Transcriptional activation of the murine *Muc5ac* mucin gene in epithelial cancer cells by TGF-β/Smad4 signalling pathway is potentiated by Sp1

Nicolas JONCKHEERE*, Maria VAN DER SLUIS†, Amélie VELGHE*, Marie-Pierre BUISINE*, Marjolein SUTMULLER†, Marie-Paule DUCOUROUBLE*, Pascal PIGNY*, Hans A. BÜLLER†, Jean-Pierre AUBERT*, Alexandra W. C. EINERHAND† and Isabelle VAN SEUNINGEN*[1]

*Unité INSERM U560, Place de Verdun, 59045 Lille Cedex, France, and †Erasmus Medical Center/Sophia Children's Hospital, Laboratory of Pediatrics, Dr. Molewaterplein 50, 3015 GE Rotterdam, The Netherlands

Changes in the expression of mucin genes in gastrointestinal cancers is thought to contribute to the development of the disease. In our laboratory we have shown previously that *MUC5AC* is aberrantly expressed in rectosigmoid villous adenomas. However, the regulatory mechanisms underlying that altered profile of expression is unknown. In order to study its regulation at the transcriptional level, we have isolated and characterized 5.5 kb of the 5′-flanking region of the mouse *Muc5ac* mucin gene. The promoter is flanked by a TATA box and a transcriptional start site is located 22 bp downstream of the TATA box. Analysis of the sequence showed a high density of binding sites for Smad4, an essential factor in the signalling cascade activated by TGF-β (transforming growth factor-β), and Sp1, an important factor in the regulation of *MUC5AC*. This led us to study *Muc5ac* regulation by TGF-β. We show that exogenous addition of TGF-β to the cells induces *Muc5ac* endogenous expression, promoter activity and Smad4 binding to the promoter. By co-transfection studies we show that Smad4 is essential for *Muc5ac* promoter activation and that it does not synergize with Smad2 or Smad3. By gel-retardation and co-transfection assays, we identified Sp1 and Sp3 as important regulators of *Muc5ac* expression and showed that Smad4 and Sp1 act in a co-operative manner to transactivate *Muc5ac* promoter activity. Altogether these results bring new insights into the molecular mechanisms of TGF-β-mediated up-regulation of *Muc5ac* and enhance our understanding as to how *Muc5ac* is regulated in certain pathologies of the gastrointestinal tract.

Key words: *Muc5ac*, mucin, Smad4, Sp1, transcription, transforming growth factor-β (TGF-β).

INTRODUCTION

Mucins have been postulated to be important molecules in maintaining epithelial homoeostasis in inflammatory diseases and cancer. Mucins are large O-glycoproteins expressed either as transmembrane proteins at the cell surface or as secreted oligomeric molecules to form a protective gel [1–4]. In the gastrointestinal tract, they play a cytoprotective role against acid and pepsin in the gastric juice and against deleterious effects of exogenous agents (pathogens, drugs) and against mechanical damage [5,6]. *MUC5AC* belongs to the family of secreted mucins that participate in mucus formation and is encoded by a gene located on the p15 arm of chromosome 11 within a cluster of four mucin genes along with *MUC2*, *MUC5B* and *MUC6* [7,8].

In normal adult, *MUC5AC* main territories of expression are the surface epithelium of the respiratory tract and stomach [9]. This expression is restricted to mucus-producing lung goblet cells and gastric pit cells. *MUC5AC* normal pattern of expression is altered in several epithelial diseases of the gastrointestinal tract. It is aberrantly expressed in Barrett's oesophagus [10,11], in gastric metaplasia in the duodenum [12], in colon adenoma and cancer [6,13]. In rectosigmoid villous adenoma, *MUC5AC* expression was found very early during the carcinogenetic sequence, in low-grade dysplasia, which makes it a valuable marker for recurrent patients [14]. Despite increasing amounts of data regarding its expression pattern in normal human tissues compared with that in disease [6,15], the precise biological role of *MUC5AC* as a key gene during sequential steps of carcinogenesis or in inflammatory processes has yet to be proven. Moreover, the molecular mechanisms that govern *MUC5AC* expression in the gastrointestinal tract are still largely unknown, thus their identification is necessary if one wants to better understand the role of MUC5AC in the pathologies of the gastrointestinal epithelium.

TGF-β (transforming growth factor-β) is a member of the superfamily of cytokines that affect a variety of cell types and elicit a wide array of cell-type-specific biological effects such as differentiation, migration, cell-cycle arrest, adhesion, extracellular matrix production and apoptosis [16]. TGF-β is also an agent involved in gastritis and development of gastric cancer [17], two pathologies in which *MUC5AC* expression is altered. TGF-β-induced signalling occurs when the TGF-β ligand binds to the type II receptor (TGF-β RII), which heterodimerizes with the type I receptor (RI). RI then phosphorylates receptor-activated Smads (Smad2, Smad3). Once activated Smad2 and Smad3 bind to Smad4 and this complex translocates to the nucleus whereupon transcription activation of the target gene occurs [18].

Recent isolation of *Muc1* and *Muc2* murine mucin genes as well as their regulatory regions has helped a great deal in defining their biological roles *in vivo* and the molecular mechanisms responsible for their regulation. The studies performed on murine *Muc1* mucin gene, which encodes a transmembrane mucin, and on $Muc1^{-/-}$ mice showed that it is overexpressed in most carcinomas, correlates with high metastatic potential and poor survival and participates in tumour progression [19]. More recently, the

Abbreviations used: EMSA, electrophoretic mobility shift assay; RT-PCR, reverse transcriptase PCR; TBST, Tris-buffered saline/0.2 % Tween 20; TGF-β, transforming growth factor-β; RACE, rapid amplification of cDNA ends; PAS, periodic acid–Schiff; NF-κB, nuclear factor κB; AP-1, activator protein 1; UTR, untranslated region.

[1] To whom correspondence should be addressed (e-mail isabelvs@lille.inserm.fr).

The nucleotide sequence of the pMS1 clone was submitted to the GenBank Nucleotide Sequence Database under accession number AF288076.

© 2004 Biochemical Society

promoter region of murine *Muc2*, which belongs to the secreted mucins, was characterized [20]. The authors showed, as for its human counterpart and other mucin genes, that Sp1 is an important factor in *Muc2* regulation [20]. Interestingly, the knockout mice for *Muc2* mucin gene allowed the authors to demonstrate for the first time a direct link between *Muc2* expression and tumour formation, as *Muc2*$^{-/-}$ mice started to develop intestinal adenomas that progressed to invasive adenocarcinoma and colorectal tumours as they were getting older [21]. From this work, the authors concluded that *Muc2* is involved in colon cancer and may be considered as a tumour-suppressor gene [21].

Murine *Muc5ac* mucin gene is partially characterized and part of its tandem repeat region was published by Shekels et al. [22]. Of interest, the authors showed that *Muc5ac* is located on murine chromosome 7; that is the syntenic chromosomal region corresponding to human chromosome 11. Thus it appears from this work and recent data released from the human and mouse genome databases (NCBI, MGI) that murine *Muc2*, *Muc5ac*, *Muc5b* and *Muc6* are clustered on murine chromosome 7, which means that this cluster of mucin genes is conserved throughout evolution [23].

Our aim is to better understand the transcriptional regulation of human and murine mucin genes in order to propose new therapeutic targets in epithelial diseases (inflammation and cancer) and better understand their role during embryonic development and differentiation of the gastrointestinal epithelium [8]. In this study, we have isolated, characterized and studied the regulation of the 5′-flanking region of murine *Muc5ac* mucin gene in order to obtain a base for future studies in animal models [6] and better understand *MUC5AC* biological role in the pathophysiology of the epithelium. We show for the first time that *Muc5ac* is regulated at the transcriptional level by TGF-β, an agent involved in gastritis and development of gastric cancer.

MATERIALS AND METHODS

Cloning and characterization of murine *Muc5ac* 5′-flanking region

The 5′-flanking region of murine *Muc5ac* mucin gene was isolated and characterized after screening a murine OLA 129 genomic DNA library made in λGEM12 (kindly provided by Dr A. Berns, The Netherlands Cancer Institute, Amsterdam, The Netherlands) with a probe spanning 750 bp of the human *MUC5AC* N-terminus (GenBank accession no. AF043909) [24]. Two positive plaques were identified and the corresponding DNA was isolated from the bacteriophages by using the Lambda DNA isolation kit (Qiagen) according to the manufacturer's protocol. Southern blot analysis and restriction mapping of both inserts identified a 5.5 kb fragment that hybridized with 750 bp of the human *MUC5AC*. The positive 5.5 kb fragment was subsequently cloned into the *Sst*I site of pBluescript II SK(+/−) (Stratagene), which resulted in clone pMS1. That fragment was then digested by *Apa*I and *Pst*I to raise smaller fragments pMS2-pMS7. The clones pMS1-pMS7 were sequenced by Eurogentec (Seraing, Belgium) using T3 and T7 primers, and in the laboratory using the DYEnamic ET* Terminator Cycle Sequencing Kit, with fluorescently labelled nucleotides (Amersham Biosciences), according to the manufacturer's protocol. Sequence reactions were analysed on an ABI-Prism 310 Genetic analyser (Perkin Elmer Applied Biosystems). The obtained sequences were aligned with the BioEdit sequence alignment editor. The derived 5.5 kb nucleotide and amino acid sequences were compared with known sequences using NCBI-Blast database. In addition, the sequence was analysed with PC/Gene software (Intelligenetics, Mountain View, CA, U.S.A.), Infobiogen database and MatInspector V2.2 and Alibaba2 software based on the Genomatix database to determine the location of putative transcription factor binding sites [25]. The nucleotide sequence of pMS1 clone was submitted to GenBank under the accession no. AF288076.

Determination of 5′-end of Muc5ac mRNA by RACE (rapid amplification of cDNA ends) PCR

The transcription start site was determined by 5′-RACE using the 5′/3′ RACE kit (Roche Molecular Biochemicals) according to the manufacturer's instructions. RACE PCR was conducted on mouse total gastric RNA (2 μg) using 5′-CAGGGAAGTAGA-AGACCTGTCC-3′ as first primer and 5′-TGTGCATTGGCT-GCAGGCCAG-3′ as nested primer for reverse transcription and amplification. Resulting amplified PCR products were cloned directly into the TA cloning vector (Invitrogen) and sequenced on an automatic LI-COR sequencer (ScienceTech, France) with T7 and RM13 primers as described below.

Animals

Adult specified pathogen-free Balb/c mice, obtained from Harlan (Zoetermeer, The Netherlands), were killed by cervical dislocation. The stomach was removed and fixed in 4 % paraformaldehyde in PBS and subsequently processed for light microscopy as described previously [26]. The animal experiments were performed with the approval of the Animal Studies Ethics Committee of the Erasmus MC (Rotterdam, The Netherlands).

Histology

Sections of mouse stomach tissue (5 μm thick) were routinely stained with haematoxylin and eosin to study the morphology, or stained with Alcian Blue/PAS (periodic acid–Schiff) reagent to stain for acidic and neutral mucins, respectively. Immunolocalization of mouse *Muc5ac* was carried out as described previously [11] using 45M1 monoclonal antibody (Novocastra).

Probe preparation for *in situ* hybridization

Total RNA was extracted from murine stomach using TRIzol (Life Technologies) following the manufacturer's protocol. RNA (1 μg) was transcribed at 42 °C into cDNA using Moloney murine leukaemia virus reverse transcriptase (Promega) in a total volume of 20 μl following the manufacturer's protocol. This was followed by a PCR reaction using 1 μl of cDNA as template in 10 mM Tris/HCl buffer, pH 8.4, containing 50 mM KCl, 2 mM $MgCl_2$, 0.01 % gelatin, 1 unit of *Taq* polymerase (Eurogentec), 0.2 mM dNTPs and 10 pmol of each primer. The primers were 5′-CCAATTGGCTAGATGGCAGT-3′ and 5′-AGATCAA-ACCCTCCTCTCG-3′, which correspond to nucleotides 374–394 and 552–572 of murine *Muc5ac* (GenBank accession no. L42292) [22]. The PCR sample (20 μl) was first denatured at 96 °C for 5 min, followed by 30 cycles at 96 °C for 1 min, 60 °C for 1 min and 72 °C for 1 min, with a final extension step at 72 °C for 2 min. The resulting 199 bp PCR product was isolated using the Qiagen gel extraction kit and ligated into the *Eco*RI site of pBluescript SK vector and subsequently sequenced. The digoxigenin-11-UTP-labelled sense and antisense *Muc5ac* RNA probes were prepared according to the manufacturer's protocol (Roche Molecular Biochemicals).

In situ hybridization

Tissue sections were deparffinized with xylene and rehydrated through RNase-free ethanol/water solutions. The non-radioactive

© 2004 Biochemical Society

Table 1 Sequences of the pairs of oligonucleotides used in PCR to produce deletion mutants covering the *Muc5ac* promoter

*Sac*I (GAGCTC) and *Mlu*I (ACGCGT) sites (italicized) were added at the end of the primers to direct subcloning into the pGL3 basic vector.

Position in the promoter	Oligonucleotide sequence	Orientation
− 199/+ 3	5′-CGC *GAG CTC* TGG GGG AGC CTC AGG GAA-3′	Sense
	5′-CGC *ACG CGT* GAA AGA CTC TAG TCA CCA-3′	Antisense
− 199/+ 132	5′-CGC *GAG CTC* TGG GGG AGC CTC AGG GAA-3′	Sense
	5′-CGC *ACG CGT* ACT TTG AGT CTT ACC TGT GCA-3′	Antisense
− 376/+ 3	5′-CGC *GAG CTC* CTC CTC TCC AGG GAA ATG-3′	Sense
	5′-CGC *ACG CGT* GAA AGA CTC TAG TCA CCA-3′	Antisense
− 376/+ 132	5′-CGC *GAG CTC* CTC CTC TCC AGG GAA ATG-3′	Sense
	5′-CGC *ACG CGT* ACT TTG AGT CTT ACC TGT GCA-3′	Antisense
− 1021/+ 3	5′-CGC *GAG CTC* CTC TCT TTC ACA CAC ACA-3′	Sense
	5′-CGC *ACG CGT* GAA AGA CTC TAG TCA CCA-3′	Antisense
− 1021/+ 132	5′-CGC *GAG CTC* CTC TCT TTC ACA CAC ACA-3′	Sense
	5′-CGC *ACG CGT* ACT TTG AGT CTT ACC TGT GCA-3′	Antisense
− 1171/+ 3	5′-CGC *GAG CTC* CAC TCC TGT GAT GTG TGA-3′	Sense
	5′-CGC *ACG CGT* GAA AGA CTC TAG TCA CCA-3′	Antisense
− 1171/+ 132	5′-CGC *GAG CTC* CAC TCC TGT GAT GTG TGA-3′	Sense
	5′-CGC *ACG CGT* ACT TTG AGT CTT ACC TGT GCA-3′	Antisense
− 1021/− 828	5′-CGC *GAG CTC* CTC TCT TTC ACA CAC ACA-3′	Sense
	5′-CGC *ACG CGT* AGA GAG GTC AAA GCT TAA-3′	Antisense
− 1171/− 828	5′-CGC *GAG CTC* CAC TCC TGT GAT GTG TGA-3′	Sense
	5′-CGC *ACG CGT* AGA GAG GTC AAA GCT TAA-3′	Antisense

in situ hybridization was essentially carried out as described previously [27]. Briefly, the riboprobes were diluted in hybridization solution [50 % deionized formamide (v/v), 10 % dextran sulphate (w/v), 2 × SSC (where 1 × SSC is 0.15 M NaCl/0.015 M sodium citrate), 1 × Denhardt's solution, 1 μg/ml tRNA and 250 μg/ml herring sperm DNA] to a concentration of 100 ng/ml, hybridized overnight at 55 °C in a humid chamber. Post-hybridization washes were performed at 45 °C using the following steps: 50 % formamide (v/v) in 2 × SSC, 50 % formamide (v/v) in 1 × SSC and 0.1 × SSC. A 15 min incubation with RNase T1 (2 units/ml in 1 mM EDTA in 2 × SSC) at 37 °C was followed by washes of 0.1 × SSC at 45 °C and 2 × SSC at room temperature. The digoxigenin-labelled hybrids were detected by incubation with anti-digoxigenin (Fab; 1:2000) conjugated to alkaline phosphatase for 2.5 h at room temperature. Thereafter, sections were washed in 0.025 % (v/v) Tween in Tris-buffered saline, pH 7.5. For staining, sections were layered with detection buffer (0.1 M Tris/HCl, pH 9.5, containing 0.1 M NaCl and 0.05 M $MgCl_2$) containing 0.33 mg/ml 4-Nitro Blue Tetrazolium chloride, 0.16 mg/ml 5-bromo-4-chloro-3-indolyl phosphate, 8 % (v/v) polyvinyl alcohol (31 000–50 000 Da; Aldrich) and 1 mM levamisol (Sigma). Development of the reaction was performed overnight in the dark and was stopped when the desired intensity of the resulting blue precipitate was reached. Finally, sections were washed in distilled water and mounted with Aquamount improved (Gurr, Brunschwig, Germany).

Muc5ac-pGL3 deletion mutant constructions

The *Muc5ac*-pGL3 deletion mutants that cover 1.2 kb of the promoter were constructed into pGL3 Basic vector (Promega) using a PCR-based method as described previously [28,29]. PCR reactions were carried out on pMS1 clone in order to subclone the promoter region of *Muc5ac*. PCR products were then subcloned into pCR2.1 vector (Invitrogen) before subcloning into *Sac*I–*Mlu*I sites of the promoterless pGL3 Basic vector. Internal deletion mutants were generated by PCR using pairs of primers bearing specific restriction sites at their 5′ and 3′ ends (Table 1). PCR products were digested, gel-purified (Qiaquick gel extraction kit; Qiagen) and subcloned into the pGL3 Basic vector that had been previously cut with the same restriction enzymes. All clones were sequenced on both strands on an automatic LI-COR sequencer using infrared-labelled RV3 and GL2 primers (Promega). Plasmids used for transfection studies were prepared using the Endofree plasmid Mega kit (Qiagen).

Cell culture

Murine rectal cancer cell line CMT-93 was a kind gift of Dr D. Podolsky (Massachusetts General Hospital, Boston, MA, U.S.A.). CMT-93 cells were cultured in Dulbecco's modified essential medium containing 10 % fetal bovine serum, 4 mM L-glutamine, penicillin (100 units/ml) and streptomycin (100 μg/ml). IEC-6 cells were purchased from the ECACC (European Collection of Animal Cell Cultures). This cell line was established from rat small-intestine crypt cells and was cultured in Dulbecco's modified essential medium containing 5 % fetal bovine serum, 2 mM L-glutamine, 10 μg/ml insulin, 50 units/ml penicillin and 50 μg/ml streptomycin. HCT116-Smad4$^{+/+}$ and HCT116-Smad4$^{-/-}$ cells were a kind gift of Dr A. Atfi (INSERM U482, Paris, France). Cells were cultured in Dulbecco's modified essential medium supplemented with 2 mM glutamine and 10 % fetal calf serum. Human gastric cancer cell line KATO-III was cultured as described previously [30]. All cells were cultured at 37 °C in a humidified 5 % CO_2 water-jacketed incubator. To study TGF-β effect, cells were incubated for 24 h with TGF-β (recombinant human TGF-β; 10 ng/ml). All reagents were from Sigma unless otherwise indicated.

RT-PCR (reverse transcriptase PCR)

Total RNAs from cultured cells and mouse tissues were prepared using the QIAamp RNA blood mini-kit and midi-kit (Qiagen), respectively. Total RNA (1.5 μg) was used to prepare first-strand cDNA (Advantage™ RT-for-PCR kit; Clontech). PCR was performed on 2 μl of cDNA using specific pairs of primers as follows: *Muc5ac* forward primer, 5′-GAGGGCCCAGTGAGCATCTCC-3′; *Muc5ac* reverse primer, 5′-TGGGACAGCAGCAGTATTCAGT-3′ (accession number AJ010792). β-Actin was used as an internal control: mouse β-actin forward primer, 5′-GTGGGCCGCTCTAGGCACCA-3′; mouse β-actin reverse primer, 5′-TGGCCTTAGGGTGCAGGGGG-3′ (accession number M12481). *Muc5ac* and β-actin PCR product sizes are 361 and 241 bp, respectively. Mouse TGF-βRII forward primer, 5′-CGTGTGGAGGAAGAACAACA-3′; reverse primer, 5′-TCTCAAACTGCTCTGAGGTG-3′ (accession number S69114). The PCR product was 560 bp long. Rat β-actin forward primer, 5′-ATATCGCTGCGCTCGTCGTCGACAA-3′; rat β-actin reverse primer, 5′-AACACAGCCTGGATGGCTACGTACAT-3′ (accession number V01217). PCR reactions were carried out in 50 μl final solutions as described in [31]. Annealing temperature was 58 °C. PCR products were analysed on 1.5 % ethidium bromide-stained agarose gels run in 1 × Tris/borate/EDTA buffer. A 100 bp DNA ladder was purchased from Amersham Biosciences.

Transfections

Transfections and co-transfections experiments were performed using Effectene® reagent (Qiagen) as described previously using 1 μg of *Muc5ac*-pGL3 deletion mutants [28]. Total cell extracts were prepared after a 48 h incubation at 37 °C using 1 × Reagent Lysis Buffer (Promega) as described in the manufacturer's instruction manual. Luciferase activity (20 μl) was measured on

© 2004 Biochemical Society

Table 2 Sequences and positions within the promoter of *Muc5ac* of the sense oligonucleotides used for EMSAs

Antisense oligonucleotides were also synthesized and annealed to the sense oligonucleotides to produce double-stranded DNA. Putative binding sites in the sequences are italicized.

Probe	Position of putative binding site	Sequence
− 57/− 34	Sp1 (− 51/− 42)	5′-GTCACT*GGGGCTGGAG*CCAGCTCT-3′
− 83/− 56	Sp1/CACCC (− 76/− 65)	5′-AGACCTG*CTCCACCCACCC*ACGTGAAG-3′
− 113/− 81	Sp1 (− 100/− 90)	5′-GCCACTGTTTACC*TGGGTGAGGGG*AACCACAGA-3′
− 92/− 71	Smad (− 84/− 80)	5′-GGGAACCA*CAGAC*CTGCTCCAC-3′
− 467/− 444	Smad (− 458/− 454)	5′-CTCACTGGG*GTCTG*GGAACTTGAA-3′
− 552/− 529	Smad (− 538/− 534)	5′-TCTGTCCATCCCAG*CAGAC*ATGAA-3′
− 645/− 622	Smad (− 641/− 637)	5′-CCTG*GTCTG*GGCAAAGGTCCATGC-3′
− 1146/− 1126	Smad (− 1131/− 1128)	5′-TCTCTCTCTCTCCTC*TCTG*TC-3′

a TD 20/20 luminometer (Turner Design). Total protein content in the extract (4 μl) was measured using the bicinchoninic acid method in 96-well plates as described in the manufacturer's instruction manual (Pierce, Bezons, France). In co-transfection experiments, 1 μg of the deletion mutant of interest was transfected with 0.25 μg of the expression plasmid encoding the transcription factor of interest. Results were expressed as fold activation of luciferase activity in samples co-transfected with the transcription factor of interest compared with the control co-transfected with the corresponding empty vector.

Nuclear extract preparation

Nuclear extracts from the different cells were prepared as described by Van Seuningen et al. [32], and kept at − 80 °C until use. Protein content (2 μl of cell extracts) was measured using the bicinchoninic acid as described above.

Western blotting

Nuclear proteins (20 μg) and prestained molecular-mass markers (8 μl; Life Technologies) were loaded on a SDS/PAGE gel (10 %) and electrophoresed for 1 h 15 min at 35 mA until Bromophenol Blue reached the bottom of the gel. The polyacrylamide gel was then electrotransferred on a 0.45 μm PVDF membrane (Millipore) for 1 h at 100 V using a Bio-Rad transfer apparatus (Bio-Rad, Ivry-sur-Seine, France). Immunostaining was performed as follows: the PVDF membrane was incubated with Tris-buffered saline/0.2 % Tween 20 (TBST) containing 20 % non-fat dry milk (w/v) for 2 h at room temperature, then washed for 3 × 5 min with TBST. Incubation with polyclonal anti-Sp1 PEP2 (sc-59) and anti-Sp3 (sc-644) antibodies (Santa Cruz Biotechnology) at 1:1275 dilution in TBST was carried out for 2 h at room temperature with shaking. The membrane was then washed for 3 × 5 min at room temperature with TBST before adding alkaline-phosphatase-conjugated anti-rabbit IgGs (Promega) for 30 min at room temperature. The membrane was washed for 5 min with TBST and for 5 min with substrate buffer (0.1 M Tris/HCl, pH 9.5, containing 0.1 M NaCl and 0.05 M $MgCl_2$). Staining was developed by incubating the membrane with 10 ml of substrate buffer containing 0.33 mg/ml 4-Nitro Blue Tetrazolium chloride and 0.16 mg/ml 5-bromo-4-chloro-3-indolyl phosphate. The enzymic reaction was stopped by soaking the membrane in distilled water.

Oligonucleotides and DNA probes

The sequences of the oligonucleotides used for gel-shift assays are indicated in Table 2. They were synthesized by MWG-Biotech (Ebersberg, Germany). Equimolar amounts of single-stranded oligonucleotides were annealed and radiolabelled using T4 polynucleotide kinase (Promega) and [γ-^{32}P]dATP. Radiolabelled probes were purified by chromatography on a Bio-Gel P-6 column (Bio-Rad).

EMSA (electrophoretic mobility shift assay)

Incubation of nuclear proteins (8 μg) with radiolabelled probes, supershift analyses and unlabelled competition ('cold competition') experiments were carried out as in [31]. Anti-Sp1, anti-Sp3, anti-Smad2, anti-Smad4 and anti-NF-κB (nuclear factor κB) p65 antibodies and consensus Smad4 oligonucleotide were purchased from Santa Cruz Biotechnology (Tebu, France). Reactions were stopped by adding 2 μl of loading buffer. Samples were loaded on to a 4 % non-denaturing polyacrylamide gel and electrophoresis conditions were as described in [29]. Gels were vacuum-dried and autoradiographed overnight at − 80 °C.

RESULTS

Cloning and characterization of murine *Muc5ac* mucin gene 5′-flanking region

In order to obtain the mouse *Muc5ac* 5′-flanking region, a mouse genomic library in λGEM12 was screened with a 750 bp DNA fragment encoding the N-terminal region of human *MUC5AC* [24]. The two positive plaques contained inserts of approx. 18 kb. After restriction mapping and Southern blot analysis, seven clones (pMS1-pMS7) were isolated and sequenced (see the Materials and methods section). Sequencing of the 5.5 kb pMS1 clone indicated that it contains the promoter region, the first three exons and the first three introns of the mouse *Muc5ac* mucin gene (Figure 1A). Sequence analysis of the pMS1 clone and alignment to human and rat *MUC5AC* 5′-flanking regions revealed a translational start codon at position 3211 identical with rat *Muc5ac* first codon and three short exons encoding a total of 72 amino acids with 52 % identity to human MUC5AC and 79 % identity with rat Muc5ac (Figure 1B).

Determination of the transcription start site of *Muc5ac*

To determine the transcription start site, 5′-RACE assays were performed on mouse gastric RNA using two nested antisense primers selected at about 300 and 70 bp downstream of the presumed start codon, respectively, based on comparison with the human *MUC5AC* gene [33]. The sequences showed 100 % homology when aligned with the 5′-flanking sequence of mouse *Muc5ac*. The 5′-end of the longest product (117 bp) corresponded to a thymidine residue located 33 nucleotides upstream of the first ATG and 22 nucleotides downstream of the TATA box (Figure 2).

© 2004 Biochemical Society

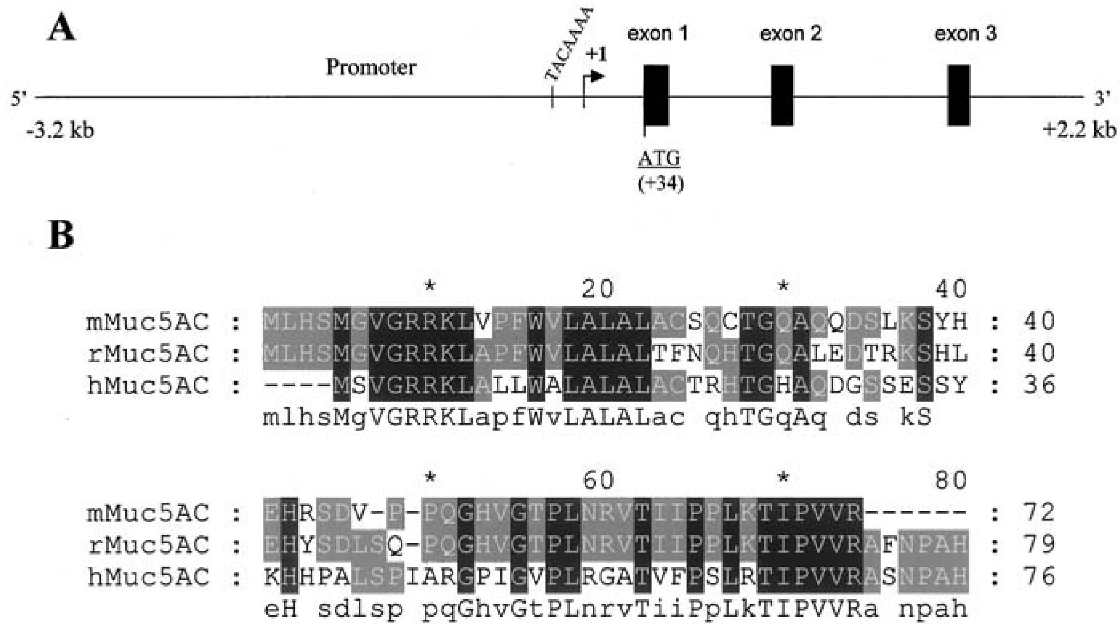

Figure 1 Isolation and characterization of a 5.5 kb genomic DNA fragment containing the 5′-flanking region of murine *Muc5ac* mucin gene

(**A**) Schematic representation of the organization of the 5′-flanking region of *Muc5ac* gene showing the promoter, the first three exons, ATG and transcription initiation site locations. (**B**) Alignment of deduced N-terminal amino acid sequences of mouse (mMuc5ac), rat (rMuc5ac) and human (hMuc5ac) MUC5AC peptides. Black boxes indicate conserved amino acid residues in the three species and grey boxes amino acid residues identical in two different species. The deduced consensus sequence is indicated at the bottom.

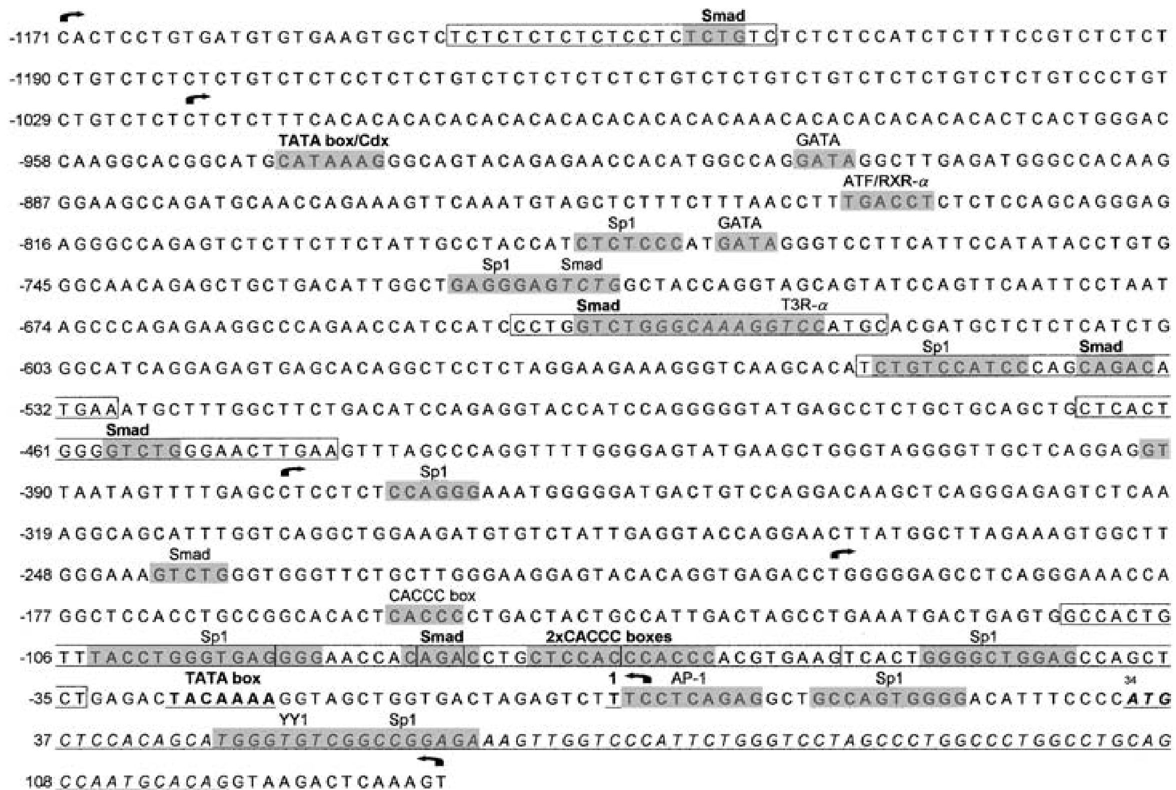

Figure 2 Sequence of the 5.5 kb genomic fragment containing the *Muc5ac* promoter

The transcription start site +1 (bold and underlined) is located 22 nucleotides downstream of the TATA box (underlined). The first ATG (+34) is bold, italicized and underlined. Nucleotides representing the first exon are italicized and underlined. Grey boxes indicate putative binding sites for transcription factors and boxed sequences indicate the sequences of oligonucleotides used in gel-shift assays. Arrows delineate the sequence of the deletion mutants used in this study.

© 2004 Biochemical Society

In conclusion, the *Muc5ac* transcription start site is a T residue located 22 nucleotides downstream of the TATA box.

Characterization of the promoter sequence of Muc5ac

The sequence found upstream of the transcription initiation site is characterized by the presence of a TATA box (TACAAAA) at −28/−22 (Figure 2). The first 150 nucleotides upstream of the TATA box are rich in GC and CACCC boxes and Sp1 binding sites. Putative binding sites for Smad factors are present throughout the promoter sequence. It is interesting to note that the Smad sites are always found in close vicinity or embedded within Sp1-binding sites and/or GC-rich sequences. More upstream (−944/−938) is found an AT-rich sequence that is a putative site representative of a TATA box. Note that this sequence may also bind the Cdx transcription factor. Two putative binding sites for GATA factors were found at −1001/−998 and −774/−771. Consensus binding sites for retinoid [ATF (activating transcription factor)/RXRα (retinoid X receptor), −837/−832] or thyroid hormone [T3R-α (thyroid hormone receptor α), −634/−626] receptors were also found. In the 5′-UTR region, which is 33 nucleotides long, putative binding sites for AP-1 (activator protein 1) and Sp1 were found at +2/+10 and +14/+23, respectively. In the first exon, a putative binding site for the YY1 transcription factor is found at +47/+56 adjoining an Sp1 binding site at +54/+63.

Expression of *Muc5ac* in mouse tissues

Expression of Muc5ac mRNA in mouse tissues was studied by RT-PCR and *in situ* hybridization. By RT-PCR, *Muc5ac* expression is only seen in stomach (Figure 3A). No expression was found in submaxillary glands, parotid glands, trachea, thymus, gallbladder, liver, small intestine, colon or kidney. *In situ* hybridization was performed to localize *Muc5ac* expression in the stomach. Alcian Blue/PAS staining of acidic mucins was found in mucus granules of surface epithelial cells (Figure 3C). Labelling with *Muc5ac* antisense probe showed that Muc5ac mRNA is only expressed in the surface epithelium of the stomach (Figure 3D). The specificity of the labelling was confirmed by the absence of signal when using a sense probe (Figure 3E) and by absence of signal in colon when using the antisense probe (Figure 3F). Immunohistochemical staining of a mouse gastric mucosa with 45M1 monoclonal anti-MUC5AC antibody confirmed the expression of Muc5ac mucin in surface gastric epithelial cells (Figure 3G). In conclusion, Muc5ac expression in normal mice is restricted to the surface epithelium of the stomach.

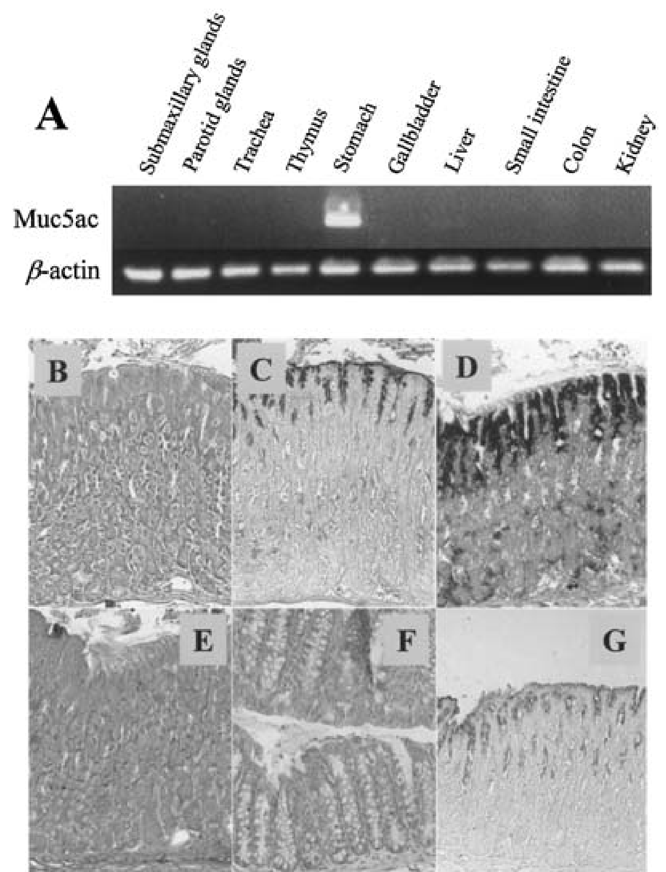

Figure 3 Expression of *Muc5ac* in mouse tissues by RT-PCR, *in situ* hybridization and immunohistochemistry

(**A**) *Muc5ac* (10 μl) and β-actin (2 μl) PCR products were separated on a 1.5 % agarose gel. *In situ* hybridization (**B**–**F**) and immunohistochemistry (**G**) studies. (**B**) Haematoxylin and eosin staining of gastric mucosa, (**C**) Alcian Blue and PAS staining of gastric mucosa, (**D**) staining of gastric mucosa with *Muc5ac* digoxigenin-labelled antisense RNA probe, (**E**) staining of gastric mucosa with *Muc5ac* digoxigenin-labelled sense probe, (**F**) staining of colon mucosa with *Muc5ac* digoxigenin-labelled antisense probe and (**G**) immunostaining of Muc5ac in gastric mucosa with monoclonal 45M1 antibody. Magnification, × 200.

Characterization of *Muc5ac* promoter activity

To study *Muc5ac* promoter activity we used two *Muc5ac*-expressing (CMT-93, IEC-6) and one *Muc5ac*-non-expressing (KATO-III) cell lines. Expression of *Muc5ac* in the murine rectal cancer cells (CMT-93, lane 3) and in the rat intestinal cell line (IEC-6, lane 6) is shown in Figure 4(A). Absence of MUC5AC expression in KATO-III cells was reported previously [8]. To define essential regions that drive transcription of *Muc5ac* promoter, 10 deletion mutants that cover 1.2 kb of the promoter were constructed in the promoterless pGL3 Basic vector (Figure 4B). Numbering refers to the transcription start site designated as +1. Deletion mutants −1021/−828 and −1171/−828 were made in order to check the functional activity of the distal TATA box found at −944/−939.

The luciferase results obtained in the three cell lines indicate that *Muc5ac* promoter activity is the strongest in *Muc5ac*-expressing cell lines (Figure 4C). In CMT-93 and IEC-6 cells, the highest luciferase activity was obtained with fragment −199/+3 (17–20-fold activation), which indicates that this region possesses essential positive regulatory elements that confer maximal activity to the promoter. In CMT-93 cells, the luciferase activity gradually decreases as the constructs include longer portions of the promoter's distal region. In IEC-6 cells, a strong decrease in luciferase activity is seen with fragment −376/+3 (6-fold activation) and this decrease is maintained in fragments −1021/+3 and −1171/+3. This indicates that inhibitory elements, active in IEC-6 cells, are present within the −376/−200 region of the promoter. Interestingly, one can note that the fragments, in which the +3/+132 region was added, have a decreased luciferase activity of about 50 to 100 %. The inhibitory activity was found in the three cell lines. That region includes the 5′-UTR (untranslated region) and first exon in which a binding site for YY1 repressor (+47/+56) is clustered between two Sp1-binding sites at +29/+38 and +54/+63, respectively. The −1021/+3 and −1171/+3 mutants that contain the distal TATA box located at −944/−939 are not active in the cell lines tested. In conclusion, essential regulatory elements for the basal activity of the promoter are found within the −199/+3 proximal region of *Muc5ac* promoter and negative elements are present within the 5′-UTR.

© 2004 Biochemical Society

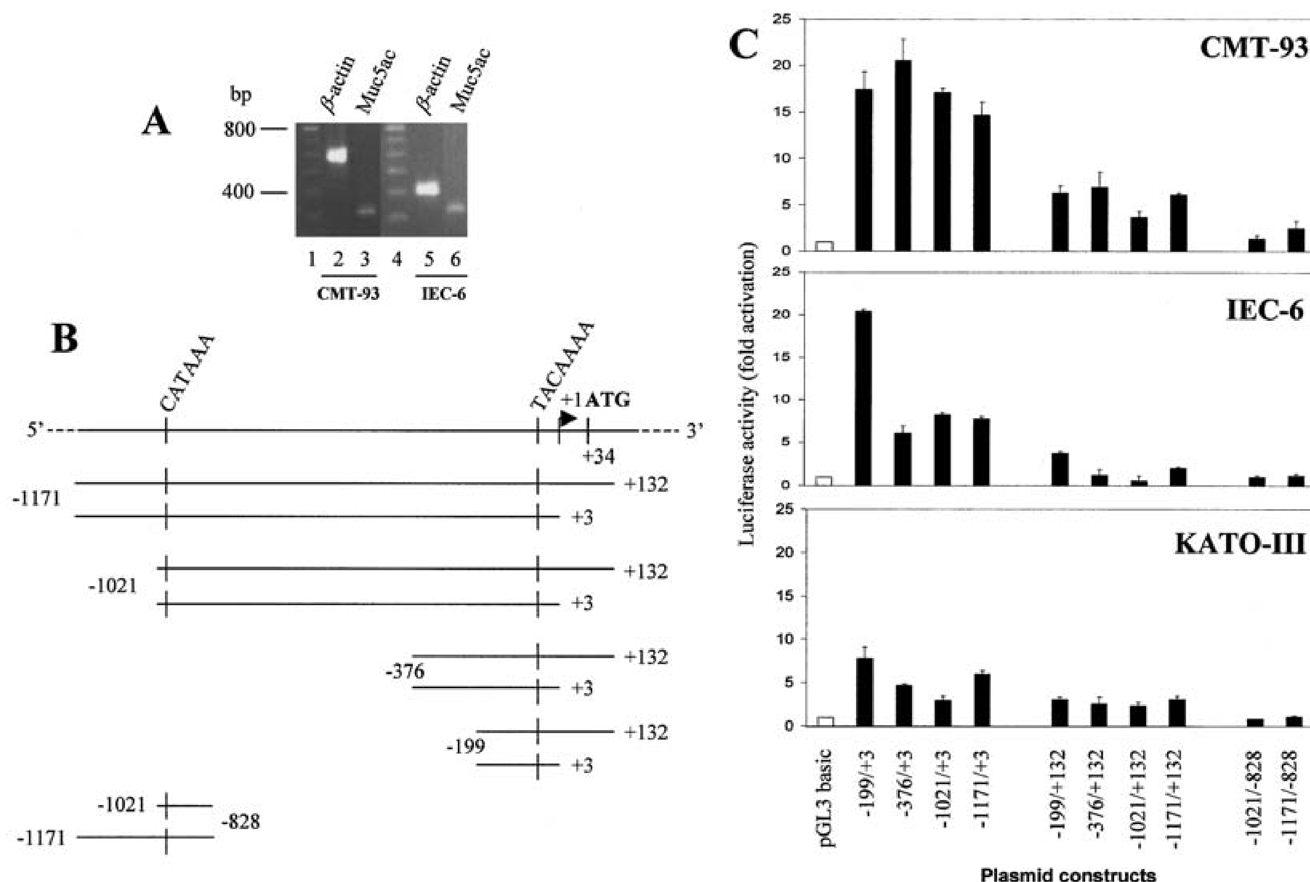

Figure 4 Characterization of *Muc5ac* promoter activity in CMT-93, IEC-6 and KATO-III cell lines by transient transfection

(**A**) Expression of *Muc5ac* mucin gene in CMT-93 (lane 3) and IEC-6 (lane 6) cells by RT-PCR. Mouse (lane 2) and rat (lane 5) β-actins (2 μl) were used as internal controls. Lanes 1 and 4, 100 bp DNA ladder. (**B**) Schematic representation of the different deletion mutants used to study *Muc5ac* promoter activity. Numbering refers to transcription initiation site designated as +1. (**C**) Luciferase activity diagrams showing *Muc5ac* promoter activity in murine rectal CMT-93, rat colon IEC-6 and human gastric KATO-III cells. Results are expressed as fold activation of luciferase activity of the deletion mutant of interest compared with the activity of empty pGL3 Basic vector (white bars). Results represent the means ± S.D. obtained in triplicate in three separate experiments.

Regulation of *Muc5ac* promoter by TGF-β and Smads transcription factors

The high number of putative binding sites for Smad transcription factors within the *Muc5ac* promoter (see Figure 2) supports a regulatory role for TGF-β on *Muc5ac* promoter activity. Smads are the main factors activated by TGF-β [18]. For this reason, and because it was previously shown that *MUC5AC* mucin gene expression is altered in inflammatory pathologies of the epithelium in which TGF-β is implicated [34], we undertook to study the regulation of *Muc5ac* promoter by TGF-β and Smad factors. TGF-β sends intracellular signals after binding to TGF-β RII on the cell membrane. Thus before studying TGF-β signalling pathway on *Muc5ac* expression in CMT-93 cells, we checked whether the cells expressed TGF-β RII. As shown in Figure 5(A), TGF-β RII mRNAs are highly expressed in CMT-93 cells. Treatment of CMT-93 cells with TGF-β substantially induced the amount of *Muc5ac* mRNA in the cells (Figure 5B). Identification of TGF-β-responsive elements within the *Muc5ac* promoter was then tested in transfection studies in which transfected cells were treated with TGF-β under the same conditions (Figure 5C). The luciferase result indicates that exogenous TGF-β induces the activity of the fragments covering the −1021/+3 region (2.0-, 2.7- and 1.8-fold, respectively). Induction was lost when the longest fragment (−1171/+3) was used. These results indicate that TGF-β-responsive elements are present within the −1021/+3 region of the promoter of *Muc5ac*, which contains seven putative Smad-binding sites (CAGAC).

In order to identify Smad-binding sites, EMSAs were performed with double-stranded oligonucleotides representative of the Smad-binding sites found at −538/−534, −641/−637, −1131/−1128, −458/−454 and −84/−80, respectively (Table 2). Incubation of the radiolabelled probes with CMT-93 nuclear extracts resulted in one shifted band characteristic of Smad–DNA complexes (Figure 6A, lanes 2, 7 and 12) when compared with the mobility of the Smad4 consensus radiolabelled probe (Figure 6A, lane 17). The specificity of the complexes were confirmed by total disappearance of the shifted bands when unlabelled competition was performed (Figure 6A, lanes 3, 8, 13 and 18). Involvement of Smad4 in the DNA–protein shifted complex was then proven by inhibition of complex formation upon addition of Smad4 antibody (Figure 6A, lanes 5, 10 and 14), whereas no effect was observed upon addition of Smad2 antibody (Figure 6A, lanes 4, 9, and 15). The same result was obtained with −552/−529 and −92/−71 probes (results not shown). Induction of Smad4 binding by TGF-β was then tested on nuclear extracts from TGF-β-treated cells. Increase of Smad4 binding to the Smad *cis*-elements was indeed observed with nuclear extracts from TGF-β-treated cells (Figure 6A, lane 21) when compared with untreated cells (Figure 6A, lane 20). These

© 2004 Biochemical Society

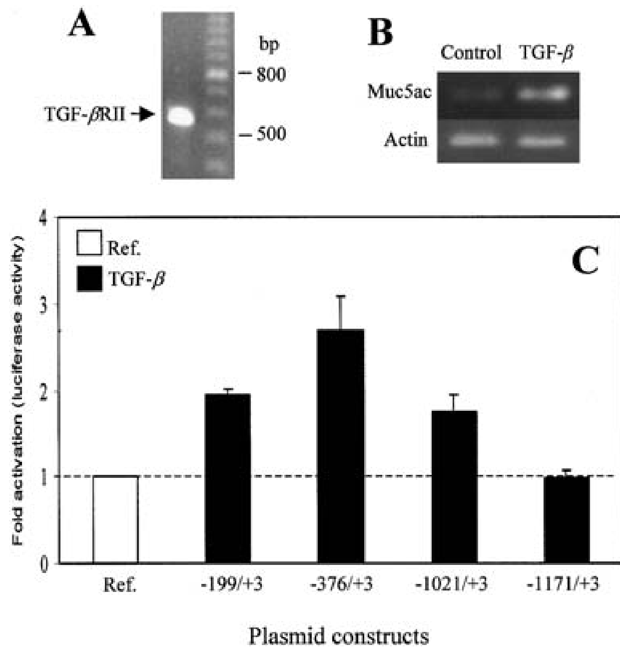

Figure 5 Regulation of *Muc5ac* mRNA expression and promoter activity by exogenous TGF-β in CMT-93 cells

(**A**) Expression of TGF-β RII in CMT-93 cells by RT-PCR. 10 μl of the PCR product was loaded. (**B**) Expression of *Muc5ac* mRNA by RT-PCR in untreated cells (control) and TGF-β-treated (TGF-β) cells. (**C**) Luciferase diagram showing effect of TGF-β on *Muc5ac* promoter activity. Results are expressed as fold activation of luciferase activity in samples treated with TGF-β (black bars) compared with the control (Ref., untreated cells, white bar).

studies indicate that Smad4 binds to five cognate *cis*-elements at −84/−80, −458/−454, −538/−534, −641/−635 and −1131/−1128 within the TGF-β-inducible region of *Muc5ac* promoter.

In order to show whether the TGF-β-responsive region that binds Smad4 is indeed regulated by Smad factors, we then performed co-transfection experiments in which *Muc5ac*-pGL3 deletion mutants were co-transfected with Smad4 alone or in combination with expression vectors encoding Smad2 or Smad3, which are co-factors of Smad4. As shown in Figure 6(B), Smad2 induces *Muc5ac* promoter activity with the strongest effect on fragment −376/+3 (10-fold). Smad3 transactivation of *Muc5ac* promoter is confined to the −376/+3 region of the promoter and is not as strong (4–5-fold activation, fragments −199/+3 and −376/+3). The transactivating effect of Smad3 is lost when co-transfected with longer fragments of the promoter (−1021/+3, −1171/+3). Smad4, as expected, strongly transactivates *Muc5ac* promoter activity throughout the sequence (12–22-fold activation). Co-transfection experiments in the presence of Smad2 and Smad4 or Smad3 and Smad4 did not lead to synergistic activation of the promoter (results not shown). The involvement of Smad4 in up-regulating *Muc5ac* transcription was then confirmed by comparing *Muc5ac* promoter activity in a cell line that either constitutively expresses active Smad4 (HCT116-Smad4$^{+/+}$) or is mutated for Smad4 (HCT116-Smad4$^{-/-}$; Figure 6C). The luciferase diagram shows that *Muc5ac* promoter activity (fragment −1021/+3) is three times more active in cells expressing Smad4 (Figure 6C, black bars) compared with cells mutated for Smad4 (Figure 6C, white bars). The same result was obtained with fragment −376/+3 (results not shown).

From these studies, it can be concluded that Smad4 is an activator of *Muc5ac* transcription, that TGF-β-responsive Smad4-binding sites are present throughout the promoter and that Smad2 and Smad3 factors do not act in synergy with Smad4 to induce *Muc5ac* transcription.

Role of Sp1 and Sp3 in the regulation of *Muc5ac* promoter

Since we showed that Smad2 and Smad3 are not the partners of Smad4 to activate *Muc5ac* transcription, we looked for other partners. Interestingly, when we looked at the location of the Smad4-binding sites in the promoter, we noticed that they were often either embedded in or neighbouring Sp1 *cis*-elements, and those transcription factors are known to synergize to activate transcription of many genes. Before testing the synergistic effect between Smad4 and Sp1, we studied the regulation of *Muc5ac* promoter by Sp1 and Sp3, since the proximal region is GC-rich and Sp1/Sp3 are important regulators of mucin gene expression [8].

EMSAs studies were performed with nuclear extracts from CMT-93 cells in which we checked the presence of Sp1 (96 kDa) and Sp3 (two isoforms, 100 and 60 kDa) proteins by Western blotting (Figure 7A). Three double-stranded radiolabelled probes (Table 2), each containing a putative Sp1 binding site (−113/−81, −57/−34) and CACCC boxes (−83/−56), were tested. Incubation of the −113/−81 probe with CMT-93 nuclear proteins did not produce any shift. Incubation with −57/−34 probe produced a strong retarded band that could not be supershifted in the presence of anti-Sp1 or anti-Sp3 antibodies (results not shown). On the other hand, incubation of the −83/−56 probe (Figure 7B), which contains CACCC box-binding sites, with CMT-93 nuclear proteins produced three shifted complexes (Figure 7B, lane 2, arrows). Specificity of these complexes was confirmed by complete inhibition of complex formation when unlabelled competition was performed with a 50× excess of the unlabelled probe (Figure 7B, lane 3). Complex 1 was completely supershifted upon addition of specific anti-Sp1 antibody in the reaction mixture (Figure 7B, lane 4, SS Sp1), whereas complex 2 was specifically supershifted upon addition of anti-Sp3 antibody (Figure 7B, lane 5, SS Sp3). No supershift was observed upon addition of irrelevant NF-κB p65 antibody (Figure 7B, lane 6). In conclusion, Sp1 and Sp3 bind to the same *cis*-element located at −76/−65 within the *Muc5ac* proximal promoter.

To study the role of Sp1 and Sp3 in regulating *Muc5ac* transcription, co-transfection experiments were carried out in CMT-93 cells in the presence of an expression vector encoding either Sp1 (Figure 7C, black bars, pCMV4-Sp1) or Sp3 (Figure 7C, white bars, pCMV4-Sp3). Sp1 strongly transactivates both the proximal (fragments −199/+3 and −376/+3, 8–9-fold activation) and distal (fragments −1021/+3 and −1171/+3, 10–15-fold activation) regions of the promoter. Addition of the 5′-UTR to the promoter construct −1171/+3 (=construct −1171/+132) inhibited the transactivating effect of Sp1. Sp3 only has a mild transactivating effect on the distal part of the promoter (2–3-fold activation). Again, addition of the 5′-UTR inhibited that effect. Altogether, these studies indicate that Sp1 is a strong transactivator of *Muc5ac* promoter activity in CMT-93 cells.

Co-operation between Sp1 and Smad4 to activate *Muc5ac* promoter

Having shown that both Smad4 and Sp1 were strong activators of *Muc5ac* transcription in CMT-93 cells, we then undertook to look at their co-operative effect on the promoter. Co-transfections in the presence of Smad4 and Sp1 were performed on fragments −199/+3, −376/+3 and −1171/+3 (Figure 8). No synergistic effect was seen on the shortest fragment −199/+3

© 2004 Biochemical Society

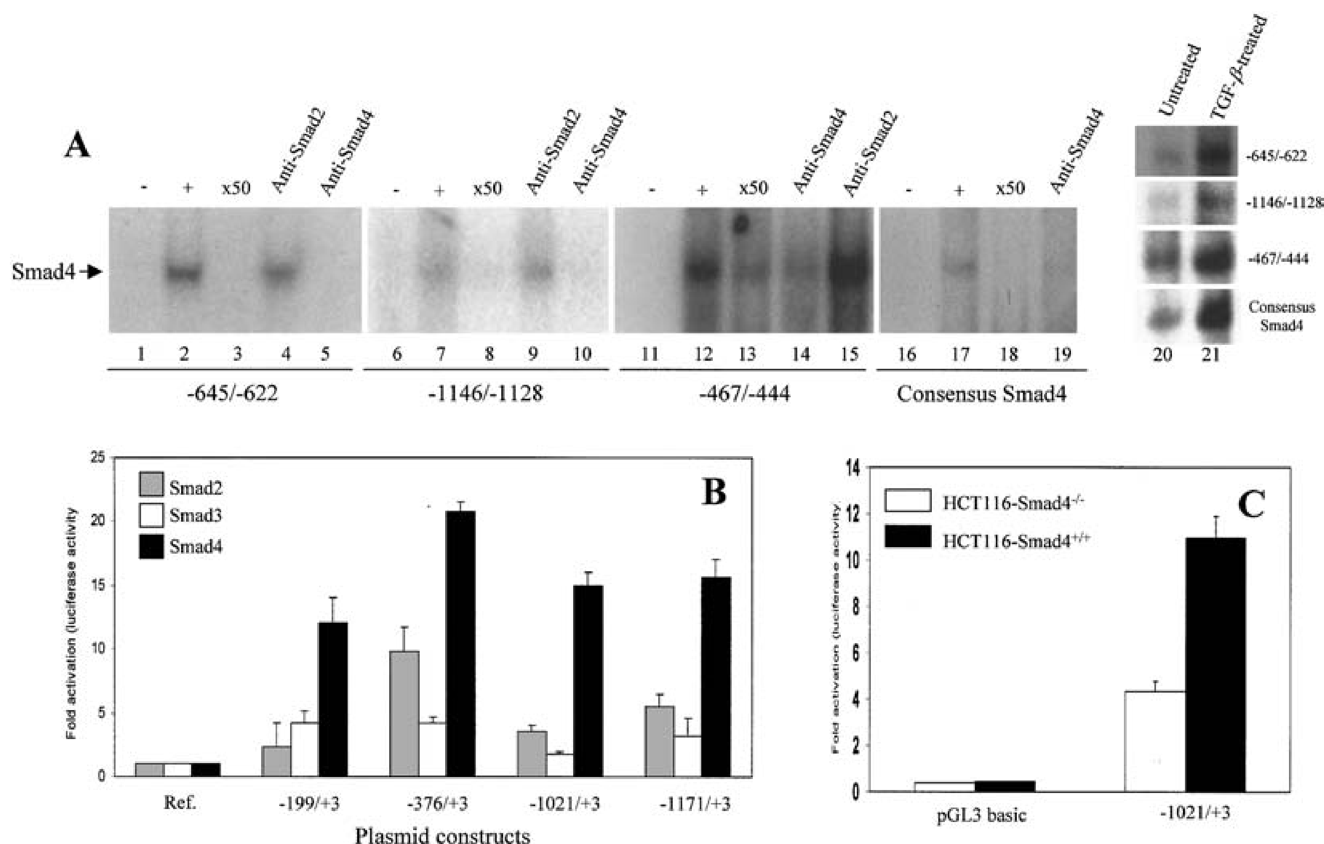

Figure 6 Identification of Smad4-binding sites within the promoter of *Muc5ac* by EMSA and regulation of the promoter by the Smads by transient transfection

(**A**) Nuclear extracts from CMT-93 cells were incubated with radiolabelled DNA probes, −645/−622 (lanes 1–5), −1146/−1126 (lanes 6–10), −467/−444 (lanes 11–15) and consensus Smad4 (lanes 16–19) probes. Radiolabelled probes alone (lanes 1, 6, 11 and 16), nuclear extract incubated with radiolabelled probe (lanes 2, 7, 12 and 17), unlabelled competition with a 50× excess of unlabelled probe (lanes 3, 8, 13 and 18), supershift analysis by preincubating nuclear extract with 1 μl of anti-Smad2 (lanes 4, 9 and 15) or anti-Smad4 (lanes 5, 10, 14 and 19) antibodies before adding the radiolabelled probe. The arrow on the left indicates the position of the DNA–protein complex engaging Smad4. Lanes 20 and 21, Smad4–DNA complex formation with nuclear extracts from untreated (lane 20) and TGF-β-treated (lane 21) cells. (**B**) Regulation of *Muc5ac* promoter by Smad2, Smad3 and Smad4 transcription factors in CMT-93 cells. Co-transfection experiments in the presence of pCMV-Smad2 (grey bars), pCMV-Smad3 (white bars) or pCMV-Smad4 (black bars) expression vectors. (**C**) Luciferase activity of *Muc5ac* promoter in HCT116-Smad4$^{-/-}$ (white bars) and in HCT116-Smad4$^{+/+}$ (black bars) cells.

(Figure 8, white bars). An additive effect was observed on the construct −376/+3 (Figure 8, black bars), whereas a synergistic transactivating effect was obtained on fragment −1171/+3 (Figure 8, grey bars). In conclusion, Sp1 appears to be the co-factor of Smad4 to activate *Muc5ac* transcription, and essential elements that convey the synergistic effect are localized within the −1171/−377 region of the promoter.

DISCUSSION

The organization, evolution and regulation of expression of the four 11p15 mucin genes, *MUC2*, *MUC5AC*, *MUC5B* and *MUC6*, have been studied extensively over the past few years [2,7,35]. *In situ* hybridization and promoter functional studies have led the investigators to the conclusion that these mucin genes are tightly regulated at the transcriptional level [8]. They often harbour an altered profile of expression in human tumours of the respiratory, gastrointestinal and urogenital tracts [2,5,6,36] and such alterations (overexpression, repression, neo-expression) are usually characteristic of genes playing important roles in carcinogenesis. In order to study mucin gene regulation in animal models, isolation of the 5′-flanking regions including promoter of the murine counterparts are now mandatory.

In this paper, we have isolated and characterized the 5′-flanking region of the murine mucin gene *Muc5ac*, including the promoter and first three exons/introns. Analysis and alignment of the first three exons with its human and rat counterparts show a high degree of homology with rat *Muc5ac* (79 %) whereas it shares less with human *MUC5AC* (52 %). The first ATG is identical with rat *Muc5ac* (GenBank accession no. AB042530), which adds four amino acid residues (MLHS) when compared with the human *MUC5AC* N-terminal amino acid sequence [33]. The TATA box sequence (TACAAAA) is conserved throughout evolution, since both mouse (this study) and human TATA boxes are identical [33]. The sequence of the TATA box in rat *Muc5ac* gene is not known. Overall, it can be stressed that the 5′ region of *Muc5ac* mucin gene is relatively well conserved between human, mouse and rat and shows typical features of genes with a cell-specific profile of expression. Indeed, in this report we found that expression of murine *Muc5ac* gene and protein was very specific and restricted to the surface epithelial cells of the gastric mucosa, like that described previously for rat *Muc5ac* [37]. No expression was found in trachea. This latter result is in contrast with expression territories in humans in which *MUC5AC* is expressed in the surface epithelium of stomach but also in goblet cells of the tracheobronchial tract [9]. This difference is most likely to be due to histological differences between rodents and humans, since the

© 2004 Biochemical Society

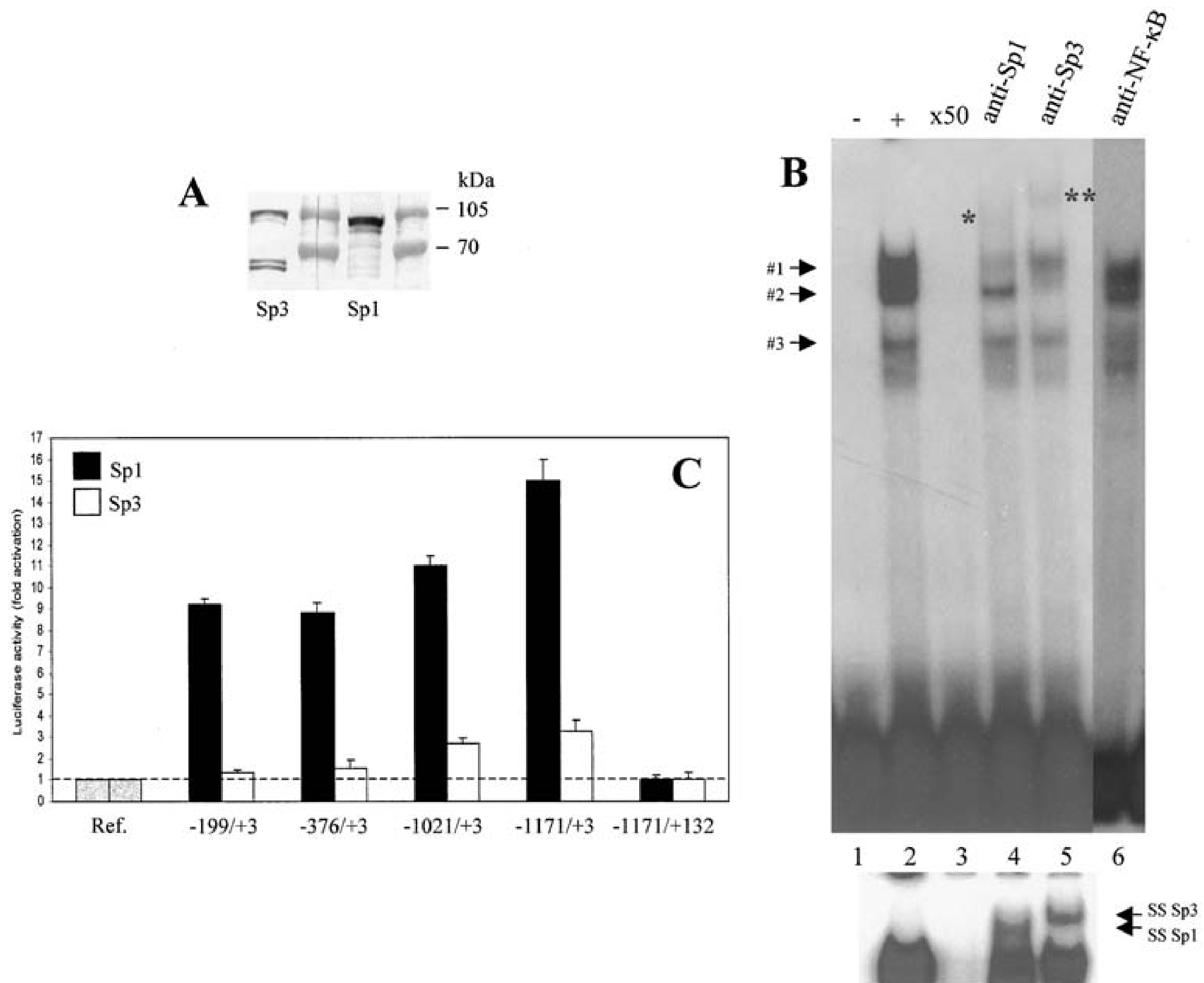

Figure 7 Regulation of *Muc5ac* promoter by Sp1 and Sp3 transcription factors in CMT-93 cells: identification of an Sp1 *cis*-element in the promoter by EMSA

(**A**) Western blotting of nuclear proteins prepared from CMT-93 cells. Prestained molecular-mass markers. Immunostaining of the membrane with polyclonal PEP-2 anti-Sp1 (*Sp1*) and with polyclonal anti-Sp3 (*Sp3*) antibodies. (**B**) Identification of an Sp1 *cis*-element by EMSA. Nuclear extracts from CMT-93 cells were incubated with the −83/−56 radiolabelled DNA probe. Radiolabelled probe alone (lane 1). Incubation of −83/−56 probe with CMT-93 nuclear proteins (lane 2), unlabelled competition with 50× excess of unlabelled probe (lane 3), supershift analysis upon addition of anti-Sp1 (lane 4) or anti-Sp3 (lane 5) antibodies. Lane 6, addition of irrelevant NF-κB p65 antibody. DNA–protein complexes are indicated by arrows on the left. Asterisks indicate the positions of Sp1 (SS Sp1) and Sp3 (SS Sp3) supershifted complexes. A longer exposure of the upper part of the same autoradiogram is shown at the bottom. (**C**) Co-transfection experiments in the presence of pCMV-Sp1 (black bars) or pCMV-Sp3 (white bars) expression vectors. Ref. refers to the normalized luciferase activity of the pGL3 deletion mutants of interest co-transfected with the empty expression vector pCMV4. Results are the means ± S.D. obtained in triplicate from three separate experiments.

epithelium of their conducting airways mostly consists of ciliated cells (upper tract) and Clara cells (lower tract), whereas mucous cells remain rare [38].

Analysis of the promoter nucleotide sequence immediately upstream of the TATA box (over the first 199 nucleotides) shows that it is GC-rich and bears numerous putative consensus Sp1-binding sites or CACCC boxes also known to bind transcription factors of the Sp family [8]. Consistent with a role in cancer for factors of the Sp family is their ability to be oncogenic themselves or to interact with oncogenes or tumour suppressors [39]. The high GC content of proximal region of promoters is a common feature in mucin genes, as it was also found in human *MUC2*, *MUC5AC*, *MUC5B* and murine *Muc2* promoters [8,20,30,40], and functional studies pointed to an important role for the transcription factors of the Sp family as mucin gene regulators. The fact that these *cis*-elements are conserved between mouse and human *MUC5AC* genes is also in favour of an important role for Sp1 in *MUC5AC* transcriptional regulation. Here, we confirmed that hypothesis by means of co-transfection experiments in which we showed that Sp1 strongly transactivates the promoter of *Muc5ac*. Having previously shown that Sp3 interferes with Sp1-mediated regulation of human *MUC5AC* promoter [28], we studied its effect on murine *Muc5ac* promoter as well. Our results indicate that Sp3 is also a regulator of *Muc5ac* transcription but with a weaker effect. Since Sp1 and Sp3 are ubiquitously expressed in the cells and recognize the same DNA motif, it also implies that Sp1 and Sp3 will compete when they are both present in equimolar amounts in the cell.

Interestingly, *Muc5ac* promoter activity was repressed by 50 % when pGL3 constructs contained the 5′-UTR (33 nucleotides long) and part of the first exon. Analysis of the sequence indicated that AP-1 and Sp1 binding sites are present within the 5′-UTR. Those factors are generally considered as activators of the transcription and could not be responsible for the repression observed. In close vicinity of the Sp1-binding site was found a second Sp1 *cis*-element overlapping a YY1-binding site. YY1

© 2004 Biochemical Society

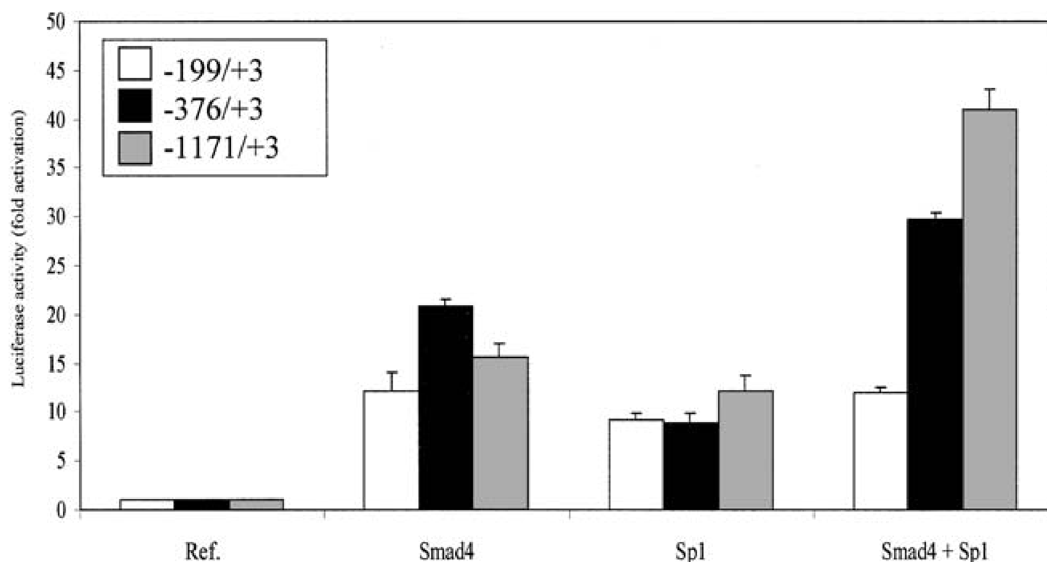

Figure 8 Co-operation between Smad4 and Sp1 to transactivate *Muc5ac* promoter

Co-transfection experiments were performed in CMT-93 cells in the presence of pCMV-Sp1, pCMV-Smad4 or both. Results are expressed as fold activation of the luciferase activity obtained in co-transfection with the expression vector encoding Smad4, Sp1 or both compared with cells transfected with corresponding empty vector (Ref.). Results are the means ± S.D. obtained in triplicate from three separate experiments.

possesses dual activity (repressor or activator), depending on the molecular context and is known to interact with Sp1 [41]. As it was recently shown for hamster *Muc1* mucin gene [42], YY1 may repress *Muc5ac* transcription via interactions with Sp1.

The *Muc5ac* promoter contains binding sites for the Smad4 transcription factor throughout its sequence. Alignment of human [33] and mouse (this study) sequences of the promoters indicates that Smad putative sites are present in both genes but not at the same location. The response of the two genes to TGF-β is different and seems to depend on the cellular situation as it was recently shown that human MUC5AC is negatively regulated when non-typeable *Haemophilus influenzae*-infected cells are treated with TGF-β [43]. Smad transcription factors are activated by growth factors of the TGF-β family in a sequential manner [18] and form either Smad2–Smad4 or Smad3–Smad4 complexes that are translocated into the nucleus where they bind to the promoter of the target gene to activate transcription. However, once bound to the promoter, Smad4 may interact with other factors to activate transcription. Our results indicate that *Muc5ac* is a target gene of TGF-β. However, they also suggest that the pathway induced by TGF-β does not imply complex formation between Smad4 and Smad2 or Smad4 and Smad3. Apart from activating Smad factors, TGF-β is also known to activate Sp1 site-dependent transcription of its target genes, Smad factors are known to activate transcription through Sp1 sites and Sp1 interacts with Smad4 to induce transcription and this mechanism is TGF-β-dependent [39,44,45]. The data presented in this article are in favour of such a positive regulatory mechanism between Smad4 and Sp1 to explain the TGF-β-mediated up-regulation of *Muc5ac* expression in cancer cells. Since TGF-β is a pleiotropic cytokine with diverse functions during development and in adult tissue homoeostasis, carcinogenesis and inflammation [46,47], it will be interesting in the future to study the correlation between TGF-β-mediated processes with that of *Muc5ac* expression and their consequences on epithelium homoeostasis.

We are indebted to Danièle Petitprez, Brigitte Hémon and Dominique Demeyer for their excellent technical help. We thank Dr J. M. Horowitz (North Carolina State University, Raleigh, NC, U.S.A.) for the kind gift of the pCMV-Sp1 and pCMV-Sp3 expression vectors. We thank Dr A. Atfi (INSERM U482, Paris, France) for pCMV-Smad2, pCMV-Smad3 and pCMV-Smad4 vectors and HCT116-Smad4$^{-/-}$ and HCT116-Smad4$^{+/+}$ cell lines. This work was supported by grants from l'Association de Recherche sur le Cancer (no. 5785, to I. V. S.), Ligue Régionale du Pas de Calais Contre le Cancer (to I. V. S.), l'Association François Aupetit (to I. V. S.), Sophia Foundation for Medical Research (to M. v. d. S. and A. W. C. E.) and from the European Regional Development Fund. N. J. is the recipient of a Conseil Régional Nord-Pas de Calais and the Centre Hospitalier & Universitaire de Lille PhD fellowship. M. S. received a fellowship from the Erasmus University of Rotterdam.

REFERENCES

1 Van Klinken, B. J., Dekker, J., Büller, H. A. and Einerhand, A. W. C. (1995) Mucin gene structure and expression: protection vs. adhesion. Am. J. Physiol. **269**, G613–G627
2 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) MUC genes: a superfamily of genes? Towards a functional classification of human apomucins. J. Soc. Biol. **193**, 85–99
3 Dekker, J., Rossen, J. W., Büller, H. A. and Einerhand, A. W. C. (2002) The MUC family: an obituary. Trends Biochem. Sci. **27**, 126–131
4 Carraway, K. L., Ramsauer, V. P., Haq, B. and Carothers Carraway, C. A. (2002) Cell signaling through membrane mucins. Bioessays **25**, 66–71
5 Corfield, A. P., Myerscough, N., Longman, R., Sylvester, P., Arul, S. and Pignatelli, M. (2000) Mucins and mucosal protection in the gastrointestinal tract: new prospects for mucins in the pathology of gastrointestinal disease. Gut **47**, 589–594
6 Einerhand, A. W. C., Renes, I. B., Makkink, M. K., van der Sluis, M., Büller, H. A. and Dekker, J. (2002) Role of mucins in inflammatory bowel disease: important lessons from experimental models. Eur. J. Gastroenterol. Hepatol. **14**, 1–8
7 Pigny, P., Guyonnet-Dupérat, V., Hill, A. S., Pratt, W. S., Galiègue-Zouitina, S., Collyn-d'Hooge, M., Laine, A., Van-Seuningen, I., Degand, P., Gum, J. R. et al. (1996) Human mucin genes assigned to 11p15.5: identification and organization of a cluster of genes. Genomics **38**, 340–352
8 Van Seuningen, I., Pigny, P., Perrais, M., Porchet, N. and Aubert, J.-P. (2001) Transcriptional regulation of the 11p15 mucin genes. Towards new biological tools in human therapy, in inflammatory diseases and cancer? Front. Biosci. **6**, D1216–D1234
9 Audié, J.-P., Janin, A., Porchet, N., Copin, M.-C., Gosselin, B. and Aubert, J.-P. (1993) Expression of human mucin genes in respiratory, digestive, and reproductive tracts ascertained by *in situ* hybridization. J. Histochem. Cytochem. **41**, 1479–1485
10 Guillem, P., Billeret, V., Buisine, M.-P., Fléjou, J. F., Lecomte-Houcke, M., Degand, P., Aubert, J.-P., Triboulet, J.-P. and Porchet, N. (2000) Mucin gene expression and cell differentiation in human normal, premalignant and malignant esophagus. Int. J. Cancer **88**, 856–861

© 2004 Biochemical Society

11 Warson, C., Van de Bovenkamp, J. H., Korteland-Van de Male, A. M., Buller, H., A., Einerhand, A. W. C., Ectors, N. L. and Dekker, J. (2002) Barrett's esophagus is characterized by expression of gastric-type mucins (MUC5AC, MUC6) and TFF peptides (TFF1 and TFF2), but the risk of carcinoma development may be indicated by the intestinal-type mucin, MUC2. Hum. Pathol. **33**, 660–668
12 Van de Bovenkamp, J. H. B., Korteland-Van Male, A. M., Büller, H. A., Einerhand, A. W. C. and Dekker, J. (2003) Gastric-type mucin and TFF-peptide expression in Barrett's oesophagus is disturbed during increased expression of MUC2. Hum. Pathol. **34**, 156–165
13 Bartman, A. E., Sanderson, S. J., Ewing, S. L., Niehans, G. A., Wiehr, C. L., Evans, M. K. and Ho, S. B. (1999) Aberrant expression of MUC5AC and MUC6 gastric mucin genes in colorectal polyps. Int. J. Cancer **80**, 210–218
14 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) Aberrant expression of a human mucin gene (MUC5AC) in rectosigmoid villous adenoma. Gastroenterology **110**, 84–91
15 Copin, M.-C., Buisine, M.-P., Devisme, L., Leroy, X., Escande, F., Gosselin, B., Aubert, J.-P. and Porchet, N. (2001) Normal respiratory mucosa, precursor lesions and lung carcinomas: differential expression of human mucin genes. Front. Biosci. **6**, D1264–D1275
16 Massagué, J. (1998) TGF-β signal transduction. Annu. Rev. Biochem. **67**, 753–791
17 Milani, S. and Calabro, A. (2001) Role of growth factors and their receptors in gastric ulcer healing. Microsc. Res. Tech. **53**, 360–371
18 Massagué, J. and Wotton, D. (2000) Transcriptional control by the TGF-β/Smad signaling system. EMBO J. **19**, 1745–1754
19 Gendler, S. J. (2001) MUC1, the renaissance molecule. J. Mamm. Gland Biol. Neopl. **6**, 339–353
20 Aslam, F., Palumbo, L., Augenlicht, L. H. and Velcich, A (2001) The Sp family of transcription factors in the regulation of the human and mouse MUC2 gene promoters. Cancer Res. **61**, 570–576
21 Velcich, A., Yang, W., Heyer, J., Fragale, A., Nicholas, C., Viani, S., Kucherlapati, R., Lipkin, M., Yang, K. and Augenlicht, L. (2002) Colorectal cancer in mice genetically deficient in the mucin Muc2. Science **295**, 1726–1729
22 Shekels, L. L., Lyftogt, C., Kieliszewski, M., Filie, J. D., Kozak, C. A. and Ho, S. B. (1995) Mouse gastric mucin: cloning and chromosomal localization. Biochem. J. **311**, 775–785
23 Desseyn, J.-L. and Laine, A. (2003) Characterization of mouse *Muc6* and evidence of conservation of the gel-forming mucin gene cluster between human and mouse. Genomics **81**, 433–436
24 van de Bovenkamp, J. H., Hau, C. M., Strous, G. J., Buller, H. A., Dekker, J. and Einerhand, A. W. C. (1998) Molecular cloning of human gastric mucin MUC5AC reveals conserved cysteine-rich D-domains and a putative leucine zipper motif. Biochem. Biophys. Res. Commun. **245**, 853–859
25 Quandt, K., Frech, K., Karas, H., Wingender, E. and Werner, T. (1995) MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. Nucleic Acids Res. **23**, 4878–4884
26 Renes, I. B., Boshuizen, J. A., Van Nispen, D. J., Bulsing, N. P., Buller, H. A., Dekker, J. and Einerhand, A. W. C. (2002) Alterations in Muc2 biosynthesis and secretion during dextran sulfate sodium-induced colitis. Am. J. Physiol. **282**, G382–G389
27 Renes, I. B., Verburg, M., Bulsing, N. P., Ferdinandusse, S., Buller, H. A., Dekker, J. and Einerhand, A. W. C. (2002) Protection of the Peyer's patch-associated crypt and villus epithelium against methotrexate-induced damage is based on its distinct regulation of proliferation. J. Pathol. **198**, 60–68
28 Perrais, M., Pigny, P., Copin, M.-C., Aubert, J.-P. and Van Seuningen, I. (2002) Induction of MUC2 and MUC5AC mucins by factors of the epidermal growth factor (EGF) family is mediated by EGF receptor/Ras/Raf/extracellular signal-regulated kinase cascade and Sp1. J. Biol. Chem. **277**, 32258–32267
29 Van Seuningen, I., Perrais, M., Pigny, P., Porchet, N. and Aubert, J.-P. (2000) Sequence of the 5′-flanking region and promoter activity of the human mucin gene MUC5B in different phenotypes of colon cancer cells. Biochem. J. **348**, 675–686
30 Perrais, M., Pigny, P., Buisine, M.-P., Porchet, N., Aubert, J.-P. and Van Seuningen-Lempire, I. (2001) Aberrant expression of human mucin gene MUC5B in gastric carcinoma and cancer cells. Identification and regulation of a distal promoter. J. Biol. Chem. **276**, 13586–13596
31 Mesquita, P., Jonckheere, N., Almeida, R., Ducourouble, M.-P., Serpa, J., Silva, E., Pigny, P., Santos Silva, F., Reis, C., Silberg, D., Van Seuningen, I. and David, L. (2003) Human MUC2 mucin gene is transcriptionally regulated by Cdx homeodomain proteins in gastrointestinal carcinoma cell lines. J. Biol. Chem. **278**, 51549–51556
32 Van Seuningen, I., Ostrowski, J., Bustelo, X. R., Sleath, P. and Bomsztyk, K. (1995) The K protein domain that recruits the interleukin 1-responsive K protein kinase lies adjacent to a cluster of c-Src and Vav SH3-binding sites. Implications that K protein acts as a docking platform. J. Biol. Chem. **270**, 26976–26985
33 Li, D., Gallup, M., Fan, N., Szymkowski, D. E. and Basbaum, C. B. (1998) Cloning of the amino-terminal and 5′-flanking region of the human MUC5AC mucin gene and transcriptional up-regulation by bacterial exoproducts. J. Biol. Chem. **273**, 6812–6820
34 Hanada, T. and Yoshimura, A. (2002) Regulation of cytokine signaling and inflammation. Cytokine Growth Factor Rev. **13**, 413–421
35 Desseyn, J.-L., Aubert, J.-P., Porchet, N. and Laine, A. (2000) Evolution of the large secreted gel-forming mucins. Mol. Biol. Evol. **17**, 1175–1184
36 Lesuffleur, T., Zweibaum, A. and Real, F. X. (1994) Mucins in normal and neoplastic human gastrointestinal tissues. Crit. Rev. Oncol. Hematol. **17**, 153–180
37 Inatomi, T., Tisdale, A. S., Zhan, Q., Spurr-Michaud, S. and Gipson, I. K. (1997) Cloning of rat Muc5AC mucin gene: comparison of its structure and tissue distribution to that of human and mouse homologues. Biochem. Biophys. Res. Commun. **236**, 789–797
38 Pack, R. J., Al-Ugaily, L. H. and Morris, G. (1981) The cells of the tracheobronchial epithelium of the mouse: a quantitative light and electron microscope study. J. Anat. **132**, 71–84
39 Black, A. R., Black, J. D. and Azizkhan-Clifford, J. (2001) Sp1 and kruppel-like factor family of transcription factors in cell growth regulation and cancer. J. Cell Physiol. **188**, 143–160
40 Gum, J. R., Hicks, J. W. and Kim, Y. S. (1997) Identification and characterization of the MUC2 (human intestinal mucin) gene 5′-flanking region: promoter activity in cultured cells. Biochem. J. **325**, 259–267
41 Thomas, M. J. and Seto, E. (1999) Unlocking the mechanisms of transcription factor YY1: are chromatin modifying enzymes the key? Gene **236**, 197–208
42 Lee, I. J., Hyun, S. W., Nandi, A. and Kim, K. C. (2003) Transcriptional regulation of the hamster Muc1 gene: identification of a putative negative regulatory element. Am. J. Physiol. **284**, L160–L168
43 Jono, H., Xu, H., Lim, D. J., Kim, Y. S., Feng, X.-H. and Li, J.-D. (2003) Transforming growth factor-β-Smad signaling pathway negatively regulates nontypeable *Haemophilus influenzae* -induced MUC5AC mucin transcription via mitogen-activated protein kinase (MAPK) phosphatase-1-dependent inhibition of p38 MAPK. J. Biol. Chem. **278**, 27811–27819
44 Datta, P. K., Blake, M. C. and Moses, H. L. (2000) Regulation of plasminogen activator inhibitor-1 expression by transforming growth factor-β-induced physical and functional interactions between Smads and Sp1. J. Biol. Chem. **275**, 40014–40019
45 Feng, X. H., Lin, X. and Derynck, R. (2000) Smad2, Smad3 and Smad4 cooperate with Sp1 to induce p15(Ink4B) transcription in response to TGF-β. EMBO J. **19**, 5178–5193
46 Moustakas, A., Pardali, K., Gaal, A. and Heldin, C.-H. (2002) Mechanisms of TGF-β signaling in regulation of cell growth and differentiation. Immunol. Lett. **82**, 85–91
47 Dünker, N. and Krieglstein, K. (2000) Targeted mutations of transforming growth factor-β genes reveal important roles in mouse development and adult homeostasis. Eur. J. Biochem. **267**, 6982–6988

Received 25 June 2003/9 October 2003; accepted 22 October 2003
Published as BJ Immediate Publication 22 October 2003, DOI 10.1042/BJ20030948

© 2004 Biochemical Society