovers, J.M., Hartmann, E., Strauss, M., van Lier, R.A., 1995. Expression cloning and chromosomal mapping of the leukocyte activation antigen CD97, a new seven-span transmembrane molecule of the secretion receptor superfamily with an unusual extracellular domain. J. Immunol. 155, 1942-1950.

Hamann, J., Vogel, B., van Schijndel, G.M., van Lier, R.A., 1996. The seven-span transmembrane receptor CD97 has a cellular ligand (CD55, DAF). J. Exp. Med. 184, 1185-1189.

Kavurma, M.M., Khachigian, L.M., 2003. Sp1 inhibits proliferation and induces apoptosis in vascular smooth muscle cells by repressing p21WAF1/Cip1 transcription and cyclin D1-Cdk4-p21WAF1/Cip1 complex formation. J. Biol. Chem. 278, 32537-32543.

Kel, A.E., Gossling, E., Reuter, I., Cheremushkin, E., Kel-Margoulis, O.V., Wingender, E., 2003. MATCH: A tool for searching transcription factor binding sites in DNA sequences. Nucleic Acids Res. 31, 3576-3579.

Koga, T., Suico, M.A., Nakamura, H., Taura, M., Lu, Z., Shuto, T., Okiyoneda, T., Kai, H., 2005. Sp1-dependent regulation of Myeloid Elf-1 like factor in human epithelial cells. FEBS Lett. 579, 2811-2816.

Kwakkenbos, M.J., Kop, E.N., Stacey, M., Matmati, M., Gordon, S., Lin, H.H., Hamann, J., 2004. The EGF-TM7 family: a postgenomic view. Immunogenetics 55, 655-666.

Liu, F., Pore, N., Kim, M., Voong, K.R., Dowling, M., Maity, A., Kao, G.D., 2006. Regulation of histone deacetylase 4 expression by the SP family of transcription factors. Mol. Biol. Cell. 17, 585-597.

Majello, B., De Luca, P., Lania, L., 1997. Sp3 is a bifunctional transcription regulator with modular independent activation and repression domains. J. Biol. Chem. 272, 4021-4026.

Owens, G.K., Kumar, M.S., Wamhoff, B.R., 2004. Molecular regulation of vascular smooth muscle cell differentiation in development and disease. Physiol Rev. 84, 767-801.

Pang, R.T., Lee, L.T., Ng, S.S., Yung, W.H., Chow, B.K., 2004. CpG methylation and transcription factors Sp1 and Sp3 regulate the expression of the human secretin receptor gene. Mol. Endocrinol. 18, 471-483.

Papadodima, O., Sergaki, M., Hurel, C., Mamalaki, A., Matsas, R., 2005. Characterization of the BM88 promoter and identification of an 88 bp fragment sufficient to drive neurone-specific expression. J. Neurochem. 95, 146-159.

Philipsen, S., Suske, G., 1999. A tale of three fingers: the family of mammalian Sp/XKLF transcription factors. Nucleic Acids Res. 27, 2991-3000.

Sato, T., Furukawa, K., 2004. Transcriptional regulation of the human beta-1,4-galactosyltransferase V gene in cancer cells: essential role of transcription factor Sp1. J. Biol. Chem. 279, 39574-39583.

Schreiber, E., Matthias, P., Muller, M.M., Schaffner, W., 1989. Rapid detection of octamer binding proteins with 'mini-extracts', prepared from a small number of cells. Nucleic Acids Res. 17, 6419.

Shimada, J., Suzuki, Y., Kim, S.J., Wang, P.C., Matsumura, M., Kojima, S., 2001. Transactivation via RAR/RXR-Sp1 interaction: characterization of binding between Sp1 and GC box motif. Mol. Endocrinol. 15, 1677-1692.

Shin, T., Sumiyoshi, H., Matsuo, N., Satoh, F., Nomura, Y., Mimata, H., Yoshioka, H., 2005. Sp1 and Sp3 transcription factors upregulate the proximal promoter of the human prostate-specific antigen gene in prostate cancer cells. Arch. Biochem. Biophys. 435, 291-302.

Stacey, M., Lin, H.H., Gordon, S., McKnight, A.J., 2000. LNB-TM7, a group of seven-transmembrane proteins related to family-B G-protein-coupled receptors. Trends Biochem. Sci. 25, 284-289.

Steinert, M., Wobus, M., Boltze, C., Schutz, A., Wahlbuhl, M., Hamann, J., Aust, G., 2002. Expression and regulation of CD97 in colorectal carcinoma cell lines and tumor tissues. Am. J. Pathol. 161, 1657-1667.

Suske, G., 1999. The Sp-family of transcription factors. Gene. 238, 291-300.

Suske, G., Bruford, E., Philipsen, S., 2005. Mammalian SP/KLF transcription factors: bring in the family. Genomics. 85, 551-556.

Wang, R.S., Zhang, X.S., Chen, L., 2007. Inferring transcriptional interactions and regulator activities from experimental data. Mol. Cells. 24, 307-315.

Wobus, M., Huber, O., Hamann, J., Aust, G., 2006. CD97 overexpression in tumor cells at the invasion front in colorectal cancer (CC) is independently regulated of the canonical Wnt pathway. Mol. Carcinog. 45, 881-886.

Wobus, M., Vogel, B., Schmucking, E., Hamann, J., Aust, G., 2004. N-glycosylation of CD97 within the EGF domains is crucial for epitope accessibility in normal and malignant cells as well as CD55 ligand binding. Int. J. Cancer. 112, 815-822.

Figures

Fig. 1: *CD97* promoter activities of wildtype and GC-box mutated luciferase reporter constructs in SK-LMS-1 cells and HbSMCs.

A) Depiction of the identified GC-boxes in the *CD97* 5' flanking region.

B) Basal *CD97* promoter activities of different 5' deletion constructs. Luciferase activity was measured in cell lysates and normalized to the corresponding pRL-null activity. Data represent means ± SEM of the fold increase in luciferase activity as compared to empty pGL-3 of at least four independent experiments.

C) Effect of mutation of each Sp1/Sp3 binding site on *CD97* promoter activities. Activities of constructs with mutated GC-sites 1 to 4 (filled box) were compared with activities of the wt –2101/+45 construct and activity of the construct with mutated GC-box 5 was compared with the wt –3101/+45 construct. Data represent the means ± SEM of at least four independent

experiments. *Asterisks* indicate significant difference ($p < 0.05$) between the wt -2101/+45 and the mutated construct.

Fig. 2: Sp1 and Sp3 bind to the *CD97* promoter in-vitro and in-vivo in SMCs as representatively shown for SK-LMS-1 cells.

A) In-vitro binding analysis by EMSA was carried out using radiolabelled oligonucleotides against a GC-consensus sequence (GC), and either the wildtype or mutant sequence of the Sp1 sites 1 to 5 in the human *CD97* promoter with 5 µg nuclear extract of SK-LMS-1 cells in the absence (-) or presence 50-fold cold competitor oligonucleotides. *Arrows* indicate specific DNA-protein complexes for Sp1 and Sp3, respectively, whereas the *asterisks* represent nonspecific binding.

B) For supershift experiments, 2 µg of antibodies complentary to Sp1 and Sp3 were preincubated with SK-LMS-1 nuclear extract before the addition of the radiolabelled wildtype GC-box 1 probe. Supershifted bands were detected for Sp1 (lane 2) as well as Sp3 (lane 3).

C) ChIP assay from SK-LMS-1 cells with Sp1 and Sp3 antibodies. PCR was performed with primers amplifying a Sp1/Sp3-binding site containing fragment at –122 in the *CD97* promoter. A 230 bp DNA fragment was obtained in the Sp1 and Sp3 precipitated samples as well as the input DNA which corresponds to the chromatin fragments before immunoprecipitation, whereas no band was apparent in the normal IgG control immunoprecipitation.

Fig. 3: *CD97* promoter activity is increased by Sp1 or Sp3 over-expression, respectively.

HEK293 cells were co-transfected with 200 ng of the *CD97* promoter reporter constructs together with 200 ng of Sp1 or Sp3 expressing pN3 vector. As an uninduced control, the reporter constructs were co-transfected with the empty vector pN3. The GC-mutation is represented as filled box in the –2101/+45 construct. Luciferase activity was measured in cell

lysates and normalized to the corresponding pRL-null activity. Data represent the means ± SEM of the fold increase in luciferase activity as compared to the empty vector pGL-3 of at least four independent experiments.

Tab. 1: Primers used for PCR procedure in the experiments

Name	**Sequence 5' –3'**
For 5' RACE	
HCD97r34	aga cac aga atg cga gaa aga
HCD97r38	tgc cgg cag gag cgg cag ggg
For generation of the 5' stepwise deletion of *CD97* promoter luciferase constructs	
pGL -3071/+45 (sense, –3071)	cgg ggt acc tca aga aag ccg cca gga at
pGL -2101/+45 (sense, –2101)	gta ggt acc aga ctt gat ttc ggc atc ct
pGL –1801/+45 (sense, –1801)	tga ggt acc cat tgc aga tgt ttt gag ga
pGL –1151/+45 (sense, –1151)	tct ggt acc taa gtt tcc cac agg ggg aa
pGL –218/+45 (sense, –218)	tgt ggt acc atg cag cgc ccc tgg gtc tgt
pGL -111/+45 (sense, -111)	aat ggt acc ttc ata aag tcc tgg cct cg
common (antisense, +45)	cta gct agc cca gcg ggg ccc caa agt a
For generation of oligonucleotides used for EMSA	
GC-site (1) wt	agc ggc tgt *ccg ccc cc*c ctc c
GC-site (1) mut	agc ggc tgt *aaa aaa aa*c ctc c
GC-site (2) wt	aag acc g*cc ccg ccc g*gc tgc c
GC-site (2) mut	aag acc *gcc cca aac g*gc tgc c
GC-site (3) wt	cac cac aac *ctc cgc ctc* ctg g
GC-site (3) mut	cac cac aac *ctc caa atc* ctg g
GC-site (4) wt	ggg act ct*g ggg ctg gg*c tgc c
GC-site (4) mut	ggg act ct*a aaa aaa aaa* tgc c

GC-site (5) wt	gtg ttc att *ccc ctc ccc* cac a
GC-site (5) mut	gtg ttc att *aaa aaa aaa* cac a
GC-consenus sequence	att cga tcg ggg cgg ggc gag c
For site-directed mutagenesis of GC-sites in *CD97* promoter luciferase constructs	
GC-site (1) sense	ggc cag cgg ctg tcc *aaa* ccc cct cct tca taa agt
GC-site (1) antisense	act tta tga agg agg ggg *ttt* gga cag ccg ctg gcc
GC-site (2) sense	agg caa gac cgc ccc *aaa* cgg ctg cca cac ctt c
GC-site (2) antisense	gaa ggt gtg gca gcc g*tt t*gg ggc ggt ctt gcc t
GC-site (3) sense	cag ctc acc aca acc tcc aaatcc tgg gtt cta tca gtc c
GC-site (3) antisense	gga ctg ata gaa ccc agg att tgg agg ttg tgg tga gct g
GC-site (4) sense	cat tgt agg gac tct gg*a aa*t ggg ctg ccc gcc tat gg
GC-site (4) antisense	cca tag gcg ggc agc cca *ttt* cca gag tcc cta caa tg
GC-site (5) sense	tgg tgt tca ttc ccc t*aa a*cc aca gcc cag gcc acc
GC-site (5) antisense	ggt ggc ctg ggc tgt gg*t tt*a ggg gaa tga aca cca
For amplification after ChIP	
GC-site (1) sense	atg cag cgc ccc tgg gtc tgt
GC-site (1) antisense	aag agt gag tgg gac agg gc
GC-site (2) sense	gct att tct gac aag acc cc
GC-site (2) antisense	tgc act cag taa gaa tgg agt
GC-site (3) sense	tgc ccg cct atg gat ct
GC-site (3) antisense	tat ttc aga ggc ctc ttc gt
GC-site (4) sense	agt aag gca cag tca tct at
GC-site (4) antisense	cag tgt gag aga aaa gat aa

GC-site (5) sense	ttc cag cgg gga gga cag tt
GC-site (5) antisense	ccc ttc ctg tcc tgg gct cc

Italic letters indicate the positions that carry mutations.

Tab. 2: Sp1/Sp3 sites in the sequence (3071 nt) upstream of the translation start site of the human *CD97* gene predicted with different methods. The position in the column head indicates the first nucleotide of the triplet mutated to 'AAA' in site mutation studies.

mutated position	-122	-392	-1392	-1888	-2317	others
sequence	TGTCCGCCCCCC	GCCCCGCCCG	CTCCCGCCTC	GGGGCTGGGC	TCCCCTCCCC	
Patch (public 1.0) database = TRANSFAC 6.0 sites searched for = vertebrates maximum number of mismatches = 0 mismatch penalty = 100 lower score boundary = 87.5	√	√	√	√	√	0
Match (public 1.0) database = TRANSFAC 6.0 matrix similarity = 0.9 core similarity = 0.75	√	√	-	√	√	7
pwmatch database = TRANSFAC 6.0 licensed motiflist = M00008, M00196 M00931, M00932, M00933, M00665						
score = 0.8	√	√	√	√	√	8
score = 0.85	√	√	-	√	√	4
pwmatch Database = JASPAR 2005 Motiflist = MA0079						
score = 0.8	√	√	-	√	√	7
score = 0.85	-	-	-	√	-	3

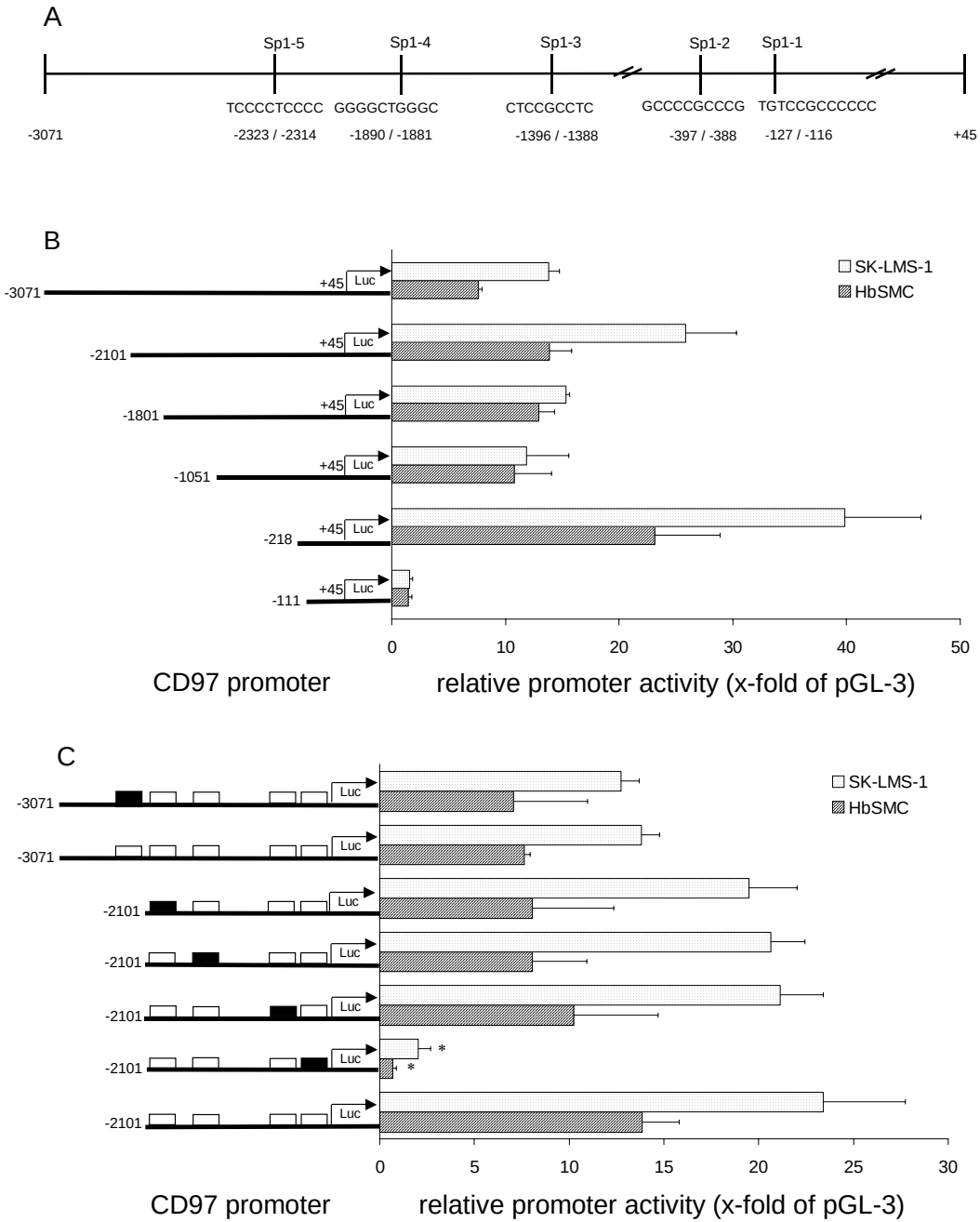
A
Sp1-5
Sp1-4
Sp1-3
Sp1-2
Sp1-1
TCCCCTCCCC
GGGGCTGGGC
CTCCGCCTC
GCCCCGCCCG
TGTCCGCCCCCC
-3071
-2323 / -2314
-1890 / -1881
-1396 / -1388
-397 / -388
-127 / -116
+45
B
SK-LMS-1
HbSMC
Luc
-2101
-1801
-1051
-218
-111
0
10
20
30
40
50
CD97 promoter
relative promoter activity (x-fold of pGL-3)
C
0
5
10
15
20
25
30
*

Fig. 1

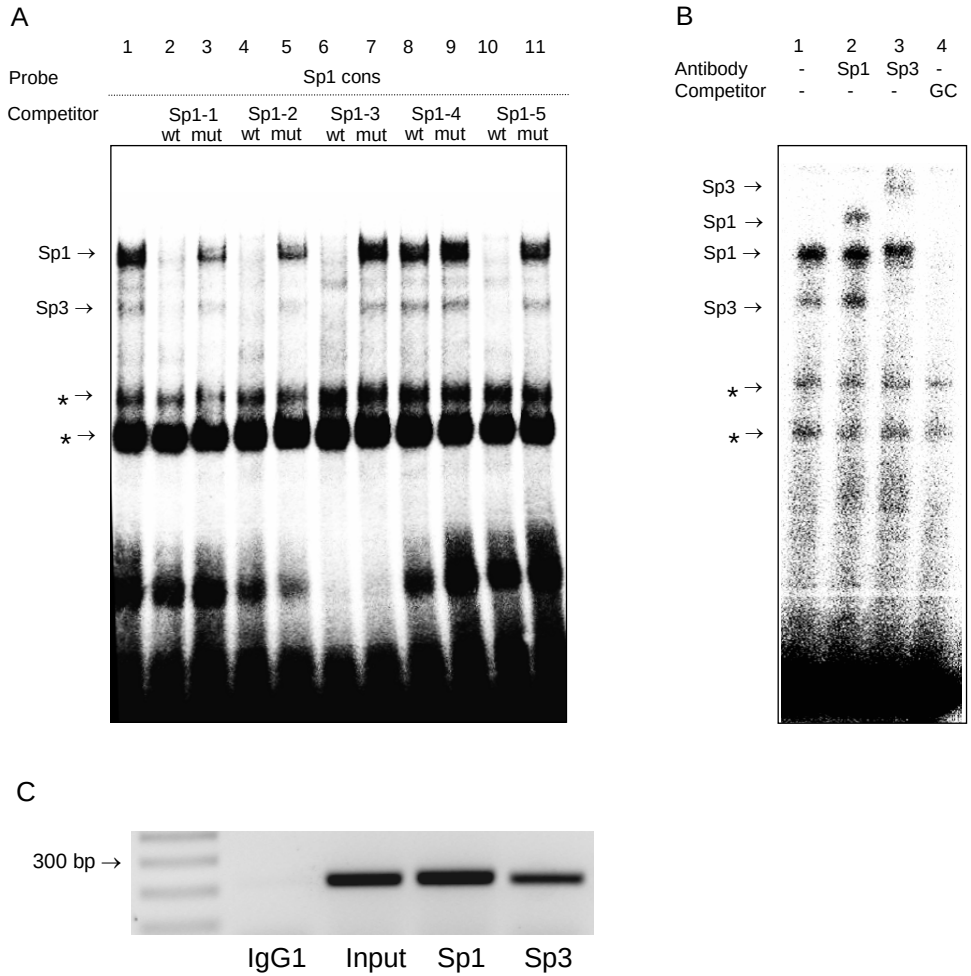
A
1 2 3 4 5 6 7 8 9 10 11
Probe
Sp1 cons
Competitor
Sp1-1 wt mut
Sp1-2 wt mut
Sp1-3 wt mut
Sp1-4 wt mut
Sp1-5 wt mut
Sp1 →
Sp3 →
* →
* →
B
1 2 3 4
Antibody - Sp1 Sp3 -
Competitor - - - GC
Sp3 →
Sp1 →
Sp1 →
Sp3 →
* →
* →
C
300 bp →
IgG1 Input Sp1 Sp3

Fig. 2

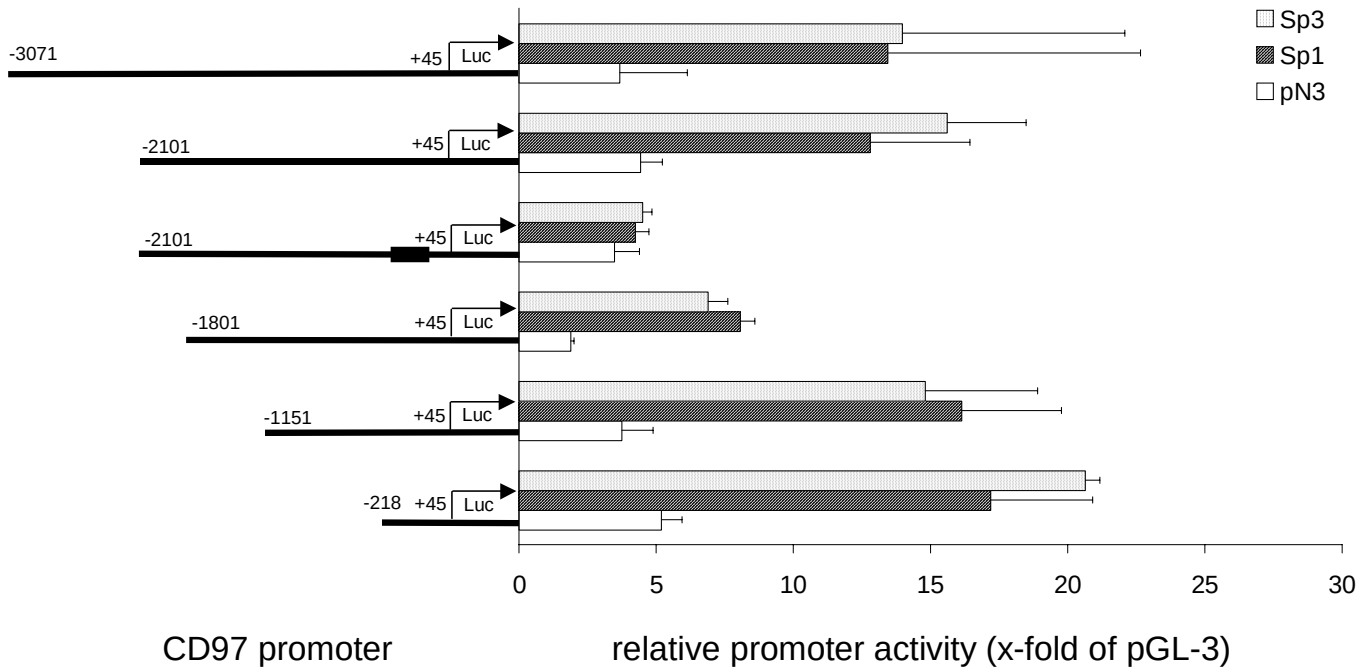
Sp3
Sp1
pN3
-3071
+45
Luc
-2101
+45
Luc
-2101
+45
Luc
-1801
+45
Luc
-1151
+45
Luc
-218
+45
Luc
0
5
10
15
20
25
30
CD97 promoter
relative promoter activity (x-fold of pGL-3)

Fig. 3