Cloning and characterization of the 5′ flanking region of the sialomucin complex/rat *Muc4* gene: promoter activity in cultured cells

Shari A. PRICE-SCHIAVI[1], Aymee PEREZ[1], Roy BARCO and Kermit L. CARRAWAY[2]

Department of Cell Biology and Anatomy (R-124), University of Miami School of Medicine, PO Box 016960, Miami, FL 33101, U.S.A.

Sialomucin complex (SMC/Muc4) is a heterodimeric glycoprotein complex consisting of a mucin subunit ascites sialoglycoprotein-1 (ASGP-1) and a transmembrane subunit (ASGP-2), which is aberrantly expressed on the surfaces of a variety of tumour cells. SMC is transcribed from a single gene, translated into a large polypeptide precursor, and further processed to yield the mature ASGP-1/ASGP-2 complex. SMC has complex spatial and temporal expression patterns in the normal rat, suggesting that it has complex regulatory mechanisms. A crude exon/intron map of the 5′ regions of the SMC/*Muc4* gene generated from clones isolated from a normal rat liver genomic DNA library reveals that this gene has a small first exon comprising the 5′ untranslated region and signal peptide, followed by a large intron. The second exon appears to be large, comprising the 5′ unique region and a large part (probably all) of the tandem repeat domain. This structure is strikingly similar to that reported for the human *MUC4* gene. Using PCR-based DNA walking, 2.4 kb of the 5′-flanking region of the SMC/*Muc4* gene was cloned and characterized. Promoter-pattern searches yielded multiple motifs commonly found in tissue-specific promoters. Reporter constructs generated from this 2.4 kb fragment demonstrate promoter activity in primary rat mammary epithelial cells (MEC), the human colon tumour cell line HCT-116, and the human lung carcinoma cell line NCI-H292, but not in COS-7 cells, suggesting epithelial cell specificity. Deletion constructs of this sequence transfected into rat MEC or HCT-116 cells demonstrate greatly varying levels of activity, suggesting that there are positive and negative, as well as tissue-specific, regulatory elements in this sequence. Taken together, these data suggest that the rat SMC/Muc4 promoter has been identified, that it is tissue- (epithelial cell-) specific, and that there are both positive and negative, as well as tissue-specific, regulatory elements in the sequence.

Key words: glycoprotein, mucin, regulatory elements, transcriptional regulation.

INTRODUCTION

Mucins are large glycoconjugates with molecular masses ranging from 300 kDa to over 40000 kDa [1–3]. The most important characteristic of these molecules is that they contain serine- and threonine-rich domains, which can be extensively O-glycosylated [3–5]. Mucins are subdivided into three categories: secretory gel-forming mucins, like MUC2 and MUC5AC, form large oligomers through linkage of protein monomers via disulphide bonds in their N- and C-terminal domains; small, soluble mucins, like MUC7 [6], have smaller polypeptide backbones and do not form disulphide-crosslinked gels; and membrane-bound mucins, such as MUC1, have a hydrophobic membrane-spanning domain and also do not form disulphide-linked oligomers [1]. Because of its large extended structure and cell surface location, MUC1 has been assigned a number of possible functions. In normal and tumour tissue, MUC1 may create a barrier to cell–cell interactions, reduce adhesion of macromolecules to the cell surface, and provide resistance to natural killer and cytotoxic T-cell killing [7].

Sialomucin complex (SMC) is another well-characterized membrane mucin. It was originally discovered as the major cell surface glycoprotein complex of the highly malignant metastatic ascites 13762 rat mammary adenocarcinoma [8]. SMC consists of a peripheral, O-glycosylated glycoprotein [8,9], ascites sialoglycoprotein-1 (ASGP-1), which is tightly, but non-covalently bound to an N-glycosylated integral membrane glycoprotein, ASGP-2 [10]. The complex is encoded on a single gene, transcribed as a 9 kb transcript [11–13], and translated into a single large (approx. 300 kDa) polypeptide precursor that is proteolytically cleaved during its transit to the cell surface [11]. Both subunits of this complex have been cloned and sequenced. The ASGP-1 sequence predicts a protein of 2172 amino acids with a potential molecular mass of > 200000 [13]. It consists of three domains: an N-terminal unique sequence, a large repeat sequence rich in serine and threonine residues which can be O-glycosylated, and a C-terminal unique sequence [13]. ASGP-2 is a 120–140 kDa glycoprotein consisting of 744 amino acid residues [12]. Its deduced amino acid sequence encodes seven putative domains: two hydrophilic N-glycosylated domains, two epidermal growth factor (EGF)-like domains, a cysteine-rich domain, a transmembrane domain, and a small cytoplasmic domain [12].

Recent studies have shown that SMC is the rat homologue of human MUC4. Human MUC4 shows substantial similarities to rat SMC at both the N- and C-terminal portions of the molecules [14]. They differ in their repeat domains in that SMC does not have the 16 amino acid repeat cloned and sequenced in the original description of MUC4 [15]. Thus the similarity between the molecules was not observed previously. However, the 70% identity between the human MUC4 analogue of ASGP-2 and rat ASGP-2 provides strong evidence that they are homologous proteins.

Abbreviations used: SMC, sialomucin complex; ASGP, ascites sialoglycoprotein; EGF, epidermal growth factor; CAT, chloramphenicol acetyltransferase; MEC, mammary epithelial cells; HCT, human colon tumour.

[1] These authors contributed equally to this work.

[2] To whom correspondence should be addressed (e-mail Kcarrawa@med.miami.edu).

Several functions have been proposed for SMC/MUC4. Through its highly O-glycosylated tandem repeat domain, ASGP-1 can act as a protective barrier to tumour cells and exposed epithelia. Indeed, SMC expression in transfected A375 tumour cells provides anti-recognition and anti-adhesive properties [8,16,17]. Furthermore, SMC expression in tumour cells reduces their killing by natural killer cells [18]. This anti-recognition property may be important to the high metastatic capacity of the 13762 ascites cells [16,18,19]. Further, the presence of SMC on exposed epithelia such as cornea, vagina and cervix suggests a protective role for SMC [20,21]. ASGP-2 has two EGF-like domains, which have all of the consensus residues present in active members of the EGF family [12]. Moreover, SMC has been shown to bind to, and modulate phosphorylation of, the receptor ErbB2 [22]. Thus the transmembrane subunit ASGP-2 is proposed to modulate signalling through the EGF family of receptors via its interaction with ErbB2 [22,23], the critical receptor for formation of active heterodimers of the class I receptor tyrosine kinases [24]. This interaction may play a role in the constitutive phosphorylation of ErbB2 in the 13762 ascites cells [25], and the rapid growth of these cells *in vivo*.

The expression patterns of both human MUC4 and rat SMC/Muc4 have been studied in detail. MUC4 is expressed in a broad range of secretory epithelial cells. By Northern blotting and/or *in situ* hybridization, human MUC4 mRNA has been detected in colon, stomach, cervix and lung [26,27], but not in normal pancreas, gall bladder and breast tissues. Western blot and immunohistochemistry reveal that rat SMC/Muc4 protein is expressed in a number of normal rat tissues including colon, small and large intestine, trachea, uterus, cornea and mammary gland [28]. SMC is overexpressed on the surface of the 13762 rat mammary adenocarcinoma ($> 10^6$ copies/cell) at a level approx. 100-fold higher than that in lactating mammary gland or approx. 10000-fold higher than that in normal virgin gland [8,16,29,30]. Human MUC4 has also been shown to be aberrantly expressed in some tumours [31]. Thus SMC/Muc4 must be tightly regulated, since aberrant expression of this protein could have deleterious consequences.

In an effort to study regulation of SMC in normal cells and tumour cells, we have cloned and characterized the 5′ flanking region of the *SMC* gene. Although the new sequence apparently does not contain a TATA box, it does contain other motifs commonly found in tissue-specific promoters. The 5′ flanking sequence shows promoter activity in both rat mammary epithelial cells (MEC) and human lung and colon epithelial cell lines, but not in COS-7 cells, suggesting tissue specificity. Moreover, deletion mutants transfected into rat MEC and human colon tumour (HCT) cells display different patterns of activity, suggesting different regulatory mechanisms in different tissues. Taken together, these data suggest that the rat SMC/Muc4 promoter has been identified, it is tissue specific, and it has different regulatory mechanisms in different cell types.

EXPERIMENTAL

Materials

COS-7, HCT-116 cells and NCI-H292 human lung carcinoma cells were obtained from A.T.C.C. Probe A2G2-#9, which comprises the 5′ 1.7 kb of SMC cDNA, has been described in [13]. Anti-ASGP-2 mouse monoclonal antibody 4F12 has been described in [30]. Unless otherwise stated, all cell culture supplies were purchased from Gibco Life Technologies. The positive control vector for the p50 subunit of human DNA polymerase delta was generously given by Dr Marietta Lee.

Cell culture

Preparation and culture of normal rat MEC has been described previously in [29]. COS-7 cells were maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum, 100 units/ml penicillin, and 100 μg/ml streptomycin. HCT-116 cells were maintained in McCoy's 5A medium supplemented with 10% fetal bovine serum, 1 mM L-glutamine, 100 units/ml penicillin, and 100 μg/ml streptomycin. Human lung carcinoma NCI-H292 cells were maintained in RPMI 1640 medium supplemented with 10% fetal bovine serum, 100 units/ml penicillin, and 100 μg/ml streptomycin.

Primer sequences

Oligonucleotide primers were synthesized from sequences in the 5′ unique region of SMC cDNA: L-19: 5′-AGTAGGGACAGGCTCACCTC-3′; A1-L: 5′-ACCTCTGTGCTGTGGAGATT-3′; R-85: 5′-CTCCCTCAGGAATCTCCACAGCACAGA-3′; A1-S: 5′-TTCTCTGGAGCCATGCGAGG-3′; R-173: 5′-GGCAAGAACACAGGCAGCTCAGACACA-3′; R-187: 5′-AGGTAGCAGGAGGAGGCAAG-3′; L-197: 5′ TCAACTACATCAGCCCCGAA-3′; R-339: 5′-GAAGTAACCATCGGTGATGC-3′.

Library screening

A normal rat liver EMBL3-SP-6/T7 library was obtained from ClonTech Laboratories (Palo Alto, CA, U.S.A.). Phage were plated on *Escherichia coli* K802 at a density of 50000 plaque forming units/150 mm plate, and incubated overnight at 37 °C. Phage were lifted on to positively charged nylon filters and denatured in 1.5 M NaCl, 0.5 M NaOH. Denatured DNA was crosslinked using a Stratalinker (Stratagene, La Jolla, CA, U.S.A.). The membranes were prehybridized for at least 2 h at 42 °C in prehybridization solution [50% formamide, 5 × SSC (where 1 × SSC is 0.15 M NaCl/0.015 M sodium citrate), 5 × Denhardt's reagent, 0.1% SDS, and 0.5 mg/ml salmon sperm DNA]. The probe A2G2-#9, a 1.7 kb probe which spans the 5′ unique region and four tandem repeats of SMC cDNA, was random primed labelled with [^{32}P]dCTP using a Random Primed Labeling kit (Boehringer Mannheim). The membranes were hybridized overnight at 42 °C in prehybridization solution containing 0.1 g/ml dextran sulphate and the labelled probe. Following hybridization, membranes were washed once at room temperature in 2 × SSC with 0.1% SDS for 15 min, twice at 50 °C in 2 × SSC with 0.1% SDS for 20 min each, and once at 50 °C in 0.1 × SSC with 0.1% SDS for 15 min. Signals were detected by exposure with Kodak XAR-5 X-ray film. Positive clones were confirmed on duplicate filters and purified to homogeneity through three more rounds of screening. Phage DNA was isolated using the manufacturer's protocol from the LambdaPhage DNA midiprep kit from Qiagen.

Southern blotting

PCR products or digested DNA samples were resolved on 1% agarose gels. Resolved DNAs were depurinated by soaking the gel in 0.25 M HCl for 15 min, then transferred to positively charged nylon membranes under denaturing conditions (0.4 M NaOH) overnight, and crosslinked using a Stratalinker. Hybridizations were carried out using the recommended protocols from the Boerhinger Mannheim Digoxigenin Labeling kit. Briefly, membranes were prehybridized in prehybridization buffer (5 × SSC, 1.0% blocking reagent, 0.1% *N*-lauroylsarcosine, 0.02% SDS) for 1 h at 63 °C, followed by hybridization for 4 h

at 63 °C to oligonucleotides which were digoxigenin end-labelled. The membranes were washed three times with 2 × SSC, 0.1 % SDS at 63 °C, and once with 0.5 × SSC, 0.1 % SDS at 63 °C. For development, the membranes were blocked for 30 min with blocking buffer [100 mM Tris/HCl (pH 7.5), 150 mM NaCl, 2 % blocking reagent], incubated with alkaline phosphatase-conjugated anti-digoxigenin antibody for 30 min in blocking buffer, and washed twice with wash buffer [100 mM Tris/HCl (pH 7.5), 150 mM NaCl]. Membranes were developed with Lumi-Phos 530 or disodium 3-(4-methoxyspiro-{1,2-dioxetane-3,2′-(5′-chloro)tricyclo[3.3.1.1.13,7]}decan-4-yl)phenyl phosphate ('CSPD', Boehringer Mannheim), and signals were detected by exposure to Kodak XLS film.

PCR-based DNA walking

A PromoterFinder PCR-based DNA walking kit was purchased from ClonTech Laboratories. High-molecular-mass genomic DNA was purified from 13762 MAT-C1 ascites cells or normal rat liver tissue, and its quality was determined by electrophoresis on a 0.5 % agarose gel. The isolated DNA was digested overnight with five different rare cutting blunt-end restriction enzymes: *Dra*I, *Eco*RV, *Pvu*II, *Sca*I and *Stu*I. After digestion, each sample of digested DNA was purified by phenol/chloroform extraction and ethanol precipitation. A portion of each purified DNA sample was ligated to an adapter provided in the PromoterFinder kit using the manufacturer's recommended protocol. The resulting enzyme-digested, adapter-ligated genomic DNA pools were then ready to be used for PCR.

The first PCR reaction was performed using adapter primer-1, a primer from the ligated adapter region, and the gene-specific primer R-173. Cycling conditions were 94 °C for 30 s (1 cycle); 94 °C for 30 s, 59 °C for 30 s, 72 °C for 3 min (40 cycles); and 72 °C for 10 min (1 cycle). The second PCR reaction was performed on 1 μl of PCR product from the first round of PCR. The second set of primers, AP-2 (from the ligated adapter region) and R-85, were 'nested' completely within the region amplified by the first set of primers. Cycling conditions for the second round of PCR were 94 °C for 30 s (1 cycle); 94 °C for 30 s, 71 °C for 3 min (40 cycles); and 71 °C for 7 min (1 cycle). PCR products were resolved on 1 % agarose gels, and the resulting bands were confirmed by hybridization to digoxigenin-labelled oligonucleotide L-19.

Constructs

Original constructs

The p10-Sal construct, which contains the entire phage insert from phage clone 10-2, was generated by excision of the entire SMC genomic insert with *Sal*I, which only cleaves in the phage multicloning region, not in the insert region. This insert was ligated into the *Sal*I-digested pSport vector (Gibco Life Technologies). The pCR2.1-EV plasmid was generated by cloning the major 600 bp PCR product produced in the *Eco*RV sample from PCR-based DNA walking into the T/A cloning vector pCR2.1 (Invitrogen, Carlsbad, CA, U.S.A.). The pCR2.1-NDI plasmid was generated by cloning the major 2400 bp PCR product from the *Dra*I sample from PCR-based DNA walking into the T/A cloning vector pCR2.1.

Reporter constructs

A *BamH*I–*Xba*I fragment from pCR2.1-EV was ligated to pCAT3Basic (Promega) digested with *Nhe*I–*Bgl*II to yield pFEcoRV. A *Kpn*I–*Xho*I fragment from pCR2.1-EV was ligated to pCAT3Basic (Promega) digested with *Kpn*I–*Xho*I to obtain pREcoRV. Similarly, *Kpn*I–*Xho*I and *Bam*HI–*Xba*I fragments from pCR2.1-NDI were ligated to pCAT3Basic digested with *Kpn*I–*Xho*I and *Nhe*I–*Bgl*II respectively, to create pF*Dra*I and pR*Dra*I. p*Sma*I was generated by inserting the *Sma*I–*Xho*I fragment from pCR2.1-NDI into pCAT3Basic, which was digested with *Sac*I, filled in with T4 DNA polymerase, and redigested with *Xho*I. pBamHI was obtained by digesting pCR2.1-NDI with *BamHI*, filling in the overhang with T4 DNA polymerase, and redigesting with *Xho*I. The resulting fragment was inserted into pCAT3Basic, which was digested with *Sac*I, filled in with T4 DNA polymerase, and redigested with *Xho*I. pEco47III was generated by inserting a *Eco*47III/*Xho*I fragment from pCR2.1-NDI into pCAT3Basic, prepared as described above. pNcoI was prepared by digesting pCR2.1-NDI with *Nco*I, filling in the overhang with T4 DNA polymerase, and redigesting with *Xho*I. The resulting fragment was ligated into pCAT3Basic prepared as described. pHindIII was constructed by digesting pCR2.1-NDI with *Hin*dIII, filling the overhang with T4 DNA polymerase and redigesting with *Xho*I. The resulting fragment was ligated into pCAT3Basic prepared as described. pEcoRI and pPstI were created by ligating *Eco*RI/T4 DNA polymerase and *Pst*I/T4 DNA polymerase fragments from pCR2.1-NDI to pCAT3Basic digested with *Sma*I and dephosphorylated with calf intestinal alkaline phosphatase.

The ligation mixtures were used to transform electro-competent INVα-F' (Invitrogen) *E. coli*, and colonies were screened by filter hybridization to digoxigenin-labelled oligonucleotide L-19, or by restriction mapping from plasmid mini-preps. Plasmid preparations were performed for the appropriate clones with a Qiagen maxi-prep kit. For reporter constructs XL-1 Blue *E. coli* (Stratagene) were chemically transformed and selected by convenient restriction digestion.

Transfections

HCT-116, NCI-H292 human lung carcinoma, or normal primary rat MEC were transfected with reporter vectors using the recommended protocol for LipofectAMINE (Gibco Life Technologies). Transfections were performed in triplicate. Cells and conditioned media were harvested 48 h after transfection. Briefly, cells were washed twice with Dulbecco's PBS without calcium. To each well 300 μl of Reporter Lysis Buffer (Promega) was added. The cells were incubated in a shaker for 15 min at room temperature and scraped with a rubber policeman. The samples were centrifuged for 10 min at 4 °C and the resulting supernatants were utilized for subsequent chloramphenicol acetyltransferase (CAT) and β-galactosidase assays. The β-galactosidase activity, used to monitor transfection efficiency per well, was performed as follows: 1 ml of 60 mM Na_2-PO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgCl_2$, 50 mM β-mercaptoethanol, and 0.2 ml of *o*-nitrophenyl galactoside (2 mg/ml in 60 mM Na_2HPO_4, 40 mM NaH_2PO_4) were added to 15 μl (HCT 116), 100 μl (NCI H292), 50 μl (COS7) and 120 μl (MEC) of cell extract. The reaction was carried out at 37 °C, until a yellow shade was obvious, and stopped by adding 0.5 ml of 1 M Na_2CO_3, so that the absorbance at 420 nm ranged from 0.2 to 0.7. β-Galactosidase activity was expressed as $D_{420} \times 100$, and extrapolated to the reaction time in hours for the total of 300 μl of cell extract. Under these conditions the β-galactosidase activity obtained ranged from 200–600 units per well of HCT 116 cells, 200–450 units per well of COS-7 cells, 50–300 units per well of NCI H292 cells, and 30–250 units per well of MEC.

CAT assays were performed with the amounts of extracts corresponding to equal β-galactosidase activity, and then brought

to 125 μl with lysis buffer (Promega); 5 μl of 25 mM acetyl-CoA and 0.5 μCi of ^{14}C-chloramphenicol (50 mCi/mmol, New England Nuclear) were added, and the reaction was performed at 37 °C. Considering that acetyl-CoA is rapidly degraded, it was added again in the reaction for MEC. The incubation time at 37 °C was usually 1 h but in the case of MEC the CAT reaction was left overnight. The reaction was stopped and the chloramphenicol was extracted with 1 ml of ethyl acetate. The solvent was dried, the pellet resuspended in 15 μl of ethyl acetate, and spotted on a silica gel thin-layer plate to separate the native chloramphenicol from its acetylated derivatives. Migration was in chloroform/methanol (19:1, v/v) for 1 h. After overnight autoradiography, the spots were cut out and counted to quantitate the amount of chloramphenicol converted into its acetylated forms.

Reporter construct results were calculated as the percent chloramphenicol converted from the normalized CAT assay. For transfection studies, results were calculated as the percent activity of the positive p50 control. For deletion mutant studies, results were calculated as the percent activity of the longest construct, pFDraI.

RESULTS

Characterization of SMC 5′ genomic clone

In an effort to isolate the promoter of the *SMC* gene, a rat liver genomic λ EMBL-3 SP-6/T-7 library was screened with ^{32}P-labelled A2G2-#9, which spans the 5′ 1.7 kb of the SMC cDNA and comprises the entire unique region and four repeats (Figure 1A). Two positive clones (10-2 and 7-1) were isolated and partially characterized. *Sal*I digestion of both clone 10-2 and clone 7-1 separated the entire insert fragment from the two phage arms of 21 kb and 9 kb. The insert sizes for clone 10-2 and clone 7-1 were approx. 18 kb and approx. 14 kb respectively, so the clones were different but overlapping. When clone 10-2 and clone 7-1 were probed with oligonucleotide R-339, which lies just 5′ of the repeat region, only clone 10-2 digests hybridized. This result suggests that the genomic insert of clone 7-1 stops within or at the 5′ end of the repeat region and that clone 10-2 contains more SMC/Muc4 5′ sequence.

For further characterization, the entire insert of clone 10-2 was subcloned into the pSport vector yielding p10-Sal. A partial restriction map of p10-Sal is shown in Figure 2. Long-range PCR between primer R-339 and T7 using p10-Sal as a template yielded a product of approx. 8.5 kb. Since primer R-339 is a right (3′-5′) primer, this result suggests that clone p10-Sal contains approx.

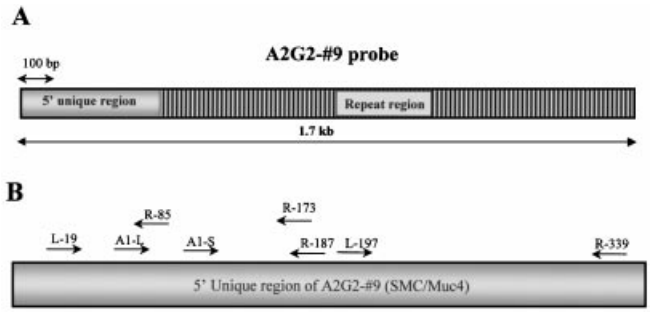

Figure 1 Probes and primers used in these studies

(**A**) Schematic representation of SMC/Muc4 probe A2G2-#9. (**B**) Relative positions of oligonucleotide primers used in these studies.

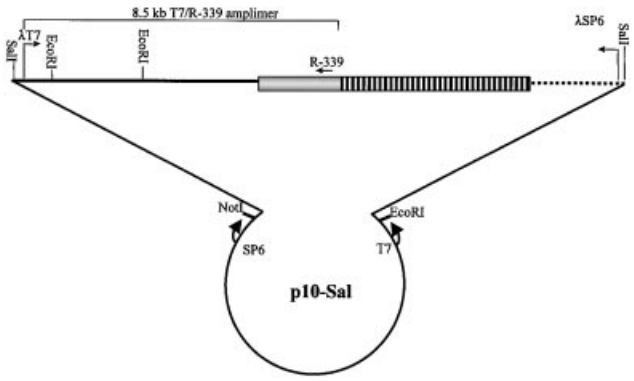

Figure 2 Characterization of phage clone 10-2

Restriction map and schematic representation of clone p10-Sal. Box indicates coding sequence, striped region indicates repeat sequences.

Figure 3 Preliminary genomic map of the 5′ regions of the rat SMC/*Muc4* gene

Striped region in exon 2 indicates repeat sequences.

8.5 kb of SMC sequence upstream of this primer (Figure 2). The area of interest is the 5′-most region of the gene. Therefore to determine if clone p10-Sal contains any of the most 5′ coding region of the SMC/*Muc4* gene, *Sac*I digests of clone p10-Sal were subjected to Southern-blot analysis with several oligonucleotides from the 5′ region of the SMC cDNA: R-339, R-187, Al-S, A1-L, and L19 (see Figure 1B for positions of the oligonucleotide probes). Clone p10-Sal did not hybridize to any of the 5′ probes tested (results not shown), suggesting that this clone does not contain any of the 5′-most sequence of the *SMC* gene.

Clone p10-Sal was sequenced in the 5′ direction from the oligonucleotide primer R-339 to determine how much known coding sequence it contained. From R-339, the sequence was identical to 120 bases of the SMC/Muc4 cDNA sequence. After 120 bases, the sequences diverged, suggesting the presence of an intron. Only 196 bp of known sialomucin coding sequence lies upstream of the putative intron. In subsequent experiments, this 196 bp region is called exon 1. Sequence generated from the 5′ end of p10-Sal failed to align with the 5′-most 196 bp of the SMC/Muc4 coding region, again suggesting that this clone does not contain SMC/Muc4 exon 1 or any 5′ flanking sequence. Clone p10-Sal was also sequenced in the 3′ direction from primer L-197. For approx. 600 bp of sequence obtained, the genomic sequence was identical to the 5′ unique region and tandem repeat region originally described for the SMC cDNA, with no divergent sequence. This result suggests that exon 2 contains the bulk of the 5′ unique sequence, and at least part of the tandem repeat domain. The presence of divergent sequence and the lack of hybridization to any of the most 5′ oligonucleotide probes suggests that the phage clone p10 contains at least part of rat SMC/Muc4 exon 2 and at least part of a large intron but none of the 5′-most coding or flanking regions (Figure 3).

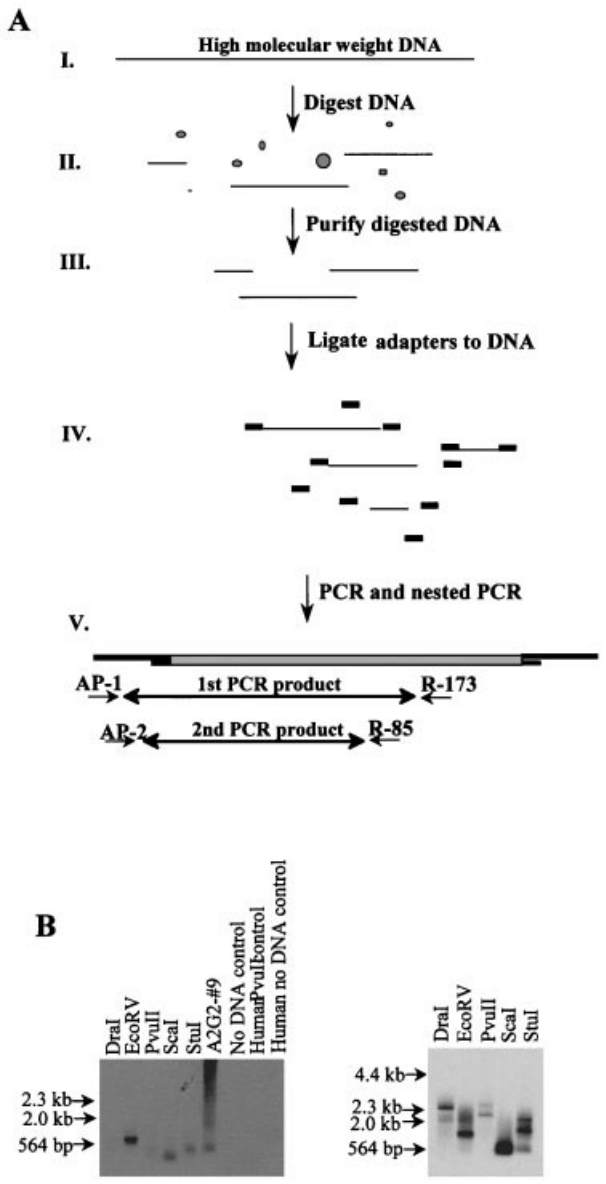

Figure 4 Cloning of SMC promoter

(**A**) Schematic representation of PromoterFinder PCR-based DNA walking. (**B**) Oligonucleotide L-19 hybridization of DNA walking products derived from ascites tumour cells (left panel) and normal rat liver (right panel) DNA. Samples are indicated at the top of the blots. Molecular mass markers are indicated to the left of the blots.

Cloning of the SMC promoter

After screening 3.5×10^6 phage with SMC/Muc4 5′ probes, no exon 1-containing phage clones were identified, suggesting that no exon 1-containing phage were present in the normal rat genomic library. Thus a new approach was used to identify the SMC promoter. Using PCR-based DNA walking with the ClonTech PromoterFinder Construction kit, 600 bp from ascites tumour DNA and 2.4 kb from normal rat liver DNA of new genomic sequence were isolated 5′ of the known SMC cDNA sequence. To accomplish this, five different genomic DNA preparations were used for PCR-based DNA walking by the scheme outlined in Figure 4(A). For each DNA preparation two rounds of PCR were performed. The first PCR reaction was performed using a primer from the ligated adapter region and a gene-specific primer. The second PCR reaction was performed using 'nested' primers for both the adapter and gene-specific regions. The gene-specific primers chosen for the two rounds of PCR for both the normal and tumour samples were 27-mers, R-173 and R-85, which are located in the putative first exon of the *SMC* gene (Figure 1B). The first round of PCR was carried out using a standard three-step protocol of 40 cycles, followed by a 7-min extension. On an ethidium bromide agarose gel, smears ranging from approx. 4 kb to 500 bp were visible for all the DNA preparations. For each DNA preparation, 1 μl of first-round PCR product was used as a template for the second-round 'nested' PCR. The second round of PCR was carried out using 40 cycles and a two-step protocol, where the annealing and extension steps were combined into one step. The T_m values of the gene-specific primers were high (72 °C), and the single step annealing and extension at 71 °C allowed for more specific product amplification.

On an ethidium bromide-stained agarose gel of the second-round PCR, discreet bands of approx. 600 bp in the *Eco*RV and 450 bp for *Sca*I samples were visible for the tumour DNA preparations. Several discreet bands, including 2.4 kb in *Dra*I, 600 bp in *Eco*RV, 2.0 kb in *Pvu*II, and 800 bp in *Stu*I samples, were visible for the normal DNA preparations. An A2G2-#9 positive control, prepared in a manner similar to that for the genomic samples, yielded a discreet product of approx. 3.0 kb for the first round of PCR, and a smear for the second round. A human positive control *Pvu*II library amplified with primers provided in the kit yielded the expected product of 1.5 kb, while all negative control samples showed no amplification. To confirm which bands were authentic, the products were probed with oligonucleotide L-19. All bands hybridized with this probe, suggesting that all the products contain 5′ SMC sequence (Figure 4B).

The 600 bp *Eco*RV band from the tumour DNA and 2.4 kb *Dra*I band from normal DNA were the largest products for each genomic DNA preparation tested. Therefore they were each cloned into the pCR2.1 vector using the Invitrogen TA cloning kit, yielding plasmid pCR2.1-EV for the tumour and plasmid pCR2.1-nDI for the normal. Plasmids pCR2.1-EV and pCR2.1-nDI were sequenced from the T7 promoter and a M13 reverse sequence contained in the vector. The insert sequences obtained for the tumour and normal samples aligned with the first 85 bases of the 5′ end of known SMC/Muc4 sequence, and contained approx. 500 bp or 2.4 kb of new sequence upstream respectively. The region between bases -700 and $+1$ has approx. 50 % GC content, while upstream the sequence gradually becomes more AT-rich. The new normal and tumour sequences are identical for the 600 bp they have in common. A pattern search yielded no TATA box close to the described transcription start site [13], but motifs were identified which are commonly found in promoters, such as a CAAT box, activator protein-1 and -2 sites, SP-1 transcription factor sites, GAS sequence, oestrogen response and glucocorticoid response half-sites (Figure 5).

Promoter activity of 600 bp and 2.4 kb clones

To determine if the 2.4 kb *Dra*I and 600 bp *Eco*RV fragments of new 5′ SMC sequence obtained from the normal rat have promoter activity, these sequences were subcloned into the reporter construct pCAT3-Basic in the forward and reverse directions to yield pFDraI and pRDraI for the 2.4 kb fragment and pFEcoRV and pREcoRV for the 600 bp fragment respectively. To test for promoter activity, these constructs along with pCAT3-Basic (with no insert as a negative control), p-βgal

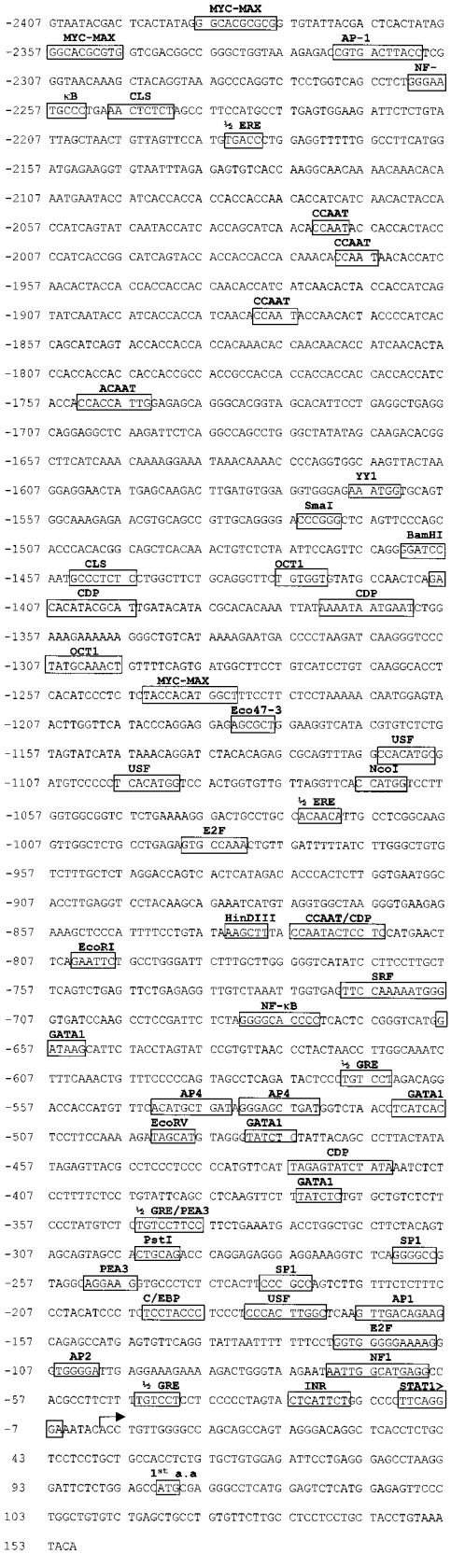

Figure 5 Sequence of 5′-flanking region of SMC/Muc4

Common promoter motifs and restriction sites are shown in **bold** above the corresponding sequence.

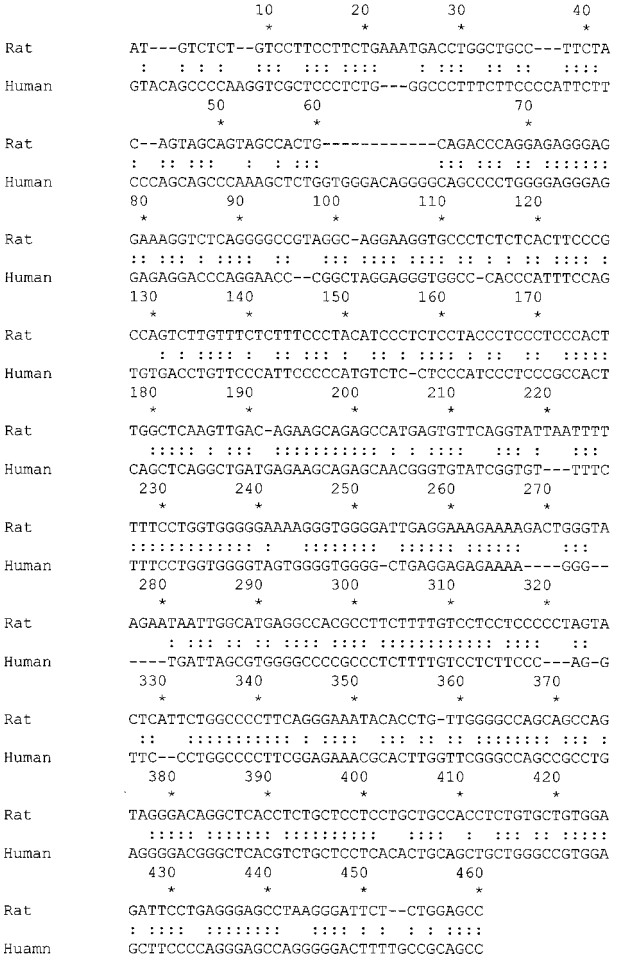

Figure 6 Comparison of rat SMC/Muc4 and human MUC4 5′-flanking regions

Table 1 CAT conversion in promoter assays

Values given as %.

Cell line	pFDral	pRDral	pFEcoRV	pREcoRV	p50+	pCAT3-Basic
HCT-116	33	0.5	53	1	90	20
NCI-H292	6	0.5	9	0.5	12	1
COS-7	8	2	2	1	45	0.5

(as transfection efficiency control), and p50 (a housekeeping gene promoter as positive control) were transfected into HCT-116 and human lung carcinoma NCI-H292 epithelial cell lines. These cell lines were chosen for initial reporter construct analysis because they have been reported to express MUC4 [33,34]. The first approx. 450 bases of the new rat 5′ flanking region is approx. 70% homologous to that described for human MUC4 [35] (Figure 6), suggesting that the rat promoter may also have activity in MUC4-expressing human cell lines. The fibroblast cell line COS-7 was also used to test tissue specificity, as SMC/MUC4 has only been reported in epithelial cells rather than stromal (fibroblastoid) cells. Normalized CAT assays were performed (Table 1), and the results were plotted. The forward constructs of

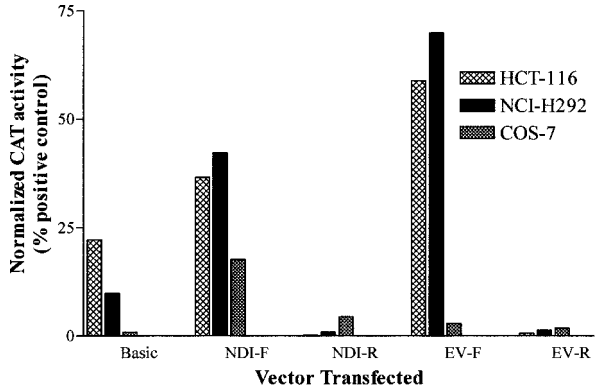

Figure 7 Promoter activity of rat SMC/Muc4 promoter sequence in transfected cell lines

Human lung carcinoma cell line NCI-H292, HCT-116 and COS-7 cells were transfected with empty vector (pCAT3Basic), positive control p50 subunit of DNA polymerase, or reporter constructs containing 2.4 kb (pFDraI and pRDraI) or 600 bp (pFEcoRV and pREcoRV) in forward and reverse orientations. Cells were co-transfected with β-galactosidase vectors to normalize for transfection efficiency. Normalized CAT assays were performed and CAT conversion for the p50 positive control was set at 100%. Results for reporter constructs were plotted as percent total activity of the positive p50 control.

both the 2.4 kb and 600 bp fragments have much higher activity than the reverse constructs in both HCT-116 and NCI-H292 cell lines (Figure 7). Promoter activity from both constructs is substantially lower in COS-7 cells, suggesting that this promoter activity is tissue specific to epithelial cells, as expected. The level of CAT conversion in cells transfected with the pCAT3Basic vector is most likely due to an enhancer element present in this vector. However, the marked lack of CAT activity in cells transfected with the reverse promoter constructs indicates that the promoter activity detected with the forward promoter constructs is authentic. Moreover, the higher levels of promoter activity in cells transfected with pFEcoRV constructs suggests that elements which confer epithelial specificity are present in the proximal promoter. Taken together, these data indicate that at least part of the SMC promoter has been identified and that it is tissue specific.

The initial pattern search yielded several positive and negative regulatory elements in the SMC/MUC4 5′ flanking region (Figure 5). To further investigate promoter activity of the SMC/Muc4 5′ flanking sequence, deletion constructs initiating from bases −2407, −1523, −1462, −1180, −1067, −834, −802, −490 and −291 were produced in the pCAT3Basic vector (Figure 8). These constructs were transfected into rat primary MEC or HCT-116 cells and assayed for promoter activity. The activity of each deletion construct was plotted as a percent activity relative to that of the largest construct, pFDraI (Figure 8). Promoter activity was different in both cells types with the deletion constructs. MEC promoter activity was greatly enhanced in the longer (pSmaI, pBamHI, pEco47-3) and shorter (pEcoRV and pPstI) constructs. In HCT116 cells activity was enhanced in the pBamHI and the three shortest constructs. These data suggest that there are positive regulatory elements in both the proximal and distal SMC/Muc4 promoter and that there are negative regulatory elements in the intermediate and far distal regions of this promoter. Further, these data suggest that there are elements in the distal promoter region which confer cell specificity for MEC and colon.

DISCUSSION

The main goal of these studies was to isolate the SMC/Muc4 promoter to begin elucidating how this gene is transcriptionally regulated in normal tissues and how its misregulation may lead to aberrant expression in tumour cells. Genomic library screening did not produce the SMC/Muc4 promoter. However, it yielded a large clone containing the repeat domain, part of the 5′ unique sequence, and part of a large intron. Interestingly, the crude exon/intron map of SMC/Muc4 is strikingly similar to that described for human MUC4 [35]. The putative first exon of the rat *SMC* gene is approx. 190 bp, and encodes a 5′ untranslated region and 27 amino acids which comprise the putative signal peptide, followed closely by a large intron. The presence of both primers used for PCR-based DNA walking on the same small genomic DNA fragments, and the absence of any other sequence

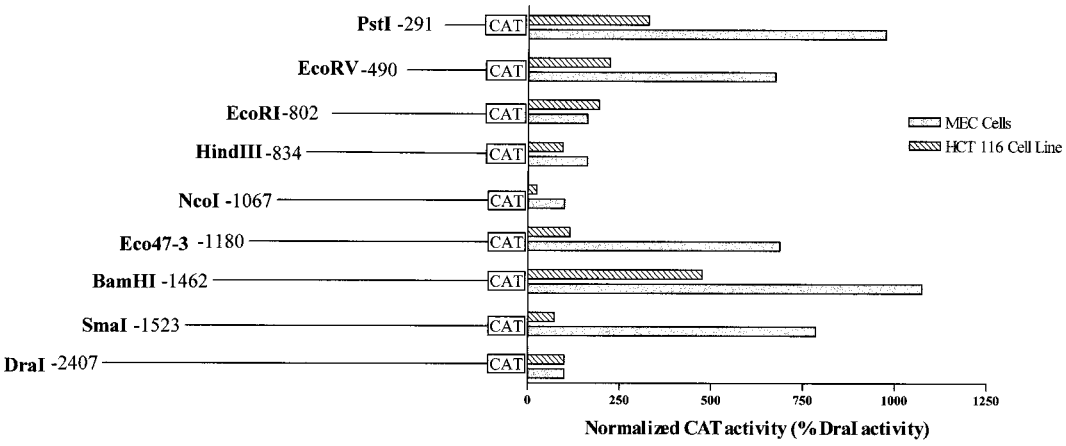

Figure 8 Activity of SMC/Muc4 promoter deletion mutants

Deletion mutants of the SMC/Muc4 promoter were generated by restriction digests of the new 5′ SMC flanking region as outlined in the Experimental section. Deletion mutants and β-galactosidase vectors (to normalize for transfection efficiency) were transfected into primary MEC or HCT-116 cells. Normalized CAT assays were performed, and CAT conversion for the largest construct pFDraI was set at 100%. Results for deletion constructs were plotted as percent total activity of the pFDraI reporter construct.

divergence within the first 85 bp of SMC cDNA and genomic DNA, suggests that this 190 bp is the first exon of the SMC/*Muc4* gene. Clone 10-2 contained ~ 8.5 kb of intron sequence and no exon 1 sequence, so the full size of rat SMC/Muc4 intron 1 is not known at this time. Like rat SMC/Muc4, the first exon of the human *MUC4* gene encodes the 5′ untranslated region and the first 27 N-terminal residues followed by a 15 kb intron [35]. The second exon of the human *MUC4* gene contains three large tandem repeats similar to those in rat SMC/Muc4, a 554 amino acid unique region, the large 16 amino acid tandem repeat domain, and another small unique sequence [35]. Sequencing in the 3′ direction from primer L-197 indicates that the second exon of the rat SMC/*Muc4* gene contains the bulk of the 5′ unique region and at least part of the tandem repeat region. Based on the striking similarity of the 5′ ends of the rat SMC/*Muc4* and human *MUC4* genes, we suggest that the second exon of the rat *SMC* gene contains the 5′ unique region, all of the tandem repeat domain, and at least part of the 3′ unique region of ASGP-1. It is interesting to note that the tandem repeat domains of several mucins are located on a single exon of their respective genes [36]. The exon/intron boundaries of the ASGP-2 subunit are as yet unknown. Thus a preliminary genomic map for rat SMC has been generated from these data, which shows strong similarity to that described for the human homologue of SMC, MUC4 [35].

Using an alternate PCR-based DNA walking protocol, PromoterFinder from ClonTech Laboratories, 600 bp and 2400 bp of new 5′ sequence were obtained from 13762 tumour cell and normal rat DNA respectively. Both of these fragments were sequenced, and for the 600 overlapping base pairs, the sequences were identical. A pattern search yielded no TATA box, but several important promoter motifs were identified. However, due to the small size of these sequences, their presence does not confirm that these DNA sequences contain the normal or tumour SMC/Muc4 promoter. Thus these sequences were cloned into the pCAT3Basic reporter vector in both forward and reverse orientations and transfected into HCT-116, NCI-H292, and COS-7 cells to test for promoter activity. The forward constructs had much higher activity than the controls and exhibited epithelial cell specificity, indicating that these fragments contain at least part of the epithelial-cell-specific SMC/Muc4 promoter. Moreover, transfection of deletion constructs into MEC and HCT-116 cells suggests complex regulation of this promoter, with positive regulatory elements in the proximal and distal promoter regions, and negative regulatory elements in the intermediate and far distal regions. This result is to be expected given the complex spatial and developmental expression patterns reported for the SMC/Muc4 protein [28]. Taken together, these data indicate that at least part of the SMC/Muc4 promoter has been identified and that it is tissue specific.

Not much is currently known about mucin gene promoters. The MUC1 promoter is the best studied of the mucin gene promoters. As in the SMC/Muc4 promoter, the MUC1 promoter contains two SP-1 sites. Mutation of these sites results in reduced transcription of reporter constructs in MUC1-expressing epithelial cell lines. However, mutation of the more proximal site increased transcription in non-epithelial cells lines, suggesting that this SP-1 site may be involved in tissue specificity [37]. The promoter of rat MUC2 also has an SP-1 site that has been proposed to be involved in tissue specificity of this gene [38]. SP-1 sites are found less frequently in tissue-specific promoters; however, other proteins, which are as yet unknown, can bind to the same core element (GGGCGG), suggesting that interactions of several factors may be involved in tissue-specific gene expression [37,39]. It remains to be seen if the SP-1 sites in the *SMC* gene promoter contribute to its tissue-specific regulation.

There are other regions of the *SMC* gene that will be of interest to study, such as the oestrogen response element and glucocorticoid response element half-sites. Ishikawa cells, a hormone-responsive human epithelial cell line, have been shown to express MUC4 (human SMC) [38]. Treatment of these cells with oestrogen or dexamethasone results in an increase in MUC4 RNA levels. Treatment with progesterone alone does not affect MUC4 RNA levels, but does inhibit MUC4 upregulation by oestrogen [38]. This is in agreement with SMC/Muc4 expression in rat uterus, which is strongly upregulated at the transcript level by oestrogen [40]. However, upregulation of SMC/Muc4 by oestrogen appears to be an indirect effect mediated by uterine stromal cells (N. Idris and K. L. Carraway, unpublished work). In addition, we have shown that transforming growth factor-β downregulates SMC/Muc4 at the transcript level directly in uterine luminal epithelial cells. We have also recently demonstrated that in mammary epithlelial cells, SMC/Muc4 is partially regulated at the transcript level by an ERK-dependent pathway (X. Zhu, S. A. Price-Schiavi and K. L. Carraway, unpublished work). Thus it will be of interest to determine which areas of the SMC/Muc-4 promoter are involved in regulation of SMC in normal tissues and how it is misregulated in tumour cells.

This research was supported in part by Grant CA52498 from the National Institutes of Health and by the Sylvester Comprehensive Cancer Center of the University of Miami.

REFERENCES

1 Corfield, A. P., Carrington, S.D., Hicks, S. J., Berry, M. and Ellingham, R. (1997) Ocular mucins: purification, metabolism and functions. Prog. Retinal Eye Res. **16**, 627–656
2 Hicks, S. J., Carrington, S.D., Kaswan, R., Adam, S., Bara, J. and Corfield, A. (1997) Demonstration of discrete secreted and membrane-bound ocular mucins in the dog. Exp. Eye Res. **64**, 597–607
3 Strous, G. J. and Dekker, J. (1992) Mucin-type glycoproteins. Crit. Rev. Biochem. Mol. Biol. **27**, 57–92
4 Carlstedt, I., Sheehan, J. K., Corfield, A. P. and Gallagher, J. T. (1985) Mucous glycoproteins: a gel of a problem. Essays Biochem. **20**, 40–75
5 Forstner, J. F. and Forstner, G. G. (1994) Gastrointestinal mucus. In Physiology of the Gastrointestinal Tract. (Johnson, L. K., ed.), pp. 1245–1283, Raven Press, New York
6 Bobek, L. A., Tsai, H., Biesbrock, A. R. and Levine, M. J. (1993) Molecular cloning, sequence, and specificity of expression of the gene encoding the low molecular weight human salivary mucin (MUC7). J. Biol. Chem. **268**, 20563–20569
7 Gendler, S. J. and Spicer, A. P. (1995) Epithelial mucin genes. Annu. Rev. Physiol. **57**, 607–634
8 Sherblom, A. P. and Carraway, K. L. (1980) A complex of two cell surface glycoproteins from ascites mammary adenocarcinoma cells. J. Biol. Chem. **255**, 12051–12059
9 Spielman, J., Rockley, N. L. and Carraway, K. L. (1987) Temporal aspects of O-glycosylation and cell surface expression of ascites sialoglycoprotein-1, the major cell surface sialomucin of 13762 mammary ascites tumor cells. J. Biol. Chem. **262**, 269–275
10 Hull, S. R., Sheng, Z., Vanderpuye, O. A., David, C. and Carraway, K. L. (1990) Isolation and partial characterization of ascites sialoglycoprotein-2 of the cell surface sialomucin complex of 13762 rat mammary adenocarcinoma cells. Biochem. J. **265**, 121–129
11 Sheng, Z., Hull, S. R. and Carraway, K. L. (1990) Biosynthesis of the cell surface sialomucin complex of ascites 13762 rat mammary adenocarcinoma cells from a high molecular weight precursor. J. Biol. Chem. **265**, 8505–8510
12 Sheng, Z., Wu, K., Carraway, K. L. and Fregien, N. (1992) Molecular cloning of the transmembrane component of the 13762 mammary adenocarcinoma sialomucin complex. A new member of the epidermal growth factor superfamily. J. Biol. Chem. **267**, 16341–16346
13 Wu, K., Fregien, N. and Carraway, K. L. (1994) Molecular cloning and sequencing of the mucin subunit of a heterodimeric, bifunctional cell surface glycoprotein complex of ascites rat mammary adenocarcinoma cells. J. Biol. Chem. **269**, 11950–11955
14 Moniaux, N., Nollet, S., Porchet, N., Degand, P., Laine, A. and Aubert, J.-P. (1999) Complete sequence of the human mucin MUC4: a putative cell membrane-associated mucin. Biochem. J. **338**, 325–333

15 Porchet, N., Nguyen, V. C., Dufosse, J., Audie, J. P., Guyonnet Duperat, V., Gross, M. S., Denis, C., Degand, P., Bernheim, A. and Aubert, J. P. (1991) Molecular cloning and chromosomal localization of a novel human tracheo-bronchial mucin cDNA containing tandemly repeated sequences of 48 base pairs. Biochem. Biophys. Res. Commun. **175**, 414–422
16 Sherblom, A. P., Buck, R. L. and Carraway, K. L. (1980) Purification of the major sialoglycoproteins of 13762 MAT-B1 and MAT-C1 rat ascites mammary adenocarcinoma cells by density gradient centrifugation in cesium chloride and guanidine hydrochloride. J. Biol. Chem. **255**, 783–790
17 Komatsu, M., Carraway, C. A. C., Fregien, N. L. and Carraway, K. L. (1997) Reversible disruption of cell-matrix and cell-cell interactions by overexpression of sialomucin complex. J. Biol. Chem. **272**, 33245–33254
18 Howard, S. C., Hull, S. R., Huggins, J. W., Carraway, C. A. C. and Carraway, K. L. (1982) Relationship between xenotransplantability and cell surface properties of ascites sublines of a rat mammary adenocarcinoma. J. Nat. Cancer Inst. **69**, 33–40
19 Carraway, K. L. and Spielman, J. (1986) Structural and functional aspects of tumor cell sialomucins. Mol. Cell. Biochem. **72**, 109–120
20 Price-Schiavi, S. A., Meller, D., Jing, X., Merritt, J., Carvajal, M. E., Tseng, S.C.G. and Carraway, K. L. (1998) Sialomucin complex at the rat ocular surface: a new model for ocular surface protection. Biochem. J. **335**, 457–463
21 Idris, N. and Carraway, K. L. (1999) Sialomucin complex (Muc4) expression in the rat female reproductive tract. Biol. Reprod. **61**, 1431–1438
22 Carraway, K. L. III, Rossi, E. A., Komatsu, M., Price-Schiavi, S. A., Huang, D., Carvajal, M. E., Guy, P. M., Fregien, N., Carraway, C. A. C. and Carraway, K. L. (1999) An intramembrane modulator of the ErbB2 receptor tyrosine kinase that potentiates neuregulin signaling. J. Biol. Chem. **274**, 5263–5266
23 Carraway, K. L., Fregien, N., Carraway, C. A. C. and Carraway, K. L. III. (1992) Tumor sialomucin complexes as tumor antigens and modulators of cellular interactions and proliferation. J. Cell Sci. **103**, 299–307
24 Carraway, K. L. III and Cantley, L. C. (1994) A neu acquaintance for erbB3 and erbB4: a role for receptor heterodimerization in growth signaling. Cell **78**, 5–8
25 Juang, S.-H., Carvajal, M. E., Whitney, M., Liu, Y. and Carraway, C. A. C. (1996) Tyrosine phosphorylation at the membrane-microfilament interface: a p185neu-associated signal transduction particle containing Src, Abl and phosphorylated p58, a membrane- and microfilament-associated retroviral gag-like protein. Oncogene **12**, 1033–1042
26 Audie, J. P., Janin, A., Prochet, 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
27 Audie, J. P., Tetaert, D., Pigny, P., Buisine, M. P., Janin, A., Aubert, J. P., Porchet, N. and Boersma, A. (1995) Mucin gene expression in the human endocervix. Hum. Reprod. **10**, 98–102
28 Carraway, K. L., Price-Schiavi, S. A., Komatsu, M., Idris, N., Perez, A., Li, P., Jepson, S., Zhu, X., Carvajal, M. E. and Carraway, C. A. C. (2000) Multiple facets of sialomucin complex/MUC4, a membrane mucin and Erbb2 ligand, in tumors and tissues (Y2K Update). Front. Biosci. **5**, 95–107
29 Price-Schiavi, S. A., Carraway, C. A. C., Fregien, N. and Carraway, K. L. (1998) Post-transcriptional regulation of a milk membrane protein, the sialomucin complex (Ascites sialoglycoprotein (ASGP)-1/ASGP-2, rat muc4), by transforming growth factor beta. J. Biol. Chem. **273**, 35228–35237
30 Rossi, E. A., McNeer, R., Price-Schiavi, S. A., Komatsu, M., Van den Brande, J. M. H., Thompson, J. F., Carraway, C. A. C., Fregien, N. L. and Carraway, K. L. (1996) Sialomucin complex, a heterodimeric glycoprotein complex. Expression as a soluble, secretable form in lactating mammary gland and colon. J. Biol. Chem. **271**, 33476–33485
31 Walsh, M. D., McGuckin, M. A., Devine, P. L., Hohn, B. G. and Wright, R. G. (1993) Expression of MUC2 epithelial mucin in breast carcinoma. J. Clin. Pathol. **46**, 922–925
32 Reference deleted
33 Berger, J. T., Voynow, J. A., Peters, K. W. and Rose, M. C. (1999) Respiratory carcinoma cell lines. MUC genes and glycoconjugates. Am. J. Respir. Cell Mol. Biol. **3**, 500–510
34 Ogata, S., Uehara, H., Chen, A. and Itzkowitz, S. H. (1992) Mucin gene expression in colonic tissues and cell lines. Cancer Res. **52**, 5971–5978
35 Nollet, S., Moniaux, N., Maury, J., Petitprezk, D., Degand, P., Laine, A., Porchet, N. and Aubert, J.-P. (1998) Human mucin gene MUC4: organization of its 5′-region and polymorphism of its central tandem repeat array. Biochem. J. **332**, 739–748
36 Gum, J. R., Ho, J. J.L, Pratt, W. S., Hicks, J. W., Hill, A. S., Vinall, L. E., Roberton, A. M., Swallow, D. M. and Kim, Y. S. (1997) MUC3 human intestinal mucin. Analysis of gene structure, the carboxyl terminus, and a novel upstream repetitive region. J. Biol. Chem. **272**, 26678–26686
37 Kovarik, A., Lu, P. J., Peat, N., Morris, J. and Taylor-Papadimitriou, J. (1996) Two GC boxes (Sp1 sites) are involved in regulation of the activity of the epithelium-specific MUC1 promoter. J. Biol. Chem. **271**, 18140–18147
38 Gollub, E. G., Waksman, H., Goswami, S. and Marom, Z. (1995) Mucin genes are regulated by estrogen and dexamethasone. Biochem. Biophys. Res. Commun. **217**, 1006–1014
39 Nogami, H., Ohmori, H. Li, J.-D., Gallup, M. Gum, J., Kim, Y. and Basbaum, C. (1997) Sp1 protein contributes to airway-specific rat MUC 2 mucin gene transcription. Gene **198**, 191–201
40 McNeer, R. R., Carraway, C. A. C., Fregien, N. and Carraway, K. L. (1998) Characterization of the expression and steroid hormone control of sialomucin complex in the rat uterus: implications for uterine receptivity. J. Cell. Physiol. **176**, 110–119

Received 8 March 2000/18 April 2000; accepted 5 May 2000