Mouse gastric mucin: cloning and chromosomal localization

Laurie L. SHEKELS,* Carolyn LYFTOGT,* Marcia KIELISZEWSKI,† Jane D. FILIE,‡ Christine A. KOZAK‡ and Samuel B. HO*§

*Department of Medicine, University of Minnesota and VA Medical Center, Minneapolis, MN 55417, U.S.A., †The Complex Carbohydrate Research Center, University of Georgia, Athens, GA 30602, U.S.A., and ‡National Institutes of Health, Bethesda, MD 20892, U.S.A.

Mucins protect gastric epithelium by maintaining a favourable pH gradient and preventing autodigestion. The purpose of this study was to clone a mouse gastric mucin which would provide a foundation for analysis of mucin gene regulation. Mucin was purified from the glandular portion of gastric specimens and deglycosylated by HF solvolysis. Antibodies against native and deglycosylated mouse gastric mucin (MGM) were raised in chickens. Screening of a mouse stomach cDNA library with the anti-(deglycosylated MGM) antibody yielded partial clones containing a 48 bp tandem repeat and 768 bp of non-repetitive sequence. The 16-amino-acid tandem repeat has a consensus sequence of QTSSPNTGKTSTISTT with 25% serine and 38% threonine. The MGM tandem repeat sequence bears no similarity to previously identified mucins. The MGM non-repetitive region shares sequence similarity with human MUC5AC and, to a lesser extent, human MUC2 and rat intestinal mucin. Northern blot analysis reveals a polydisperse message beginning at 13.5 kb in mouse stomach with no expression in oesophagus, trachea, small intestine, large intestine, caecum, lung or kidney. Immunoreactivity of antibodies against deglycosylated MGM and against a synthetic MGM tandem repeat peptide was restricted to superficial mucous cells, antral glands and Brunner's glands in the pyloric–duodenal region. DNA analysis shows that MGM recognizes mouse and rat DNA but not hamster, rabbit or human DNA. The MGM gene maps to a site on mouse chromosome 7 homologous to the location of a human secretory mucin gene cluster on human chromosome 11p15. Due to sequence similarity and predominant expression in the stomach, the MGM gene may be considered a MUC5AC homologue and named *Muc5ac*.

INTRODUCTION

The epithelial cells of the gastrointestinal tract reside in a hostile environment where they may suffer damage from foreign agents, pathogenic bacteria, physical abrasion or dehydration. Gastrointestinal epithelium is protected by a layer of mucus which acts as a physical barrier preventing contact of potentially damaging exogenous agents with the underlying epithelial cells. Within the stomach, the mucous layer must provide additional protection from the deleterious effects of hydrochloric acid and pepsin. The mucous layer predominantly consists of mucin (large heavily glycosylated proteins). Biochemical and microscopic analysis has revealed that mucins exist as long polymers linked together by disulphide bonds. The extended length and high carbohydrate content of mucins confer thickness and viscosity to the mucous gel, resulting in an effective barrier preventing the back-diffusion of luminal acid [1,2].

Characterization of the mucin peptide core has been limited due to its large size and abundant glycosylation. However, the cloning of numerous mucins has provided new insight into their structure. To date, seven unique human mucins have been identified. Each mucin contains an extended domain of tandem repeats which is rich in serine and/or threonine; however, each mucin can be distinguished by the length and amino acid sequence of its tandem repeat unit [3]. Thus the human mucins are designated MUC1 to MUC7 in order of the discovery of their unique tandem repeats. To either side of the tandem repeat domains exist non-repetitive sequences which are characterized by either cysteine-rich regions or a membrane-spanning domain.

Two classes of mucins have been defined; the membrane-bound mucins and the secreted mucins. Currently the MUC1 gene product is the only identified membrane-bound mucin and is ubiquitously expressed on the apical membrane of a wide variety of epithelial tissues. It is also the only mucin for which a full-length clone has been obtained for the human and mouse homologues [4,5]. The human MUC1 tandem repeat is 20 amino acids in length while the mouse MUC1 tandem repeat varies between 20 and 21 amino acids. Despite sharing only 34% similarity in the tandem repeats, the mouse and human MUC1 homologues exhibit 87% and 74% sequence similarity in their non-repetitive and promoter regions respectively [5].

In contrast to the ubiquitous expression of the membrane-bound mucin, the secreted mucins exhibit tissue-specific expression. Human MUC2 mucin is expressed in small intestine and colon and human MUC3 is primarily expressed in small intestine, colon and gallbladder [6]. MUC4 gene expression is observed in bronchial tissue and colon [7] and MUC5 gene expression is found in gastric and bronchial tissue [8]. The MUC6 gene is primarily expressed in gastric tissue [9], and MUC7 in the salivary glands [10]. Non-human secretory mucin genes for which partial clones have been isolated include bovine and porcine submaxillary mucin [11,12], canine tracheobronchial mucin [13], frog integumentary mucin, rainbow trout egg mucin [14] and rat intestinal mucin [15]. While the rat intestinal mucin tandem repeat is only six amino acids in length and the human intestinal MUC2 mucin tandem repeat is 23 amino acids, the rat intestinal mucin represents the rat homologue of MUC2 based on the complete lack of serine in the tandem repeats and

Abbreviations used: MGM, mouse gastric mucin; KLH, keyhole limpet hemocyanin; TBS, Tris-buffered saline containing 25 mM Tris, pH 7.4, 140 mM NaCl, 3 mM KCl; GAP, glyceraldehyde phosphate dehydrogenase; CRRI and CRRII, cysteine-rich regions I and II respectively; PEG, poly(ethylene glycol).

§To whom correspondence should be addressed at Department of Gastroenterology (111-D), VA Medical Center, 1 Veterans Drive, Minneapolis, MN 55417, U.S.A.

The nucleotide sequence reported in this paper has been submitted to the EMBL/GenBank/DDB J Nucleotide Sequence Databases under the accession no. L42292.

73% and 80% sequence similarity in a portion of the amino and carboxyl non-repetitive domains respectively [16,17]. Sequence identity in the C-termini of MUC2, MUC5, rat intestinal mucin-like protein, and bovine and porcine submaxillary mucins ranges from 18% to 29% [18]. However, this limited similarity occurs with a striking conservation of cysteine residues (from 64% to 90%), suggesting a functional importance for these residues.

It is becoming increasingly apparent that aberrant expression of the secretory mucin genes occurs in certain disease states and in conditions associated with an elevated risk of carcinoma. Alterations in the quantity and type of mucin are found in *Helicobacter pylori*-induced acute and chronic gastritis, intestinal metaplasia and gastric adenocarcinoma [19,20]. In normal gastric tissue, MUC5 and MUC6 show high levels of expression with no detectable expression of the intestinal mucins MUC2 and MUC3 [6,8]. However, the expression of gastric-specific mucins is lost in intestinal metaplasia and gastric carcinoma where there is a conversion to colonic and intestinal mucin gene expression [6].

Because mucin glycoproteins play an important role in the protective nature of the mucous barrier, determination of the factors responsible for the modulation of mucin expression is of great diagnostic and therapeutic interest. Currently, little is known regarding the regulation of mucin gene expression and no rodent homologues of gastric mucin genes have been identified. To establish a foundation for further analysis of mucin regulation through the use of animal models of gastrointestinal disease, we report here the cloning of a mouse gastric MUC5AC homologue.

EXPERIMENTAL

Materials

Peroxidase-conjugated rabbit anti-(chicken IgG) antibody, 4-chloro-1-naphthol, diaminobenzidine and isopropylthio-β-D-galactoside were obtained from Sigma (St. Louis, MO, U.S.A.). Biotinylated rabbit anti-(chicken IgG) antibody was purchased from Zymed Laboratories (South San Francisco, CA, U.S.A.). $\alpha[^{32}P]$dCTP and $\alpha[^{35}S]$dATP were purchased from Amersham.

Mucin purification

Mucin was purified by gel filtration and CsCl density-gradient centrifugation from the soluble fraction of mouse gastric mucosa as has been described previously for the purification of mucin from tissue-culture cells [21]. Briefly, the mucosa of freshly harvested mouse stomachs was scraped into 0.1 M NH_4HCO_3, 0.5 M NaCl, 0.1 mM PMSF on ice. The sample was then homogenized followed by centrifugation for 45 min at 45000 *g*. After removing the lipid layer, the supernatant was centrifuged as before and then dialysed overnight against 10 mM Tris, pH 8.0. The protein was size-fractionated on a 2.5 cm × 70 cm Sepharose CL-4B column equilibrated in 10 mM Tris, pH 8.0. The void volume of the column was collected and dialysed against water, lyophilized and digested for 2 h at room temperature with RNase A and DNase I (1:100 protein ratio) in PBS, 1 mM $MgSO_4$, 0.1 mM PMSF, 0.2% NaN_3. Following digestion, the sample was centrifuged and the supernatant dialysed overnight against PBS. CsCl was added to the dialysed supernatant to a final concentration of 0.54 g/ml and then the sample was centrifuged for 72 h at 160000 *g*. Fractions of 1 ml were collected and the density, protein and hexose content of each fraction were measured. Fractions of high density (> 1.35 g/ml) and with a hexose:protein ratio of 2:1 were pooled, dialysed against water and then applied to a 1 cm × 43 cm Sepharose CL-4B column equilibrated in 10 mM Tris, pH 8.0. The void volume containing the purified mucin was dialysed against water and lyophilized. The purity of the mucin was assessed by SDS/PAGE [22] and periodic acid–silver nitrate staining [23].

Amino acid analysis and mucin deglycosylation

Amino acid analysis of the purified mouse gastric mucin (MGM) was performed by the University of Minnesota Microchemical Facility. An aliquot of the purified protein was deglycosylated by treatment with anhydrous HF. Deglycosylation was performed in 2 ml Sarstedt screw-cap microtubes with anhydrous HF at a concentration of 20 μg of protein/μl of HF and containing 10% methanol. After 3 h at room temperature the reaction mixture was quenched by freezing it in liquid N_2 and by adding ice-cold water to bring the final HF concentration to 10%. The sample was dried under nitrogen, washed three times with water and again dried under nitrogen to concentrate the protein and remove the residual HF. The deglycosylated protein was resuspended in water and lyophilized.

Antibody production

White Leghorn chickens, aged 22–24 weeks, were injected with either 50 μg of native MGM or 15 μg of deglycosylated MGM emulsified in complete Freund's adjuvant. Two boosters of 25 μg of native MGM or 7.5 μg of deglycosylated MGM emulsified in incomplete Freund's adjuvant were given at 2-week intervals. Ten days after the final booster, eggs were collected for antibody purification. Polyclonal chicken IgY was purified as described by Goueli et al. [24]. Briefly, the egg yolks were separated and mixed with an equal volume of buffer A (10 mM potassium phosphate, pH 7.5, 0.1 M NaCl and 0.1% NaN_3). A vol. of 10.5% poly-(ethylene glycol) (PEG) in buffer A, equivalent to the total egg yolk volume, was added and stirred for 30 min at room temperature. The mixture was centrifuged at 12000 *g* for 20 min. Lipids were removed by addition of silicon dioxide to the supernatant to a final concentration of 5 g/100 ml. The mixture was stirred for 20 min and then allowed to stand at room temperature for 10 min before centrifugation. A solution of 42% PEG in buffer A was added to yield a final concentration of 12% PEG and stirred for 30 min at 4 °C. The precipitated proteins were collected by centrifugation and the pellet was dissolved in a minimal volume of buffer A. An equivalent volume of 4 M $(NH_4)_2SO_4$, pH 7.0, was added and the sample was stirred for 30 min at 4 °C. The partially pure antibodies were collected by centrifugation. Further purification by ion-exchange chromatography did not result in an increase in specificity; therefore antibodies were used prior to ion-exchange chromatography.

A 17-amino-acid peptide corresponding to the MGM tandem repeat (KQTSSPNTGKTSTISTT) was synthesized by the University of Minnesota Microchemical Facility using 9-fluorenylmethoxycarbonyl/benzotriazolyl-*N*-oxytris(dimethylamino)-phosphonium hexafluorophosphate/1-hydroxybenzotriazole ('FMOC/BOP/HOBt') chemistry on a Milligen Biosearch 9600 peptide synthesizer. The N-terminal lysine was included in the peptide sequence to aid in coupling the peptide to keyhole limpet haemocyanin (KLH) [25]. The KLH-conjugated peptide (250 μg) was injected into chickens followed by two 125 μg additional injections on the immunization schedule described above.

ELISA and Western blot analysis

Antigens (10 ng) were plated on 96-well ELISA plates for 2 h at room temperature. Plates were blocked with 5% BSA in Tris-

Table 1 Amino acid compositions of mouse, rat and human gastric mucins

Amino acid compositions are given as percentages. Potential O-linked glycosylation sites are indicated in **bold**.

Amino acid	Mouse	Rat*	Human*
Asx	9.6	7.4	5.4
Thr	**13.5**	**15.4**	**21.1**
Ser	**12.2**	**14.8**	**13.6**
Glx	10.4	10.5	7.4
Pro	7.4	8.9	11.6
Gly	8.4	10.2	6.2
Ala	4.3	7.3	7.4
Cys	4.0	1.2	4.1
Val	6.4	5.6	4.6
Met	1.4	0.3	0
Ile	3.2	1.7	2.5
Leu	5.4	3.6	3.7
Tyr	2.5	0.3	1.2
Phe	2.6	1.4	1.7
His	2.3	3.2	2.9
Lys	3.9	3.8	2.5
Arg	2.3	0.9	4.1

*From [41].

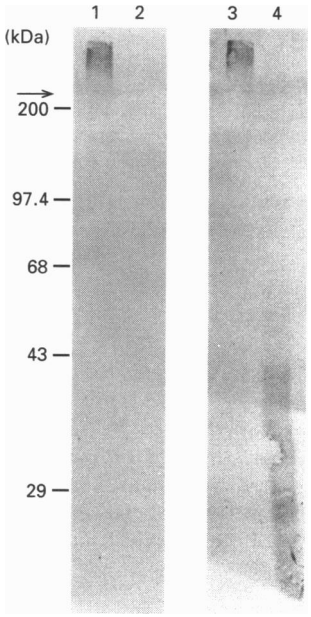

Figure 1 Western blot analysis of native and deglycosylated MGM

Native MGM (lanes 1 and 3) and deglycosylated MGM (lanes 2 and 4) were subjected to SDS/PAGE on a 4% stacking, 5% separating polyacrylamide gel. The proteins were transferred to nitrocellulose and incubated with anti-(native MGM) (lanes 1 and 2) or anti-(deglycosylated MGM) (lanes 3 and 4). Reactive proteins were visualized with 4-chloronaphthol. The arrow marks the interface between the stacking and separating gels and molecular masses are indicated in kDa.

buffered saline (TBS; 25 mM Tris, pH 7.4, 140 mM NaCl, 3 mM KCl) overnight at 4 °C. The plates were washed with 0.02% Tween-20 in TBS and incubated with the primary antibody for 3 h at room temperature. Following washing as before, peroxidase-conjugated rabbit anti-(chicken IgY) (1:2000) was added for 1.5 h. Colour development was performed with

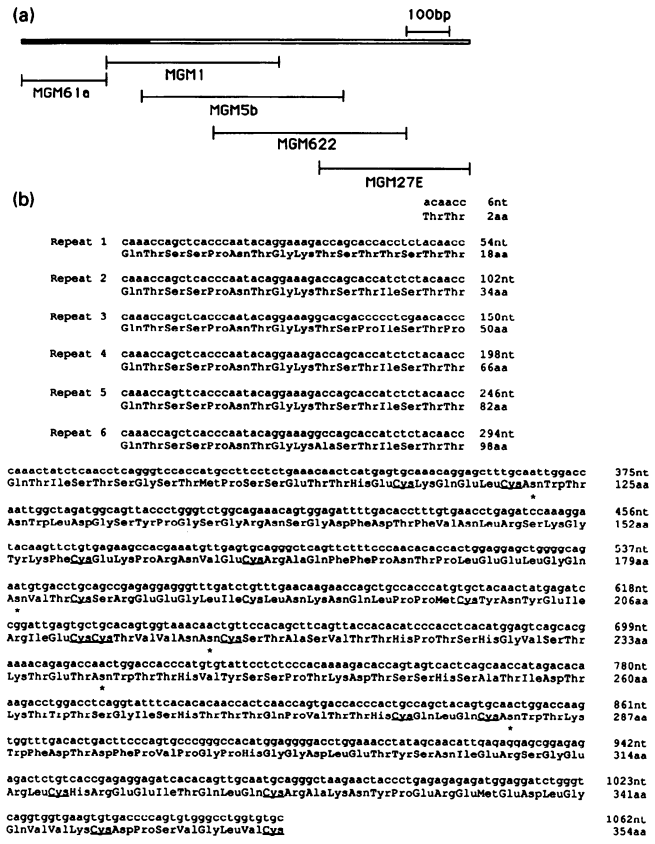

Figure 2 Sequence of mouse gastric mucin

(**a**) Diagram of the partial MGM clones. Shaded regions indicate the presence of tandem repeats. (**b**) The nucleotide and amino acid sequences of the mouse gastric mucin are shown. The six complete tandem repeats are numbered along the left-hand side. Potential sites for N-glycosylation are marked with an asterisk and cysteine residues are underlined.

3,3′,5,5′-tetramethylbenzidine and quenched with 8 M H_2SO_4. Bound antibody was quantified by measuring the absorbance at 450 nm with a TiterTek spectrophotometer. Preimmune antibodies were used as negative controls.

For Western blot analysis, proteins were separated on a 4% stacking/5% separating polyacrylamide gel and transferred to nitrocellulose as described [26]. The nitrocellulose filter was blocked with 3% BSA in TBS and incubated with the primary antibody. The membrane was then washed with 0.05% Tween-20 in TBS and incubated with peroxidase-conjugated rabbit anti-(chicken IgY) (1:2000) for 1 h. Following two additional washes, colour development was performed with 4-chloro-1-naphthol as described [27].

Immunohistochemistry

The streptavidin–peroxidase technique was used as described previously [6]. Mucin expression was determined using frozen and formalin-fixed specimens of normal mouse tissues. Antibodies against native gastric mucin and the synthetic peptide were reactive with both formalin-fixed and frozen-ethanol-fixed sections, whereas antibody against deglycosylated mucin was only reactive with frozen tissue sections which were unfixed or fixed in ethanol. Briefly, tissue sections were deparaffinized, rehydrated, incubated with fresh 3% hydrogen peroxide in methanol for 10 min, and then washed with PBS. Normal rabbit

QTSSPNTGKISTtSTT QTSSPNTGKISTISTT QTSSPNTGKvSTpSTp hTSSPNTGKTSTISTT	N-terminus of MGM6a
QTSSPNTGKTSTISTT QTSSPNTGKgSTpSTp QTSSPNTGKTSTISTT	C-terminus of MGM6a
QTSSPNTGKTSTtSTT QTSSPNTGKTSTISTT QTSSPNTGKTSpISTp QTSSPNTGKTSTISTT	MGM6al
QTSSPNTGKTSTISTT QTSSPNTGKaSTISTT	MGM1
(Q) TSSPNTGK (T) ST (I) ST (T)	Consensus sequence

Figure 3 Tandem repeat sequences in the MGM clones

The amino acid sequences of the tandem repeat units found in the MGM clones are shown. Lowercase letters indicate residues which differ from the consensus sequence. Residues which are not totally conserved appear in parentheses in the consensus sequence.

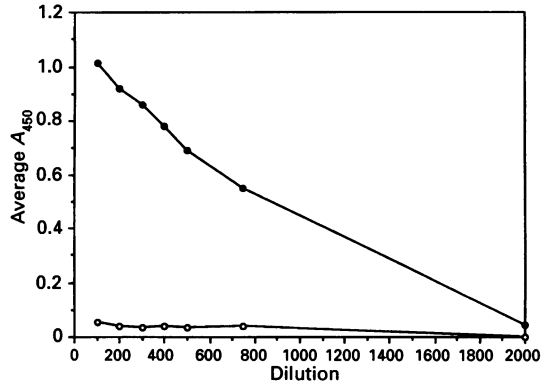

Figure 4 ELISA analysis

Wells were coated with 10 ng of the synthetic tandem repeat peptide followed by exposure to the anti-(deglycosylated MGM) antibody (●) or preimmune serum (○). The reaction was visualized with peroxidase-conjugated secondary antibody and 3,3′,5,5′-tetramethylbenzidine (TMB) and quantified by determining the absorbance at 450 nm.

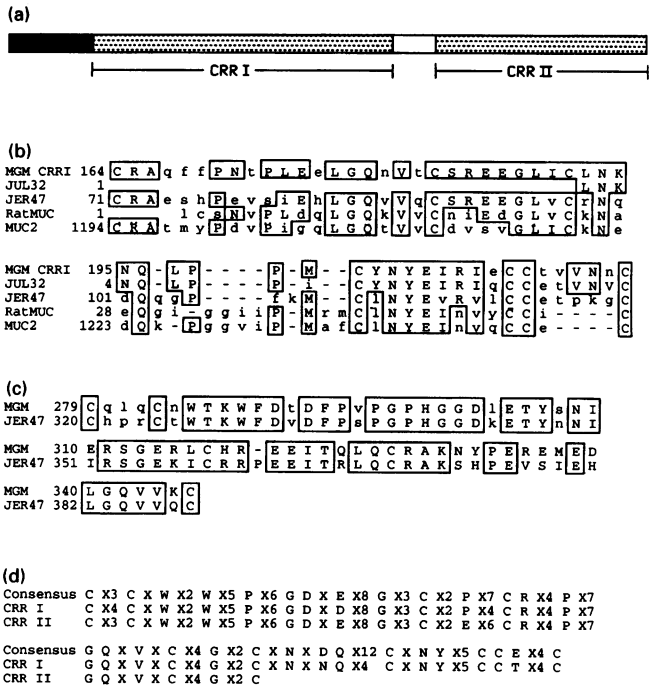

Figure 5 Sequence similarity between MGM and human MUC2, MUC5AC and rat intestinal mucin

(**a**) Schematic of the MGM structure. The black box represents the tandem repeat domain; the hatched boxes are the non-repetitive cysteine-rich regions, CRRI and CRRII, and the white box represents the non-repetitive serine/threonine-rich domain. (**b**) Sequence similarity between MGM CRRI and the MUC5AC clones JUL32 and JER47 [49], human intestinal MUC2 [17] and rat intestinal mucin [58]. The single-letter amino acid code is used with identical residues capitalized. (**c**) Sequence similarity between MGM CRRII and the MUC5AC clone JER47. (**d**) Alignment of the consensus cysteine-rich sequence with the MGM cysteine-rich regions CRRI and CRRII.

serum (5%, v/v) was applied for 20 min and removed by blotting. Next the sections were incubated with the primary antibody for 90 min at the following dilutions: anti-MGM, 1:5000 (2 μg/ml); anti-(deglycosylated MGM), 1:5000 (3 μg/ml), or anti-(synthetic MGM), 1:4000 (7 μg/ml). The sections were then washed and incubated with the biotinylated rabbit anti-(chicken IgY) antibody (1:75 dilution in PBS) for 20 min. After washing, the sections were incubated with streptavidin–peroxidase conjugate (10 μg/ml) for 30 min followed by repeated washing. Next the sections were incubated with diaminobenzidine in 0.03% hydrogen peroxide for 10 min, washed, counterstained with haematoxylin, rinsed in tap water, and mounted. Preimmune chicken IgY (2 μg/ml) was substituted for the primary antibodies as a negative control.

Isolation of an MGM cDNA

Mouse gastric RNA was isolated by the acid guanidinium thiocyanate–phenol–chloroform extraction method [28]. The RNA preparation was further enriched for poly(A) RNA by oligo-dT chromatography for use in the construction of a cDNA library in the bacteriophage λZAP II by Stratagene. Identification of MGM clones was performed by screening the expression library with anti-(deglycosylated MGM). Positive clones were visualized with peroxidase-conjugated rabbit anti-(chicken IgY) antibody using 4-chloro-1-naphthol as the substrate. Hybridizing plaques were purified by successive rounds of screening.

DNA sequencing and sequence analysis

The cDNAs were sequenced by the Sanger dideoxy-mediated chain termination method [29] using Sequenase Version 2.0 (United States Biochemical Corporation). Both strands were sequenced for each clone discussed. The University of Wisconsin Genetics Computer Group software was used to analyse DNA sequence information [30].

RNA and DNA analysis

RNA was isolated from various mouse and rat tissues by the acid guanidinium thiocyanate–phenol–chloroform extraction method [28]. Aliquots (10 μg) of each RNA were separated on 1.2% agarose gels. The gels were stained with ethidium bromide to assess RNA integrity and then the RNA was transferred to Nytran nylon membranes. Following prehybridization, the filters were hybridized in the presence of radiolabelled cDNA probes

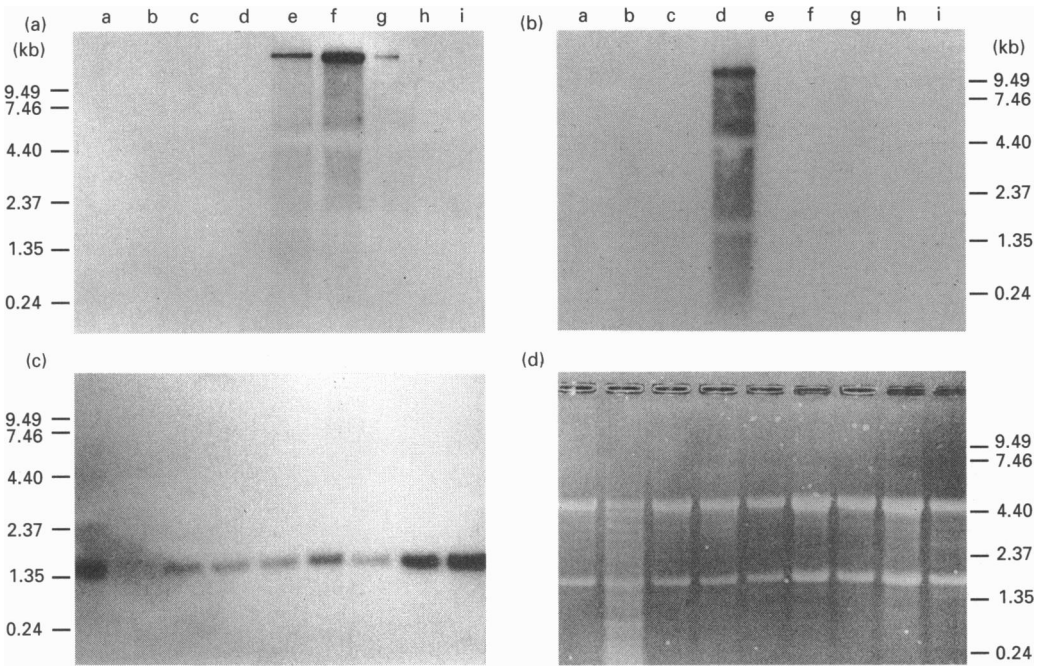

Figure 6 Northern blot analysis of mouse RNA

Mouse RNA (10 μg) was separated on an agarose gel and transferred to a nylon membrane. Following prehybridization, the blot was hybridized with (**a**) rat intestinal mucin cDNA or (**b**) MGM1 cDNA. (**c**) This shows the same blot probed with rat GAP cDNA to demonstrate RNA integrity. Markers along the side indicate size in kb. (**d**) This shows the ethidium bromide staining of the gel prior to transfer to assess equal sample loading. Lanes a, kidney RNA; lanes b, pancreas RNA; lanes c, lung RNA; lanes d, stomach RNA; lanes e, caecum RNA; lanes f, large intestine RNA; lanes g, small intestine RNA; lanes h, trachea RNA; and lanes i, oesophagus RNA.

which had been prepared by the random primer method [31]. In addition to the MGM clones, rat intestinal mucin (RMUC176) [15], rat glyceraldehyde phosphate dehydrogenase (GAP) [32] and human MUC5 and human gastric mucin MUC6 [8,9] cDNAs were also used as probes. The membranes were washed twice with 2 × SSC/0.1 % SDS (SSC: 0.15 M NaCl, 0.015 M sodium citrate) at room temperature for 30 min, once with 0.1 × SSC/0.1 % SDS for 1 h at room temperature and finally with 0.1 × SSC/0.1 % SDS at 55 °C for 30 min.

DNA was isolated from mouse, hamster, rabbit, rat and human using the Puragene DNA isolation kit for human and animal tissue (Gentra Systems, Inc). For Southern blot analysis, 10 μg of purified DNA was digested with *Bam*HI or *Pst*I overnight at 37 °C and the digested DNA was separated on a 1.2 % agarose gel. Following denaturation of the gel, the DNA was transferred to a Nytran nylon membrane. The membrane was prehybridized, hybridized and washed as described for Northern blot analysis of RNA.

Chromosomal localization

The MGM gene was mapped by analysis of the progeny of two genetic crosses: (NFS/N or C58/J × *Mus musculus musculus*) × *M.m. musculus* [33] and (NFS/N × *Mus spretus*) × *M. spretus* or C58/J [34]. DNA from parental mice and the progeny of both crosses were typed by Southern blotting for restriction enzyme polymorphisms using MGM1 as the probe. The progeny of these crosses was also typed for inheritance of over 700 markers which map to all 19 autosomes and the X chromosome including the Chr 7 markers *Zp2* (zona pellucida protein 2), *Oat* (ornithine aminotransferase), *Cyp2e1* (cytochrome *P*-450 2e1), *Hras1* (Harvey *ras* oncogene 1), *Fgf3* (fibroblast growth factor 3, formerly *Int2*) and *Mtv35* (mammary tumour virus 35). Probes and enzymes used to type *Zp2*, *Oat*, *Cyp2e1*, *Fgf3* and *Hras1* have been described previously [35–37]. *Mtv35* was typed as a 12.7 kb *Eco*RI *spretus* fragment using as probe a 1.4 kb *Pst*I fragment of the C3H MMTV [38]. Percentage recombination and standard errors between loci were calculated as described by Green [39]. Data were stored and analysed using the program LOCUS prepared by C. E. Buckler (NIAID, NIH, Bethesda, MD, U.S.A.).

RESULTS

Mucin purification and antibody production

MGM was purified by gel filtration and CsCl density-gradient centrifugation as has been previously done for human and rat mucins [15,21,40]. Because the purification is based on the high molecular mass and high density of mucin, the final preparation may contain more than one type of mucin if the mouse stomach expresses multiple mucins. The stomachs from 150 mice yielded 395 μg of mucin. Analysis of the mucin preparation demonstrates a high carbohydrate content, with a hexose to protein ratio of 2:1. Amino acid analysis reveals that the protein is rich in hydroxyl amino acids with 13.5 % threonine and 12.2 % serine (Table 1) as is characteristic of mucin glycoproteins. The threonine and serine compositions of the mouse mucin preparation are similar to that found for the rat gastric mucin [41]; however, the human gastric mucin possesses a higher threonine content than either of the rodent proteins [41].

Antibodies were prepared using both the native and deglycosylated MGM as antigens. Due to the extensive glycosylation, much of the sample weight is lost following

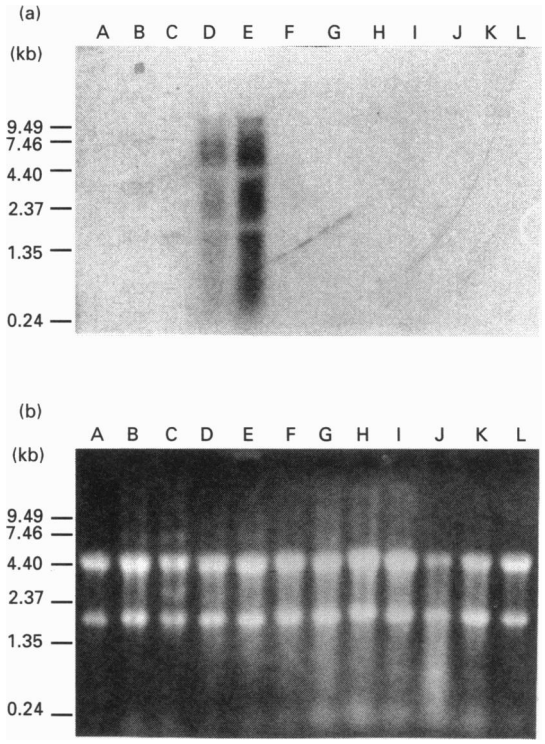

Figure 7 Northern blot analysis of rat RNA

Rat RNA (10 μg) was separated on an agarose gel and transferred to a nylon membrane. Following prehybridization, the blot was hybridized with MGM1. (**a**) Markers along the side indicate size in kb. (**b**) This shows the ethidium bromide staining of the gel prior to transfer in order to assess equal sample loading. Lanes A, small intestine RNA; lanes B, large intestine RNA; lanes C, caecum RNA; lanes D, fundus RNA; lanes E, antrum RNA; lanes F, oesophagus RNA; lanes G, kidney RNA; lanes H, lung RNA; lanes I, trachea RNA; lanes J, pancreas RNA; lanes K, liver RNA; and lanes L, spleen RNA.

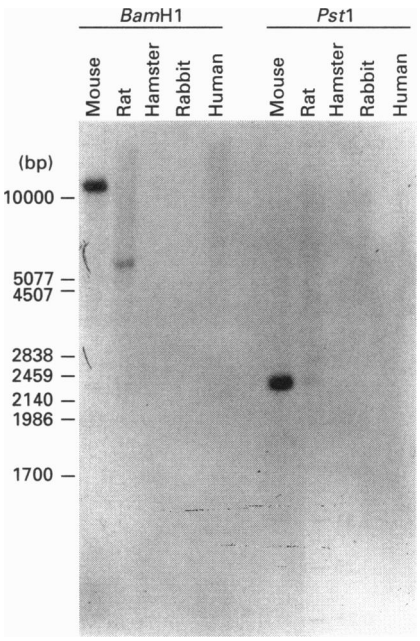

Figure 8 Southern blot analysis

Spleen DNA (10 μg) from the indicated species was digested with either *Bam*H1 or *Pst*1 and separated on an agarose gel. The DNA was transferred to a nylon membrane, prehybridized and then probed with MGM1. Markers along the side indicate size in bp.

deglycosylation, resulting in limited material for further analysis. Therefore, antibodies were raised in chickens, given the chicken's ability to produce a high titre following immunization with microgram amounts of antigen [42,43]. Characterization of the antibodies by ELISA and Western blot analysis showed that the anti-(native MGM) antibodies cross-reacted with only the native mucin, recognizing a high-molecular-mass species that does not enter the separating gel (Figure 1). The ability of anti-(native MGM) to react only with native MGM indicates that it predominantly recognizes carbohydrate epitopes which are lost following deglycosylation. The anti-(deglycosylated MGM) antibody cross-reacted with both the native and deglycosylated antigens (Figure 1). The cross-reactivity between the anti-(deglycosylated MGM) antibody and the native mucin suggests that the polyclonal antibody preparation recognizes protein epitopes which are devoid of carbohydrate moieties in the native mucin. The increase in migration of the deglycosylated protein results from the loss of carbohydrate groups which may contribute as much as 50% of the mucin molecular mass. The anti-(deglycosylated MGM) antibody recognizes the deglycosylated mucin as an extensive smear, which may indicate the degradation of protein epitopes has occurred or may be the result of non-enzymic deamidation [44,45]. While there is no definitive consensus sequence identifying positions at which this occurs, the presence of hydroxyl amino acids in the vicinity of either asparagine or glutamine (as in the case of mucins) appears to make these residues more susceptible to deamidation reactions.

cDNA isolation and sequencing

Approximately 500000 recombinants were screened with the anti-(deglycosylated MGM) antibody from which one positive clone was obtained. The clone MGM1 contained an insert of 390 bp and sequence analysis revealed two tandem repeats of 48 bp at the 5′ end of the clone (Figure 2b). The tandem repeats encode a 16-amino-acid peptide rich in both threonine and serine. The region 3′ to the tandem repeats is non-repetitive.

MGM1 was then used to screen the library for longer clones. After screening 160000 recombinants, four positive related clones were purified: MGM5b, MGM622, MGM6a and MGM61a (Figure 2a). Clones MGM5b and MGM622 overlapped with the non-repetitive region of MGM1. An additional clone MGM27E was identified by screening with a fragment of MGM5b. These partial clones provided the sequence for 768 bp of the non-repetitive region located 3′ to the MGM tandem repeat domain (Figure 2b). The non-repetitive sequence of MGM contains five potential sites for N-glycosylation. In addition there are 16 cysteine residues found in the non-repetitive portion of MGM which may function in disulphide bond formation for the oligomerization of the mouse gastric mucin.

Clones MGM6a and MGM61a, which were isolated by screening with MGM1, consisted solely of tandem repeats. Clone MGM6a was approximately 2000 bp and appeared to consist entirely of tandem repeats based on sequence analysis of the 5′ and 3′ ends. However, as has been encountered with other mucin clones composed solely of repeat units [15,46], this clone was unstable in pBluescript and further characterization was not attempted. MGM61a contained 198 bp and consisted of four complete tandem repeats plus 6 bp of an incomplete tandem repeat unit (Figure 3).

Comparison of the complete tandem repeat units found in

Table 2 Immunohistochemical reactivity of antibodies against native and deglycosylated MGM and against the MGM tandem repeat

Abbreviation: ND, not determined.

Tissue specimen	Anti-MGM	Anti-(deglycosylated MGM)	Anti-(synthetic MGM)
Forestomach	–	–	±
Glandular stomach			
Mucinous cells	+ + + Luminal content + + Cytoplasmic	+ + Cytoplasmic	+ + Cytoplasmic
Mucous neck cells	+ + Luminal content	+ + Cytoplasmic	+ + Cytoplasmic
Parietal cells	–	–	–
Chief cells	–	–	–
Antral glands	+ + Cytoplasmic	+ + Cytoplasmic	+ + Cytoplasmic
Duodenum	+ Goblet cells luminal content	ND	+ Glands at crypt base + + + Brunner's glands
Jejunum	+ Goblet cells luminal content	–	–
Ileum	+ Goblet cells luminal content	–	–
Caecum	+ Goblet cells luminal content	–	–
Rectum	+ Goblet cells luminal content	–	–
Kidney	–	–	–
Liver	–	–	–
Bronchus	–	–	–
Oesophagus	ND	ND	–
Pancreas	ND	ND	–
Gall-bladder	ND	ND	–

MGM1, MGM6a and MGM61a yields a consensus sequence of (Q)TSSPNTGK(T)ST(I)ST(T) where the residues in parentheses are not entirely conserved (Figure 3). To confirm that the MGM cDNA represents a gastric mucin found in the initial mucin preparation, a peptide was made based on the tandem repeat consensus sequence. ELISA analysis (Figure 4) illustrates that indeed the anti-(deglycosylated MGM) antibody recognizes the synthetic peptide. The consensus sequence illustrates the high serine and threonine content of the MGM with 25% and 38% similarity respectively. The presence of proline at the conserved position 5 and often at position 13 may be important for recognition by glycosyltransferases for O-linked glycosylation [47]. No significant sequence similarity exists between the MGM tandem repeat and the tandem repeats of other currently known animal or human mucins found in the GENEMBL sequence database.

Despite the lack of sequence similarity between the tandem repeat units of the mucins, the non-repetitive region of MGM shares sequence similarity with human MUC5AC. Two independent laboratories have isolated MUC5 clones from tracheobronchial libraries and designated them MUC5A, MUC5B, MUC5C and NP3a [18,48]. Recently Guyonnet Duperat et al. demonstrated that MUC5A and MUC5C are derived from the same gene which is now referred to as MUC5AC, while MUC5B originates from a different gene [49]. Clone NP3a contains sequences found within MUC5AC, suggesting NP3a and MUC5AC are part of the same gene. In addition, a MUC5 cDNA has been identified in a gastric cDNA library which contains the same tandem repeats as does MUC5AC [8]. Analysis of the MUC5AC cDNAs JER47 and JUL32 reveals an alternating structure of threonine- and serine-rich tandem repeat domains and cysteine-rich domains [49]. A similar structure has also been observed for MUC2 with two tandem repeat domains separated and flanked by non-repetitive cysteine-rich regions [50]. The MGM cDNA can be illustrated in a similar pattern with the tandem repeat domain followed by a 133-amino-acid cysteine-rich non-repetitive region (CRRI), a 63-residue non-repetitive serine/threonine-rich domain and then a second cysteine-rich region (CRRII) (Figure 5a). Comparison of the cysteine-rich regions of MGM shows that CRRI shares 81% and 70% similarity with the first cysteine-rich domains of the MUC5AC clones JUL32 and JER47 respectively (Figure 5b). MGM CRRII is 76% similar to the second cysteine-rich domain of JER47 and 77% similar to the second partial cysteine-rich domain of JUL32 (Figure 5c). These domains are separated by a tandem repeat domain in JER47 and by a non-repetitive 58-amino-acid serine/threonine-rich region in JUL32. Between the MGM CRRs is found a non-repetitive 63-amino-acid serine/threonine-rich region which shares little sequence similarity (45%) with the serine/threonine-rich region of JUL32. MGM CRRI and CRRII are only 38% similar; however, this occurs with a striking conservation of cysteine residues. Guyonnet Duperat et al. [49] derived a consensus sequence found in the cysteine-rich regions of human MUC5AC and MUC2. The CRRs of MGM also conform to the consensus sequence (Figure 5d). The similar spatial arrangement of the thiol groups suggests a functional importance for these residues.

Sequence similarity is also found to a lesser extent between the non-repetitive cysteine-rich portion of MGM and (i) the human intestinal MUC2 and (ii) the rat intestinal mucin, with 57% and 54% similarity respectively (Figure 5b). These cysteine-rich regions are found in the two MUC2 tandem repeat domains and within a non-repetitive region of the rat intestinal mucin. No significant similarity exists between any of the non-repetitive MGM sequence and 800 bp of the 5′ non-repetitive MUC6 sequence and 1600 bp of the 3′ non-repetitive MUC6 sequence determined to date (N. Toribara, personal communication). The high degree of sequence similarity between MGM and MUC5AC suggests that MGM may represents a homologue of the human MUC5AC.

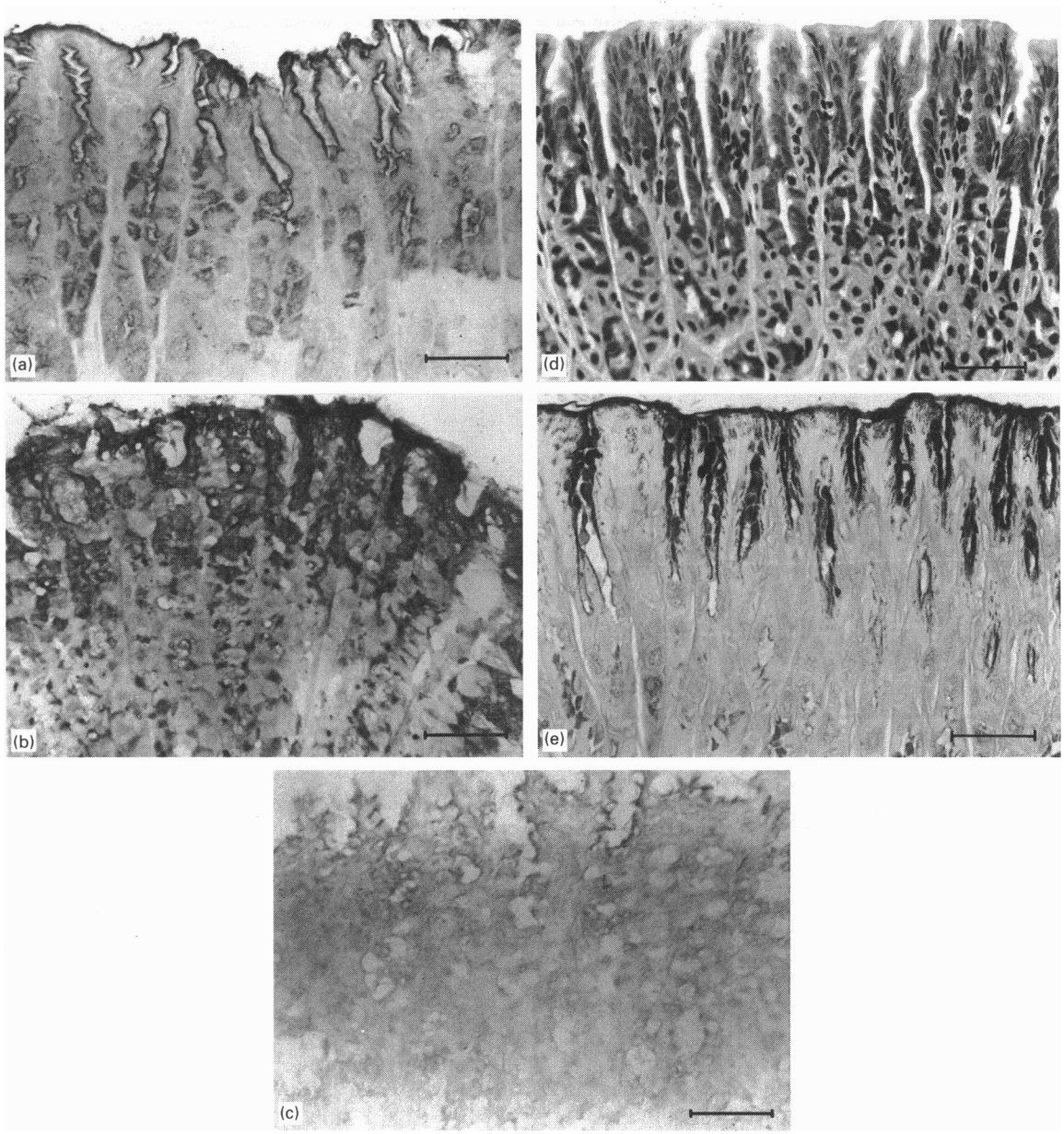

Figure 9 Immunohistochemical analysis of mouse stomach

Mouse gastric tissues were stained with (**a**) anti-(native MGM), (**b**) anti-(deglycosylated MGM) or (**c**) preimmune antibodies as described in the Experimental section. Serial sections were stained with (**d**) haematoxylin and eosin in order to demonstrate morphology and (**e**) Alcian Blue/periodic acid/Schiff. Mucin carbohydrate is stained with Alcian Blue/periodic acid/Schiff and corresponds with cells reactive with mucin antibodies. Bars represent 50 μm.

RNA and DNA analysis

Analysis of the tissue and species specificity of MGM was performed by Northern and Southern blot analysis. When mouse RNA is probed with a rat intestinal mucin probe, strong hybridization is observed with RNA isolated from mouse intestine and colon (Figure 6a). Using the MGM clones, hybridization is strictly confined to the mouse stomach, recognizing a polydisperse smear beginning at approximately 13.5 kb which is slightly smaller than the intestinal mucin of approximately 15.7 kb (Figure 6b). Each of the MGM clones yielded a similar hybridization pattern. The polydisperse pattern has been observed in Northern blot analysis of previously cloned secretory mucins. The cause of this polydisperse pattern is unknown; however, it may be due to rapid turnover of mucin mRNA or to instability of very long mRNA molecules. Degradation of RNA samples has been ruled out by demonstrating the integrity of the RNA with ethidium bromide staining and also by the presence of a discrete band when the blot is probed with a GAP cDNA (Figures 6c and 6d). Analysis of rat RNA with MGM1 shows weak cross-hybridization with rat stomach RNA, with slightly more message in the antrum than the fundus (Figure 7a). The polydisperse message is slightly smaller (ranging from near 11.3 kb down to several hundred base pairs) than that observed with mouse RNA, indicating that there exists a related but not identical rat gastric mucin. Again, the

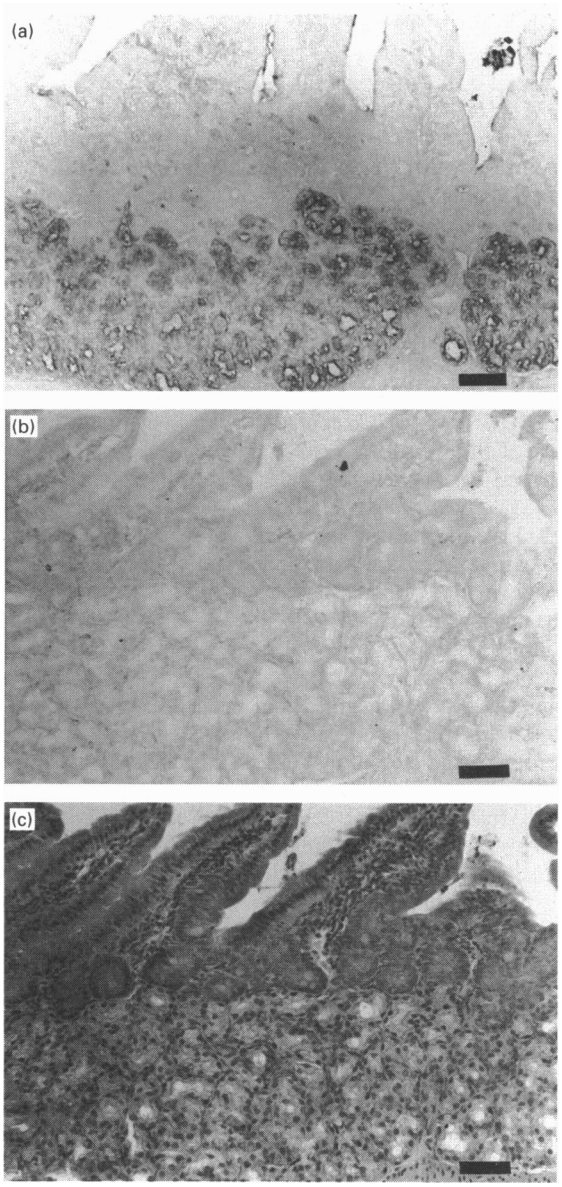

Figure 10 Immunohistochemical analysis using the anti-(tandem repeat) antibody

(**a**) Mouse duodenal tissue was stained with the antibody raised against the synthetic MGM tandem repeat peptide showing postive reactivity in Brunner's glands. (**b**) Duodenal mucosa stained with preimmune antibody. (**c**) Same tissue stained with haematoxylin and eosin to demonstrate morphology. Bars represent 50 μm.

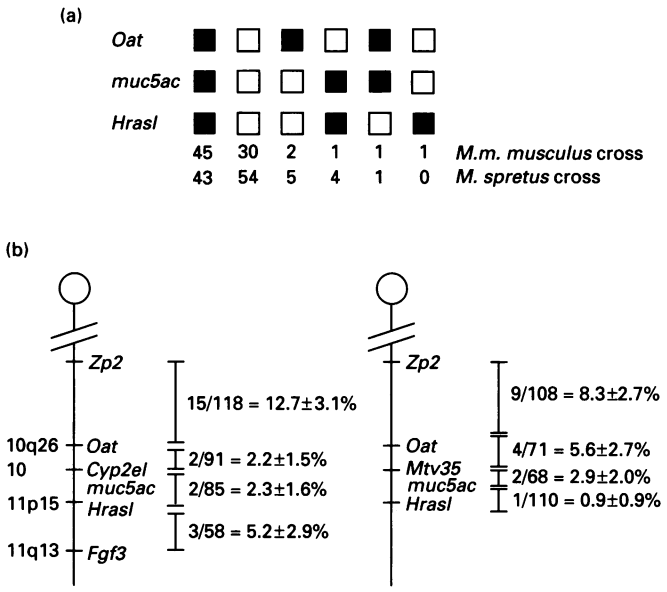

Figure 11 Inheritance of *muc5ac* with markers on Chr 7

(**a**) Black squares represent heterozygous mice, open squares represent homozygous mice. Numbers at the bottom of each column represent the number of mice in each cross with the indicated genotype. (**b**) Abbreviated genetic maps of distal mouse Chr 7 indicating the map location of *Muc5ac* with respect to *Oat* and *Hras1* as well as the additional markers *Zp2, Mtv35, Cyp2e1* and *Fgf3*. The map locations for the human homologues of the underlined genes are listed on the left-hand side.

polydisperse hybridization pattern does not indicate RNA degradation due to demonstration of intact RNA by ethidium bromide staining and the presence of a single band when the blot is probed with a GAP cDNA (results not shown). No hybridization was observed with the mouse RNA and probes containing either a portion of the human gastric MUC6 tandem repeat or the human MUC5 tandem repeat (results not shown), indicating that mouse homologues with close similarity to the MUC5 and MUC6 tandem repeats do not exist.

Southern blot analysis of mouse DNA cleaved with *Bam*H1 reveals a hybridizing signal of greater than 10 kb when using the MGM1 clone (Figure 8). A related gene is present in rat as indicated by the hybridization of MGM1 to rat DNA (Figure 8). No cross-reactivity was found between the MGM clone and hamster, rabbit or human DNA under the indicated wash conditions.

Immunohistochemical analysis

Determination of the cellular locale and distribution of gastric mucin was performed by immunohistochemical analysis. Antibodies against native and deglycosylated MGM stained the apical cytoplasm and luminal content of superficial gastric mucous cells, the cytoplasm of cells from the mucous neck region of gastric fundus, and antral gland cells (Table 2, Figure 9). The chief and parietal cells were negative with both antibody preparations. The luminal content and goblet cell vacuoles in the small intestine and colon demonstrated weak staining with the anti-(native MGM) but were negative for staining with the anti-(deglycosylated MGM). No reactivity was observed for either antibody in tissue from the kidney, liver or bronchus.

Because the initial mucin preparation may contain multiple mucins, an antibody directed against a synthetic MGM tandem repeat was prepared to define further the expression pattern of MGM. The purified antibody had a high titre towards the tandem repeat peptide as measured by ELISA. Immunohistochemical analysis confirmed the tissue-specific expression of the MGM suggested by the anti-(deglycosylated MGM) antibody. The anti-(MGM peptide) antibody strongly stained surface mucous cells of the mouse fundus, antral gland cells (Figure 9) and Brunner's glands of the duodenum (Table 2, Figure 10). No reactivity was seen with trachea, lung, oesophagus, pancreas, kidney, small intestine, caecum or rectum.

Chromosomal localization

Southern blot analysis identified a 3.4 kb *Pst*I fragment in NFS/N and a 3.7 kb fragment in *M.m. musculus*. *Bgl*II digestion produced a 9.5 kb fragment in NFS/N and a 7.4 kb fragment in *M. spretus*. Inheritance of the inbred strain fragment and the *M. spretus* fragment was followed in the progeny of two genetic crosses and compared with inheritance of over 700 markers in the two crosses. By convention, mouse genes that are thought to be homologous with known human genes receive the same name, hence, *Muc5ac*. As shown in Figure 11, the gene for mouse gastric mucin, *Muc5ac*, was linked to markers on distal mouse Chr 7. The closest marker was identified in the *M.m. musculus* cross. No recombinants were found between *Muc5ac* and *Cyp2e1* in the 87 mice typed for both markers. At the 95% confidence level, this indicates that these genes are separated by no greater than 3.4 cM [39].

DISCUSSION

Direct amino acid sequencing of mucins is complicated by the large size and heavy glycosylation of the protein and by the fact that multiple distinct mucins may be present in a single tissue and cannot be separated using traditional biochemical techniques. Expression cloning provides a convenient alternative to determining the primary sequence of mucins. By screening a mouse gastric cDNA library with the anti-(deglycosylated MGM) antibody, we have isolated an MGM cDNA clone. Taken together, data from sequence analysis, Northern and Southern blots, chromosomal location and immunohistochemical analysis indicate that this cDNA is a tissue-specific mucin representing a mouse homologue of MUC5AC. First, confirmation that our clone represents a mucin is provided by the presence of tandem repeat units rich in threonine and serine. A domain of tandem repeat arrays enriched in hydroxyl amino acids characterizes all known mucins. The abundance of threonine and serine residues provides many potential sites for O-glycosylation. The MGM tandem repeat sequence is recognized by the anti-(deglycosylated MGM) antibody, confirming that this sequence was present in a protein found in our initial mucin preparation. The MGM tandem repeat of 48 bp encodes for a peptide of 16 amino acids with 38% threonine and 25% serine. The tandem repeats found in other cloned mucins range from 6 to 23 amino acids with the exceptions of the human gastric mucin (169-amino-acid repeat) [9] and the porcine submaxillary mucin (81-amino-acid repeat) [12]. Of additional note regarding the MGM tandem repeat is the presence of at least one proline residue. Proline is an important recognition factor for use by GalNAc transferases in the identification of sites for O-glycosylation [47].

A search of the sequence database GENEMBL indicates that the MGM tandem repeat shares no similarity with any previously identified mucin tandem repeat. The tandem repeats of the human gastric mucins MUC5AC and MUC6 share no sequence similarity with those of MGM, and the MUC6 tandem repeat differs greatly in length. The content of serine within the human gastric MUC5AC and MUC6 tandem repeats and the MGM tandem repeat is similar (18%, 25% and 25% respectively). The MUC5AC tandem repeat contains a higher proportion of threonine than do MUC6 and MGM (50% compared with 31% and 38% respectively) [9]. Further evidence that the mouse gastric tandem repeat is unique to the mouse is provided by the lack of hybridization between probes specific for the human gastric MUC5AC and MUC6 tandem repeats and mouse RNA.

A second feature of our MGM cDNA which identifies it as a mucin clone is the polydisperse pattern of hybridization observed on Northern blot analysis. Analysis of mouse stomach RNA revealed a polydisperse message beginning at 13.5 kb and extending to 200 bp. The presence of a 'smear' extending from close to 10 kb down to several hundred base-pairs is a characteristic feature of each of the human secreted mucins [51]. The cause of the polydisperse message is unknown. Possibilities include a rapid turnover of mucin message, degradation due to enhanced instability of long tandem repeats, or incomplete or alternative splicing [51,52].

In addition, the tissue distribution of MGM expression is as expected for a MUC5AC homologue. RNA analysis showed that MGM was exclusively expressed in normal mouse stomach. Normal human stomach possesses high levels of MUC5AC and MUC6 mRNA and immunoreactive protein [48,53]. No MUC5B expression is detected [48]. By immunohistochemical comparison, gastric MUC5 and MUC6 expression can be distinguished by the presence of MUC5 in the surface mucous cells, whereas MUC6 is expressed by mucous neck cells, antral glands and Brunner's glands [53]. These mucins are not unique to the gastric mucosa as MUC5 is expressed by bronchial epithelium and low levels of MUC6 can be found in ileum and colon. The pattern of MGM expression in the stomach, as shown by immunoreactivy with anti-(deglycosylated gastric mucin) and anti-(MGM peptide), has characteristics of both MUC5AC and MUC6; however, expression of MGM occurs only in the stomach, without concomitant expression in normal bronchial or colonic tissue. The possibility that MGM may be expressed by bronchial tissue has not been ruled out. We performed RNA and immunohistochemical analyses on tissue from specific pathogen-free rodents. Mucin-producing cells in the bronchial tissue of these animals are very rare and difficult to detect. Tracheobronchial mucin isolated from an asthmatic individual [18,54] contains a peptide corresponding to a portion of the MUC5AC sequence [49]. Examination of apparently normal bronchial mucosa failed to demonstrate any MUC5AC expression in the glandular acini with inconsistent expression found throughout the respiratory tree [48]. Mucin expression can be induced in bronchial tissue by exposure to an irritant such as sulphur dioxide [55]. Taken together these data suggest that expression of MUC5AC in bronchial tissue is induced upon progession to a diseased or cancerous state. Studies are currently underway to determine whether MGM expression can be induced in rodent airways using irritants.

The MGM gene has been localized to mouse chromosome 7. It is noteworthy that the region of chromosome 7 distal to *Cyp2e1* is homologous to human chromosome 11p15. Previous studies have shown that the human genes for MUC2 [56], MUC5 [57] and MUC6 [9] are clustered on chromosome 11p15. While there may be additional, as yet unidentified, mucin genes residing in this region, the proximity of the MUC2, MUC5 and MUC6 genes suggests that they comprise a multigene family whose members encode secretory mucins. It is tempting to speculate that our MGM belongs to an analogous multigene family of mouse secretory mucin genes.

We have presented here the first report of a mouse MUC5AC homologue. Mouse models provide a powerful tool to investigate the control of mucin gene expression. The availability of the mouse MUC5AC homologue will greatly enhance the ability to investigate the interrelationship of mucin gene expression and gastric and respiratory disease.

Note added in proof (received 21 August 1995)

Recently a pig gastric mucin cDNA has been sequenced and found to have a 16-amino-acid tandem repeat which shares no similarity to the MGM tandem repeat [59].

We would like to thank Alan Davis, Dr. Kahlil Ahmed and Dr. Steffan N. Ho for invaluable discussions during the course of this project. We would like to thank Dr. Neil Toribara for providing the human gastric MUC6 cDNA and Drs. James R. Gum and Young S. Kim for generously providing the rat intestinal cDNA probe. We also thank Leone and Kenneth Henry for providing chickens for antibody production. This work was supported by a Veterans Affairs Merit Review Award (to S.B.H.), a National Research Service Award (to L.L.S.) and the Research Service of the Veterans Affairs Medical Center.

REFERENCES

1 Neutra, M. and Forstner, J. (1987) in Physiology of the Digestive Tract (Johnson, L. and Johnson, L. S., eds.), pp. 975–1009, Raven Press, New York
2 Bhaskar, K. R., Garik, P., Turner, B. S., et al. (1992) Nature (London) **360**, 458–461
3 Strous, G. J. and Dekker, J. (1992) Crit. Rev. Biochem. Mol. Biol. **27**, 57–92
4 Gendler, S., Lancaster, C., Taylor-Papdimitriou, J., et al. (1990) J. Biol. Chem. **265**, 15286–15293
5 Spicer, A. P., Parry, G., Patton, S. and Gendler, S. J. (1991) J. Biol. Chem. **266**, 15099–15109
6 Ho, S. B., Niehans, G. A., Lyftogt, C., et al. (1993) Cancer Res. **53**, 641–651
7 Porchet, N., Van Cong, N., Dufosse, J., et al. (1991) Biochem. Biophys. Res. Commun. **175**, 414–422
8 Ho, S. B., Roberton, A. M., Shekels, L. L., Lyftogt, C. T., Niehans, G. A. and Toribara, N. W. (1995) Gastroenterology **109**, 735–747
9 Toribara, N. W., Roberton, A. M., Ho, S. B., et al. (1993) J. Biol. Chem. **268**, 5879–5885
10 Bobek, L. A., Tsai, H., Biesbrock, A. R. and Levine, M. J. (1993) J. Biol. Chem. **268**, 20563–20569
11 Bhargava, A., Woitach, J., Davidson, E. and Bhavanandan, V. (1990) Proc. Natl. Acad. Sci. U.S.A. **87**, 6798–6802
12 Eckhardt, A., Timpte, C., Abernethy, J., Toumadje, A., Johnson, W. J. and Hill, R. (1987) J. Biol. Chem. **262**, 11339–11344
13 Verma, M. and Davidson, E. A. (1993) Proc. Natl. Acad. Sci. U.S.A. **90**, 7144–7148
14 Sorimachi, H., Emori, Y., Kawasaki, H., et al. (1988) J. Biol. Chem. **263**, 17678–17684
15 Gum, J., Hicks, J., Lagace, R., et al. (1991) J. Biol. Chem. **266**, 22733–22738
16 Ohmori, H., Dohrman, A. F., Gallup, M., et al. (1994) J. Biol. Chem. **269**, 17833–17840
17 Gum, J., Hicks, J., Toribara, N., Rothe, E.-M., Lagace, R. and Kim, Y. (1992) J. Biol. Chem. **267**, 21375–21383
18 Meerzaman, D., Charles, P., Daskal, E., Polymeropoulos, M. H., Martin, B. M. and Rose, M. C. (1994) J. Biol. Chem. **269**, 12932–12939
19 Slomiany, B. L., Sarosiek, J. and Slomiany, A. (1987) Dig. Dis. **5**, 125–145
20 Correa, P. (1988) Cancer Res. **48**, 3554–3560
21 Byrd, J., Nardelli, J., Siddiqui, B. and Kim, Y. (1988) Cancer Res. **48**, 6678–6685
22 Laemmli, U. (1970) Nature (London) **227**, 680–685
23 Dubray, G. and Bezard, G. (1982) Anal. Biochem. **119**, 325–329
24 Goueli, S. A., Hanten, J., Davis, A. and Ahmed, K. (1990) Biochem. Int. **21**, 685–694
25 Avrameas, S. (1969) Immunochemistry **6**, 43–52
26 Maniatis, T., Fritsch, E. and Sambrook, J. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
27 Harlow, E. and Lane, D. (1988) Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
28 Chomczynski, P. and Sacchi, N. (1987) Anal. Biochem. **162**, 156–159
29 Sanger, F., Nicklen, S. and Coulson, A. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 5463–5467
30 Devereux, P., Haeberli, P. and Smithies, O. (1984) Nucleic Acids Res. **12**, 387–391
31 Feinberg, A. and Vogelstein, B. (1983) Anal. Biochem. **132**, 6–13
32 Tso, J. Y., Sun, X. H., Kao, T. H., Reece, K. S. and Wu, R. (1985) Nucleic Acids Res. **13**, 2485–2502
33 Kozak, C. A., Peyser, M., Krall, M., et al. (1990) Genomics **8**, 519–524
34 Adamson, M. C., Silver, J. and Kozak, C. A. (1991) Virology **183**, 778–781
35 Danciger, M., Farber, D. B. and Kozak, C. A. (1993) Genomics **16**, 361–365
36 Lunsford, R. D., Jenkins, N. A., Kozak, C. A., et al. (1990) Genomics **6**, 184–187
37 Ramesh, V., Cheng, S. V., Kozak, C. A., et al. (1992) Mammal. Genome **3**, 17–22
38 Major, J. E. and Varmus, H. (1983) J. Virol. **47**, 495–504
39 Green, E. L. (1981) in Genetics and Probability in Animal Breeding Experiments, pp. 77–113, Macmillian, NY
40 Gum, J., Hicks, J., Swallow, D., et al. (1990) Biochem. Biophys. Res. Commun. **171**, 407–415
41 Dekker, J., Aelmans, P. H. and Strous, G. J. (1991) Biochem. J. **277**, 423–427
42 Gassmann, M., Thommes, P., Weiser, T. and Hubscher, U. (1990) FASEB J. **4**, 2528–2532
43 Polson, A. and von Wechmar, M. B. (1980) Immunol. Commun. **9**, 475–493
44 Kieliszewski, M. J., Kamyab, A., Leykam, J. F. and Lamport, D. T. A. (1992) Plant Physiol. **99**, 538–547
45 Wright, H. T. (1991) Crit. Rev. Biochem. Mol. Biol. **26**, 1–52
46 Toribara, N. W., Gum, J. R., Culhane, P. J., et al. (1991) J. Clin. Invest. **88**, 1005–1013
47 Briand, J. P., Andrews, S. P., Cahill, E., Conway, N. A. and Young, J. D. (1981) J. Biol. Chem. **256**, 12205–12207
48 Audie, J. P., Janin, A., Porchet, N., Copin, M. C., Gosselin, B. and Aubert, J. P. (1993) J. Histochem. Cytochem. **41**, 1479–1485
49 Guyonnet Duperat, V., Audie, J.-P., Debailleul, V., et al. (1995) Biochem. J. **305**, 211–219
50 Gum, J., Hicks, J., Toribara, N., Siddiki, B. and Kim, Y. (1994) J. Biol. Chem. **269**, 2440–2446
51 Gum, J. R. (1992) Am. J. Respir. Cell Mol. Biol. **7**, 557–564
52 Rose, M. (1992) Am. J. Physiol. **263**, L413–L429
53 Ho, S. B., Shekels, L. L., Toribara, N. W., Lyftogt, C. T. and Niehans, G. A. (1994) Gastroenterology **106**, A94
54 Rose, M., Kaufman, B. and Martin, B. (1989) J. Biol. Chem. **264**, 8193–8199
55 Basbaum, C., Gallup, M., Gum, J., Kim, Y. and Jany, B. (1990) Biorheology **27**, 485–489
56 Griffiths, B., Matthews, D. J., West, L., et al. (1990) Ann. Hum. Genet. **54**, 277–285
57 Van Cong, N., Aubert, J. P., Gross, M. S., Porchet, N., Degand, P. and Frezal, J. (1990) Hum. Genet. **86**, 167–172
58 Hansson, G., Baeckstrom, D., Carlstedt, I. and Klinga-Levan, K. (1994) Biochem. Biophys. Res. Commun. **198**, 181–190
59 Turner, B. S., Bhaskar, K. R., Hadzopoulou-Cladaras, M., Specian, R. D. and LaMont, J. T. (1995) Biochem. J. **308**, 89–96

Received 22 March 1995/15 May 1995; accepted 8 June 1995