THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 276, No. 18, Issue of May 4, pp. 15386–15396, 2001

Printed in U.S.A.

Aberrant Expression of Human Mucin Gene *MUC5B* in Gastric Carcinoma and Cancer Cells

IDENTIFICATION AND REGULATION OF A DISTAL PROMOTER*

Received for publication, November 21, 2000, and in revised form, January 29, 2001
Published, JBC Papers in Press, January 30, 2001, DOI 10.1074/jbc.M010534200

Michaël Perrais‡§, Pascal Pigny‡¶‖, Marie-Pierre Buisine‡‖, Nicole Porchet‡‖**, Jean-Pierre Aubert‡**, and Isabelle Van Seuningen-Lempire‡ ‡‡**

*From the ‡Unité INSERM 377, Place de Verdun, 59045 Lille Cedex, France and Laboratoires **de Biochimie et de Biologie Moléculaire and ¶d'Endocrinologie de l'Hôpital C. Huriez, Centre Hospitalier Régional et Universitaire, 59037 Lille Cedex, France, and ‖Faculté de Médecine, Université de Lille II, 59045 Lille France*

In gastric cancer, altered expression of *MUC1, MUC2, MUC5AC*, and *MUC6* mucin genes has already been described. We show in this report by the means of *in situ* hybridization, reverse transcriptase-polymerase chain reaction, and transfection assays that *MUC5B* is also abnormally expressed in gastric carcinomatous tissues and cell lines. We thus undertook to elucidate the molecular mechanisms that regulate the transcription of *MUC5B* in gastric cancer cells. To this end, high expressing (KATO-III) and low expressing (AGS) gastric cancer cell lines were chosen to study human mucin gene *MUC5B* expression and promoter activity. Sequencing of the promoter region revealed a distal TATA box located 1 kilobase upstream of the proximal TATA box. Functional activity of the promoter was addressed by using deletion mutants covering 2044 nucleotides upstream of the *MUC5B* transcription start site. We identified a distal promoter 10 times more active than the proximal promoter in KATO-III cells. In AGS cells, both promoters, much less active, showed the same range of activity. Binding assays allowed us to show that the transcription factor ATF-1 binds to a *cis*-element present in the distal promoter. Sp1, which binds to both promoters specifically transactivates the proximal promoter. Treatment of transfected cells with PMA, cholera toxin A subunit, and calcium ionophore A23187 showed that only PMA led to a substantial activation of the distal promoter. *MUC5B* 5′-flanking region having a high GC content, influence of methylation on the *MUC5B* expression was assessed. Our results indicate that repression of *MUC5B* expression visualized in AGS cells is due in part to the presence of numerous methylated cytosine residues throughout the 5′-flanking region. Altogether these results demonstrate that *MUC5B* expression in gastric cancer cells is governed by a highly active distal promoter that is up-regulated by protein kinase C and that repression is under the influence of methylation.

Mucins are high molecular weight *O*-glycoproteins synthesized by epithelial cells as large secreted or membrane-bound glycoproteins (1). So far, eight mucin genes have been well characterized (*MUC1*, *MUC2*, *MUC3*, *MUC4*, *MUC5AC*, *MUC5B*, *MUC6*, and *MUC7*) (1), and cDNAs have been proposed for *MUC8*, *MUC9*, *MUC11*, and *MUC12* (1, 2). Numerous studies have now demonstrated that the expression of mucin genes is tissue- and cell-specific and that their expression is altered during the pathogenesis of several diseases, which suggests that human mucin gene expression is tightly regulated and that they may play important roles during cell differentiation and carcinogenesis (3–8).

In normal stomach, *MUC5AC* is expressed at the surface/foveolar epithelium and *MUC6* in the mucous neck cells and in the antral glands (9–11). Other mucin genes expressed in normal gastric mucosae are *MUC1* and to a lesser extent *MUC2*, *MUC3*, and *MUC4*. In gastric carcinomas, a decrease of *MUC1, MUC5AC* and *MUC6* expression and an increase of *MUC2*, *MUC3*, and *MUC4* expression has already been demonstrated (11–15).

Human mucin gene *MUC5B* has been extensively studied in our laboratory. Studies of *MUC5B* genomic sequence (39.1 kb)[1] showed that it encodes a high molecular weight polypeptide (627,000) (16–20). *MUC5B* mucin gene is localized on chromosome 11 (band p15) and is clustered with three other mucin genes: *MUC2*, *MUC6*, and *MUC5AC* (21). In normal adult, *MUC5B* is essentially expressed in trachea, bronchi, submaxillary glands, pancreas, gallbladder, and endocervix (4, 22–24). In cancer, *MUC5B* has been shown to be highly expressed in colon carcinoma (5), in HT-29 treated with methotrexate (19, 20, 25), and LS174T (20, 24, 26) mucus-secreting colon cancer cell lines.

We recently characterized the first 956 nucleotides located upstream of *MUC5B* transcription start site and studied the promoter functional activity in colon cancer cells (20). The region is characterized by the presence of a TATA box and numerous putative binding sites for ubiquitous (Sp1) and specific transcription factors (NF-κB, c-Myc). The high expression of *MUC5B* was correlated with the mucus-secreting phenotype of the LS174T colon cancer cell line. Introns 1 and 37 of *MUC5B* have also been studied in our laboratory because they contain tandemly repeated GA- and GC-rich sequences that

* This work was supported by l'Association de Recherche contre le Cancer Grant 5785 and a grant from le Comité du Nord de la Ligue Nationale contre le Cancer. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AJ012453.

§ Recipient of a CHRU de Lille-Région Nord-Pas de Calais Ph.D. fellowship.

‡‡ To whom correspondence should be addressed. Tel.: 33-320-29-88-65; Fax: 33-320-53-85-62; E-mail: vanseuni@lille.inserm.fr.

[1] The abbreviations used are: kb, kilobase(s); AP-1, activating protein 1; ATF-1, activating transcription factor 1; CREB, cAMP-responsive element binding protein; CTA, cholera toxin A subunit; EMSA, electrophoretic mobility shift assay; PCR, polymerase chain reaction; PKC, protein kinase C; PMA, phorbol 12-myristate 13-acetate; RT, reverse transcriptase; SSC, saline-sodium citrate buffer; bp, base pair(s).

bind the transcription factors Sp1 (20) and NF1-MUC5B, respectively (19).

In this report, we show that *MUC5B* is abnormally expressed in gastric carcinomatous tissues and cell lines. Computer analysis of the genomic sequence upstream of the *MUC5B* transcription start site revealed the presence of a distal TATA box. Numerous putative binding sites for Sp1, CREB, ATF-1, and AP-1 transcription factors were found adjacent to this TATA-box. Binding assays showed that the nuclear factors Sp1 and ATF-1 bind to the distal promoter. The functional activity of the two promoters of *MUC5B* was studied in two gastric cancer cell lines that either show a high (KATO-III) or a low (AGS) level of *MUC5B* mRNA expression. From these studies, a distal region highly active in KATO-III cells was identified. Moreover, this region was shown to be highly methylated in AGS cells in accordance to the low level of *MUC5B* mRNA transcripts found in these cells.

MATERIALS AND METHODS

In Situ Hybridization—Specimens of tumoral and normal gastric mucosae were obtained from six patients undergoing gastrectomy for gastric carcinoma (cardia, well differentiated (n = 2); fundus, mucous type (n = 1); fundus, well differentiated (n = 1); antrum, well differentiated (n = 1); and antrum, moderately differentiated (n = 1)). Each specimen was immediately immersed in 4% paraformaldehyde in phosphate buffer and further embedded in paraffin. Sections 3-μm-thick were cut and mounted onto gelatin covered slides. Adjacent sections from the same blocks were systematically stained with hematoxylin-eosin-safran and Astra blue for a histological control. *In situ* hybridization was performed using a specific *MUC5B* ^{35}S-labeled oligonucleotide probe. The 48-mer oligonucleotide antisense probe (5′-TGTGGTCAGCTCTGTGAGGATCCAGGTCGTCCCCGAGTGGAGAGGG-3′) was chosen in the tandem repeat domain of *MUC5B* (27). The labeling of the probe and the hybridization steps were as described in Ref. 4. Controls consisted in a treatment of tissue sections with a large excess of unlabeled oligonucleotide identical or distinct from the *MUC5B* radiolabeled probe.

Cell Lines and Cell Culture—The KATO-III and AGS gastric adenocarcinoma cell lines were purchased from European Collection of Cell Culture (Salisbury, UK) (28, 29). KATO-III cells were cultured in RPMI 1640 medium supplemented with 20% fetal calf serum (Roche Diagnostics, Meylan, France). AGS cells were cultured in Ham's F-12 medium supplemented with 10% fetal calf serum. Both cell lines were maintained in a 37 °C incubator with 5% CO_2. The inhibitor of methylation, 5-aza-2′-deoxycytidine (5 μM) (Sigma Saint-Quetin Fallavier, France), was added to confluent cells, and cells were cultured in the presence of the chemical reagent for 3 more days before being harvested in appropriate buffer to prepare total RNA.

Cloning—Inserts were prepared using the restriction map of the cosmid clone called ELO9 (30), which covers the 5′-flanking region of *MUC5B*. Gel purified fragments (QIAquick gel extraction kit, Qiagen, Courtaboeuf, France) were subcloned into the promoterless pGL3 Basic vector (Promega, Charbonnières, France). Internal deletion mutants were generated by PCR using pairs of primers bearing specific restriction sites at their 5′ and 3′ ends (see Table I). PCR products were digested, gel purified and subcloned into the pGL3 vector that had been previously cut with the same restriction enzymes. All clones were sequenced on both strands on an automatic LI-COR sequencer (ScienceTec, Les Ulis, France) using infrared labeled RV3 and GL2 primers (Promega). Plasmids used for transfection studies were prepared using the Endofree plasmid Mega kit (Qiagen).

RT-PCR—Total RNAs from gastric cancer cells were prepared using the RNeasy midi-kit from Qiagen. Cells were harvested at 70% of confluence, and 1.5 μg of total RNA was used to prepare cDNA (Advantage™ RT-for-PCR kit, Clontech, Ozyme, France). PCR was performed on 5 μl of cDNA using specific pairs of primers for *MUC5B* mucin gene (*MUC5B* forward primer: 5′-CTGCGAGACCGAGGTCAACATC-3′; *MUC5B* reverse primer: 5′-TGGGCAGCAGGAGCACGGAG-3′ (nucleotides 9057–9078 and nucleotides 10108–10127; accession number Y09788) (17). The PCR product expected size is 415 bp. Single-stranded oligonucleotides were synthesized by MWG-Biotech, Germany. Glyceraldehyde-3-phosphate dehydrogenase was used as an internal control. PCR reactions and PCR product analyses were carried out as previously described (20).

Oligonucleotides and DNA Probes—The oligonucleotides used for gel shift assays are indicated in Table II. They were synthesized by MWG-Biotech (Ebersberg, Germany). Equimolar amounts of single-stranded oligonucleotides were annealed and radiolabeled using T4 polynucleotide kinase (Promega) and [γ-^{32}P]dATP. DNA fragments 1450 and 1896 cloned into pGL3 vector used as probes were first digested with *Sac*I-*Mlu*I *(1450)* and *Kpn*I-*Mlu*I (1896) to obtain the insert. The fragments were then gel purified (QIAquick gel extraction kit, Qiagen) and labeled with [α-^{32}P]dCTP at the 3′ end by a fill-in reaction using the Klenow fragment (Roche Diagnostics). Oligonucleotides and DNA fragments radiolabeled probes were separated from free nucleotides on Bio-Gel P-6 and Bio-Gel P-30 columns, respectively (Bio-Rad Marnes la Coquette, France).

Primer Extension—Primer extension reactions were performed using 25 μg of total RNAs prepared from KATO-III and AGS cells as above and from human trachea (Clontech). Annealing and labeling of the exon 1 (5BOAS): 5′-TGCCTGCGGCACCACGAGCATG-3′ and NAU 647: 5′-TCCCTGGTCACCAGCGTCCTG-3′ reverse primers and extension reactions were performed as previously described (20). ϕX174 DNA/*Hin*fI dephosphorylated markers (Promega) were radiolabeled with [γ-^{32}P]dATP just before use. Manual sequencing of DNA fragment 1429 was performed using the T7 Sequenase version 2.0 kit (Amersham Pharmacia Biotech, Orsay, France). Samples were denatured for 10 min at 90 °C before loading on a 6% sequencing gel (Sequagel-6, National diagnostic, Prolabo, France). The gel was then vacuum dried and autoradiographed for 3–4 days at −80 °C.

Transfections—Transfections were performed using Effectene® reagent (Qiagen). Cells were passed the day before the transfection. Transfection conditions were optimized to 1 μg of DNA, 5 μl of Effectene®, and 0.5 × 10^6 cells/well in a 6-well plate. Transfected cells were then incubated for 48 h at 37 °C. Total cell extracts were prepared using 1× reagent lysis buffer (Promega) as described in the manufacturer's instruction manual. Results were corrected for transfection efficiency by cotransfecting 0.1 μg of pSV-βGal vector (Promega). β-Galactosidase activity was measured in 96-well plates as described in the manufacturer's instruction manual using 10 μl of cell extracts (Promega). Luciferase activity was measured on a Berthold 9501 luminometer on 20 μl of cell extracts using luciferase assay reagent (Promega). The luciferase activity is expressed as fold of induction of the test plasmid activity compared with that of the corresponding control vector (pGL3 control vector, Promega) after correction for transfection efficiency by dividing by β-galactosidase activity. Each plasmid was assayed in duplicate in at least three separate experiments. Cotransfection studies with pCMV-Sp1 and pCMV-Sp3 expression vectors were performed as previously described (20). PMA (100 nM) or CTA (1 μg/ml) were added to the cells 24 h after cell transfection and left for another 24 h before harvesting cells to measure luciferase activity. Calcium ionophore A23187 (250 nM) was added to the cells 1 h prior to harvesting cells. This corresponds to the optimized conditions to obtain the maximum effect of each reagent on *MUC5B* promoter activity. All reagents were from Sigma unless otherwise indicated.

Nuclear Extract Preparation—Nuclear extracts from cell lines of interest were prepared as described in (31) and kept at −80 °C until use. Protein content (5 μl of cell extracts) was measured using the bicinchoninic acid method in 96-well plates as described in the manufacturer's instruction manual (PERBIO Science, Bezons, France).

Electrophoretic Mobility Shift Assays—Nuclear proteins (5 μg) were preincubated for 20 min on ice in 20 μl of binding buffer with 2 μg of poly(dI-dC) (Sigma) and 1 μg of sonicated salmon sperm DNA. Radiolabeled DNA probe was added (120,000 cpm/reaction), and the reaction was left for another 20 min on ice. For super-shift analyses, 1 μl of the antibody of interest (anti-Sp1, anti-Sp2, anti-ATF-1, anti-CREB-1, and anti-HoxD9; TEBU, Le Perray en Yvelines, France) was added to the proteins and left for 1 h on ice before adding the radiolabeled probe. Cold competition were performed by preincubating the nuclear proteins with an excess (×50) of the cold oligonucleotide for 20 min before adding the radiolabeled probe. Negative controls were carried out using 1 μl of irrelevant antibody in the reaction mixture. Reactions were stopped by adding 2 μl of loading buffer and loaded onto a 4% nondenaturing polyacrylamide gel, and electrophoresis conditions were as described in Ref. 20. Gels were vacuum dried and autoradiographed overnight at −80 °C.

Preparation of Genomic DNA for Methylation Studies—Genomic DNA was prepared using a blood and cell culture DNA mini kit (Qiagen). 20 μg of genomic DNA was submitted to an overnight digestion with *Bam*HI (50 units) at 37 °C. To study methylation, *Bam*HI-digested DNA was ethanol precipitated and submitted to either a *Hpa*II (methylation-sensitive, 40 units) or a *Msp*I (methylation insensitive, 40 units) digestion overnight at 37 °C. Digested DNA was then loaded on a 2%

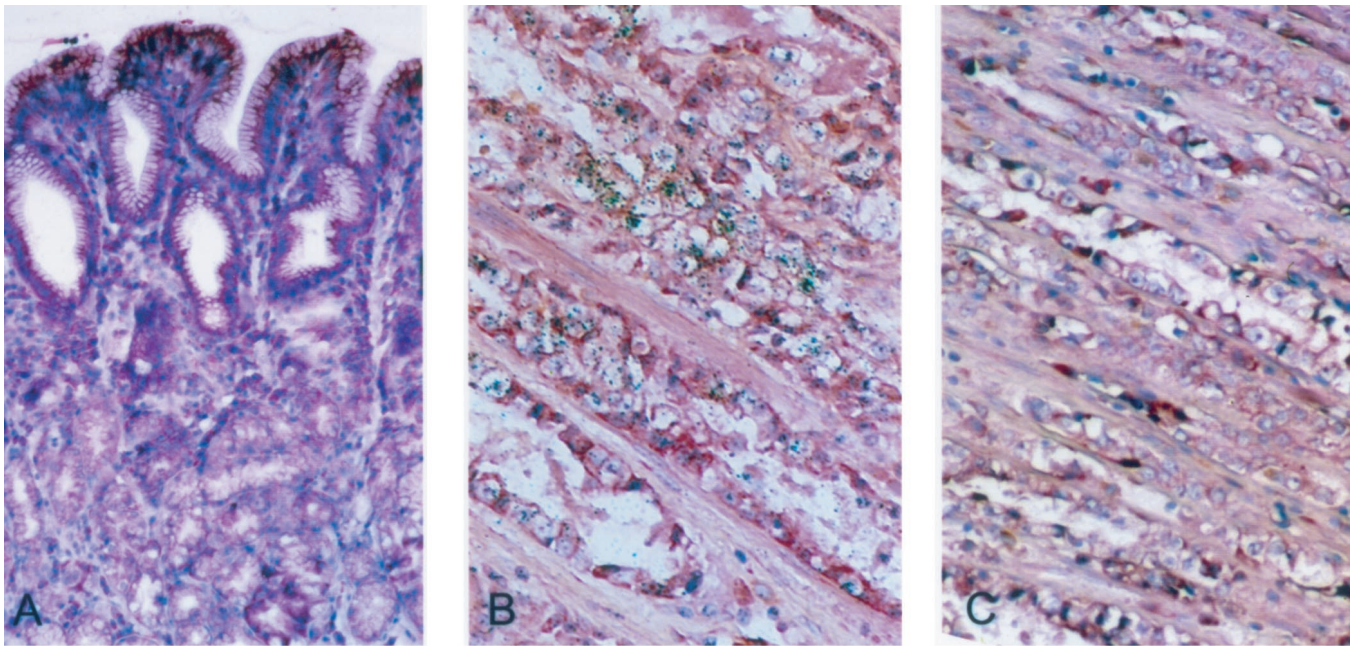

FIG. 1. ***In situ* hybridization for *MUC5B* mRNAs in normal stomach and gastric carcinoma.** *A*, normal stomach (antrum) with the *MUC5B* probe with methyl green pyronin counterstain. The hybridization signal is absent. Magnification, ×200. *B* and *C*, gastric carcinoma (antrum, well differentiated) with the *MUC5B* probe. *B*, with the ^{35}S-labeled *MUC5B* probe, the signal is distributed throughout tumoral glands. *C*, with the ^{35}S-labeled *MUC5B* probe and a large excess of unlabeled *MUC5B* probe, the hybridization signal is absent. Methyl green pyronin counterstain was used. Magnification, ×200.

agarose gel. Electrophoresis was run in 1× Tris-borate-EDTA buffer. After electrophoresis, denatured DNA was transferred onto a nylon membrane (Biotrans +, 0.45 μm; ICN, Orsay, France) in 20× SSC buffer overnight and UV cross-linked for 4 min. The membrane was first incubated in prehybridization buffer (6× SSC, 5× Denhardt's solution, 0.5% SDS) for 3 h at 65 °C followed by a 3-h incubation at 65 °C with 1450 and 1896 DNA probes (1×10^6 cpm/lane) in hybridization buffer (6× SSC, 5× Denhardt's solution, 0.1% SDS, 10% dextran sulfate (w/v), 0.25 mg/ml herring sperm DNA). Excess of the probe was washed off with 10 ml of 0.1× SSC, 0.1% SDS for 15 min at 65 °C, and the wash was repeated once. The blot was then rinsed with 3× SSC and autoradiographed for a few days at −80 °C.

DNA Sequence and Transcription Factor Binding Site Analyses—DNA sequences were analyzed using PC-GENE software, and the TRANSFAC 4.0 data base was used to define potential transcription factor binding sites within the clones of interest. The search was conducted using MatInspector V2.2 software (32).

RESULTS

Expression of MUC5B in Gastric Adenocarcinoma and Gastric Cancer Cell Lines—Expression of *MUC5B* was analyzed in gastric adenocarcinoma and normal resection margins from six patients undergoing gastric resection using *in situ* hybridization. *MUC5B* was not detected in the specimens of normal gastric mucosa (6/6) (Fig. 1*A*). In contrast, a signal was detected with the *MUC5B* probe in four of six specimens of gastric carcinomas (all well differentiated). The labeling was distributed heterogeneously throughout the tumoral glands (Fig. 1*B*). The hybridization procedure was repeated several times. Competition studies checked the validity of the signal. The labeling disappeared when a large excess of unlabeled *MUC5B* oligonucleotide was added to the ^{35}S-labeled *MUC5B* probe (Fig. 1*C*).

The expression of *MUC5B* mRNA was also studied by RT-PCR in two gastric cancer cell lines (Fig. 2*A*). A high expressing (KATO-III) and a low expressing (AGS) cell line were chosen. As shown in Fig. 2*A*, the expected 415-bp PCR product was found in a large amount in KATO-III cells (*lane 3, KATO-III*), whereas a very faint band was observed in AGS cells (*lane 3, AGS*).

DNA Sequence and Characterization of MUC5B 5′-Flanking Region—Having shown that *MUC5B* is abnormally expressed in gastric adenocarcinoma tissues and cancer cell lines, further investigation of *MUC5B* 5′-flanking region DNA sequence was conducted to identify new regulatory regions. The first 956 nucleotides upstream of the transcription start site were previously described (20). Numerous Sp1 binding sites were found clustered in the close vicinity of the TATA box. In this report, further sequencing of the region located upstream of DNA fragment 1896 over 1.1 kb was conducted, and analysis of the DNA sequence using PC-GENE software revealed the presence of a second TATA-box like sequence (TAAATAAAA). The distal TATA box is located 1.1 kb upstream of the proximal TATA box (Fig. 3). Using the TRANSFAC 4.0 data base, we found out that the region adjacent to this second TATA box is characterized by the presence of two clustered putative binding sites for CREB/ATF and AP-1 transcription factors. Further upstream were located potential binding sites for Sp1, glucocorticoid receptor, thyroid transcription factor-1, retinoid orphan receptor-α1, TGT3, Wilm's tumor-1 transcription factor (KTS), and insulin receptor factor-2.

Identification of a Distal Transcription Unit in MUC5B 5′-Flanking Region—The presence of a distal putative TATA box suggests that a distal transcription start site may exist in this region. To address this question, a reverse primer called NAU 647 was designed and chosen 131 bp downstream of the putative distal TATA box. Primer extension experiments were also performed using the reverse primer located in exon 1 (5BOAS) (20). The extension product obtained with the 5BOAS oligonucleotide is 124 bp long as expected (Fig. 2*B*). The intensity of the band is about the same in KATO-III (Fig. 2*B*, *lane 5*) and AGS cells (Fig. 2*B*, *lane 6*). The positive control with RNA from human trachea, a tissue in which *MUC5B* is expressed, also produced a 124-bp extension product (Fig. 2*B*, *lane 7*). No extension product was observed in the negative control (Fig. 2*B*, *lane 8*). The extension with NAU 647 confirmed the presence of a transcription start site in the distal part of the 5′-flanking region of *MUC5B* in both cell lines. The extension product is 109 bp long (Fig. 2*B*, *lanes 2–4*) and starts at a cytosine residue located 23 bp downstream of the distal TATA box (Fig. 3). One can note that the intensity of the band is far

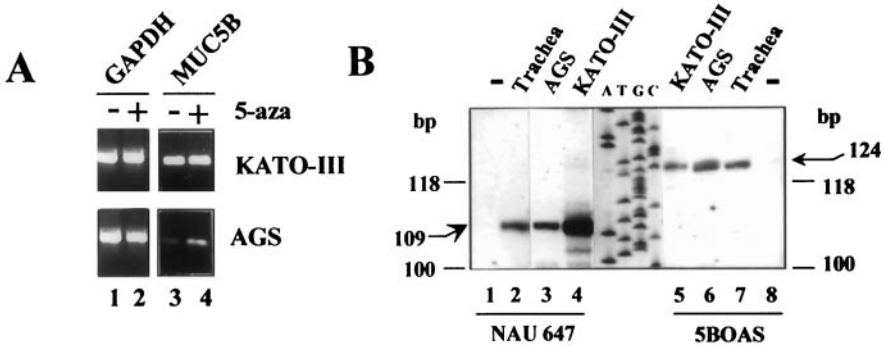

FIG. 2. **Expression of *MUC5B* in KATO-III and AGS cells.** *A*, RT-PCR on 1.5 μg of RNA used for cDNA synthesis. Glyceraldehyde-3-phosphate dehydrogenase (*lanes 1* and *2*) and *MUC5B* (*lanes 3* and *4*) PCR products are 980 and 415 bp long, respectively. PCR products were separated on a 2% agarose gel. Untreated (*lanes 1* and *3*) and 5-aza-2′-deoxycytidine-treated (*lanes 2* and *4*) KATO-III and AGS cells. *B*, primer extension on 25 μg of total RNA from human trachea (*lanes 2* and *7*), AGS (*lanes 3* and *6*), and KATO-III (*lanes 4* and *5*) cells. *Lanes 1* and *8*, no RNA. Two extension products of 109 bp (*lanes 2–4*) and 124 bp (*lanes 6–8*) were produced when using reverse primers located downstream of the distal TATA box (NAU 647) and in exon 1 of *MUC5B* (5BOAS), respectively. ϕX174 DNA/*Hin*fI dephosphorylated markers previously radiolabeled and denatured are indicated on each side of the gel. The sequence of the fragment 1429 is shown.

more intense in KATO-III (*lane 4*) than in AGS cells (Fig. 2*B*, *lane 3*) or human trachea (Fig. 2*B*, *lane 2*). This result confirms the data obtained by RT-PCR (Fig. 2*A*) in which a high amount of *MUC5B* mRNAs was found in KATO-III cells and indicates that the highly active distal transcription unit may thus be responsible for the high expression of *MUC5B* in KATO-III cells. RT-PCR experiments on cDNA prepared from KATO-III total RNA were conducted with pair of primers covering the −1117/−1 region and showed that the whole region is transcribed (not shown). Altogether these results indicate that two active transcription units are present in the 5′-flanking region of *MUC5B*.

MUC5B Promoter Activity in KATO-III and AGS Gastric Cancer Cells—To identify the DNA sequences involved in *MUC5B* transcriptional activity, constructs were generated in the promoterless pGL3 Basic vector and analyzed for transcriptional activity after cell transfection. Insert sequences were confirmed by infrared sequencing of both strands and aligned with ELO9 cosmid sequence, which covers the 5′-flanking region of *MUC5B*.

The 11 deletion mutants used in the transfection experiments cover 2044 nucleotides upstream of the proximal transcription start site (Fig. 4*A* and Table I). The fragments located upstream of the proximal TATA box are 1916, 1896, 1597, 1596, 1895, 1595 and 1598. They cover 956 nucleotides upstream of the proximal transcription start site and represent the proximal promoter of *MUC5B* that was previously characterized in our laboratory (20). The fragments 1916, 1896, 1597 and 1596 contain the TATA box-like sequence (TACATAA), the three Sp1 binding sites and the CACCC box. Fragments 1916 and 1597 contain the 5′-untranslated region segment long of 56 bp. The luciferase activity diagram indicates that the active transcription region in AGS cells is included in fragments 1916, 1896, 1597 and 1596 (Fig. 4*B*). In these cells, the luciferase activity is four times greater than the control vector (pGL3 basic). On the other hand, in KATO-III cells the luciferase activity was only present in fragment 1596 (2-fold activation). Thus, these results indicate that the first 223 bp (fragment 1596) adjoining the proximal transcription start site suffice to drive basal promoter activity of the luciferase reporter gene both in AGS and KATO-III cells. Fragments 1895, 1595, and 1598, which cover the upstream 734 nucleotides do not possess any luciferase activity and act as inhibitory domains in both cell lines.

The fragments located upstream of the distal TATA box are 1599, 1600, 1634 and 2140 and they cover 0.9 kb of DNA sequence. The fragments 1599 and 1600, 212 and 209 bp in length, respectively, contain the TATA box-like sequence (TAAATAAAA) that was characterized using the TRANSFAC 4.0 data base. In AGS cells, both fragments are active (five times more than the pGL3 basic vector) and are slightly more active (20%) than fragment 1596 of the proximal promoter. In KATO-III cells, these two fragments show a very strong activity (10 times more than the pGL3 basic vector) that is 2-fold higher than in AGS cells. The fragment further upstream of 1600, that is fragment 1634, which covers 714 nucleotides is inhibitory in both cell lines. Fragment 2140, which covers the 1599 + 1634 region of 927 nucleotides possesses luciferase activity but is not as active as 1599. This latter result suggests that inhibitory *cis*-elements are present in the DNA fragment 1634.

From these studies, it can be stressed that *MUC5B* promoter activity in gastric cancer cells is driven by two different DNA segments of the 5′-flanking region. In KATO-III cells, the fragment that contains the distal TATA box is by far the most active region, whereas in AGS cells the two regions containing a TATA box have about the same range of transcriptional activity. In AGS cells, the activity of the distal region is much less important than in KATO-III cells. In conclusion, a distal region in *MUC5B* 5′-flanking region was identified that showed high transcription activity in KATO-III cells that may account for the high amount of *MUC5B* mRNA found in these cells.

Binding Studies of MUC5B 5′-Flanking Region with Nuclear Proteins—To characterize *cis*-elements and *trans*-nuclear factors that could account for the cell-specific activity of the promoter of *MUC5B* in gastric cancer cells, DNA-protein binding studies were carried out using the EMSA technique. In Fig. 5 is shown the autoradiogram of the gel shifts performed with nuclear proteins prepared from KATO-III and AGS cells incubated with different DNA probes. Two double-stranded oligonucleotides located in the proximal region (T20 and T33) were chosen from our computer studies with the TRANSFAC 4.0 data base. T20 covers the bases −203/−180 and contains a CACCC box and a Sp1 binding site. T33 covers the bases −137/−111 and contains a Sp1 putative binding site (Fig. 3 and Table II). When incubated with nuclear proteins, the probe T20 led to one low mobility shifted band in both cell lines much more intense in KATO-III cells (compare *lanes 2* and *5*). The complex was totally supershifted when the anti-Sp1 antibody was added in the reaction mixture (*lanes 3* and *6*). Anti-Sp2 antibody used as a negative control did not produce any supershift (*lanes 4* and *7*). With the T33 probe (*lanes 8–12*), a very strong band was visualized in both cell lines (*lanes 9* and *11*), which was totally supershifted upon the addition of the anti-Sp1 antibody in the reaction mixture (*lanes 10* and *12*).

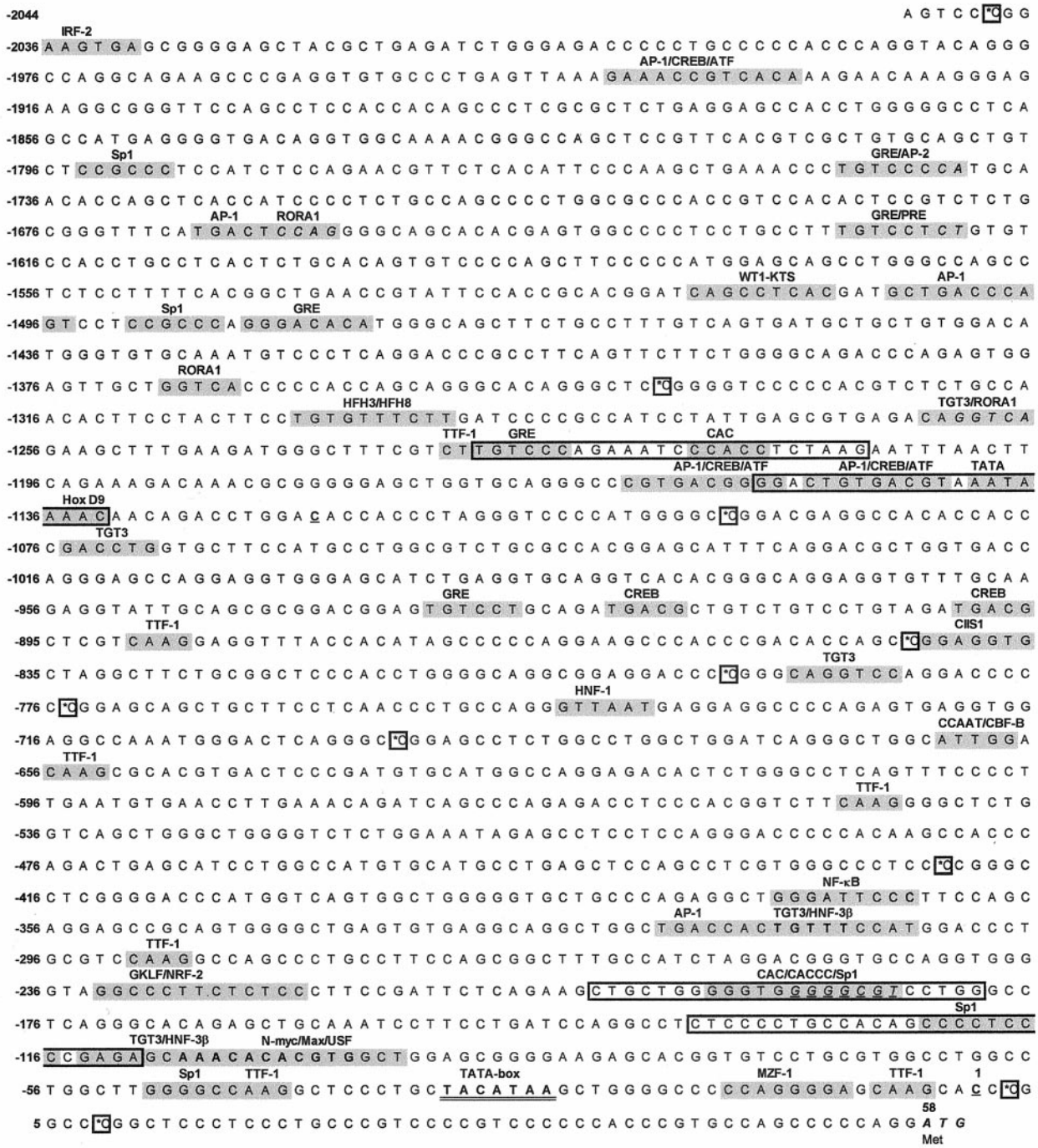

FIG. 3. **DNA sequence of the human *MUC5B* promoter.** The proximal TATA box (TACATAA) at −32/−26 is *double-underlined*, and potential binding sites for known transcription factors are *shaded*. The distal TATA-like sequence at −1142/−1134 (TAAATAAAA) is also *double-underlined*. The proximal transcription start site is designated as +1, and the first ATG is *bold* and *italicized*. The distal transcription start site at −1120 is *bold* and *underlined*. *Hpa*II (C*CGG) potential methylation sites and sequences of the oligonucleotides used in gel shift assay experiments are *boxed*.

From our computational studies, two other probes (T23 and T17) were designed from sequences located in the distal region of the promoter. T23 probe is located at −1230/−1207 and contains a CACC box and a glucocorticoid receptor element. The T17 probe is located at −1153/−1133 and contains a putative ATF-1/CREB-1/Hox D9 binding site. With T23, three major shifted bands were visualized with KATO-III (*lane 14*) and AGS (*lane 16*). The most intense shifted band that is also characterized by the lowest mobility was totally supershifted when the anti-Sp1 antibody was added to the reaction mixture, which demonstrates that Sp1 binds to that *cis*-element (*lanes 15* and *17*). With the probe T17, only one shifted band could be visualized in both cell lines (*lanes 19* and *23*). When using specific antibodies against the three transcription factors of interest, a total supershift specifically occurred upon the addition of the anti-ATF-1 antibody (*lanes 20* and *24*). Anti-CREB-1 (*lanes 21* and *25*) and anti-Hox D9 (*lanes 22* and *26*) antibodies had no effect. Specificity of the binding of the T17 probe to ATF-1 was confirmed by performing cold competitions by preincubating nuclear proteins with 50-, 150-, and 300-fold excess of the cold T17 probe before adding the radioactive probe. The shifted banded totally disappeared when X50 excess of T17 was used (not shown). Altogether these results show that Sp1 binds to two sites in the proximal region of the promoter, whereas it

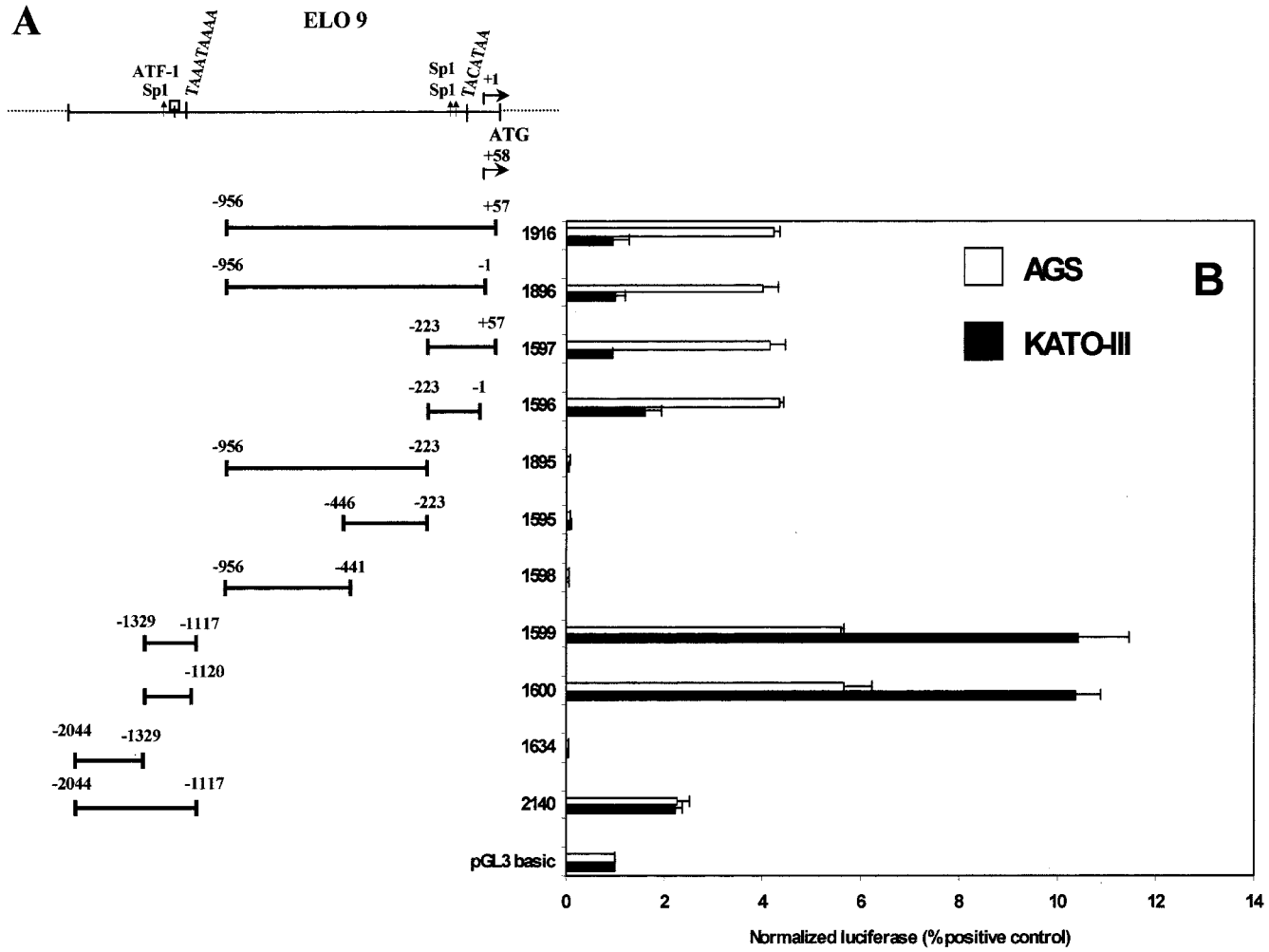

FIG. 4. **Position and transcriptional activity of the pGL3 deletion mutants in *MUC5B* promoter.** *A*, schematic representation of the localization of the different pGL3 deletion mutants covering 2044 nucleotides upstream of the first ATG. *Numbering* refers to the proximal transcription start site designated +1. The TATA box locations are indicated as well as the Sp1 and ATF-1 binding sites found in the 5′-flanking region of the gene. *B*, transcriptional activity of the deletion mutants was studied in KATO-III (*black bars*) and AGS (*white bars*) cell lines. Background activity of pGL3 basic promoterless vector used to subclone *MUC5B* fragments is shown. The results are the means ± S.D. and represent more than three different experiments in duplicate for each fragment.

TABLE I

Sequences of the pairs of oligonucleotides used in PCR to produce deletion mutants in MUC5B 5′-flanking region

*Kpn*I (GGTACC), *Sac*I (GAGCTC), and *Mlu*I (ACGCGT) sites were added at the end of the primers to direct subcloning and are italicized and underlined. Positions of the DNA fragments relative to the proximal transcription start site are indicated (see Fig. 3).

pGL3 deletion mutant	Position of the DNA fragment	Orientation	Oligonucleotides used for PCR (5′ → 3′)
1916	−956/+57	S	CGC*GGTACC*GAGGTATTGCAGCGC
		AS	CGC*ACGCGT*CCTGGGGGCTGGCAC
1896	−956/−1	S	CGC*GGTACC*GAGGTATTGCAGCGC
		AS	CGC*ACGCGT*TGCTTGCTCCCCTGGGGGC
1597	−223/+57	S	CGC*GAGCTC*TCCCTTCCGATTCTCAGA
		AS	CGC*ACGCGT*CCTGGGGGCTGGCAC
1596	−223/−1	S	CGC*GAGCTC*TCCCTTCCGATTCTCAGA
		AS	CGC*ACGCGT*TGCTTGCTCCCCTGGGGGC
1895	−956/−223	S	CGC*GGTACC*GAGGTATTGCAGCGC
		AS	CGC*ACGCGT*AGAGAAGGGCCTACCCCA
1595	−446/−223	S	CGC*GAGCTC*CAGCCTCGTGGGCCCTCC
		AS	CGC*ACGCGT*AGAGAAGGGCCTACCCCA
1598	−956/−441	S	CGC*GGTACC*GAGGTATTGCAGCGC
		AS	CGC*GAGCT*CAGGCATGCACATGGC
1599	−1329/−1117	S	CGC*GAGCTC*CACGTCTCTGCCAACACT
		AS	CGC*ACGCGT*GGTGTCCAGGTCTGTTGT
1600	−1329/−1120	S	CGC*GAGCTC*CACGTCTCTGCCAACACT
		AS	CGC*ACGCGT*GTCCAGGTCTGTTGTTTT
1634	−2044/−1329	S	CGC*GAGCTC*AGTCCCGGAAGTGAGCGG
		AS	CGC*ACGCGT*GGGGACCCGGAGCCCTGT
2140	−2044/−1117	S	CGC*ACGCGT*AGTCCCGGAAGTGAGCGG
		AS	CGC*ACGCGT*GGTGTCCAGGTCTGTTGT

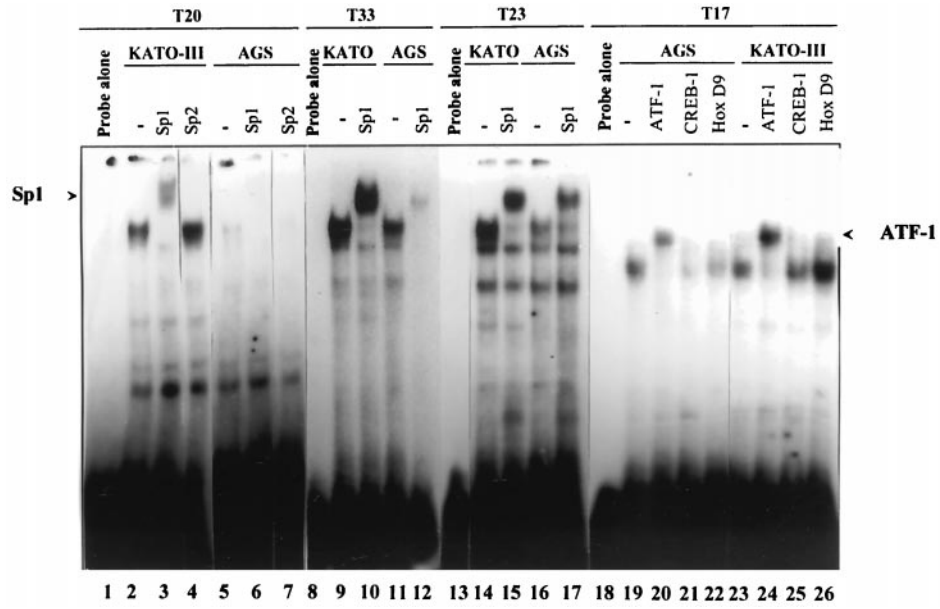

FIG. 5. **Binding of Sp1 and ATF-1 transcription factors to the promoter of *MUC5B*.** Autoradiogram of the EMSA performed with 5 μg of nuclear proteins isolated from KATO-III and AGS cell lines. Nuclear proteins were incubated with the radiolabeled DNA probes as indicated (T20 and T33: proximal promoter; T23, T17: distal promoter). Super-shift experiments were carried out by adding 1 μl of the antibodies of interest (Sp1, Sp2, ATF-1, CREB-1, and Hox D9). Radiolabeled probe alone were loaded in the first lane of each series. Sp1 binding was visualized with T20, T33, and T23 and ATF-1 with T17. On the *left-hand side*, the *arrow* indicates the position of the DNA-protein complex engaging Sp1, and the *arrow* on the *right-hand side* indicates the position of the DNA-protein complex with ATF-1.

TABLE II
Sequences of the sense oligonucleotides used for gel shift assay experiments

Antisense complementary oligonucleotides were also synthesized and annealed to the sense oligonucleotide to produce double-stranded DNA. Positions of the DNA fragments relative to the proximal transcription start site are indicated.

Oligonucleotide name (sense orientation)	Sequence (5′ → 3′)
T20 (Sp1 site), −203/−180	CTGCTGGGGGTGGGGGCGTCCTGG
T33 (Sp1 site), −137/−111	CTCCCCTGCCACAGCCCCTCCCCGAGA
T23 (CAC box site), −1230/−1207	TGTCCCAGAAATCCCACCTCTAAG
T17 (ATF/CREB site), −1153/−1133	GGACTGTGACGTAAATAAAAC

engages once in the distal region where the ATF-1 transcription factor was also shown to bind to its cognate *cis*-element.

Role of Sp1 in MUC5B Promoter Activity—Based on the above EMSA results and previously published data (20), we hypothesized that Sp1 may play a regulatory role on *MUC5B* promoter activity in gastric cancer cells. To test this hypothesis, cotransfection experiments were carried out in the presence of pCMV4, pCMV-Sp1, or pCMV-Sp3 expression vectors (Fig. 6). The luciferase diagram indicates that Sp1 strongly transactivates the proximal region (1896) in both cell lines (3.0- and 2.6-fold activation in KATO-III and AGS, respectively), whereas it has no effect on the distal region (2140). On the other hand Sp3, another member of the Sp family known to compete with Sp1 for the same binding sites did not have any effect on the proximal promoter but strongly inhibited the distal promoter in both cell lines. One can conclude from these results that Sp1 strongly transactivates the proximal promoter of *MUC5B*, whereas Sp3 inhibits the activity of the distal promoter.

Signaling Pathways Involved in MUC5B Promoter Regulation—ATF-1 is a transcription factor that binds to cAMP response elements and that is known to be activated through the cAMP protein kinase cascade. Because it binds to the distal region of *MUC5B* promoter (Fig. 5), we looked whether this signaling pathway was able to transactivate this region. To this end, transfected cells were treated for 24 h with CTA, a cAMP-dependent protein kinase activator, before measuring luciferase activity (Fig. 7). The luciferase diagram clearly indicates that CTA activates transcription of both promoters (approximately 2.5-fold) in KATO-III cells, whereas in AGS cells, cholera toxin A subunit effect was very weak.

Although each of the ATF/CREB proteins appears capable of binding cAMP response elements in its homodimeric form, certain of these proteins also bind as heterodimers with members of the AP-1 transcription factor family to induce gene transcription (33). We thus hypothesized that ATF-1 may heterodimerize with AP-1 in the distal promoter of *MUC5B* and that PKC would then be the signaling pathway used to transactivate this region. To test this hypothesis, transfected cells were treated for 24 h with PMA, a strong PKC activator. As it is shown in Fig. 7, one can see that PMA indeed strongly induced the transcription activity of the fragment 2140, which contains the distal TATA box (2.7- and 3.8-fold activation in KATO-III and AGS cells, respectively). The same PMA treatment was much less effective on the proximal promoter (1896) (2-fold activation in AGS cells). Finally, as it had already been described in the literature that increase of intracellular calcium induces mucin secretion and transcription, we tested whether that signaling pathway had an effect on *MUC5B* promoter activity. As shown in Fig. 7, calcium ionophore A23187 (250 nM for 1 h) effect on *MUC5B* transcription was mild (1.8-fold activation) and restricted to the proximal promoter in AGS cells and to the distal promoter in KATO-III cells. Altogether these results show that cAMP-dependent protein kinase signaling pathway leads to the activation of both promoters in KATO-III cells and that PKC induces a strong activation of the distal promoter in both cell lines.

Role of Methylation in MUC5B Transcription—Having

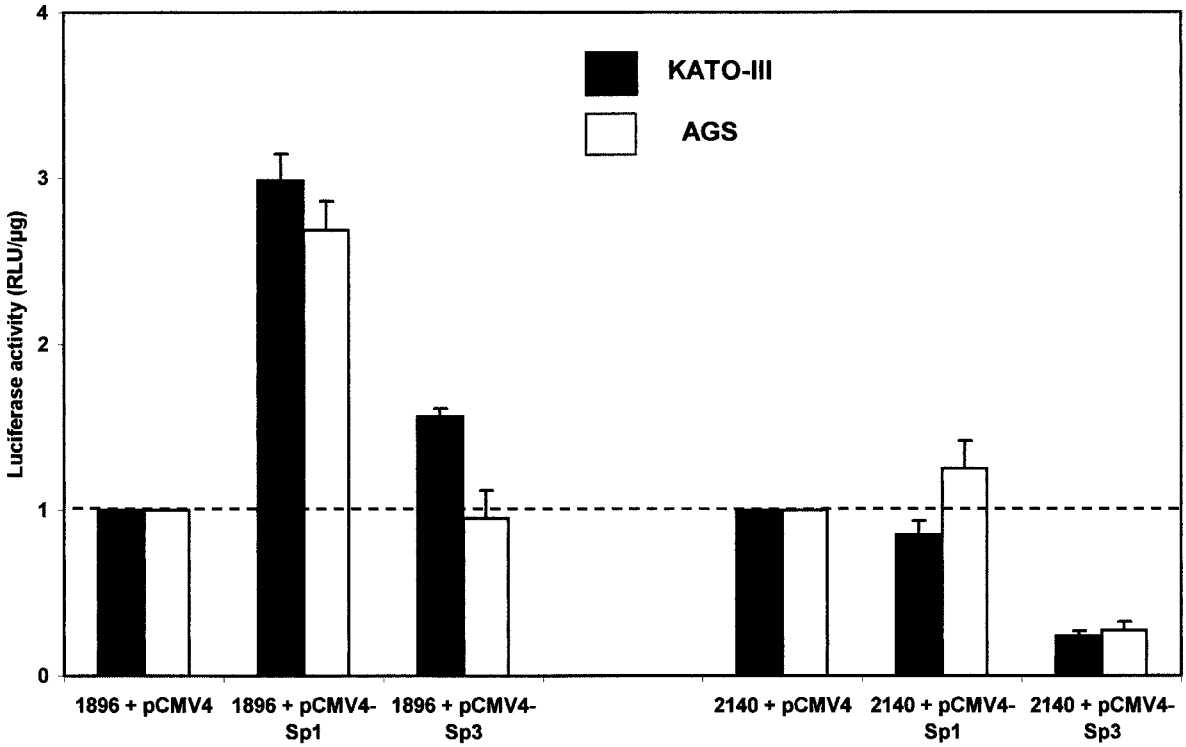

FIG. 6. **Transactivation of *MUC5B* promoter by Sp1.** Cotransfection experiments were performed with 1 µg of *MUC5B* deletion mutants (1896 and 2140) and 0.25 µg of pCMV4, pCMV-Sp1 or pCMV-Sp3 expression vectors in KATO-III (*black bars*) and AGS (*white bars*) cells. The results are the means ± S.D. and represent more than three different experiments in duplicate for each fragment.

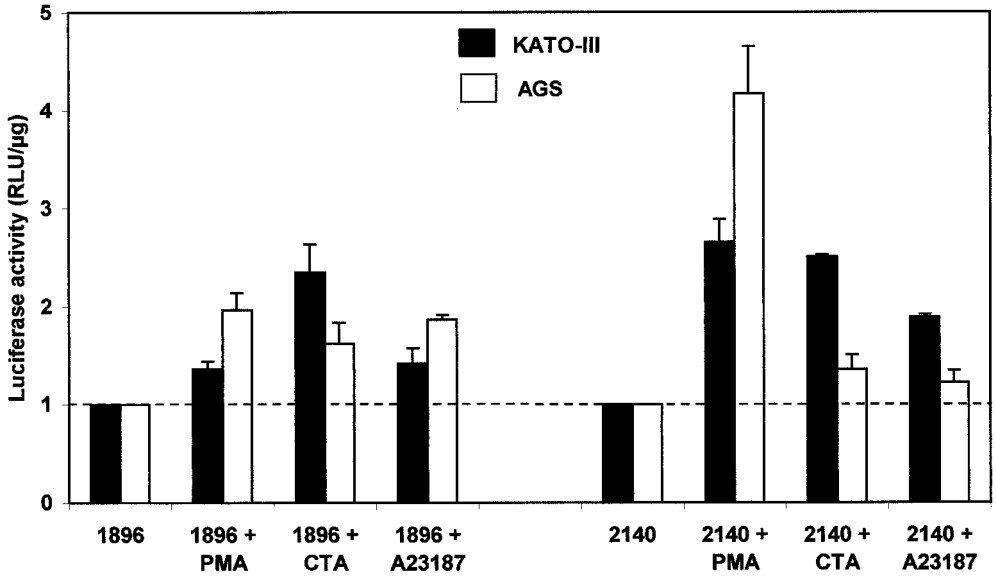

FIG. 7. **Effects of the activation of different signaling pathways on the activity of *MUC5B* promoter.** pGL3 mutant-transfected KATO-III (*black bars*) and AGS (*white bars*) gastric cancer cells were either treated with 100 nM PMA for 24 h, 1 µg/ml CTA for 24 h, or 250 nM calcium ionophore A23187 for 1 h before harvesting cells. The results are the means ± S.D. and represent more than three different experiments in duplicate for each fragment.

shown that Sp1 binds and modulates the activity of the promoter region of *MUC5B* and knowing that Sp1 elements in promoters hampers methylation of mammalian genes and thus modulate transcription activity (34), we undertook to study the level of methylation of *MUC5B* promoter and the effect of methylation on *MUC5B* transcriptional rate in both cell lines.

In the first set of experiments (Fig. 2*A*), cells were treated with the methylation inhibitor 5-aza-2′-deoxycytidine for 72 h after cells became confluent. Total RNA was prepared and RT-PCR performed on untreated and 5-aza-treated cells. The result presented in Fig. 2*A* shows that in KATO-III cells, expression of *MUC5B* is not affected by the treatment of cells with the methylation inhibitor agent (KATO-III, *lanes 3* and *4*). On the contrary, in AGS cells, where *MUC5B* is expressed at a low level in untreated cells (AGS, *lane 3*), its expression is increased by 4-fold after the treatment with 5-aza-2′-deoxycytidine (AGS, *lane 4*). Thus, this experiment confirms the fact that methylation of the promoter of *MUC5B* is one of the mechanisms used to repress *MUC5B* expression in AGS gastric cancer cells.

Having shown that methylation of *MUC5B* promoter occurs in AGS cells, we then undertook to map the cytosine residues

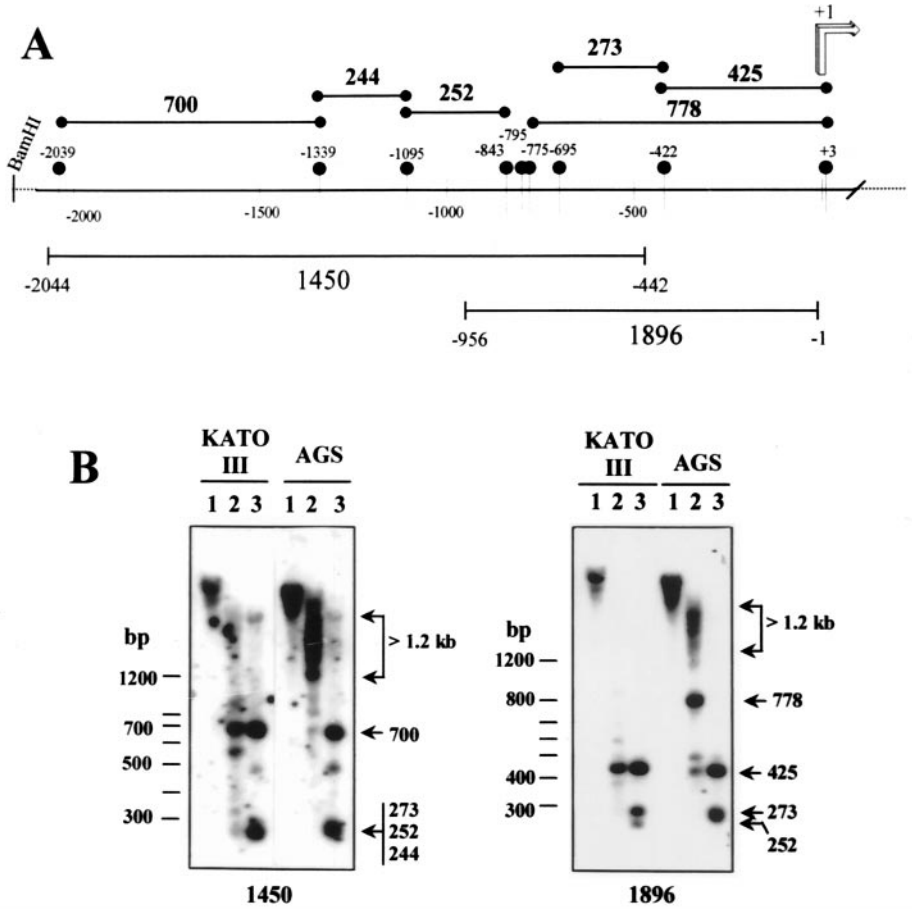

FIG. 8. **Mapping of methylation sites within *MUC5B* promoter in gastric cancer cells.** *A*, schematic representation of *MUC5B* 5′-flanking region. *Black dots* indicate the location of the nine putative methylation sites. Lengths of fragments between two methylation sites are shown. Fragment 1896 covers the −956/−1 region of the 5′-flanking region and fragment 1450 covers the −2044/−442 region. *B*, Southern blotting of 20 μg of genomic DNA from KATO-III and AGS cells probed with 1450 and 1896 DNA fragments. Genomic DNA was digested with *Bam*HI (*lanes 1*), *Bam*HI + *Hpa*II (*lanes 2*), and *Bam*HI + *Msp*I (*lanes 3*). Molecular weight markers are indicated on the *left side* of the autoradiograms. *Arrows* on the *right sides* indicate the position and size of the visualized fragments.

potentially methylated within the promoter region. Potential *Hpa*II-methylation sites (C*CGG) were mapped after analysis of the DNA sequence of the 5′-flanking region of *MUC5B*. Methylation pattern of the promoter was obtained using two DNA probes, 1450 and 1896, that cover the 5′-flanking region of *MUC5B* (Fig. 8*A*). Nine putative methylation sites were found throughout the sequence covering the −2044/+3 bases and are shown on the schematic representation of *MUC5B* 5′-flanking region (Fig. 8*A*). The methylation status of genomic DNA covering the 5′-flanking region of *MUC5B* from the two cell lines is shown in Fig. 8*B*. In KATO-III cells, one major band of 700 bp in length was recognized by the 1450 probe when genomic DNA was digested with *Bam*HI-*Hpa*II (KATO-III, *lane 2*). The spot seen on the left part of that same *lane 2* above 1.2 kb is nonspecific. When the same genomic DNA was digested with *Bam*HI-*Msp*I, the latter enzyme being insensitive to methylation, another strong wide band appeared besides the 700-bp band; this wide band most likely comprises the three expected fragments long of 273, 252, and 244 bp (*KATO-III, lane 3*) that cover the 3′-end region of the genomic region recognized by the probe 1450. This results indicates that, in KATO-III, the cytosines between −1095/−695 are methylated. When the same *Bam*HI-*Hpa*II treatment was applied to genomic DNA from AGS cells (*AGS, lane 2*), bands shifted toward the high molecular weights when compared with KATO-III cells (*KATO-III, lane 2*). The absence of the 700-bp band thus indicates that all the cytosines in the −2044/−442 region are methylated. This was confirmed when the DNA was digested with *Bam*HI-*Msp*I because the bands with molecular weights higher than 1200 bp, visualized after digestion of the DNA with *Bam*HI-*Hpa*II (*AGS, lane 2*), totally disappeared and led to the appearance of the 700-bp band as well as the wide band comprising the 244-, 252-, and 273-bp fragments when DNA was digested with *Bam*HI-*Msp*I (*AGS, lane 3*).

With the 1896 DNA probe, which covers the proximal region of the promoter, the presence of the 425-bp band when DNA was digested either with *Bam*HI-*Hpa*II (*KATO-III, lane 2*) or *Bam*HI-*Msp*I (*KATO-III, lane 3*) in KATO-III cells, indicates that the cytosine at −422 is not methylated. The presence of the 273- and 252-bp bands in *lane 3* but not in *lane 2* indicates that cytosines between −1095/−695 are methylated (Fig. 8*A*). In AGS cells, a series of high molecular weight bands larger than 1.2 kb as well as a 778-bp band were visualized on genomic DNA digested with *Bam*HI-*Hpa*II (*AGS, lane 2*), indicating that the cytosine at −422 is methylated in AGS cells. After digestion with *Bam*HI-*Msp*I enzymes (*AGS, lane 3*), all the high molecular weight bands disappeared to leave place to the expected bands of 425, 273, and 252 bp. Altogether these results indicate that methylation occurs in KATO-III cells at cytosines located between −1095/−695, whereas in AGS cells methylation occurs throughout the 5′-flanking region.

DISCUSSION

Mucins are expressed in a cell- and tissue-specific manner in normal human tissues (1, 3–5, 11) and in normal stomach mucosae, *MUC1*, *MUC5AC*, and *MUC6* are the main mucin genes (9, 10). Altered expression of mucin genes in carcinomas have extensively been described in the literature (3, 5, 11–15), and in gastric cancers, a decrease of the expression of *MUC5AC* and *MUC6* mRNAs and increased levels of *MUC2, MUC3* and *MUC4* mRNAs have been demonstrated (10, 11). Other studies have focused on *MUC1* and *MUC2* mucin gene expression in gastric carcinomas because *MUC1* is often overexpressed in various carcinomas and *MUC2* expression is correlated with the intestinal metaplasia observed during the development of gastric carcinoma (13, 15, 35). All these studies suggest that

carbohydrate and peptide moieties modification on mucins may be valuable markers of gastric neoplastic and preneoplastic states (11, 36–38).

However, control of gene expression in gastric cells remains poorly understood, and an understanding of the regulatory network of nuclear proteins that direct transcriptional initiation of mucin genes is mandatory to decipher the mechanisms of normal development and differentiation as well as disease processes such as neoplasia. These molecules, which can either be secreted to form the mucus or be included in the membrane architecture, play roles of receptors and are involved in cell-cell, cell-substratum, and cell-immune system interactions (1, 36–38). In cancers, such interactions can be altered to allow the tumor cells to migrate and induce metastasis. It is thus clear that any change in their expression will affect all these functions and modify the behavior of cancer cells.

In our laboratory, mucin gene expression has been extensively studied for many years using *in situ* hybridization, and a method to detect all mucin genes from the same sample by RT-PCR was recently developed (20, 39). The results have pointed out that it is important to look at mucin genes that are not or weakly expressed in normal tissues. Buisine *et al*. (6), for example, showed that *MUC5AC* transcripts absent in normal adult colon are re-expressed in rectovillous adenocarcinoma. *MUC5AC* being also expressed in fetal colon, it was thus concluded that it corresponds to a typical oncofetal expression pattern (7). Another study from Balagué *et al*. (40) on *MUC4* showed the same pattern of expression in pancreatic adenocarcinomas.

Very recently, it was shown in our laboratory that *MUC5B* is expressed in embryonic and fetal gastric tissues (41). Thus, to define whether this expression during the early stages of development could correspond to an oncofetal pattern of expression of *MUC5B* in gastric mucosae, we studied *MUC5B* expression in adult normal and carcinomatous gastric mucosae. In this report, we show that *MUC5B* is indeed expressed in both gastric carcinomatous tissues and cell lines. Thus, for a better understanding of the molecular mechanisms that prevail to this abnormal expression, the regulation of the transcriptional activity of *MUC5B* promoter was studied in KATO-III (high expression of *MUC5B*) and AGS (low expression of *MUC5B*) gastric cancer cells.

In a previous work, *MUC5B* promoter activity was studied in colon cancer cell lines with different phenotypes and was shown to be regulated, in part, by the ubiquitous transcription factor Sp1 (20). It was then suggested that Sp1 would be the transcription factor responsible for the basal activity of *MUC5B* in the cells expressing that gene. This hypothesis was confirmed in this report as we showed that Sp1 binds and transactivates the proximal promoter of *MUC5B* in both gastric cancer cell lines. To explain the high abundance of *MUC5B* transcripts in KATO-III cells, it was thus hypothesized that another highly active DNA segment, yet to be characterized, was present in the 5′-flanking region. We thus undertook to further sequence and analyze the DNA region located upstream of 1896. Interestingly enough, a highly active transcription unit containing a TATA box like sequence flanked by two clustered AP-1/ATF/CREB putative binding sites was characterized. The presence of a distal transcription unit in *MUC5B* 5′-flanking region is not unique in mucin genes. Recently, such a regulatory region was described for *MUC1* and was demonstrated to be responsible for the high expression of this gene in breast cancers (42).

Cell transfections with a panel of pGL3 deletion mutants and gel retardation assays confirmed that a highly active distal promoter is present within the 5′-flanking region of *MUC5B*. It contains an active TATA box, binds ATF-1 and Sp1 transcription factors, and is activated by cAMP-dependent protein kinase and PKC signaling pathways. In this report we showed that CTA, which has already been shown to activate mucin secretion in colon cancer cell (43), is capable of specifically inducing *MUC5B* promoter activity in KATO-III but not in AGS cells. Thus, we can postulate that *MUC5B* promoter activation via cAMP signaling in KATO-III cells involves activation of adenyl cyclase through activation of G_s regulatory proteins. Both the fact that ATF-1 can heterodimerize with AP-1 transcription factor family and the fact that G_s regulatory proteins can induce the PKC signaling pathway suggest that PKC may also be involved in *MUC5B* regulation. In this report we showed that PMA induced a strong activation of *MUC5B* distal promoter in gastric cancer cells. This mechanism may be specific to gastric cancer cells because it was recently shown that PMA did not induce *MUC5B* expression in T84 and HT-29/A1 colon cancer cells (44). Increase of intracellular calcium content within mucus-secreting cells is also a pathway that induces mucin secretion (45). In this work, treatment of cells with calcium ionophore A23187 did not have a significant effect on *MUC5B* promoter activity. Thus, PKC signaling pathway seems to be the pathway of choice to induce *MUC5B* promoter activity in gastric cancer cells.

MUC5B is located on chromosome 11p15.5 and is part of a mucin gene cluster comprising *MUC6-MUC2-MUC5AC-MUC5B* (21). One of the aims in this work was to provide a better understanding of *MUC5B* regulation and promoter activity as a member of this cluster. The cluster is 400 kb long and is rich in CpG islands. Among the four genes, promoter sequence is known for *MUC2*, *MUC5AC*, and *MUC5B* but not for *MUC6*. The first three genes are transcribed in the same orientation (46), whereas *MUC6* is transcribed in the opposite way (30). The human 11p15 region displays a high density of CpG islands and contains a cluster of 9–10 genes, such as imprinted *H19* and *IGFII* (insulin growth factor II) genes and Wilm's tumor 1 tumor suppressor gene that have already been shown to be regulated by methylation (47–49). Methylation is an epigenetic mechanism that is commonly used by cells to shut off the expression of a gene (50, 51) and that has profound effects in cancers (51–53). As a central event in the evolution of cancers, along with genetics events, methylation changes constitute potentially sensitive molecular markers to define risk states, monitor prevention strategies, achieve early diagnosis, and track the prognosis of cancer (53). In gastric carcinomas, association between aberrant methylation and CpG island methylator phenotype was recently published (54). Interestingly enough *MUC2,* the major mucin expressed in normal colon, was also shown to be repressed by methylation in colon cancer cells (55). From our results, methylation appears now to control the expression of another gene of the 11p15 mucin gene cluster in gastric cancer cells, that is *MUC5B*. Investigations about the methylation status of the four 11p15 mucin genes are now in progress in our laboratory and tend to show that these genes are indeed regulated by methylation in various cancer cell lines. In the two cell lines studied in this report, *MUC2* and *MUC6* were found to be repressed by methylation in KATO-III cells but not *MUC5AC*.[2] In AGS cells, *MUC2* and *MUC5B* (this report) were found to be repressed by such mechanism. From these studies, it is clear that methylation is a common mechanism used to control the transcription of these four mucin genes and that it could participate besides to the transcription factors to the specific pattern of mucin gene expression in cancer cell lines and tissues. Finally, it is known that Sp1 elements inter-

[2] I. Van Seuningen-Lempire, unpublished observation.

fere with methylation of promoters and thus affects their activity (34). Our laboratory and others (56) have already suggested that the regulation of the 11p15 mucin genes is complex and most likely involves components or genetic mechanisms that are responsible for the tissue- and cell-specific expression of the four genes. Sp1 seems a good candidate because it has now been shown to be involved in the regulation of the first three genes so far described, that is *MUC2* (57), *MUC5AC* (data not shown), and *MUC5B* (20). Studies of the relationship between the binding activity of Sp1 to the promoters and the methylation status of the cluster will certainly help into the understanding of how this region is regulated in cancers.

In conclusion, this work demonstrates that abnormal expression of *MUC5B* visualized in well differentiated gastric carcinoma is due to the presence of a highly active distal transcription unit that is up-regulated by PKC. The transcription factor Sp1, on the other hand, would be responsible for the basal expression of *MUC5B* by transactivating the proximal promoter. Besides this regulation by transcription factors, *MUC5B* also appears to be regulated in gastric cancer cells by methylation. The deciphering of the molecular mechanisms that control the transcription of mucin genes in gastro-intestinal diseases is mandatory to identify transcription factors that target mucin genes during cell differentiation and proliferation and consider mucin genes as potential molecular markers in carcinogenesis.

Acknowledgments—We are indebted to Danièle Petitprez, Pascal Mathon, Marie-Paule Ducourouble, and Isabelle Clavereau for excellent technical help and Claude Vandeperre for the photographs. We thank the members of our E.U. consortium CEEBMH4-3222 for stimulating discussions.

REFERENCES

1. Porchet, N., Buisine, M.-P., Desseyn, J.-L., Moniaux, N., Nollet, S., Degand, P., Pigny, P., Van Seuningen, I., Laine, A., and Aubert, J.-P. (1999) *J. Soc. Biol.* **193,** 85–99
2. Williams, S. J., McGuckin, M. A., Gotely, D. C., Eyre, H. J., Sutherland, G. R., and Antalis, T. M. (1999) *Cancer Res.* **59,** 4083–4089
3. Ho, S. B., Niehans, G. A., Lyftogt, C., Yan, P. S., Cherwitz, D. L., Gum, E. T., Dahyia, R., and Kim, Y. S. (1993) *Cancer Res.* **53,** 641–651
4. Audié, J.-P., Janin, A., Porchet, N., Copin, M.-C., Gosselin, B., and Aubert, J.-P. (1993) *J. Histochem. Cytochem.* **41,** 1479–1485
5. Lesuffleur, T., Zweibaum, A., and Real, F. X. (1994) *Crit. Rev. Oncol. Hematol.* **17,** 153–180
6. Buisine, M.-P., Janin, A., Maunoury, V., Audié, J.-P., Delescaut, M.-P., Copin, M.-C., Colombel, J.-F., Degand, P., Aubert, J.-P., and Porchet, N. (1996) *Gastroenterology* **110,** 84–91
7. Buisine, M.-P., Devisme, L., Savidge, T. C., Gespach, C., Gosselin, B., Porchet, N., and Aubert, J.-P. (1998) *Gut* **43,** 519–524
8. Buisine, M.-P., Desreumaux, P., Debailleul, V., Gambiez, L., Geboes, K., Ectors, N., Delescaut, M.-P., Degand, P., Aubert, J.-P., Colombel, J.-F., and Porchet, N. (1999) *Inflam. Bowel. Dis.* **5,** 24–32
9. Ho, S. B., Roberton, A. M., Shekels, L. L., Lyftogt, C. T., Niehans, G. A., and Toribara, N. W. (1995) *Gastroenterology* **109,** 735–747
10. Bartman, A. E., Buisine, M.-P., Aubert, J.-P., Niehans, G. A., Toribara, N. W., Kim, Y. S., Kelly, E. J., and Ho, S. B. (1998) *J. Pathol* **186,** 398–405
11. Ho, S. B., Shekels, L. L., Toribara, N. W., Kim, Y. S., Lyftogt, C., Cherwitz, D. L., and Niehans, G. A. (1995) *Cancer Res.* **55,** 2681–2690
12. Reis, C. A., David, L., Nielsen, P. A., Clausen, H., Mirgorodskaya, K., Roepstorff, P., and Sobrinho-Simoes, M. (1997) *Int. J. Cancer* **74,** 112–121
13. Reis, C. A., David, L., Correa, P., Carneiro, F., de Bolos, C., Garcia, E., Mandel, U., Clausen, H., and Sobrinho-Simoes, M. (1999) *Cancer Res.* **59,** 1003–1007
14. Filipe, M. I., Linehan, J. M., Durrant, L. G., Price, M. R., Smeeton, N. C., Pathak, S., and Swallow, D. M. (1996) *Eur. J. Cancer Prev.* **5,** 287–295
15. Reis, C. A., David, L., Carvalho, F., Mandel, U., de Bolos, C., Mirgorodskaya, E., Clausen, H., and Sobrinho-Simoes, M. (2000) *J. Histochem. Cytochem.* **48,** 377–388
16. Desseyn, J.-L., Guyonnet-Dupérat, V., Porchet, N., Aubert, J.-P., and Laine, A. (1997) *J. Biol. Chem.* **272,** 3168–3178
17. Desseyn, J.-L., Aubert, J.-P., Van Seuningen, I., Porchet, N., and Laine, A. (1997) *J. Biol. Chem.* **272,** 16873–16883
18. Desseyn, J.-L., Buisine, M.-P., Porchet, N., Aubert, J.-P., and Laine, A. (1998) *J. Biol. Chem.* **273,** 30157–30164
19. Pigny, P., Van Seuningen, I., Desseyn, J.-L., Nollet, S., Porchet, N., Laine, A., and Aubert, J.-P. (1996) *Biochem. Biophys. Res. Commun.* **220,** 186–191
20. Van Seuningen, I., Perrais, M., Pigny, P., Porchet, N., and Aubert, J.-P. (2000) *Biochem. J.* **348,** 675–686
21. Pigny, P., Guyonnet-Dupérat, V., Hill, A. S., Pratt, W. S., Galiegue-Zoutina, S., Collyn, D'Hooge, M., Laine, A., Van Seuningen, I., Degand, P., Gum, J. R., Kim, Y. S., Swallow, D. M., Aubert, J.-P., and Porchet, N. (1996) *Genomics* **38,** 340–352
22. Aubert, J.-P., Porchet, N., Crépin, M., Duterque-Coquillaud, M., Vergnes, G., Mazzuca, M., Debuire, B., Petitprez, D., and Degand, P. (1991) *Am. J. Respir. Cell Mol. Biol.* **5,** 178–185
23. Vandenhaute, B., Buisine, M.-P., Debailleul, V., Clément, B., Moniaux, N., Dieu, M.-C., Degand, P., Porchet, N., and Aubert, J.-P. (1997) *J. Hepatol.* **27,** 1057–1066
24. Van Klinken, B. J.-W., Dekker, J., Van Gool, S. A., Van Marle, J., Büller, H. A., and Einerhand, A. W. C. (1998) *Am. J. Physiol.* **274,** G871–G878
25. Lesuffleur, T., Roche, F., Hill, A. S., Lacasa, M., Fox, M., Swallow, D. M., Zweibaum, A., and Real, F. X. (1995) *J. Biol. Chem.* **270,** 13665–13673
26. Van Klinken, B. J. W., Oussoren, E., Weenink, J. J., Strous, G. J., Büller, H. A., Dekker, J., and Einerhand, A. W. C. (1996) *Glycoconj. J.* **13,** 757–768
27. Dufossé, J., Porchet, N., Audié, J.-P., Guyonnet-Dupérat, V., Laine, A., Van Seuningen, I., Marrackchi, S., Degand, P., and Aubert, J.-P. (1993) *Biochem. J.* **293,** 329–337
28. Sekigushi, M., Sakakibara, K., and Fujii, G. (1978) *Japan J. Exp. Med.* **48,** 61–68
29. Barranco, S. C., Townsend, C. M. J.r., Casartelli, C., Macik, B. G., Burger, N. L., Boerwinkle, W. R., and Gourley, W. K. (1983) *Cancer Res.* **43,** 1703–1709
30. Buisine, M.-P., Desseyn, J.-L., Porchet, N., Degand, P., Laine, A., and Aubert, J.-P. (1998) *Biochem. J.* **332,** 729–738
31. Van Seuningen, I., Ostrowski, J., Bustelo, X. R., Sleath, P., and Bomsztyk, K. (1995) *J. Biol. Chem.* **270**, 26976–26985
32. Quandt, K., Frech, K., Karas, H., Wingender, E., and Werner, T. (1995) *Nucleic Acids Res.* **23,** 4878–4884
33. Hai, T., and Curran, T. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88,** 3720–3724
34. Brandeis, M., Frank, D., Keshet, I., Siegfried, Z., Mendelsohn, M., Nemes, A., Temper, V., Razin, A., and Cedar, H. (1994) *Nature* **371,** 435–438
35. Machado, J. C., Nogueira, A. M. M. F., Carneiro, F., Reis, C. A., and Sobrinho-Simoes, M. (2000) *J. Pathol.* **190,** 437–443
36. Kim, Y. J., Borsig, L., Han, H.-L., Varki, N. M., and Varki, A. (1999) *Am. J. Pathol.* **155,** 461–472
37. Agrawal, B., Gendler, S. J., and Longenecker, B. M. (1998) *Mol. Med. Today* **4,** 397–403
38. Ho, J. J. L., and Kim, Y. S. (1995) *Int. J. Oncol.* **7,** 913–926
39. Choudhury, A., Singh, R. K., Moniaux, N., El-Metwally, T. H., Aubert, J.-P.and Batra, S. K. (2000) *J. Biol. Chem.* **275,** 33929–33936
40. Balagué, C., Audié, J.-P., Porchet, N., and Real, F. (1995) *Gastroenterology* **109,** 953–964
41. Buisine, M.-P., Devisme, L., Maunoury, V., Deschodt, E., Gosselin, B., Copin, M.-C., Aubert, J.-P., and Porchet, N. (2000) *J. Histochem. Cytochem.* **48,** 1657–1665
42. Zaretsky, J. Z., Sarid, R., Aylon, Y., Mittelman, L. A., Wreschner, D. H., and Keydar, I. (1999) *FEBS Lett.* **461,** 189–195
43. Lencer, W. I., Reinhardt, F. D., and Neutra, M. R. (1990) *Am. J. Physiol.* **21,** G96–G102
44. Hong, D.-H., Petrovics, G., Anderson, W. B., Forstner, J., and Forstner, G. (1999) *Am. J. Physiol.* **277,** G1041–G1047
45. Forstner, G. (1995) *Annu. Rev. Physiol.* **57,** 585–605
46. Velcich, A., Palumbo, L., Selleri, L., Evans, G., and Augenlicht, L. (1997) *J. Biol. Chem.* **272,** 7968–7976
47. Hiltunen, M. O., Koistinaho, J., Alhonen, L., Myöhänen, S., Marin, S., Kosma, V.-M., Pääkkönen, M., and Jänne, J. (1997) *Br. J. Cancer* **76,** 1124–1130
48. Yamamura, H., Masuda, H., Ikeda, W., Tokuyama, T., Takagi, M., Shibata, N., Tatsuta, M., and Takahashi, K. (1997) *J. Biochem.* **122,** 157–167
49. Schwienbacher, C., Gramantierr, L., Scelfo, R., Veronese, A., Calin, G. A., Bolondi, L., Croce, C. M., Barbanti-Brodzano, G., and Negrini, M. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 5445–5449
50. Jirtle, R. L., Sander, M., and Barrett, J. C. (2000) *Env. Health Perspect.* **108,** 271–278
51. Toyota, M., and Issa, J.-P. J. (2000) *Electrophoresis* **21,** 329–333
52. Tycko, B. (2000) *J. Clin. Invest.* **105,** 401–407
53. Baylin, S. B., and Herman, J. G. (2000) *Trends Genet.* **16,** 168–174
54. Toyota, M., Ahuja, N., Suzuki, H., Itoh, F., Ohe-Toyota, M., Imai, K., Baylin, S. B., and Issa, J.-P. J. (1999) *Cancer Res.* **59,** 5438–5442
55. Hanski, C., Riede, E., Gratchev, A., Foss, H. D., Böhm, C., Klussmann, E., Hummel, M., Mann, B., Buhr, H. J., Stein, H., Kim, Y. S., Gum, J., and Riecken, E. O. (1997) *Lab. Invest.* **77,** 685–695
56. Gum, J. R., Hicks, J. W., Gillepsie, A.-M., Carlson, E. J., Kömüves, L., Karnik, S., Hong, J. C., Epstein, C. J., and kim, Y. S. (1999) *Am. J. Physiol.* **276,** G666–G676
57. Gum, J. R., Kicks, J. W., and Kim, Y. S. (1997) *Biochem. J.* **325,** 259–267

Aberrant Expression of Human Mucin Gene*MUC5B* in Gastric Carcinoma and Cancer Cells: IDENTIFICATION AND REGULATION OF A DISTAL PROMOTER

Michaël Perrais, Pascal Pigny, Marie-Pierre Buisine, Nicole Porchet, Jean-Pierre Aubert and Isabelle Van Seuningen-Lempire

J. Biol. Chem. 2001, 276:15386-15396.
doi: 10.1074/jbc.M010534200 originally published online January 30, 2001
