THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 276, No. 21, Issue of May 25, pp. 18327–18336, 2001
© 2001 by The American Society for Biochemistry and Molecular Biology, Inc.
Printed in U.S.A.

MUC13, a Novel Human Cell Surface Mucin Expressed by Epithelial and Hemopoietic Cells*

Received for publication, September 28, 2000, and in revised form, February 20, 2001
Published, JBC Papers in Press, February 26, 2001, DOI 10.1074/jbc.M008850200

Stephanie J. Williams‡, Daniel H. Wreschner§, Mai Tran‡, Helen J. Eyre¶, Grant R. Sutherland¶, and Michael A. McGuckin‡‖

From the ‡Mater Medical Research Institute, Level 3, Aubigny Place, Mater Misericordiae Hospitals, Raymond Terrace, South Brisbane, Queensland 4101, Australia, the §Department of Cell Research and Immunology, George S. Wise Faculty of Life Sciences, Tel Aviv University, Tel Aviv 69978, Israel, and the ¶Centre for Medical Genetics, Department of Cytogenetics and Molecular Genetics, Women's and Children's Hospital, 72 King William Road, North Adelaide, South Australia 5006, Australia

Transmembrane mucins are glycoproteins involved in barrier function in epithelial tissues. To identify novel transmembrane mucin genes, we performed a tblastn search of the GenBank™ EST data bases with a serine/threonine-rich search string, and a rodent gene expressed in bone marrow was identified. We determined the cDNA sequence of the human orthologue of this gene, *MUC13*, which localizes to chromosome band 3q13.3 and generates 3.2-kilobase pair transcripts encoding a 512-amino acid protein comprised of an N-terminal mucin repeat domain, three epidermal growth factor-like sequences, a SEA module, a transmembrane domain, and a cytoplasmic tail (GenBank™ accession no. AF286113). *MUC13* mRNA is expressed most highly in the large intestine and trachea, and at moderate levels in the kidney, small intestine, appendix, and stomach. *In situ* hybridization in murine tissues revealed expression in intestinal epithelial and lymphoid cells. Immunohistochemistry demonstrated the human MUC13 protein on the apical membrane of both columnar and goblet cells in the gastrointestinal tract, as well as within goblet cell thecae, indicative of secretion in addition to presence on the cell surface. MUC13 is cleaved, and the β-subunit containing the cytoplasmic tail undergoes homodimerization. Including MUC13, there are at least five cell surface mucins expressed in the gastrointestinal tract.

The epithelial mucins are a family of secreted and cell-surface glycoproteins expressed by ductal and glandular epithelial tissues. Members of this family are characterized by a tandem repeat structure, which comprises most of the protein backbone and is the scaffold for a large number of complex *O*-linked carbohydrate side chains. Twelve human epithelial mucin genes have been identified to date, and it is now clear that there are two structurally and functionally distinct classes of mucins: secreted gel-forming mucins and transmembrane mucins.

Gel-forming mucins are secreted by goblet cells and are the major constituent of mucus, which protects and lubricates epithelial surfaces, particularly those of the gastrointestinal, respiratory and reproductive tracts. These mucins undergo homo-oligomerization and are clearly the critical contributors to the biophysical properties of mucus (1). The gel-forming mucins are encoded by a cluster of four mucin genes (*MUC2*, *MUC5AC*, *MUC5B*, and *MUC6*) on chromosome band 11p15.5 (2). The *MUC7* gene encodes a relatively small secreted salivary mucin (3), and MUC9 is a relatively small secreted mucin expressed only in the Fallopian tube (4). The *MUC8* cDNA, expressed in respiratory tissue, has not been well characterized (5).

It has recently emerged that there is a large family of epithelial transmembrane mucins, many members of which are highly expressed in the gastrointestinal tract, particularly in the large intestine. The *MUC1* gene encodes a transmembrane mucin expressed by almost all human glandular epithelial tissues and throughout all regions of the gastrointestinal tract (6). The *MUC3* and *MUC4* genes were initially thought to encode secreted mucins; however, it has recently been shown that these genes encode transmembrane mucins as their dominant isoforms (7, 8). There is a newly identified cluster of transmembrane mucin genes located on chromosome band 7q22 (9). Members of this cluster include the closely related *MUC3* and *MUC12* genes (10, 9). *MUC3* is most highly expressed in the small intestine, whereas *MUC12* is most highly expressed in the colon (9). Further genes may also lie in this cluster, and the genomic organization of these genes should be clarified shortly by the Human Genome Project. The *MUC4* gene located on chromosome band 3q29 appears to be related to genes in the 7q22 cluster, as the mucin it encodes contains a similar domain structure and, although most highly expressed in the tracheobronchus, is also present in the intestine (8, 11).

Common structural features shared by all the identified transmembrane mucins include C-terminal cytoplasmic domains with motifs suggestive of a role in signal transduction, a SEA module in the extracellular domain (12), and a large mucin domain at the N terminus of the extracellular subunit. The two mucins studied biochemically to date, human MUC1 and rat Muc4, undergo cleavage and non-covalent reassociation during biosynthesis, with disruption of this association allowing shedding of the mucin domain from the cell surface (13, 14). Thus, the ability to shed a large mucin protein from the cell surface and the ability to transmit signals have both been heavily conserved. In addition, MUC3, MUC4, and

* This work was supported by Grants 99/0339 and 98/1291 (to M. A. M.) from the National Health and Medical Research Council of Australia, grants (to D. H. W.) from the Susan Komen Breast Cancer Foundation and the Middle East Cancer Consortium, and a grant (to G. R. S.) from the Wellcome Trust. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact. *The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AF286113.*

‖ To whom correspondence should be addressed. Tel.: 61-7-38402568; Fax: 61-7-38402550; E-mail: mmcguckin@mmri.mater.org.au.

Downloaded from http://www.jbc.org/ at UQ Library on September 1, 2016

MUC12 all contain two cysteine-rich EGF[1]-like domains, one to either side of the SEA module. The function of EGF-like motifs in these mucins is unclear, although there is evidence implicating these domains in epithelial growth modulation (15).

In the present investigation, we have used data base mining to identify a novel human transmembrane mucin gene that has been designated *MUC13* by the HUGO Gene Nomenclature Committee. The rodent orthologues of *MUC13* have been previously identified and were thought to be restricted to expression in hemopoietic precursor cells in bone marrow (16). We describe the cDNA sequence of human *MUC13* and show that the putative protein has the same domain organization as MUC3, MUC4, and MUC12, and similarities to MUC1. We also demonstrate that both the human and mouse genes are expressed predominantly in epithelial tissues, particularly those of the gastrointestinal tract, and demonstrate the human protein on the epithelial apical membrane and within goblet cell thecae of gastrointestinal epithelia. Initial biochemical characterization indicates that MUC13 is cleaved into two subunits and that the subunit containing the cytoplasmic tail undergoes homodimerization.

EXPERIMENTAL PROCEDURES

Identification and Cloning of the Human MUC13 cDNA—A serine/threonine-rich search string, TT*X*S*X*SPGS*X*SPX*X*STTT, was employed in an advanced parameter tblastn search of the National Center of Biotechnology Information human and mouse EST data bases. A match in the mouse EST data base led to the identification of a known full-length cDNA, named 114/A10. Examination of the human EST data base identified partial sequences of the putative human orthologue of this gene. Using a combination of sequencing ESTs and RT-PCR, a 2.9-kb cDNA composite of the human gene was obtained (see "Results").

RNA Expression Analyses—RNA was extracted from cell lines and human tissue specimens using Trizol (Life Technologies, Inc., Mulgrave, Victoria, Australia) as per the manufacturer's instructions. Northern blot analysis was performed using standard protocols (17) on total RNA isolated from five colorectal cancer cell lines (LIM2463, LoVo, LS513, Caco-2, and SW116) and from 12 paired normal colonic mucosa and tumor tissues and 7 liver metastases, Dukes' A (n = 4), Dukes' B (n = 3), Dukes' C (n = 4), and Dukes' D (n = 7). Informed consent was obtained from each subject after approval by the Princess Alexandra Hospital Ethics Committee. A *MUC13* probe (base pair sequence 673–1756 in Fig. 1) was random primer-labeled using a Megaprime DNA labeling system (Amersham Pharmacia Biotech, Aylesbury, United Kingdom) and hybridized to Northern blots at 65 °C in buffer containing 7% SDS, 0.26 M Na_2HPO_4, 1 mM EDTA, and 1% bovine serum albumin.

A human RNA Multiple Tissue Expression Array (CLONTECH, Palo Alto, CA, catalogue no. 7775-1) was used to examine *MUC13* expression in normal human tissues (RNA from 71 adult tissues, 7 fetal tissues, and 8 cancer cell lines). The *MUC13* cDNA probe (as above) was hybridized as per the manufacturer's instructions. Levels of expression were quantitated using a scanning densitometer and Molecular Analyst software (Bio-Rad, Regents Park, New South Wales, Australia). Signal intensity was expressed relative to the sample with highest levels of expression since loading of RNA samples on the array had previously been normalized using eight housekeeping genes by the manufacturer.

Mouse intestinal cDNA was amplified by RT-PCR to yield a 768-base pair PCR product (corresponding to nucleotides 182–950 in GenBank™ accession no. J04634) that was cloned into pGEM-T and sequenced. Riboprobes were made by *in vitro* transcription of DNA with T7 and SP6 RNA polymerases and incorporation of digoxigenin-labeled uridine triphosphate (Roche, Castle Hill, New South Wales, Australia). Approval for collection of animal tissues was obtained from the University of Queensland Animal Experimentation Ethics Committee. Murine epithelial tissues and spleen were fixed for 4 h in fresh 4% paraformaldehyde in 10 mM phosphate-buffered saline, pH 7.2 (PBS), embedded in paraffin, and sectioned. Hybridization and washing conditions were modified from that described previously (18). Antisense and sense probes at 1 μg/ml in ULTRAhyb (Ambion, Austin, TX) were hybridized to sections overnight at 48 °C. The reaction was visualized with 4-bromo-5-chloro-3-indolyl phosphate and nitro blue tetrazolium substrate (Roche) also overnight. Sections were counterstained with nuclear fast red, dehydrated, cleared, and mounted in DePeX.

Chromosomal Localization of MUC13 by Fluorescence in Situ Hybridization—The *MUC13* cDNA probe (as above) was nick-translated with biotin-14-dATP and hybridized *in situ* at a final concentration of 10 ng/μl to metaphases from two normal males. The fluorescence *in situ* hybridization method was modified from that described previously (19), in that chromosomes were stained before analysis with both propidium iodide as counterstain and 4,6-diamidino-2-phenylindole for chromosome identification. Images of metaphase preparations were captured by a cooled CCD camera using the CytoVision Ultra image collection and enhancement system (Applied Imaging Corp., Sunderland, United Kingdom).

Production of MUC13-reactive Polyclonal Antibodies—Three synthetic peptides were synthesized (Auspep, Parkville, Australia) corresponding to three hydrophilic sequences of the putative MUC13 amino acid sequence, each with an terminal cysteine residue added: an extracellular domain epitope (peptide A, DPEEKHSMAYQDLHSEC (amino acids 229–244 in Fig. 1)) and two cytoplasmic tail epitopes (peptide B, CRSNNTKHIEEENLID (amino acids 446–461 in Fig. 1); and peptide C, CMQNPYSRHSSMPRPDY (amino acids 497–512 in Fig. 1)). These peptides were conjugated to bovine serum albumin using glutaraldehyde (20). Six-week-old female BALB/c mice were immunized intraperitoneally with 25 μg of conjugated peptide initially in 0.2 ml of complete Freund's adjuvant/PBS 1:1 (Life Technologies Inc.), and then at 3-week intervals in incomplete Freund's adjuvant. Blood samples were obtained via cardiac puncture under terminal anesthesia, clotted, and serum stored at −20 °C. Reactivity of serum with the peptides was assessed using the specific and irrelevant unconjugated peptides as solid phase in an enzyme-linked immunosorbent assay. All animal experiments were approved by the University of Queensland Animal Experimentation Ethics Committee.

Immunohistochemical Detection of Human MUC13 in Epithelial Tissues—Polyclonal MUC13 peptide-reactive mouse sera were used to detect MUC13 in paraffin sections of normal and diseased human gastrointestinal epithelial tissues. Immunohistochemical techniques were as described previously (21) with peptide A and B reactive sera diluted 1/100 and tissue sections subject to antigen retrieval by boiling in 0.1 M citric acid, pH 6, and peptide C reactive sera diluted 1/400 with antigen retrieval not used.

Biochemical Characterization of the MUC13 Protein—Colorectal cancer cell lines LIM2463 and LS513 were maintained in RPMI 1640 containing 10% fetal bovine serum, and routinely passaged by trypsin digestion. Monolayers of cells were washed with ice-cold PBS and lysed, with rocking for 30 min, using a variety of ice-cold lysis buffers including: RIPA (50 mM Tris, 150 mM NaCl, 0.1% SDS, 0.5% sodium deoxycholate, 1 mM PMSF, pH 7.5), Nonidet P-40 (50 mM Tris, 150 mM NaCl, 0.5% Nonidet P-40, 1 mM PMSF, pH 7.5), BRIJ97 (50 mM Tris, 150 mM NaCl, 0.5% Brij97, 1 mM aprotinin, pH 7.6), and TX100 (50 mM Tris, 150 mM NaCl, 0.5% Triton X-100, 1 mM PMSF, pH 7.5). Cell surface proteins were biotinylated by incubating cell monolayers with 0.5 mg/ml biotinamidocaproic acid 3-sulfo-*N*-hydroxysuccinimide (Sigma) in PBS with 0.1 mM $CaCl_2$ and 1 mM $MgCl_2$, followed by lysis as above. MUC13 was immunoprecipitated with polyclonal antisera against peptides A (1/100 dilution), B (1/100), or C (1/400) or with antisera raised against irrelevant conjugated peptides as a control, using previously described methods (22). For removal of *N*-linked glycans, immunoprecipitates were denatured in 0.5% SDS, 0.1 M 2-mercaptoethanol for 5 min at 100 °C, reconstituted to 0.15 M Tris, 0.16% SDS, 1.6% Nonidet P-40, pH 7.5, and 100 units/ml PNGase F (Roche) for 24 h at 32 °C. Crude lysates, immunoprecipitated biotinylated cell surface proteins and deglycosylated immunoprecipitates were prepared in SDS-PAGE buffer containing 2% SDS, with and without 5% 2-mercaptoethanol, subjected to electrophoresis either in 4–20% gradient or 10% resolving gels with 3% stacking gels, and electrotransferred to polyvinylidene difluoride or nitrocellulose membranes as described (22). Benchmark™ prestained M_r markers were included on each gel (Life Technologies, Inc.). Western blotting was performed as described previously (22). Briefly, membranes were blocked with 10% skim milk powder and MUC13 was detected using polyclonal antisera against peptides A (1/400 dilution), B (1/400), or C (1/4000) followed by donkey anti-mouse horseradish peroxidase conjugate (Jackson Laboratories) and ECL detection (Amer-

[1] The abbreviations used are: EGF, epidermal growth factor; CF, cystic fibrosis; PBS, phosphate-buffered saline; PCR, polymerase chain reaction; EST, expressed sequence tag; PAGE, polyacrylamide gel electrophoresis; RT, reverse transcription; kb, kilobase pair(s); PMSF, phenylmethylsulfonyl fluoride; PNA, *A. hypogaea* lectin; HPA, *H. pomatia* lectin; RIPA, radioimmune precipitation buffer; PNGase F, Peptide: N-glycosidase F.

Downloaded from http://www.jbc.org/ at UQ Library on September 1, 2016

```
1 CCACGCGTCCGAGCAAGAACAGCTAAAATGAAAGCCATCATTCATCTTAC
                  M  K  A  I  I  H  L  T   8
51 TCTTCTTGCTCTCCTTTCTGTAAACACAGCCACCAACCAAGGCAACTCAG
     L  L  A  L  L  S  V  N  T  A  T  N  Q  G  N  S   24
101 CTGATGCTGTAACAACCACAGAAACTGCGACTAGTGGTCCTACAGTAGCT
     A  D  A  V  T  T  T  E  T  A  T  S  G  P  T  V  A   41
151 GCAGCTGATACCACTGAAACTAATTTCCCTGAAACTGCTAGCACCACAGC
     A  A  D  T  T  E  T  N  F  P  E  T  A  S  T  T  A   58
201 AAATACACCTTCTTTCCCAACAGCTACTTCACCTGCTCCCCCCATAATTA
     N  T  P  S  F  P  T  A  T  S  P  A  P  P  I  I   74
251 GTACACATAGTTCCTCCACAATTCCTACACCTGCTCCCCCCATAATTAGT
     S  T  H  S  S  S  T  I  P  T  P  A  P  P  I  I  S   91
301 ACACATAGTTCCTCCACAATTCCTATACCTACTGCTGCAGACAGTGAGTC
     T  H  S  S  S  T  I  P  I  P  T  A  A  D  S  E  S  108
351 AACCACAAATGTAAATTCATTAGCTACCTCTGACATAATCACCGCTTCAT
     T  T  N  V  N  S  L  A  T  S  D  I  I  T  A  S   124
401 CTCCAAATGATGGATTAATCACAATGGTTCCTTCTGAAACACAAAGTAAC
     S  P  N  D  G  L  I  T  M  V  P  S  E  T  Q  S  N   141
451 AATGAAATGTCCCCCACCACAGAAGACAATCAATCATCAGGGCCTCCCAC
     N  E  M  S  P  T  T  E  D  N  Q  S  S  G  P  P  T   158
501 TGGCACCGCTTTATTGGAGACCAGCACCCTAAACAGCACAGGTCCCAGCA
     G  T  A  L  L  E  T  S  T  L  N  S  T  G  P  S   174
551 ATCCTTGCCAAGATGATCCCTGTGCAGATAATTCGTTATGTGTTAAGCTG
     N  P  C  Q  D  D  P  C  A  D  N  S  L  C  V  K  L   191
601 CATAATACAAGTTTTTGCCTGTGTTTAGAAGGGTATTACTACAACTCTTC
     H  N  T  S  F  C  L  C  L  E  G  Y  Y  Y  N  S  S   208
651 TACATGTAAGAAAGGAAAGGTATTCCCTGGGAAGATTTCAGTGACAGTAT
     T  C  K  K  G  K  V  F  P  G  K  I  S  V  T  V   224
701 CAGAAACATTTGACCCAGAAGAGAAACATTCCATGGCCTATCAAGACTTG
     S  E  T  F  D  P  E  E  K  H  S  M  A  Y  Q  D  L   241
751 CATAGTGAAATTACTAGCTTGTTTAAAGATGTATTTGGCACATCTGTTTA
     H  S  E  I  T  S  L  F  K  D  V  F  G  T  S  V  Y   258
801 TGGACAGACTGTAATTCTTACTGTAAGCACATCTCTGTCACCAAGATCTG
     G  Q  T  V  I  L  T  V  S  T  S  L  S  P  R  S   274
851 AAATGCGTGCTGATGACAAGTTTGTTAATGTAACAATAGTAACAATTTTG
     E  M  R  A  D  D  K  F  V  N  V  T  I  V  T  I  L   291
901 GCAGAAACCACAAGTGACAATGAGAAGACTGTGACTGAGAAAATTAATAA
     A  E  T  T  S  D  N  E  K  T  V  T  E  K  I  N  K   308
951 AGCAATTAGAAGTAGCTCAAGCAACTTTCTAAACTATGATTTGACCCTTC
     A  I  R  S  S  S  S  N  F  L  N  Y  D  L  T  L   324
1001 GGTGTGATTATTATGGCTGTAACCAGACTGCGGATGACTGCCTCAATGGT
     R  C  D  Y  Y  G  C  N  Q  T  A  D  D  C  L  N  G   341
1051 TTAGCATGCGATTGCAAATCTGACCTGCAAAGGCCTAACCCACAGAGCCC
     L  A  C  D  C  K  S  D  L  Q  R  P  N  P  Q  S  P   358
1101 TTTCTGCGTTGCTTCCAGTCTCAAGTGTCCTGATGCCTGCAACGCACAGC
     F  C  V  A  S  S  L  K  C  P  D  A  C  N  A  Q   374
1151 ACAAGCAATGCTTAATAAAGAAGAGTGGTGGGGCCCCTGAGTGTGCGTGC
     H  K  Q  C  L  I  K  K  S  G  G  A  P  E  C  A  C   391
1201 GTGCCCGGCTACCAGGAAGATGCTAATGGGAACTGCCAAAAGTGTGCATT
     V  P  G  Y  Q  E  D  A  N  G  N  C  Q  K  C  A  F   408
1251 TGGCTACAGTGGACTCGACTGTAAGGACAAATTTCAGCTGATCCTCACTA
     G  Y  S  G  L  D  C  K  D  K  F  Q  L  I  L  T   424
1301 TTGTGGGCACCATCGCTGGCATTGTCATTCTCAGCATGATAATTGCATTG
     I  V  G  T  I  A  G  I  V  I  L  S  M  I  I  A  L   441
1351 ATTGTCACAGCAAGATCAAATAACAAAACGAAGCATATTGAAGAAGAGAA
     I  V  T  A  R  S  N  N  K  T  K  H  I  E  E  E  N   458
1401 CTTGATTGACGAAGACTTTCAAAATCTAAAACTGCGGTCGACAGGCTTCA
     L  I  D  E  D  F  Q  N  L  K  L  R  S  T  G  F   474
1451 CCAATCTTGGAGCAGAAGGGAGCGTCTTTCCTAAGGTCAGGATAACGGCC
     T  N  L  G  A  E  G  S  V  F  P  K  V  R  I  T  A   491
1501 TCCAGAGACAGCCAGATGCAAAATCCCTATTCAAGACACAGCAGCATGCC
     S  R  D  S  Q  M  Q  N  P  Y  S  R  H  S  S  M  P   508
1551 CCGCCCTGACTATTAGAATCATAAGAATGTGGAACCCGCCATGGCCCCCA
     R  P  D  Y  *   512
1601 ACCAATGTACAAGCTATTATTTAGAGTGTTTAGAAAGACTGATGGAGAAG
1651 TGAGCACCAGTAAAGATCTGGCCTCCGGGGTTTTTCTTCCATCTGACATC
1701 TGCCAGCCTCTCTGAATGGAAGTTGTGAATGTTTGCAACGAATCCAGCTC
1751 ACTTGCTAAATAAGAATCTATGACATTAAATGTAGTAGATGCTATTAGCG
1801 CTTGTCAGAGAGGTGGTTTTCTTCAATCAGTACAAAGTACTGAGACAATG
1851 GTTAGGGTTGTTTTCTTAATTCTTTTCCTGGTAGGGCAACAAGAACCATT
1901 TCCAATCTAGAGGAAAGCTCCCCAGCATTGCTTGCTCCTGGGCAAACATT
1951 GCTCTTGAGTTAAGTGACCTAATTCCCCTGGGAGACATACGCATCAACTG
2001 TGGAGGTCCGAGGGGATGAGAAGGGATACCCACCACCTTTCAAGGGTCAC
2051 AAGCTCACTCTCTGACAAGTCAGAATAGGGACACTGCTTCTATCCCTCCA
2101 ATGGAGAGATTCTGGCAACCTTTGAACAGCCCAGAGCTTGCAACCTAGCC
2151 TCACCCAAGAAGACTGGAAAGAGACATATCTCTCAGCTTTTTCAGGAGGC
2201 GTGCCTGGGAATCCAGGAACTTTTTGATGCTAATTAGAAGGCCTGGACTA
2251 AAAATGTCCACTATGGGGTGCACTCTACAGTTTTTGAAATGCTAGGAGGC
2301 AGAAGGGGCAGAGAGTAAAAAACATGACCTGGTAGAAGGAAGAGAGGCAA
2351 AGGAAACTGGGTGGGGAGGATCAATTAGAGAGGAGGCACCTGGGATCCAC
2401 CTTCTTCCTTAGGTCCCCTCCTCCATCAGCAAAGGAGCACTTCTCTAATC
2451 ATGCCCTCCCGAAGACTGGCTGGGAGAAGGTTTAAAAACAAAAAATCCAG
2501 GAGTAAGAGCCTTAGGTCAGTTTGAAATTGGAGACAAACTGTCTGGCAAA
2551 GGGTGCGAGAGGGAGCTTGTGCTCAGGAGTCCAGCCGTCCAGCCTCGGGG
2601 TGTAGGTTTCTGAGGTGTGCCATTGGGGCCTCAGCCTTCTCTGGTGACAG
2651 AGGCTCAGCTGTGGCCACCAACACACAACCACACACACACAACCACACAC
2701 ACAAATGGGGGCAACCACATCCAGTACAAGCTTTTACAAATGTTATTAGT
2751 GTCCTTTTTTATTTCTAATGCCTTGTCCTCTTAAAAGTTATTTTATTTGT
2801 TATTATTATTTGTTCTTGACTGTTAATTGTGAATGGTAATGCAATAAAGT
2851 GCCTTTGTTAGATGGTGAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

FIG. 1. **Complete cDNA sequence and predicted amino acid sequence of human MUC13 (GenBank™ accession no. AF286113).** *Numbering* of nucleotides is given on the *left* and amino acids on the *right*. The signal peptide is *underlined*, and seven potential *N*-glycosylation sites are *double underlined*. All cysteine residues and the entire transmembrane domain are enclosed in a *box*. A non-consensus and consensus polyadenylation signal in the 3′-untranslated region are *underlined*.

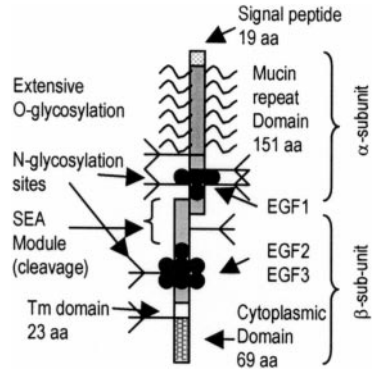

FIG. 2. **Schematic representation of the putative domain organization of the human MUC13 protein.** The number of amino acids (*aa*) in each domain is shown.

sham Pharmacia Biotech). Specific carbohydrates were detected using 2 μg/ml *Helix pomatia* and *Arachis hypogaea* lectins (Sigma). Immunoprecipitated biotinylated proteins and biotinylated lectins were detected with streptavidin-horseradish peroxidase conjugate (Zymed Laboratories Inc., San Francisco, CA) diluted 1/2000 followed by ECL detection.

RESULTS

Identification and Cloning of the Human MUC13 cDNA—Serine/threonine-rich search strings based on known mucin protein sequences were employed in tblastn searches of the NCBI human and mouse EST data bases in a search for previously unidentified mucin genes. Using one such string (TT*XXSXS*PGS*X*SP*X*STTT), a number of matches with the mouse data base revealed the previous identification and cloning of a rodent gene encoding a bone marrow glycoprotein designated 114/A10, or mouse cell surface antigen. 114/A10 appeared similar to known epithelial transmembrane mucins, containing a series of eight highly conserved serine/threonine-rich 27 amino acid repeats leading into three EGF-like domains, a transmembrane region, and a cytoplasmic tail (GenBank™ accession no. J04634) (16). Partial cDNA sequences encoding protein sequences similar to 114/A10 were also evident in the human EST data base (*e.g.* GenBank™ accession nos. AI949096, AI821099, and AI431674). The domain organization of 114/A10 was very similar to the human *MUC3*, *MUC4*, and *MUC12* transmembrane mucin genes, and preliminary RT-PCR indicated expression in human epithelial tissues. Therefore, we pursued complete human cDNA cloning and expression analysis of this gene in human and rodent tissues.

The cDNA sequence of the human mucin gene was extended in both 5′ and 3′ directions by a combination of EST sequencing and RT-PCR. The HUGO Gene Nomenclature Committee has designated this gene *MUC13*. The complete cDNA sequence and the putative amino acid sequence of MUC13 are shown in Fig. 1. The putative human precursor polypeptide comprises 512 amino acids with a predicted molecular mass of 54,703 Da and has an identical domain structure (exemplified in Fig. 2) to the previously described murine protein (16). At the N terminus is a signal peptide for the secretory pathway with cleavage predicted between residues 19 and 20 (23). The signal peptide is followed by a serine- and threonine-rich domain likely to involve extensive *O*-glycosylation (amino acid residues 20–170) consisting of 10 degenerate tandem repeats. Following this mucin domain are two distinct cysteine-rich domains containing EGF-like motifs. Separating the two cysteine-rich domains are 115 amino acids comprising a SEA module (amino acid residues 212–328). The first cysteine-rich domain contains one EGF-like sequence EGF1 (amino acid residues 177–210), and the second larger cysteine-rich domain contains two EGF-like

Downloaded from http://www.jbc.org/ at UQ Library on September 1, 2016

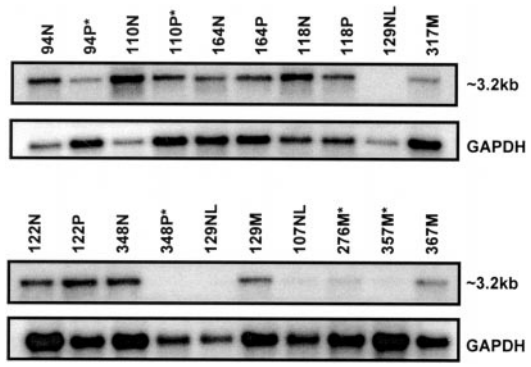

FIG. 3. **Northern blot analysis of *MUC13* gene expression in matched normal colon (*N*) and primary colorectal cancers (*P*), in individual Dukes' stage D tumors (*M*) and normal liver (*NL*).** The 3.2-kb MUC13 band is shown in the *upper panels*, and glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) is shown as a loading control in each *lower panel*. Cancers showing down-regulation are indicated by an *asterisk*.

sequences, EGF2 (amino acid residues 326–360) and EGF3 (amino acid residues 367–403). EGF3 contains a type II EGF signature (amino acid residues 389–403) and is followed closely by a 23-amino acid transmembrane domain and a 69-amino acid cytoplasmic tail. In addition to extensive potential *O*-glycosylation sites in the mucin domain, there are six extracellular and one intracellular consensus motifs for *N*-glycosylation. The cytoplasmic tail contains a protein kinase C consensus phosphorylation motif (amino acid residues 444–447), and eight serine residues and two tyrosine residues that may undergo phosphorylation and regulate *MUC13* signaling.

Expression of MUC13 in Colorectal Cancers and Colorectal Cancer Cell Lines—Northern blot analysis of colonic tissues showed a single band indicating the size of the human mRNA to be ~3.2 kb (see Fig. 3). Analysis of 12 paired samples of normal colonic and primary tumor tissues and seven liver metastases revealed *MUC13* expression in most colorectal cancers examined (Fig. 3). There was evidence for down-regulation of expression (<20% of normal colonic expression level) in three of four Dukes' C tumors and two of seven Dukes' D tumors. There was a 2.5-fold increase in expression of *MUC13* in one Dukes' stage B tumor. *MUC13* was expressed at varying levels in the six colorectal cancer cell lines under investigation, with highest levels of expression in LIM2463 cells and comparatively low levels of expression observed in the Caco-2 cells (data not shown).

Both the human and rodent genes contain a series of CA dinucleotide repeats and an alternative polyadenylation signal in the 3′-untranslated region. *mMuc13* has previously been shown, in some IL-3 dependent cell lines, to produce two transcripts, of ~2.2 and 3 kb, generated by use of two distinct polyadenylation signals (22). Northern blot analyses of *MUC13* in colorectal cancer cell lines and colonic tissues only revealed a single band of about 3.2 kb. However, an EST (GenBank™ accession no. AI431674) appears to have been derived from alternative use of the non-consensus polyadenylation signal, suggesting that smaller human *MUC13* transcripts do exist, but are probably rare.

Fluorescence in Situ Hybridization—Chromosomal localization studies using fluorescence *in situ* hybridization mapped the human gene to chromosome band 3q13.3 (Fig. 4*A*).

Comparison of Human MUC13 with Rat and Mouse Orthologues—Amino acid alignment of the C terminus of the human MUC13 mucin with its rodent orthologues revealed 52% identity to its mouse and rat counterparts (see Fig. 5). The N-terminal mucin domain shows significant divergence from the rodent proteins, which possess a near perfect repeat structure compared with the very degenerate repeats seen in the human protein, although the mucin domains are approximately the same size. The seven C-terminal amino acids of the cytoplasmic domain, including a terminal tyrosine residue, are perfectly conserved in all three species.

Expression of MUC13 mRNA in Human Tissues—Hybridization of MUC13 using a multiple tissue expression array demonstrated that *MUC13* is expressed at highest levels in the large intestine and the trachea, and is also expressed at relatively high levels in the kidney, small intestine, appendix, and stomach (see Fig. 6). After long exposures, low levels of expression were also identified in pancreas, prostate, uterus, lung, liver, adrenal gland, spleen, peripheral blood leukocytes, lymph node, bone marrow, testis, and ovary. *MUC13* was not detected in 21 different neural tissues, eight cardiovascular tissues, esophagus, thyroid gland, salivary gland, mammary gland, skeletal muscle, placenta, or urinary bladder. Fetal kidney expressed moderately high levels of *MUC13*, and low levels were seen in fetal spleen, lung, thymus, and liver. Fetal heart and brain did not express *MUC13*, and other fetal tissues, such as gastrointestinal tissues, were not represented on this array. Two of eight cancer cell lines, one derived from a colorectal cancer, SW620, and the other from a lung cancer, A549, showed moderately high levels of *MUC13* expression.

Cellular Localization of MUC13 in Normal Murine Tissues—In situ hybridization using digoxigenin-labeled riboprobes in murine tissues revealed highest levels of expression in the large intestine. In the large intestine and rectum, *MUC13* mRNA was seen in both columnar and goblet cells deep in the crypts, with lower levels of expression in some cells of the surface epithelium (Fig. 4, *B–E*). A large proportion of cells within lymphoid tissue associated with the distal large intestine and rectum also showed high levels of expression (Fig. 4*F*). *MUC13* was also detected in the villi of the murine small intestine, in gastric glands and the surface epithelium of the stomach, and in the squamous epithelium of the esophagus (data not shown). Pancreatic acini and ducts also expressed low levels of *MUC13*, as did pneumocytes in the lung and both bronchiolar and tracheal epithelium (data not shown). However, *MUC13* mRNA was not detected in the liver or kidney. Clusters of lymphoid cells in both the red pulp and, to a lesser extent, the follicles of the spleen expressed *MUC13* (data not shown).

Antisera Reactive with the Human MUC13 Protein—Murine antisera were produced that were reactive in enzyme-linked immunosorbent assay with the A, B, and C peptide epitopes based on the putative MUC13 sequence. The highest titer was achieved against the C peptide with sera reactive at greater than 1/50,000 dilution. These antisera also reacted with MUC13 by immunohistochemistry with paraffin sections, in immunoprecipitation and Western blotting (see below), and by flow cytometry (data not shown).

Cellular Localization of MUC13 in Human Gastrointestinal Tissues—Polyclonal antisera raised against peptides A, B, and C reacted similarly with paraffin sections of fixed human gastrointestinal tissues, with the peptide C-reactive antisera showing the best reactivity and not requiring antigen retrieval of sections. Surprisingly, given the lack of reactivity on the human mRNA master blot, intense MUC13 staining was observed in the cytoplasm of squamous epithelial cells of the esophagus, with staining absent or very weak in the basal cell layer (see Fig. 4*G*). In the body of the stomach, MUC13 was expressed on the apical membrane surface of cells of the surface epithelium, the gastric pits, and the more peripheral glands but only occasionally in the deep glands (see Fig. 4*H*). In addition, in the gastric glands, occa-

Downloaded from http://www.jbc.org/ at UQ Library on September 1, 2016

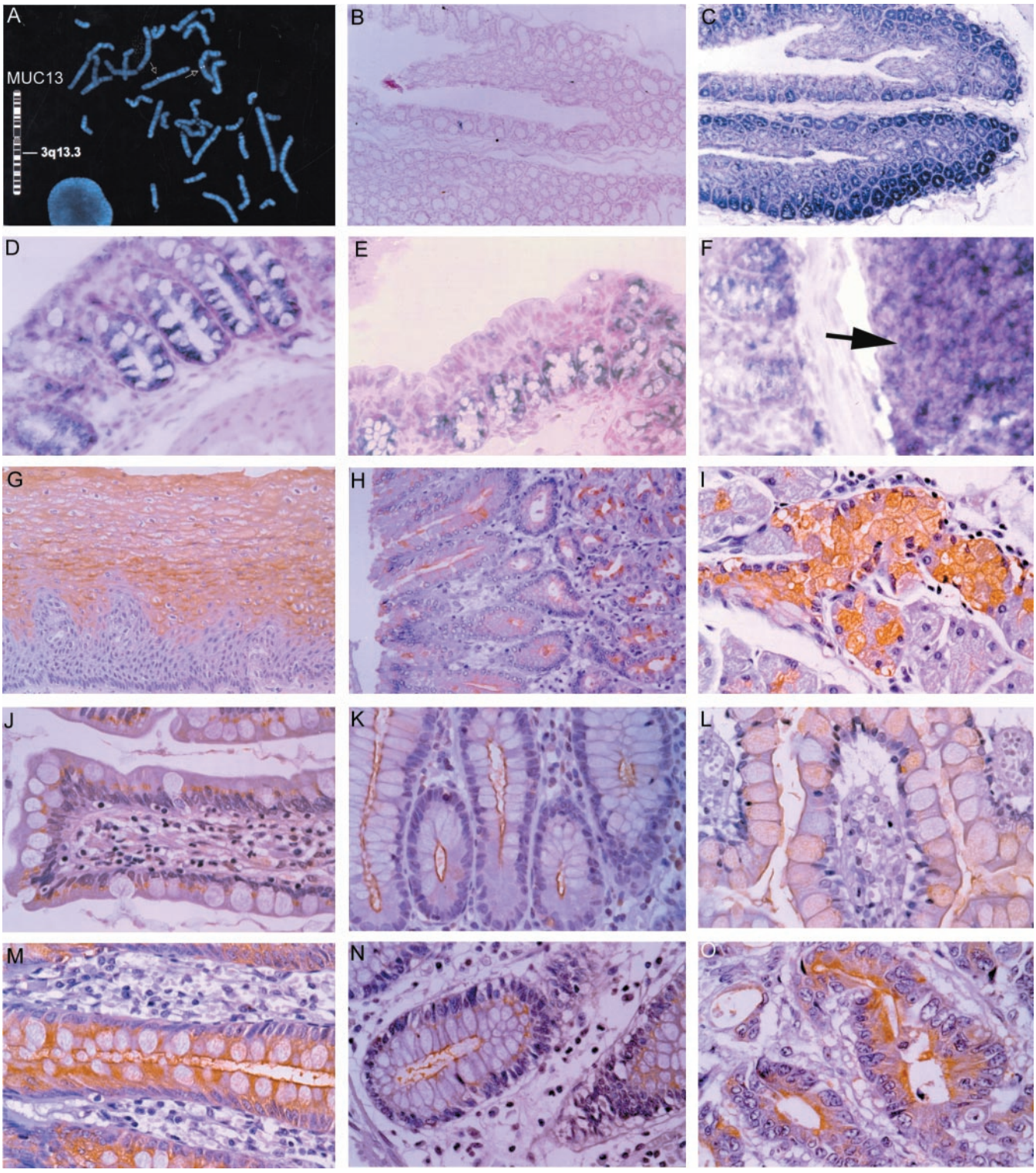

FIG. 4. ***In situ* hybridization and immunohistochemistry.** *A*, chromosomal metaphase spread showing localization of *MUC13* to chromosome band 3q13.3. *B–F*, murine tissue sections hybridized with *MUC13* sense (*B*) and antisense (*C–F*) digoxigenin-labeled probes detected using 4-bromo-5-chloro-3-indolyl phosphate and nitro blue tetrazolium (*blue/purple*) and counterstained with nuclear fast red. Tissues shown: rectum (*B* and *C*), distal large intestine (*D*), proximal large intestine (*E*), rectum and associated lymphoid tissue denoted by *arrow* (*F*). *G–O*, human tissues stained with mouse anti-sera reactive with the C peptide epitope of the MUC13 cytoplasmic tail detected using diaminobenzidine (*brown*) and counterstained with hematoxylin (*blue*; see "Experimental Procedures"). Tissues shown are esophagus (*G*), surface epithelium, gastric pits, and peripheral glands of the body of the stomach (*H*), deep glands of the pyloric stomach (*I*), villi of the duodenum (*J*), deep glands of the terminal ileum (*K*), villi of the terminal ileum (*L*), appendix (*M*), deep glands of the colon (*N*), and a colorectal carcinoma liver metastasis (*O*). Original magnifications were as follows: *B* and *C*, ×55; *D–F*, ×220; *G* and *H*, ×110; *I–O*, ×220.

sional mucus neck cells showed intense granular cytoplasmic staining (see Fig. 4*H*). Some mucus cells of the surface epithelium and gastric pits also showed moderate cytoplasmic reactivity, and in some cells supranuclear staining was also observed. In the pyloric and cardiac stomach, some of the deep glands showed moderate to strong cytoplasmic expression of MUC13, whereas adjacent glands were often negative (Fig. 4*I*). In the duodenum, MUC13 was detected in supranu-

Downloaded from http://www.jbc.org/ at UQ Library on September 1, 2016

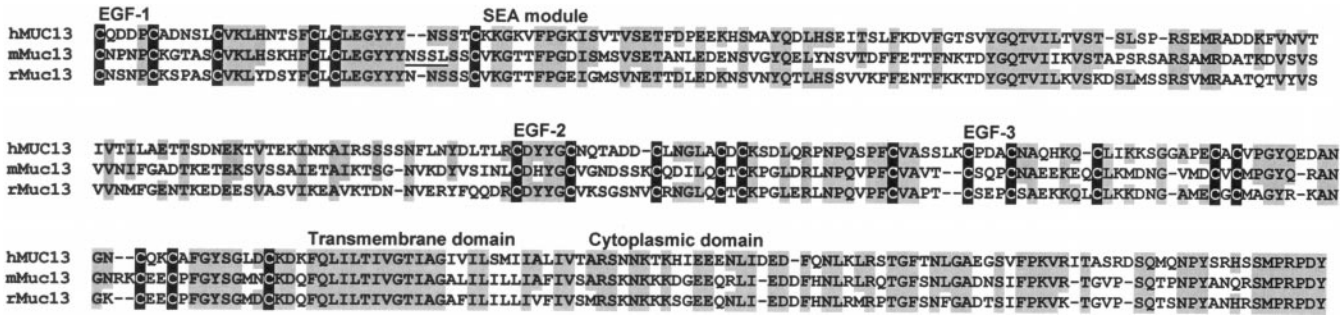

FIG. 5. **Alignment of the carboxyl termini of the human, mouse, and rat Muc13 transmembrane mucins.** *Light shading* indicates identical amino acids, and *dark shading* highlights all cysteine residues. *Hyphens* indicate spaces added to optimize the alignment.

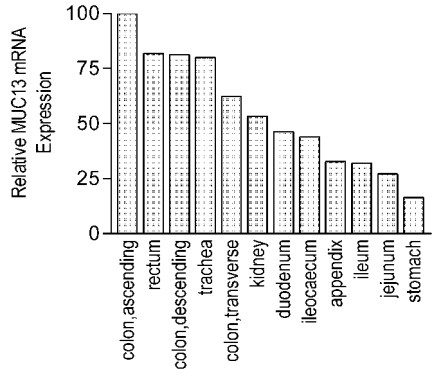

FIG. 6. **Relative mRNA expression of *MUC13* in human tissues determined using a multiple tissue expression array.** Hybridized arrays were subjected to densitometric analysis, and results are expressed as a percentage of the level in the highest expressing tissue, distal large intestine. Only those tissues showing expression greater than 5% of that in distal large intestine are shown.

clear vacuoles within all absorptive cells, consistent with detection in the Golgi region; however, apical membrane staining was not seen (see Fig. 4*J*). MUC13 was also expressed on the apical membrane of epithelial cells lining pancreatic ducts (data not shown).

In the terminal ileum of the small intestine, MUC13 was detected as intense staining of the apical membrane of all cells deep in the crypts (see Fig. 4*K*) and less frequent apical membrane staining of cells of the surface epithelium (see Fig. 4*L*). Both goblet and columnar cells appeared to express cell surface MUC13; however, staining was more intense in columnar cells, and, at high power, a microvillous-type pattern was observed. Secreted material in the crypt lumen also stained. In addition, moderate to weak staining in mesh network and punctate patterns was also seen within the thecae of goblet cells in the villi (see Fig. 4*L*) but only rarely deep in the crypts (see Fig. 4*K*). The staining of goblet cell thecae was more pronounced using peptide A-reactive antisera than C-reactive antisera (data not shown). Appendix showed strong MUC13 immunoreactivity both in the cytoplasm and on the cell surface of both goblet and columnar cells, with very strong cytoplasmic reactivity in columnar cells (see Fig. 4*M*). In the colon, MUC13 was, similarly to the terminal ileum, highly expressed on the apical membrane surface of both columnar and goblet cells deep in the crypts (see Fig. 4*N*). In addition, supranuclear vacuolar staining, like that seen in the duodenum, was also observed in these cells together with reactivity with secreted material. Occasional columnar cell apical membrane staining was observed on the colonic surface epithelium together with goblet cell thecal staining, although this was less intense than that seen in the terminal ileum.

In a small series of colorectal cancers examined, MUC13 was expressed in the cytoplasm and on the cell surface of cancer cells; however, expression of MUC13 was often low relative to normal tissue and was typically heterogeneous in nature (see Fig. 4*O*). Only very rarely did non-epithelial cells in the stroma or lymphoid patches of intestinal tissues show weak membrane staining for MUC13 (data not shown).

Biochemical Characterization of the Human MUC13 Protein—Lysates of LIM2463 and LS513 colorectal cancer cells, shown to have high *MUC13* mRNA, were lysed with RIPA and subjected to Western blotting with antisera reactive with peptides A, B, and C. The two cell lines and three antisera showed identical band patterns. Under non-reducing conditions, two bands of ~47 and 93 kDa were observed. The 47-kDa band was more intense than the 93-kDa band in both cell lines. However, following reduction a single band of 58 kDa was observed (see Fig. 7). This is consistent with the 58-kDa band representing the 47-kDa non-reduced MUC13 protein migrating more slowly following denaturation by reduction of the EGF-like cysteine-rich domains likely to be involved in intrastrand disulfide bonds. The 93-kDa non-reduced band may represent association of the MUC13 47-kDa protein with another protein. However, it is more likely, given that it is almost twice the mass of the 47-kDa protein, that the 93-kDa band represents a MUC13 homodimer. This association is clearly dependent on either intra- or intermolecular disulfide linkages. Identical reactivity of all three antisera demonstrates that the A, B, and C epitopes are present in the same MUC13 polypeptide. Cell surface-biotinylated MUC13 in LIM2463 and LS513 cells was immunoprecipitated by A, B, or C peptide-reactive antisera and showed identical reduced and non-reduced banding patterns to those shown in direct Western blotting, confirming MUC13 as a cell surface protein (see Fig. 8).

Because it was suspected that MUC13 may be cleaved into two subunits and because these subunits may disassociate from each other in harsh detergents, as has been demonstrated for MUC1 (22), we also immunoprecipitated MUC13 in several lysis buffers utilizing milder detergents. Using the milder detergents under non-reducing conditions, in addition to the 47- and 93-kDa bands, a prominent high molecular mass band at ~520 kDa was observed (see Fig. 9). This band was not present following reduction; however, an intense band at 72 kDa and a weak band at 120 kDa (Fig. 9*B*, *arrow*) were present in addition to the 58-kDa band seen previously in direct Western blotting.

When immunoprecipitates from non-biotinylated BRIJ97 lysates were probed with the GalNAc-reactive *Helix pomatia* (HPA) and β-Gal(1,3)GalNAc-reactive *Arachis hypogaea* (PNA) lectins, a single band at 120 kDa was seen with both lectins in both LIM2463 and LS513 cells (see Fig. 10*A*). Treatment of immunoprecipitated reduced biotinylated MUC13 with PNGase F resulted in no change in the 70-kDa band. However, the 55-kDa band disappeared and was replaced by bands at 49 and 38 kDa, demonstrating that *N*-glycans contributed signif-

Downloaded from http://www.jbc.org/ at UQ Library on September 1, 2016

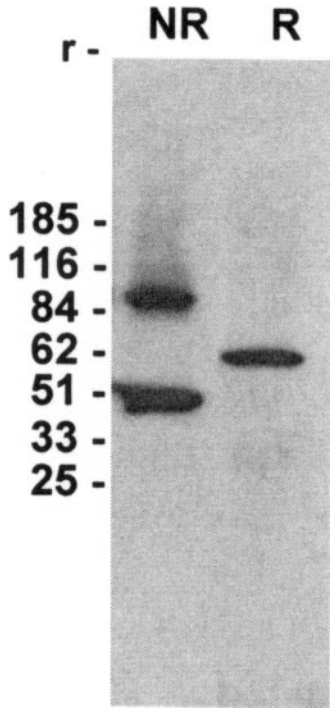

FIG. 7. **Western blot showing bands immunoreactive with MUC13 peptide C-reactive antisera.** LIM2463 colorectal cancer cells were lysed with RIPA, and subjected to SDS-PAGE in a 4–20% gradient gel both non-reduced (*NR*) and following reduction (*R*) (see "Experimental Procedures"). M_r markers are shown at the *left*. The top of the resolving gel is denoted *r*.

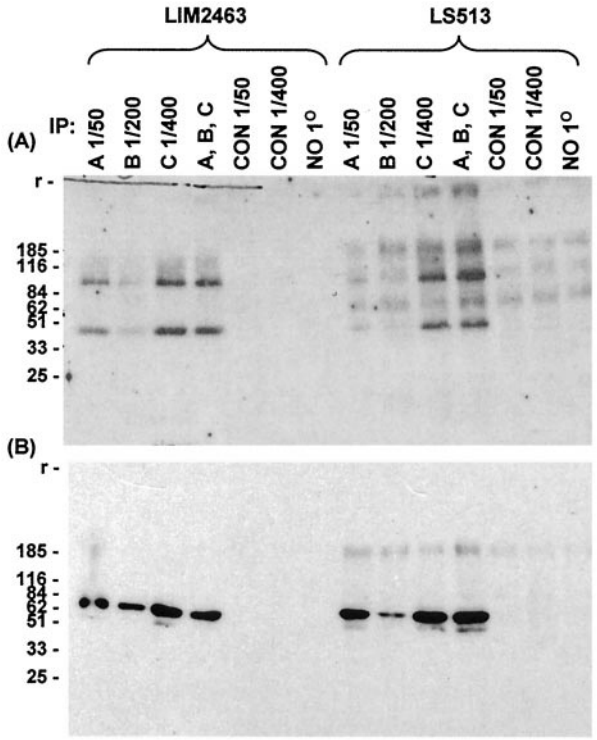

FIG. 8. **Immunoprecipitation of cell surface biotinylated MUC13 from LIM2463 and LS513 colorectal cancer cells with antisera reactive with the A, B, or C MUC13 peptides, a combination of all three antisera, irrelevant antisera (*CON*), or no primary antibody (*NO 1°*).** Immunoprecipitates were prepared either non-reduced (*A*) or reduced (*B*) and subjected to SDS-PAGE in 4–20% gradient gels (see "Experimental Procedures"). M_r markers are shown at the *left*. The top of the resolving gel is denoted *r*.

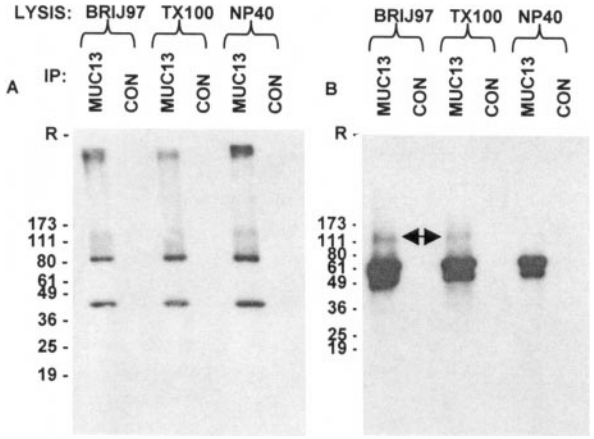

FIG. 9. **Immunoprecipitation of cell surface biotinylated MUC13 from LIM2463 colorectal cancer cells lysed with mild detergents.** Immunoprecipitation was performed with antisera reactive with the MUC13 C peptide (*MUC13*) or irrelevant antisera (*CON*). Immunoprecipitates were prepared either non-reduced (*A*) or reduced (*B*) and subjected to SDS-PAGE in 4–20% gradient gels (see "Experimental Procedures"). M_r markers are shown at the *left*. The top of the resolving gel is denoted *r*.

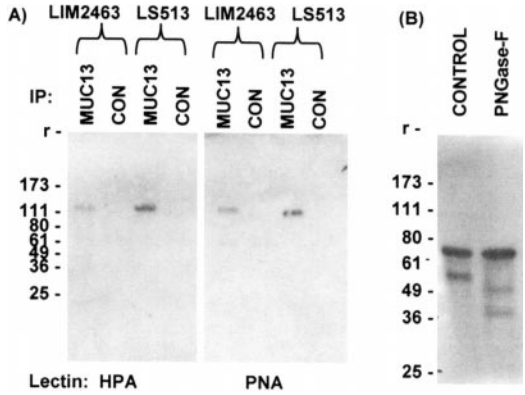

FIG. 10. **Glycosylation of MUC13.** *A*, LIM2463 and LS513 colorectal cancer cells were lysed with BRIJ96, immunoprecipitated with MUC13 C peptide (*MUC13*) or irrelevant antisera (*CON*). Immunoprecipitates were reduced and subjected to SDS-PAGE in 3–15% gradient gels, transferred to membranes, and probed with the HPA and PNA lectins (see "Experimental Procedures"). *B*, LIM2463 colorectal cancer cells were biotinylated, lysed with Nonidet P-40 buffer, and immunoprecipitated with MUC13 C peptide-reactive antisera. Immunoprecipitates were denatured and treated with PNGase F or in identical buffer lacking enzyme (*CONTROL*) for 24 h (see "Experimental Procedures"). Samples were reduced and subjected to SDS-PAGE in a 10% resolving gel. M_r markers are shown at the *left*. The top of the resolving gel is denoted *r*.

icantly to the apparent M_r of the MUC13 subunit containing the A, B, and C epitopes in LIM2463 cells (see Fig. 10*B*).

DISCUSSION

The human *MUC13* gene encodes an epithelial and hemopoietic transmembrane mucin. MUC13 is the orthologue of a murine glycoprotein initially called 114/A10 or mouse cell surface antigen and herein referred to as mMuc13. mMuc13 was initially identified by expression cloning using an antibody produced against a murine bone marrow-derived multipotential cell line (24). mMuc13 had previously been found to be highly expressed in primary myeloid progenitor cells, the myelomonocytic leukemia cell line WEHI-3, and various interleukin 3-dependent cell lines (22). Spleen, liver, muscle, and brain tissues were assessed as not expressing *mMuc13*; however, epithelial tissues were not examined.

The *MUC13* gene is most highly expressed in epithelial tissues, particularly those of the gastrointestinal and respiratory

Downloaded from http://www.jbc.org/ at UQ Library on September 1, 2016

tracts. In the large intestine, *MUC13* is expressed on the apical cell surface of both goblet and columnar cells deep in the crypts, confirming its role as a transmembrane mucin. Presence on the apical membrane of crypt epithelium mirrors the expression of two other cell surface mucins, MUC1 and MUC4, which also predominate in the crypts and are expressed by both major cell types (25). In contrast, *MUC3* (26), and *MUC12*[2] appear to be most highly expressed by cells of the surface epithelium. MUC13 also appears to be secreted, as we have demonstrated its presence within goblet cell thecae and in secreted material in the large intestine and stomach, and within supranuclear vacuoles, but not on the apical membrane surface, of absorptive cells in the small intestine. This is consistent with the previous demonstration, at both light and electron microscope levels, of MUC1 associated with goblet cell secretory granules in the intestine (27).

It is now apparent that there is considerable multiplicity of expression of these cell surface glycoproteins, especially in the gastrointestinal tract and, in particular, the large intestine. In fact, MUC1 and MUC13 appear to be co-expressed, both intracellularly and on the same cell surfaces, in many places of the gastrointestinal tract. However, there is evidence of differential regulation of cell surface mucins throughout the gastrointestinal tract. For example, although MUC1 and MUC13 appear to be expressed similarly in the colon (27) and in the stomach (28), staining of serial sections with a MUC1-reactive antibody revealed differing patterns of expression of MUC13 and MUC1 in gastric tissue (data not shown). As an additional example, in the surface epithelium *MUC3* is expressed most highly in the small intestine and decreases distally with only low expression in the distal large intestine (26). *MUC12* expression increases progressively from small intestine to rectum[2]; similarly, *MUC13* mRNA expression increases from small intestine to rectum.

In addition to multiplicity of expression, it is now apparent that there is considerable dynamism of cell-surface mucin expression in the gastrointestinal tract. For example, as cells in the large intestine migrate from deep in the crypts to the surface, they must undergo a major switch in their dominant cell surface mucin proteins from MUC1, MUC4, and MUC13 to MUC3 and MUC12. *MUC13* is also very highly expressed in the trachea, where it is also now apparent that there is considerable multiplicity of expression with *MUC1* and *MUC4* also being expressed. The expression of multiple cell-surface mucins with similar functional roles may explain why *Muc1* null mice show no obvious signs of epithelial pathology (29), although these mice are also yet to be tested comprehensively with epithelial challenges. Interestingly, when *Muc1* null mice were removed from a pathogen-free environment, they developed ocular and reproductive tract bacterial infections, implicating cell surface mucins in protection from epithelial bacterial infections (30, 31).

In addition to expression in epithelial cells, *in situ* hybridization in murine tissues demonstrated substantial expression of *MUC13* in lymphoid tissue associated with the large intestine. Lymphoid tissue expression could have contributed to *MUC13* RNA detection in human RNA samples derived from normal and malignant intestinal tissues. However, at the protein level, there was only limited histochemical reactivity of MUC13 in human gastrointestinal-associated lymphoid tissue, demonstrating clearly that protein expression levels of MUC13 are highest in epithelial cells.

Our initial biochemical characterization in two colorectal cancer cell lines demonstrates that MUC13 is predominantly a cell surface molecule, that it undergoes *N*- and *O*-glycosylation, and that the mucin domain is cleaved from the transmembrane subunit, as occurs for both MUC1 and MUC4 (13, 14). The C-terminal subunit containing the cytoplasmic tail is hereto referred to as the β-subunit using the terminology introduced for MUC4 (8). The β-subunit appears to undergo homodimerization that is dependent on intrachain disulfide bonds. The predicted mass of the entire MUC13 polypeptide minus the signal peptide is ~53 kDa. Western blotting shows that the 55–58-kDa β-subunit contains the A, B, and C peptide epitopes. The predicted mass of a minimum C-terminal polypeptide containing these three epitopes is 31 kDa and contains three potential sites for *N*-glycosylation. The immunoprecipitated and reduced MUC13 β-subunit was ~55 kDa and was reduced to 49- and 38-kDa bands following PNGase F treatment to remove *N*-linked sugars. Taken together, this suggests the 55-kDa band represents the *N*-glycosylated β-subunit of MUC13 containing the cytoplasmic domain, transmembrane domain, the C-terminal EGF-like domains, and at least part of the SEA module (see Fig. 2 for a diagrammatic representation).

The apparent cleavage of MUC13 needs to be confirmed by biosynthetic studies. The 120-kDa immunoprecipitated band reactive with HPA and PNA is most likely to represent the N-terminal mucin domain containing α-subunit of MUC13. This band was only just visible in immunoprecipitates of biotinylated cell surface proteins, perhaps due to a tendency for the α-subunit to disassociate from the β-subunit in detergents, as has been shown for MUC1 (22). The α-subunit may also show poor binding to nitrocellulose due to extensive *O*-glycosylation. Furthermore, the α-subunit would not be highly biotinylated inasmuch as, compared with 14 lysine residues in the extracellular domain C-terminal of the A peptide, there are only five lysine residues N-terminal to the A peptide, and these lie between EGF1 and the A peptide. The observation that the α-subunit is weakly biotinylated suggests the cleavage occurs between the lysine in EGF1 (amino acid 190) and the A peptide. The predicted mass of an N-terminal MUC13 α-subunit terminating at the A epitope is 22 kDa, the majority of which comprises the mucin domain. If the 120-kDa band is this N-terminal α-subunit, then it must be over 80% carbohydrate by weight, which is not unexpected for a mucin glycoprotein. Other post-translational modifications could contribute to the mass of this subunit; for example, a YYY motif near EGF1 (amino acids 203–205) may represent a site of tyrosine sulfation.

The 520-kDa immunoprecipitated band observed under non-reducing conditions may represent oligomers of intact MUC13 molecules. These complexes also appear to contain another cell surface protein as when reduced a 70–72-kDa biotinylated band was observed. The 70-kDa protein was PNGase F-resistant (not *N*-glycosylated) and did not react with the HPA or PNA lectins (not *O*-glycosylated) or with the A, B, or C MUC13 antisera (not the MUC13 β-subunit). The 70-kDa protein may be an important constituent of the 520-kDa protein complex, and requires further investigation. Additional antibodies reactive with α-subunit epitopes need to be generated to confirm these conclusions. However, it is clear that the MUC13 β-subunit undergoes homo-oligomerization into dimers and that oligomerization is dependent on intrastrand disulfide bonds.

The role of cell surface mucins in epithelial tissues is not clearly understood. The complexity and level of expression is greatest in epithelial surfaces exposed to antigenically and chemically complex luminal environments. Based on knowledge of their structure and expression patterns, it is likely that these molecules play an important role in barrier function. In

[2] S. J. Williams and M. A. McGuckin, unpublished observations.

Downloaded from http://www.jbc.org/ at UQ Library on September 1, 2016

addition to presenting large extended heavily glycosylated mucin domains on the cell surface, highly conserved structural features suggest that an ability to shed the extracellular mucin domain and to report via cytoplasmic domain signaling are important facets of these cell-surface molecules. The signaling pathways triggered by cell surface mucins such as MUC13 remain to be elucidated, although it appears that MUC1 can activate the *ras* pathway via Grb2 (32) and interact with β-catenin, an important modulator of cell adhesion and growth (33, 34). Interactions between secreted mucins and bacteria have been well characterized (35), and it is possible that interaction with bacteria is an important part of cell surface mucin function. In fact, a recent study utilizing CHO cells stably expressing Muc1 mucin (36) has demonstrated phosphorylation of the Muc1cytoplasmic domain following adhesion of *Pseudomonas aeruginosa* to the extracellular domain of this mucin.[3]

It has been proposed that cell surface EGF-like domain containing mucins can modulate epithelial cell growth via modulation of the c-erbB family of growth factor receptors. Rat Muc4 has been shown to interact with the c-erbB-2 receptor and promote receptor kinase activity, and cell proliferation, both in the presence and absence of the c-erbB-2 ligand, neuregulin-1 (15). *MUC4*, like *MUC13*, is expressed deep in the crypts of the colon where cell proliferation occurs.

MUC13 mRNA was expressed by most colorectal cancers examined, although generally at equivalent or lower levels of expression than adjacent normal tissue. Immunohistochemistry demonstrated heterogeneous expression of MUC13 within individual cancers. *MUC13* may influence tumor growth and metastasis, as has been shown for both *MUC1* (6) and *MUC4* (37). Expression of MUC13 on the surface of colorectal cancer cells may influence growth characteristics via interactions with c-erbB growth factor receptors, modulate adhesion and interfere with immune recognition. It has been suggested that mMuc13 expressed on platelets is involved in interaction of platelets with endothelial cells (38), and, by the same mechanism, MUC13 on cancer cells may interact with endothelial ligands during metastasis. MUC1 is shed/secreted into the blood by cancers, particularly those of the breast, lung, and ovary, and is the antigen measured in several different clinically utilized serum diagnostic tests such as Truquant BR, CA15.3, MCA, CA549, and CASA (39). MUC13, like MUC1, may be released from the surface of cancer cells and may therefore be a useful serum diagnostic target in patients with gastrointestinal cancers.

MUC13 may also play an important modulatory role in epithelial responses to damage and infection, and in non-malignant epithelial diseases where barrier function and/or mucin secretion are important, such as inflammatory bowel diseases, cystic fibrosis (CF), and chronic respiratory diseases. Interestingly, the MUC1 mucin has been demonstrated in animal models to have a critical influence on gastrointestinal mucus obstruction in CF (40). Previously, it had been thought that the gel-forming mucins were the major contributors to mucus accumulation in CF. *MUC13* has a similar gastrointestinal expression pattern as MUC1 and is also highly expressed in tracheal tissue and therefore also warrants investigation as a potential modulator of mucus accumulation in CF.

It is now apparent that some epithelial transmembrane mucins, like MUC13, are expressed in hemopoietic cells. MUC1 (CD227) is expressed by activated T cells (41), activated dendritic cells (42), and bone marrow mononuclear precursor cells (43). *Muc1* null mice have major disturbances to T cell development and function.[4] Previous studies demonstrated rodent Muc13 expression in bone marrow precursor cells, and we have shown expression in spleen, consistent with expression in hemopoietic precursor cells. It has been shown that spleen and bone marrow-derived mMuc13-positive cells selectively proliferate *in vitro* in response to pokeweed mitogen-stimulated spleen cell-conditioned medium or recombinant IL-3 (24). Mouse Muc13 was also identified as being associated with megakaryocyte maturation and platelet formation (37). Upregulation of mMuc13 in a murine megakaryoblastic cell line, L8057, when differentiating into a megakaryocytic lineage, suggested mMuc13 could be involved with cell differentiation and maturation. Overexpression of mMuc13 was found to inhibit cell adhesion to fibronectin, suggesting it could act as a negative regulator of cell adhesion in the megakaryocytic lineage. We also observed strong expression of *mMuc13* in intestine-associated lymphoid tissue, and weaker expression in lymphoid cells in both the red pulp and follicles of the spleen. *MUC13* is clearly expressed by mature lymphoid cells, and this requires further characterization.

In summary, we have determined the complete cDNA sequence of a novel human mucin gene, *MUC13*, and shown that this gene encodes a cell surface and secreted mucin that is most highly expressed in epithelial tissues but is also found in some hemopoietic cells. Our future research will concentrate on the role of the MUC13 mucin protein in non-malignant and malignant epithelial diseases, particularly those of the gastrointestinal tract.

Acknowledgments—We extend our thanks to Professor Jeremy Jass (Department of Pathology, University of Queensland, Herston, Australia) for helping assess the *in situ* hybridization and immunohistochemistry, and to Dr. David Thornton (Wellcome Trust Center for Cell Matrix Research, University of Manchester, Manchester, United Kingdom) for comments on the manuscript.

REFERENCES

1. Herrmann, A., Davies J. R., Lindell, G., Martensson, S., Packer, N. H., Swallow, D. M., and Carlstedt, I. (1999) *J. Biol. Chem.* **274,** 15828–15836
2. Pigny, P., Guyonnet Duperat, V., Hill, A. S., Pratt, W. S., Galiegue-Zouitina, S., D'Hooge, M. C., Laine, A., Van Seuningen, I., Degand, P., Gum, J. R., Kim, Y. S., Swallow, D. M., Aubert, J. P., and Porchet, N. (1996) *Genomics* **38,** 340–352
3. Bobek, L. A., Tsai, H., Biesbrock, A. R., and Levine, M. J. (1993) *J. Biol. Chem.* **268,** 20563–20569
4. Arias, E. B., Verhage, H. G., and Jaffe, R. C. (1994) *Biol. Reprod.* **51,** 685–694
5. Shankar, V., Gilmore, M. S., Elkins, R. C., and Sachdev, G. P. (1994) *Biochem. J.* **300,** 295–298
6. Patton, S., Gendler, S. J., and Spicer, A. P. (1995) *Biochim. Biophys. Acta* **1241,** 407–423
7. Williams, S. J., Munster, D. J., Quin, R. J., Gotley, D. C., and McGuckin, M. A. (1999) *Biochem. Biophys. Res. Commun.* **261,** 83–89
8. Moniaux, N., Nollet, S., Porchet, N., Degand, P., Laine, A., and Aubert, J. P. (1999) *Biochem. J.* **338,** 325–333
9. Williams, S. J., McGuckin, M. A., Gotley, D. C., Eyre, H. J., Sutherland, G. R., and Antalis, T. M. (1999) *Cancer Res.* **59,** 4083–4089
10. 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) *J. Biol. Chem.* **272,** 26678–26686
11. Audie, J. P., Janin, A., Porchet, N., Copin, M. C., Gosselin, B., and Aubert, J. P. (1993) *J. Histochem. Cytochem.* **41,** 1479–1485
12. Bork, P., and Patthy, L. (1995) *Protein Sci.* **4,** 1421–1425
13. Hilkens, J., and Buijs, F. (1988) *J. Biol. Chem.* **263,** 4215–4222
14. Sheng, Z. Hull, S. R., and Carraway, K. L. (1990) *J. Biol. Chem.* **265,** 8505–8510
15. Carraway, K. L., Rossi, E. A., Komatsu, M., Price-Schiavi, S. A., Huang, D., Guy, P. M., Carvajal, M. E., Fregien, N., and Carraway, C. A. (1999) *J. Biol. Chem.* **274,** 5263–5266
16. Dougherty, G. J., Kay, R. J., and Humphries, R. K. (1989) *J. Biol. Chem.* **264,** 6509–6514
17. Sambrook J., Fritsch E. F., and Maniatis T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
18. Rex, M., and Scotting, P. J. (1994) *Biochemica* **3,** 24–26
19. Callen, D. F., Baker, E., Eyre, H. J., Chernos, J. E., Bell, J. A., and Sutherland, G. R. (1990) *Ann. Genet.* **33,** 219–221
20. Zegers, N., Gerritze, K., Deen, C., Boersma, W., and Claasen. E. (1990) *J. Immunol. Meth.* **130,** 195–200

[3] K. C. Kim, personal communication.

[4] S. J. Gendler, personal communication.

Downloaded from http://www.jbc.org/ at UQ Library on September 1, 2016

21. Walsh, M. D., Luckie, S. M., Cummings, M. C., Antalis, T. M., and McGuckin, M. A. (1999) *Breast Cancer Res. Treat.* **58,** 255–266
22. Quin, R. J., and McGuckin, M. A. (2000) *Int. J. Cancer* **87,** 499–506
23. Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997) *Protein Eng.* **10,** 1–6
24. Dougherty, G. J., Dougherty S. T., Kay R. J., Landsdorp, P., and Humphries, R. K. (1989) *Exp. Hematol.* **17,** 877–882
25. Biemer-Huttmann, A., Walsh, M. D., McGuckin, M. A., Ajioka, Y., Watanabe. H., Leggett, B. A., and Jass, J. R. (1999) *J. Histochem. Cytochem.* **47,** 1039–1047
26. Weiss, A. A., Babyatsky, M. W., Ogata, S., Chen, A., and Itzkowitz, S. H. (1996) *J. Histochem. Cytochem.* **44,** 1161–1166
27. Winterford, C. M., Walsh, M. D., Leggett, B. A., and Jass, J. R. (1999) *J. Histochem. Cytochem.* **47,** 1063–1074
28. Carrato, C., Balague, C., Debolos, C., Gonzalez, E., Gambus, G., Planas, J., Perini, J. M., Andreu, D., and Real, F. X. (1994) *Gastroenterology* **107,** 160–172
29. Spicer, A. P., Rowse, G. J., Lidner, T. K., and Gendler, S. J. (1995) *J. Biol. Chem.* **270,** 30093–30101
30. Kardon, R., Price, R. E., Julian, J., Lagow, E., Tseng, S. C. G., Gendler, S. J., and Carson, D. D. (1999) *Invest. Ophthalmol. Visual Sci.* **40,** 1328–1335
31. DeSouza, M. M., Surveyor, G. A., Price, R. E., Julian, J., Kardon, R., Zhou, X., Gendler, S., Hilkens, J., and Carson, D. D. (1999) *J. Reprod. Immunol.* **45,** 127–158
32. Pandey, P., Kharbanda, S., and Kufe, D. (1995) *Cancer Res.* **55,** 4000–4003
33. Yamamoto, M., Bharti, A., Li, Y., and Kufe, D. (1997) *J. Biol. Chem.* **272,** 12492–12494
34. Li, Y. Q., Bharti, A., Chen, D. S., Gong, J. L., and Kufe, D. (1998) *Mol. Cell. Biol.* **18,** 7216–7224
35. Slomiany, B. L., Murty, V. L. N., Piotrowski, J., and Slomiany, A. (1996) *Gen. Pharmacol.* **27,** 761–771
36. Lillehoj, E. P., Hyun, S. W., Kim, B. T., Zhang, X. G., Lee, D. I., Rowland, S., and Kim, K. C. (2001) *Am. J. Physiol.* **280,** L181–L187
37. Komatsu, M., Tatum, L., Altman, N. H., Carraway, C. A., and Carraway, K. L. (2000) *Int. J. Cancer* **87,** 480–486
38. Duan, Y., Naruse, T., Nakamura, M., Yamaguchi, Y., Kawashima, T., Morikawa, Y., Kitamura, T., and Suda, T. (1998) *Biochem. Biophys. Res. Commun.* **253,** 401–406
39. McGuckin, M. A., and Devine, P. L. (1995) *Tumor Diag. Ther.* **16,** 1–6
40. Parmley, R. R., and Gendler, S. J. (1998) *J. Clin. Invest.* **102,** 1798–1806
41. Agrawal, B., Krantz, M. J., Parker, J., and Longenecker, B. M. (1998) *Cancer Res.* **58,** 4079–4081
42. McGuckin, M. A., MacDonald, K. P. A., Tran, M., Wykes, M., and Hart, D. N. J. (2001) In *Leukocyte Typing VII* (Mason, D., ed) Oxford University Press, Oxford, UK
43. Brugger, W., Buhring, H. J., Grunebach, F., Vogel, W., Kaul, S., Muller, R., Brummendorf, T. H., Ziegler, B. L., Rappold, I., Brossart, P., Scheding, S., and Kanz, L. (1999) *J. Clin. Oncol.* **17,** 1535–1544

Downloaded from http://www.jbc.org/ at UQ Library on September 1, 2016

MUC13, a Novel Human Cell Surface Mucin Expressed by Epithelial and Hemopoietic Cells

Stephanie J. Williams, Daniel H. Wreschner, Mai Tran, Helen J. Eyre, Grant R. Sutherland and Michael A. McGuckin

J. Biol. Chem. 2001, 276:18327-18336.
doi: 10.1074/jbc.M008850200 originally published online February 26, 2001

Access the most updated version of this article at doi: 10.1074/jbc.M008850200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at
http://www.jbc.org/content/276/21/18327.full.html#ref-list-1

Downloaded from http://www.jbc.org/ at UQ Library on September 1, 2016