[CANCER RESEARCH 52, 1954–1960, April 1, 1992]

Tissue-specific Expression of a Human Polymorphic Epithelial Mucin (MUC1) in Transgenic Mice

Nigel Peat, Sandra J. Gendler, El-Nasir Lalani, Trevor Duhig, and Joyce Taylor-Papadimitriou[1]

Imperial Cancer Research Fund, P. O. Box 123, Lincoln's Inn Fields, London WC2A 3PX, England

ABSTRACT

The human *MUC1* gene codes for the core protein of a mucin which is expressed by glandular epithelia and the carcinomas which develop from these tissues. The core protein is aberrantly glycosylated in cancers, and some antibodies show specificity in their reactions with the cancer-associated mucin, which also contains epitopes recognized by T-cells from breast and pancreatic cancer patients. For evaluating the potential use of mucin-reactive antibodies and mucin-based immunogens in cancer patients, a mouse model, expressing the *MUC1* gene product PEM (polymorphic epithelial mucin) as a self antigen, would be extremely useful. To this end, we have developed transgenic mouse strains expressing the human *MUC1* gene product in a tissue-specific manner. The TG4 mouse strain was established using a 40-kilobase fragment containing 4.5 kilobases of 5′ and 27 kilobases of 3′ flanking sequence. The TG18 strain was developed using a 10.6-kilobase *Sac*II fragment from the 40-kilobase fragment; this fragment contained 1.6 kilobases of 5′ sequence and 1.9 kilobases of 3′ flanking sequence. Both strains showed tissue specificity of expression of the *MUC1* gene, which was very similar to the profile of expression seen in human tissues. The antibody SM-3 is directed to a core protein epitope, which is selectively exposed in breast cancers and which shows a more restricted distribution on normal human tissues. It was established that the distribution of the SM-3 epitope of PEM in the tissues of the transgenic mice is similar to that seen in humans. The transgenic mouse strains described here should form the basis for the development of a preclinical model for the evaluation of PEM-based antigens and of antibodies directed to PEM in cancer therapy.

INTRODUCTION

Attention has been focused on the epithelial-associated mucins because many antibodies developed against normal or malignant epithelial cells and their products have been found to react with these complex molecules (for a review see Ref. 1). One such mucin is produced in abundance by the lactating mammary gland and expressed by many simple epithelial cells lining glands or ducts and the adenocarcinomas which develop from them. The complete amino acid sequence of the core protein of the mammary mucin has been obtained by gene cloning and shows the mucin to be a transmembrane protein with a large extracellular domain made up largely of 20 amino acid tandem repeats and a cytoplasmic tail of 69 amino acids (2–4). Since the gene and its expressed product show a high degree of length polymorphism due to the presence of different numbers of tandem repeats (2, 5), we have referred to the gene product as the polymorphic epithelial mucin (2, 6). Partial complementary DNA clones coding for two gastrointestinal mucin core proteins (7, 8) and a tracheobronchial mucin-like protein (9) have recently been isolated and sequenced. We therefore now refer to the gene as *MUC1* to distinguish it from the other three human mucin genes (which are referred to as *MUC2*, *MUC3*, and *MUC4*), but we continue to refer to the product as PEM.[2]

Although PEM is expressed on both normal and malignant epithelial cells, the carbohydrate side chains of the cancer-associated mucin are shorter (10, 11), resulting in the expression of new carbohydrate epitopes and in the exposure of novel epitopes in the core protein, which show a restricted distribution on normal cells. Antibodies to PEM can therefore show some degree of specificity in their reactions with tumors. The antibody SM-3 (12) reacts with a core protein epitope (13) which is selectively exposed in breast and other carcinomas (13, 14) and has been successfully used for imaging ovarian cancer (15). The difference in glycosylation also results in the exposure of core protein epitopes (overlapping the SM3 epitope) which are recognized by T-cells isolated from cancer patients (16, 17), suggesting that the PEM antigen could be effective in active specific immunotherapy.

For evaluating drug or toxin conjugates of antibodies to PEM and for assessing whether the antigen itself may be useful in the therapy of some adenocarcinomas, it is essential to have a model system for preclinical studies. In this context, the xenograft model is of limited value, since the antigen is not expressed on the normal tissues of the mouse and the immune system of the nude mouse is defective. For testing the immunogenicity and efficacy of PEM-based immunogens in tumor rejection, we have developed a syngeneic mouse model where a transplantable tumor expresses PEM (18). Although this model is useful for comparing different immunogens, the fact that the PEM antigen is recognized as foreign limits its use for predicting the immune response of human subjects where PEM is expressed as a self antigen. Thus a mouse expressing the *MUC1* gene product as a self antigen in a tissue-specific manner would provide an extremely important model for preclinical testing of both antibody conjugates and PEM-based immunogens. With this in mind we have developed transgenic mouse strains where PEM is expressed in the same epithelial tissues as in humans.

The glycosylation of PEM can vary with the tissue, resulting in the production of final mucin products with different antigenic profiles. In particular, the epitope recognized by SM3 is not found on all the expressing tissues. We have therefore also looked at the reaction of the different tissues with the antibody SM3. Our results show that, as is seen in humans (12, 14, 19), the epitope recognized by this antibody shows a more restricted distribution, suggesting that the glycosylation pattern of PEM in the tissues of the transgenic mice is similar to that seen in humans.

MATERIALS AND METHODS

Development of Transgenic Mice Expressing PEM. A cosmid clone GPEM-1 was isolated from the human genomic library pCOS2EMBL made from peripheral blood lymphocytes (kindly donated by Dr. A. M. Frischauf, ICRF, London) (20). The 40-kilobase insert was prepared by digesting the cosmid DNA, which was prepared with a Quiagen

Received 9/20/91; accepted 1/27/92.

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.

[1] To whom requests for reprints should be addressed.

[2] The abbreviations to be used are: PEM, polymorphic epithelial mucin; PBS, phosphate-buffered saline.

column (Hybaid, Teddington, Middlesex, England) with *SalI* and *ClaI* restriction endonucleases (New England Biolabs, Beverley, MA), which cut into the vector sequence at about 300 and 400 base pairs, respectively, from the insert site. The insert was electroluted from the agarose gel, ethanol precipitated, redissolved in water, and centrifuged through a 1-ml Sepharose G25 column. DNA was filter sterilized for injection by centrifuging it through a 0.22 μm cellulose acetate spin-x Costar filter unit using TE buffer made from molecular biology-grade water (BDH Eastleigh, Hertfordshire, England).

The *SacII* genomic fragment was prepared by digesting the GPEM-1 clone and isolating the 10.6-kilobase piece of DNA as described above.

DNA was diluted to a final concentration of 1 mg/ml for injection into the pronuclei of fertilized CBA × C57Blk F_1 mouse eggs, which were then tranferred to pseudopregnant CBA × C57 Blk F_1 females. All manipulations were as described by Hogan *et al.* (21). Transgenic mouse strains derived from the 40-kilobase fragment are referred to as TG4 and TG5, and that from the *SacII* fragment as TG18.

Genomic DNA Isolation and Southern Blot Analysis. Genomic DNA was prepared from T47D cells (human mammary carcinoma; Ref. 22) with an Applied Biosystems (Foster City, CA) 340A DNA extractor. Mouse tail genomic DNA was prepared according to the method of Hogan *et al.* (21). Samples (10–15 μg) were digested with the appropriate restriction endonucleases (New England Biolabs, Beverley, MA) under conditions recommended by the manufacturer, prior to electrophoresis through a 0.7% agarose gel.

Southern blotting onto nylon membranes (Pall Biodyne, Glen Cove, NY) and subsequent hybridization/washing procedures were carried out according to the manufacturer's instructions. The human probe pMUC7, corresponding to ~500 base pairs of tandem repeat domain (5), was utilized. The probe was labeled using random priming (23) in the presence of [α-^{32}P]dCTP (Amersham International plc) to a specific activity of >1 × 10^8 dpm/μg.

RNA Analysis. Total RNA was prepared from tissue culture cells and mouse mammary gland (lactating and nonlactating) by guanidinium isothiocynate extraction (24), the tissue having first been pulverized with a Braun Mikro-dismembranator II (Melsuugen, Germany). Northern analysis was carried out as previously described (2) using [α-^{32}P]dCTP-labeled probes to the human tandem repeat (pMUC7) and a mouse mucin probe (pMuc10), corresponding to ~1.2 kilobases of 3′ complementary DNA, from the mouse repeat domain to the polyadenylated tail (25). For mapping of the transcriptional start site total RNA from T47D, ICRF 23 (human embryonic lung fibroblast cell line), TG4, TG5, and nontransgenic lactating mammary gland were isolated and primer extension analysis was carried out as described previously (2).

Staining of Tissue Sections. Transgenic mice (4–6 weeks old or at lactation) were sacrificed and dissected. Tissues were divided and either stored in liquid nitrogen prior to RNA extraction or fixed in methacarn (methanol:chloroform:acetic acid, 60:30:10) and paraffin embedded for staining. Dewaxed paraffin sections were incubated with 50% fetal calf serum in PBS for 30 min to prevent nonspecific binding of antibodies. Blocking solution was then drained off, replaced by neat hybridoma culture supernatant, and incubated for 60 min at room temperature. After three 5-min washes in PBS the sections were incubated for 1 h with peroxidase-conjugated rabbit anti-mouse immunoglobulin antiserum (Dako; diluted 1:50 in PBS with 15% fetal calf serum). The slides were then washed three to four times in PBS with intermittent agitation. The substrate solution consisting of 0.03% hydrogen peroxide in PBS and 1 mg/ml diaminobenzidine (Sigma) was added, and the reaction was allowed to continue for 5–8 min. After washing in PBS the slides were counterstained with hematoxylin and mounted in gelvatol.

RESULTS

Identification of *MUC1* Transgenic Mice. We have recently reported the full sequence of the *MUC1* gene, together with 800 base pairs of 5′ flanking sequence (20). The sequence was obtained from cosmid clones, one of which, GPEM1, was obtained from a pCos2EMBL library (20). The full insert of ≃40 kilobases together with ≃700 base pairs of vector sequence could be obtained from GPEM1 by cutting with the enzymes *SalI* and *ClaI* (Fig. 1). This fragment contains the sequences for the *MUC1* gene with 4.5 kilobases of 5′ and ≃27 kilobases of 3′ flanking sequences, and was used to develop transgenic mice. In later experiments a 10.6-kilobase fragment obtained from a *SacII* digest of GPEM1 was also successfully used to develop a PEM-expressing strain of mouse. This fragment contained only 1.6 kilobases of 5′ sequence and 1.9 kilobases of 3′. Fig. 1 shows diagrammatically the restriction enzyme sites in the large fragment and the intron-exon structure of the gene contained within the 10.6-kilobase *SacII* fragment.

The 40-kilobase or 10.6-kilobase fragments, isolated from GPEM1, were microinjected into fertilized mouse ova, transplanted into pseudopregnant mice, and DNA prepared from tail snips of the litters. The DNA was digested with *Eco*RI and analyzed on Southern blots using the probe pMUC7. pMUC7 contains eight of the 60-base pair tandem repeats which make up a large part of exon 2 of the *MUC1* gene (see Fig. 1) and no unique sequence. The mouse gene corresponding to the human *MUC1* has recently been cloned (25), and although it also has been found to contain tandem repeats of 20–21 amino acids, the sequence is only 56% homologous (at the nucleotide level)

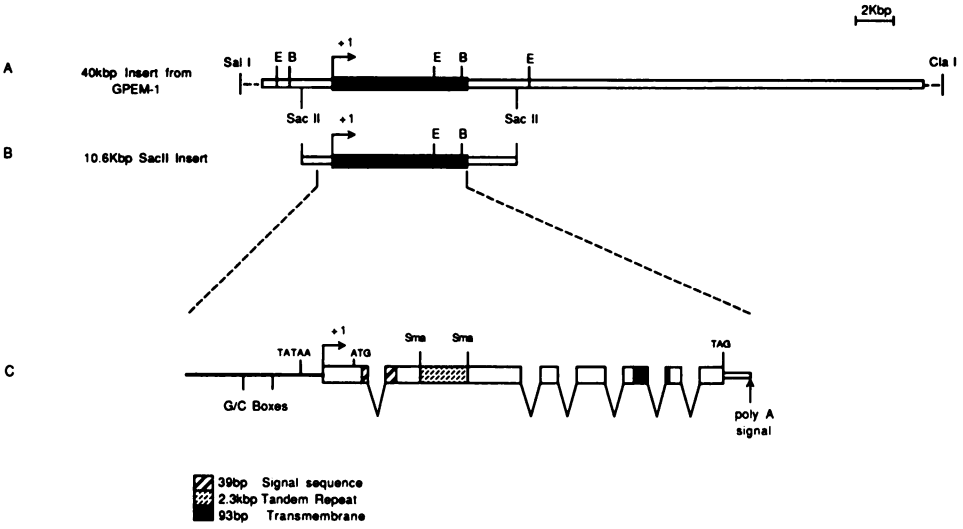

Fig. 1. DNA fragments from pCos2EMBL containing the *MUC1* gene used for injection into fertilized eggs. *A*, 40-kilobase *SalI* and *ClaI* fragment; *B*, 10.6-kilobase *SacII* fragment; *C*, intron-exon structure and major features of the gene. *E*, *Eco*RI sites; *B*, *Bam*HI sites. *A* and *B* are drawn to scale; *C* is not.

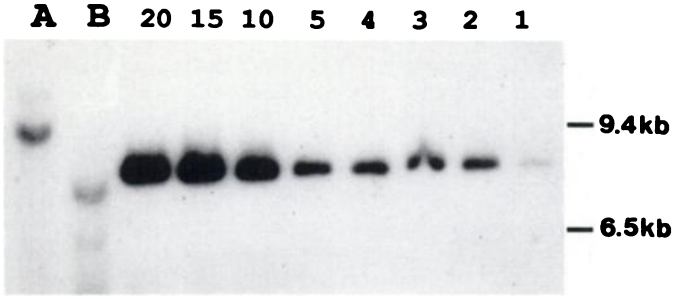

Fig. 2. Southern blot of *Eco*RI digest of DNA from the transgenic mouse strains TG18 (*Lane A*) and TG4 (*Lane B*) probed with pMUC7. Different amounts of an *Eco*RI digest of the pcos2EMBL cosmid were applied in the other lanes. *Numbers along the top*, numbers of copies which would have to be present in the 20 μg of total genomic DNA applied to *Lanes A* and *B* to give the signal shown in the cosmid digests.

to that seen in the human gene. pMUC7 therefore does not hybridize to the mouse gene and can be used to selectively detect the presence of the human *MUC1* in the mouse genome. Using the pMUC7 probe, three mice carrying the 40-kilobase fragment in their genome were identified, one of which did not transmit. The other two, TG4 and TG5, did show transmission and were analyzed further. Two mice were identified as carrying the 10.6-kilobase fragment, and one of these (TG18) was also analyzed.

Fig. 2 shows a Southern blot of *Eco*RI digests of DNA from two mouse strains (TG4 and TG18) as well as different concentrations of the cosmid clone GPEM1, also digested with *Eco*RI. The *Eco*RI fragment from the TG4 mouse strain is the same size as the fragment released by *Eco*RI digestion from GPEM1 (8.3 kilobases), while in the case of the TG18 mice derived with the 10.6-kilobase *Sac*II fragment, the *Eco*RI fragment reactive with the pMUC7 probe is around 10.6 kilobases. [The same size fragment shows a different mobility when run as picograms of pure DNA (cosmid digests) or when digested from 15 μg of genomic DNA.] A comparison of the intensity of the signals from the main band in each of the DNA digests indicates that the copy number is low, of the order of 1–3 in both lines. There is only one *Eco*RI site in the 10.6-kilobase *Sac*II fragment, indicating that there must be at least two copies integrated in tandem in this line, to produce the dominant band of 10.6 kilobases in the *Eco*RI digest. The faint bands observed in the digests probably correspond to a fragment containing some flanking host DNA together with the part of the PEM gene including the tandem repeat domain. The pattern of bands seen in Fig. 2 was consistently observed in progeny obtained over four generations, suggesting that the gene was integrated into both strains at a single site.

Developmental Regulation of Expression of the Transgene. Immunohistochemical analysis of a range of human tissues has shown PEM to be expressed by several epithelial cell types (see Table 1). In the mammary gland, however, there is a dramatic increase in expression seen at lactation. A similar increase in staining with PEM-specific antibodies was also seen in the lactating mammary glands of the TG4 and TG18 mice but not in the TG5 mice. The increase in expression of PEM could be seen at the RNA level, and Fig. 3*A* shows a Northern blot of RNA extracted from the sources indicated in the legend, probed with pMUC7. Although expression can clearly be detected in the lactating gland of TG4 mice (Fig. 3*A*, *Lane 3*), the level in the nonlactating females (Fig. 3*A*, *Lane 4*) is below that detectable by this technique. A similar result was seen with the TG18 mice (data not shown), but in the TG5 strain, where no positive

Table 1 *Comparison of expression of the MUC1 gene in humans and in transgenic mice detected by immunohistochemical staining*

All neuroectodermal and mesenchymal tissues and cells of the lymphoreticular system showed no reaction with the antibodies.

	HMFG1		HMFG2		SM3	
Tissue	Human[a]	TG4 mice	Human[a]	TG4 mice	Human[b]	TG4 mice
Breast						
Resting breast						
Acini and ducts	+[c]	+	+	+	−	−
Lactating breast						
Acini and ducts	+	+	+	+	−/+	−/+
Female reproductive system						
Ovary						
Surface epithelium, oocytes, follicular epithelium	−	−	−	−	−	−
Interstitial cells	ND	−	ND	−	−	−
Fallopian tubes						
Epithelium	+	+/−	+	+	+/−	+/−
Uterus						
Endometrium and endocervical uterine glands	+	+	+	+	−/+	−
Myometrium	−	−	−	−	−	−
Male reproductive system						
Testis						
Capsule	−	−	−	−	ND	−
Sertoli and Leidig cells	−	−	−	−	−	−
Spermatids	−	−	−	+/−	ND	−
Spermatocytes	−	+	−	+/−	ND	−
Epididymis[d]						
Columnar epithelium with stercocilia	+	+	+	+/−	ND	−
Sperm mass	ND	−	ND	−	ND	−
Interstitial tissue	ND	−	ND	−	ND	−
Urinary system						
Kidney						
Glomeruli and proximal tubules	−	−	−	−	−	−
Distal tubules and collecting ducts	+	+	+	+	+	+
Gastrointestinal tract						
Esophagus						
Squamous epithelium	+	−	+	−	(T)−	−
Stomach						
Mucus-secreting cells	−	−/+	+	−/+	−/+	−
Mucus neck cells	ND	+	ND	−/+	−/+	−
Parietal cells	+	+/−	+/−	+/−	+/−	+
Peptic cells	+	+	−	+	+/−	+
Small intestine						
Enterocytes	−	−	+	−	−	−
Goblet cells	ND	−	ND	−	−	−
Pancriatico-biliary system						
Pancreas						
Acinar cells[e]						
Male	+	+	+	+	−	−
Female	−	−	−	−		
Acinar ducts	+	+	+	+	+	+/−
Excretory ducts	ND	−	ND	+/−	−	−
Islets of Langerhans	−	−	−	−	−	−
Liver						
Hepatocytes	−	−	−	−	−	−
Bile ducts	+	−	+	−	−	−
Respiratory system						
Lung						
Respiratory epithelium	+	+/−	+	+	−/+	+/−
Clara cells	ND	+	ND	+	(+/−)	+/−
Pneumocytes	+	−/+	+	−/+	+/−	−

[a] See Ref. 28. For lactating gland see Ref. 14.
[b] See Refs. 14 and 19.
[c] +, uniform +ve; −, uniform −ve; +/−, focally positive but heterogenous in distribution; −/+, very weak, barely detectable; ND, not done.
[d] For staining of epididymis see Ref. 38.
[e] Refers to intracytoplasmic staining.

Table 1 *Continued*

Tissue	HMFG1 Human[a]	HMFG1 TG4 mice	HMFG2 Human[a]	HMFG2 TG4 mice	SM3 Human[b]	SM3 TG4 mice
Salivary glands						
Parotid gland						
Serous acini	+	+/−	+	+	−	−
Excretory ducts	ND	+	ND	+	+/−	+/−
Intercalated ducts	+	+/−	+	+	−	−
Sublingual gland						
Mucus acini	+	−	+	+	−	−
Excretory ducts	ND	+/−	ND	+	−	−
Intercalated ducts	+	−	+	+	−	+
Myoepithelial cells	−	−	−	−	−	ND
Submandibular gland						
Mucus acini	+	−	+	−	−	−
Serous acini	+	+/−	+	+/−	−	−
Excretory ducts	ND	−	ND	+	+/−	−
Intercalated ducts	+	+	+	+	−	+/−

staining with antibodies was seen even on the lactating gland, expression could not be detected in either the lactating or nonlactating glands (Fig. 3*A*, *Lanes 5* and *6*). Using primer extension analysis, we were able to show that transcription of the human *PEM* gene began from the same start site as seen in humans (see Fig. 4).

To compare the expression of the transgene with the endogenous gene, a probe specific for the mouse gene (pMuc10) was used to detect mRNA transcribed from the mouse *Muc1* gene (25). Fig. 3*B* shows that using the mouse probe, a transcript of 2.3 kilobases was detected in RNA from lactating mammary glands of both transgenic and nontransgenic mice, but not in the human breast cancer cell line T47D or the human fibroblast strain ICRF23. Expression of the mouse mucin was not detectable on Northern blots of the nonlactating mammary gland. Thus, the increased level of expression of the human gene at lactation followed the increased level of expression of the mouse gene seen at this stage of development.

Tissue-specific Expression of the *MUC1* Gene Product in Transgenic Mice. Many of the antibodies developed against human PEM react with core protein epitopes found in the tandem repeat domain between potential glycosylation sites. These epitopes are not present on the mouse mucin because the sequence of the tandem repeat is different, which probably explains why this region is so immunodominant in the rodent in the production of both monoclonal antibodies and polyclonal antisera. To analyze the level of expression of the transgene, therefore, a polyclonal antiserum and the monoclonal antibodies HMFG-1 and HMFG-2 (26), directed to abundant core protein epitopes (13, 27), were used. A similar profile of expression was seen with the polyclonal antiserum and with the monoclonal antibodies HMFG-1 and HMFG-2. Zotter and colleagues (28) have made an extensive survey of the expression of PEM by human tissues, using a bank of monoclonal antibodies, including HMFG-1 and HMFG-2. We have therefore compared our results, obtained from staining the tissues of transgenic mice with these two antibodies, with those of Zotter and colleagues. The detailed comparison for the TG4 strain is listed in Table 1, which shows the results of analyzing five males, five nonlactating females, and five lactating females of the TG4 strain. A similar staining pattern was seen with the TG18 mice. Since in the TG5 mice expression was not seen in the lactating mammary gland (although it was seen in the stomach and kidney), these mice were not analyzed in detail.

The results of staining the various tissues of the TG4 mice with HMFG-1 and HMFG-2 showed that PEM expression was limited to epithelial tissues, and the pattern of expression closely resembled the pattern of expression seen in normal human tissues (Table 1 and Fig. 5). Mucin expression was mainly seen in epithelial cells lining glands or ducts. In most glandular epithelia such as the mammary gland and the salivary gland, PEM was expressed on the luminal surface (Fig. 5*A*, *B*, and *K*). However, in the pancreas of female mice it was difficult to ascertain whether the expression on the luminal aspect of the acinar cell reflected acinar expression or expression by the intercalating duct (see Fig. 5*G*). In the male mice, however, intracytoplasmic staining of the acinar cells was seen. As in humans the tissues showing the strongest staining with HMFG-1 and HMFG-2 were the lactating mammary gland (Fig. 5, *A* and *B*), the proximal and collecting ducts of the kidney, the peptic cells of the stomach (Fig. 5, *H* and *I*), and lung epithelial cells (Fig. 5, *D–F*). The tissues which did not appear to express PEM were also similar in the TG4 mice and humans. These were the small and large intestine (occasional goblet cells in the TG4 mice showed positive staining), ovary, mesenchyme (blood vessels, all muscle types, connective tissue, including fat), the lymphoreticular system, including spleen, lymph nodes, and thymus (epithelioid cells in the medulla were positive), the central nervous system, and the peripheral nervous system.

There were some minor differences in the staining patterns of HMFG-1 and HMFG-2 in the transgenic mice as compared to that reported by Zotter and colleagues for human tissues. Thus we found some staining of the squamous epithelium of the ectocervix in the female transgenic mice and in the spermatocytes and spermatids, which was not reported by Zotter and colleagues for the corresponding human tissue. Moreover, the squamous epithelium of the human esophagus was reported to be stained by HMFG-2, whereas in the mouse there was a clear demarcation between the negative staining of the esophagus and the positive staining of the stomach in the transgenic mouse (Fig. 5*I*). Also, some staining of the human small and large intestine and of the bile ducts was reported by Zotter and colleagues, although this was not seen by Girling and colleagues (14, 19) using the same antibodies (19). These may reflect minor differences in the pattern of expression of the *MUC1* gene in humans and mice. However, minor quantitative differences between the two studies could reflect differences in the fixatives used and the time after death before tissues were processed.

Reaction of the Antibody SM3 with the Human Mucin Expressed on Tissues of TG4 Mice. Glycosylation of PEM in humans can vary with the tissue where it is expressed, thus resulting in differences in the antigenic profile of the final mucin product. In particular, core protein epitopes recognized by the antibody SM3 are exposed where the carbohydrate side chains are short or masked where the side chains are longer and/or branched (12, 13). Since the SM3 epitope is selectively exposed in tumors (14) and may overlap a T-cell epitope (17), its accessibility in the mucin molecules produced by the different tissues is of particular interest. To get some idea as to whether the differences in glycosylation detected by SM3 reactivity which were observed in humans were also seen in the transgenic mice, the reactivity of the antibody SM3 with the different tissues was also examined. Table 1 shows the results of this analysis and for comparison lists the findings of Girling and colleagues (14, 19), who examined the expression of the SM-3 epitope on normal human tissues. Again, the correlation with SM3 staining of human tissues was good, and in particular the

lack of staining of the lactating mammary gland (Fig. 5*C*) was very striking, since this tissue produces high levels of PEM. As with the human tissues, the most consistent expression of the SM3 epitope was seen in the lung (Fig. 5*F*), the kidney, the stomach, and the fallopian tube. Other epithelial tissues showed weak or focal staining with the antibody (see Table 1). The weak focal staining seen in the intercalating ducts of the salivary gland of the transgenic mouse was not reported by Girling and colleagues for the corresponding human tissue.

DISCUSSION

In this report, we describe the development of transgenic mouse strains which express the human *MUC1* gene product PEM in a tissue-specific manner. The pattern of expression in the tissues of the adult transgenic mice shows remarkably good agreement with the pattern of expression of the gene seen in human tissues. The comparison was made on the basis of immunohistochemical data using antibodies to the core protein epitopes recognized by HMFG-1 and HMFG-2. The immunohistochemical analysis also showed that the human gene was found to be dramatically up-regulated in the lactating mammary gland, and the increased expression could be detected for both the mouse and the human gene at the RNA level.

There are minor differences in the tissue specificity of expression seen in the transgenic mice and in humans. These could reflect differences in transcription factors present and active in tissues in question. A complete survey of the tissue specificity of expression of the mouse homologue could help shed light on this possibility, and such a study is under way. It is also likely that where expression is seen in the mouse and not in the corresponding human tissue, the more rapid time of processing which was possible with the mouse tissue could be an important factor. Overall, however, the expression of the *MUC1* gene in the transgenic mouse parallels very closely the expression of the gene seen in the human tissues. Of particular interest is the observation that PEM is expressed cytoplasmically in the acinar cells of the pancreas of the male transgenics but only on the luminal aspect of these cells in the female transgenics. A more detailed analysis of the expression of *MUC1* in the human pancreas is warranted to see if expression in this tissue could be influenced by the hormonal environment.

Although the *MUC1* gene product is expressed in most glandular epithelia, the glycosylation of the core protein can vary widely in different tissues, resulting in differences in the antigenic profile of the final mucin product. The core protein epitope (pro-asp-thr-arg-pro) recognized by the antibody SM-3 lies between potential glycosylation sites within the tandem repeat sequence and is selectively exposed when the carbohydrate side chains are short (13). Although the epitope is abundantly expressed in breast and other carcinomas, it shows a more restricted distribution on normal tissues than other core protein epitopes, such as those recognized by antibodies HMFG-1 and HMFG-2 (14). It was important therefore to ascertain whether the distribution of the SM-3 epitope on the transgenic mouse tissues showed the same tissue distribution as in humans. Immunohistochemical analysis of the transgenic mouse tissues showed that the profile of SM-3 reactivity did indeed resemble that seen by Girling and colleagues in humans. This suggests that glycosylation patterns in the mouse tissues resemble those seen in the human and that the tissue specificity of expression applies not merely to the expression of the core protein but to the antigenic profile of the final mucin product. It would seem, therefore, that the transgenic mouse expressing the human *MUC1* gene will indeed provide an appropriate preclinical model for assessing the toxicity of conjugates of antibodies directed to PEM and for investigating the immune response to antigens based on PEM.

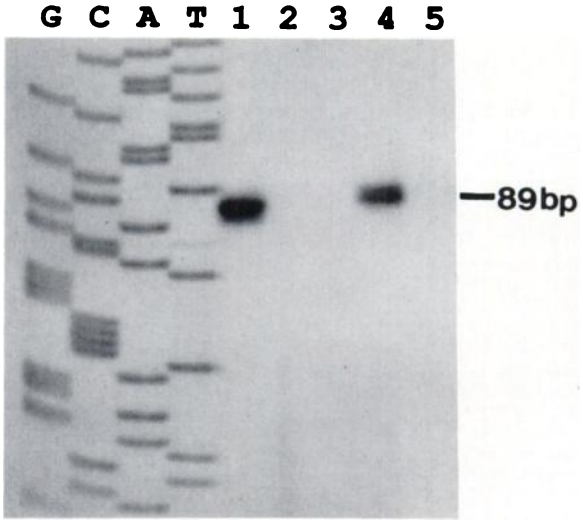

Fig. 4. Transcriptional start site of human *PEM* gene in lactating mammary gland of TG4 mouse, mapped by primer extension. A synthetic oligonucleotide (21 mer) complementary to nucleotides +73 to +93 of the human *PEM* gene was end labeled and used as primer with reverse transcriptase and total RNA from T47D cells (*Lane 1*), lactating mammary gland of TG5 mice (*Lane 2*), lactating mammary gland of nontransgenic mouse (*Lane 3*), lactating mammary gland of TG4 mouse (*Lane 4*), and ICRF-23 cells (*Lane 5*). The products were separated on a 6% sequencing gel, showing the primary extension product to be 89 base pairs, indicating that the transcriptional start site is 68 base pairs 5′ to the ATG translational start site. (Due to a numerical error this was reported to be 72 base pairs in Ref. 2.)

Fig. 3. Northern blot of total RNA from transgenic and nontransgenic mammary gland probed with pMUC7 specific for the human PEM (*A*) or pMuc10 specific for the mouse PEM (*B*). Fifteen μg of total RNA from the cells or tissues indicated were subject to gel electrophoresis and blotted onto Biodyne membranes (see "Materials and Methods") before probing with pMUC7 or pMUC10. *Lanes 1*, human breast cancer cell line T47D; *Lanes 2*, human fibroblast strain, ICRF-23; *Lanes 3*, TG4 lactating mammary gland; *Lanes 4*, TG4, nonlactating mammary gland; *Lanes 5*, TG5, nonlactating mammary gland; *Lanes 6*, TG5, lactating mammary gland; *Lanes 7*, nontransgenic mouse lactating mammary gland.

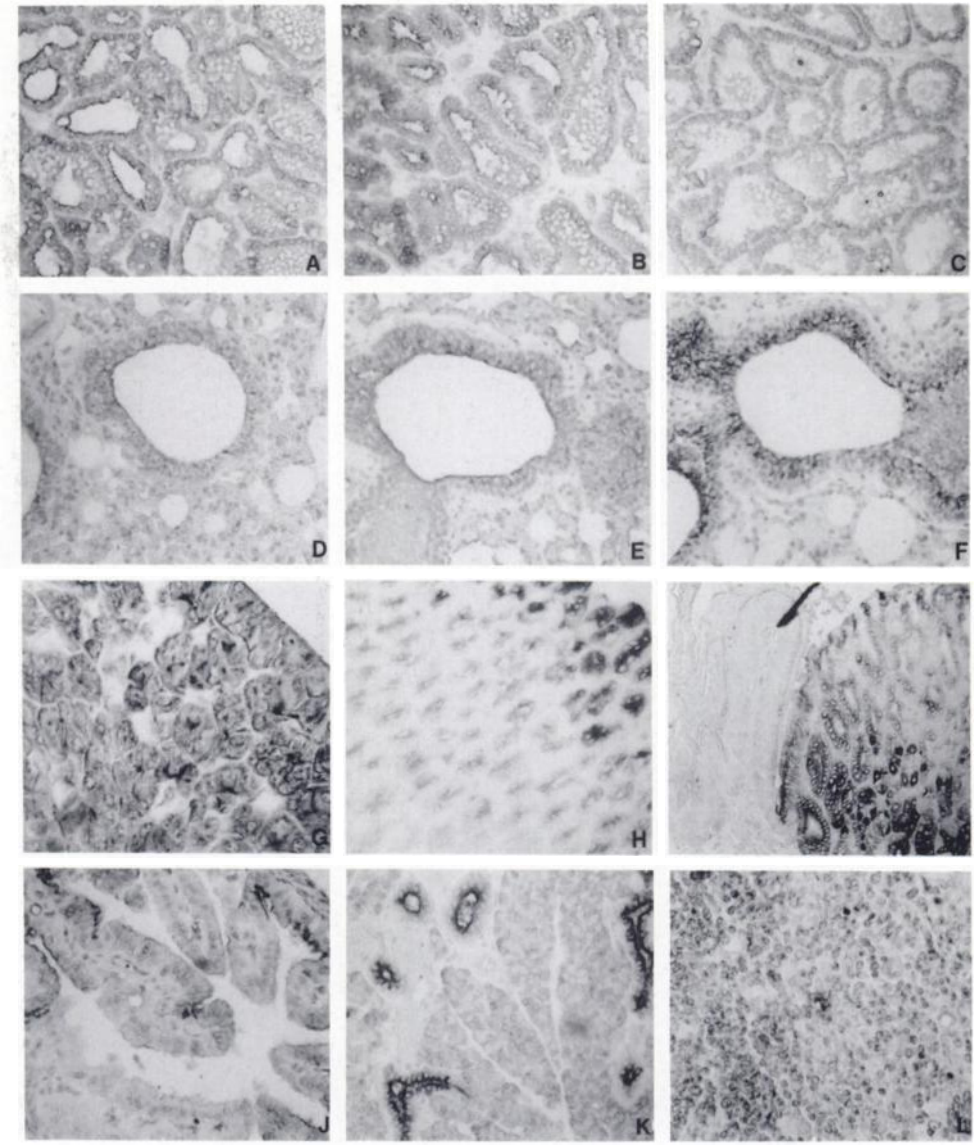

Fig. 5. Immunohistochemical staining of tissues from the transgenic mouse strain TG4 using antibodies HMFG-1 (*A, D, G, J*), HMFG-2 (*B, E, H, I, K, L*), and SM3 (*C, F*). *A–C*, lactating mammary gland; *D–F*, lung; *G*, pancreas of a female TG4 mouse; *H*, stomach; *I*, stomach/esophageal junction; *J*, fallopian tube; *K*, salivary gland; *L*, thymus.

In addition to establishing their potential as appropriate recipients for antibodies and antigens, the tissue-specific expression seen in the transgenic mice described here suggests that the regulatory sequences required for tissue-specific expression of PEM are close to the coding sequences. The TG4 mouse was developed using a 40-kilobase genomic fragment, but the TG18 mouse, which showed a similar tissue specificity in expression, contained only a 10.6-kilobase fragment, 1.6 kilobases of which were 5′ flanking sequences. Although it is possible that sequences within the introns or 3′ to the gene may be involved in the regulation of tissue-specific expression, there are many cases where tissue-specific expression of a reporter transgene has been obtained with only a limited amount of 5′ flanking sequence (29–37). We have been able to show tissue-specific expression of a reporter gene in tissue culture cells using less than 1.6 kilobases of 5′ flanking sequence of the *PEM* gene,[3] and we are now developing transgenic mice with these constructs. Should it be possible to obtain tissue-specific expression from a defined 5′ flanking sequence, this sequence could be used to direct expression of an oncogene in a double transgenic derived from the TG4 or TG18 strain. Such mice, which would be expected to develop tumors in PEM-expressing tissues, would provide an excellent model for examining the use of PEM-based antibodies and antigens in cancer therapy.

[3] D. Wilson, personal communication.

REFERENCES

1. Burchell, J., and Taylor-Papadimitriou, J. Antibodies to human milk fat globule molecules. Cancer Invest., *17:* 53–61, 1989.
2. Gendler, S., Lancaster, C., Taylor-Papadimitriou, J., Duhig, T., Peat, N., Burchell, J., Pemberton, L., Lalani, E-N., and Wilson, D. Molecular cloning and expression of human tumor-associated polymorphic epithelial mucin. J. Biol. Chem., *265:* 15286–15293, 1990.
3. Wreschner, D. H., Hareuveni, M., Tsarfaty, I., *et al.* Human epithelial tumor antigen cDNA sequences. Eur. J. Biochem., *189:* 463–473, 1990.
4. Ligtenberg, M., Vos, H., Gennissen, A., and Hilkens, J. Episialin, a carcinoma-associated mucin, is generated by a polymorphic gene encoding splice variants with alternative amino termini. J. Biol. Chem., *265:* 5573–5578, 1990.
5. Gendler, S. J., Burchell, J. M., Duhig, T., Lamport, D., White, R., Parker, M., and Taylor-Papadimitriou, J. Cloning of partial cDNA encoding differentiation and tumor-associated mucin glycoproteins expressed by human mammary epithelium. Proc. Natl. Acad. Sci. USA, *84:* 6060–6064, 1987.
6. Taylor-Papadimitriou, J., and Gendler, S. J. Molecular aspects of mucins. Cancer Rev., *11–12:* 1–24, 1988.
7. Gum, J. R., Byrd, J., Hicks, J., Toribara, N., Lamport, D., and Kim, Y. Molecular cloning of human intestinal mucin cDNAs. J. Biol. Chem., *264:* 6480–6487, 1989.
8. Gum, J. R., Hicks, J., Swallow, D., *et al.* Molecular cloning of cDNAs

derived from a novel human intestinal mucin gene. Biochem. Biophys. Res. Commun., *171:* 407–415, 1990.
9. Porchet, N., van Cong, N., Dufosse, J., Audie, J. P., Guyonnet-Duperat, V., Gross, M. S., Denis, C., Degand, P., Bernheim, A., and Aubert, J. P. Molecular cloning and chromosomal localizatin of a novel human tracheobronchial mucin cDNA containing tandemly repeated sequences of 48 base pairs. Biochem. Biophys. Res. Commun., *175:* 414–422, 1991.
10. Hanisch, F-G., Uhlenbruck, G., Peter-Katalinic, J., Egge, H., Dabrowski, J., and Dabrowski, U. Structures of neutral *O*-linked polylactosaminoglycans on human skim milk mucins. A novel type of linearly extended poly-*N*-acetyl-lactosamine backbones with Gal 3b(1–4) GlcNAc 3b(1–6) repeating units. J. Biol. Chem., *264:* 872–883, 1989.
11. Hull, S., Bright, A., Carraway, K., Abe, M., Hayes, D., and Kufe, D. Oligosaccharide differences in the DF3 sialomucin antigen from normal human milk and the BT20 human breast carcinoma cell line. Cancer Commun., *1:* 261–267, 1989.
12. Burchell, J., Gendler, S., Taylor-Papadimitriou, J., Girling, A., Lewis, A., Millis, R., and Lamport, D. Development and characterization of breast cancer reactive monoclonal antibodies directed to the core protein of the human milk mucin. Cancer Res., *47:* 5476–5482, 1987.
13. Burchell, J., Taylor-Papadimitriou, J., Boshell, M., Gendler, S., and Duhig, T. A short sequence, within the amino acid tandem repeat of a cancer-associated mucin, contains immunodominant epitopes. Int. J. Cancer, *44:* 691–696, 1989.
14. Girling, A., Bartkova, J., Burchell, J., Gendler, S., Gillett, C., and Taylor-Papadimitriou, J. A core protein epitope of the polymorphic epithelial mucin detected by the monoclonal antibody SM-3 is selectively exposed in a range of primary carcinomas. Int. J. Cancer, *43:* 1072–1076, 1989.
15. Granowska, M., Mather, S. J., Jobling, T., Naeem, M., Burchell, J., Taylor-Papadimitriou, J., Shepherd, J., and Britton, K. E. Radiolabelled stripped mucin, SM-3, monoclonal antibody for immunoscintigraphy of ovarian tumours. Int. J. Biol. Markers, *5:* 89–96, 1990.
16. Barnd, D., Lan, M., Metzgar, R., and Finn, O. Specific, MHC-unresticted recognition of tumor-associated mucins by human cytotoxic T cells. Proc. Natl. Acad. Sci. USA, *86:* 7159–7163, 1989.
17. Jerome, K. R., Barnd, D. L., Bendt, K. M., Boyer, C. M., Taylor-Papadimitriou, J., McKenzie, I. F. C., Bast, R. C., Jr., and Finn, O. J. Cytotoxic T-lymphocytes derived from patients with breast adenocarcinoma recognize an epitope present on the protein core of a mucin molecular preferentially expressed by malignant cells. Cancer Res., *51:* 2908–2916, 1991.
18. Lalani, E-N., Berdichevsky, F., Boshell, M., Shearer, M., Wilson, D., Stauss, H., Gendler, S. J., and Taylor-Papadimitriou, J. Expression of the gene coding for a human mucin in mouse mammary tumor cells can affect their tumorigenicity. J. Biol. Chem., *266:* 15420–15426, 1991.
19. Girling, A. Evaluation of the monoclonal antibody SM3, spectrum of reactivity in the breast and at other sites and value as a clinical imaging agent. Ph.D. thesis, London University, 1990.
20. Lancaster, C., Peat, N., Duhig, T., Wilson, D., Taylor-Papadimitriou, J., and Gendler, S. Structure and expression of the human polymorphic epithelial mucin gene: an expressed VNTR unit. Biochem. Biophys. Res. Commun., *173:* 1019–1029, 1990.
21. Hogan, B., Costantini, F., and Lacy, E. Manipulating the Mouse Embryo. A Laboratory Manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory, 1986.
22. Keydar, I., Chen, L., Karby, S., Weiss, F. R., Delanei, J., Radu, M., Chaicik, M., and Brenner, M. J. Establishment and characterization of a cell line of human breast carcinoma origin. Eur. J. Cancer, *15:* 659–670, 1979.
23. Feinberg, A. P., and Vogelstein, B. A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem., *137:* 266–267, 1984.
24. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., and Rutter, W. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry, *18:* 5294–5299, 1979.
25. Spicer, A. P., Parry, G., Patton, S., and Gendler, S. J. Molecular cloning and analysis of the mouse homologue of the tumor-associated mucin, MUC1, reveals conservation of protential *O*-glycosylation sites, transmembrane and cytoplasmic domains and a loss of minisatellite-like polymorphism. J. Biol. Chem., *266:* 15099–15109, 1991.
26. Burchell, J., Durbin, H., and Taylor-Papadimitriou, J. Complexity of expression of antigenic determinants recognized by monoclonal antibodies HMFG-1 and HMFG-2, in normal and malignant human mammary epithelial cells. J. Immunol., *131:* 508–513, 1983.
27. Gendler, S., Taylor-Papadimitriou, J., Duhig, T., Rothbard, J., and Burchell, J. A highly immunogenic region of a human polymorphic epithelial mucin expressed by carcinomas is made up of tandem repeats. J. Biol. Chem., *263:* 12820–12823, 1988.
28. Zotter, S., Hageman, P. C., Lossnitzer, A., Mooi, W. J., and Hilgers, J. Tissue and tumor distribution of human polymorphic epithelial mucin. Cancer Rev., *11–12:* 55–101, 1988.
29. Chamberlain, J. W., Nolan, J. A., Conrad, P. J., Vasavada, H. A., Vasavada, H. H., Ganguly, S., Janeway, C. A., Jr., and Weissman, S. M. Tissue-specific and cell surface expression of human major histocompatibility complex class I heaby (HLA-B7) and light (β_2-microglobulin) chain genes in transgenic mice. Proc. Natl. Acad. Sci. USA, *85:* 7690–7694.
30. Idzerda, R. L., Behringer, R. R., Theisen, M., Huggenvik, J. I., McKnight, G. S., and Brinster, R. L. Expression from the transferrin gene promoter in transgenic mice. Mol. Cell. Biol., *9:* 5154–5162, 1989.
31. Robinson, M. O., McCarrey, J. R., and Simon, M. I. Transcriptional regulatory regions of testis-specific PGK2 defined in transgenic mice. Proc. Natl. Acad. Sci. USA, *86:* 8437–8441, 1989.
32. Yamada, K., Noguchi, T., Miyazaki, J., Matsuda, T., Takenaka, M., Yamamura, K., and Tanaka, T. Tissue-specific expression of rat pyruvate kinase L/chloramphenicol acetyltransferase fusion gene in transgenic mice and its regulation by diet and insulin. Biochem. Biophys. Res. Commun., *171:* 243–249.
33. Ledent, C., Parmentier, M., and Vassart, G. Tissue-specific expression and methylation of a thyroglobulin-chloramphenicol acetyltransferase fusion gene in transgenic mice. Proc. Natl. Acad. Sci. USA, *87:* 6176–6180, 1990.
34. Sigmund, C. D., Jones, C. A., Fabian, J. R., Mullins, J. J., and Gross, K. W. Tissue and cell specific expression of a renin promoter-reporter gene construct in transgenic mice. Biochem. Biophys. Res. Commun., *170:* 344–350, 1989.
35. McNab, A. R., Andrus, P., Wagner, T. E., Buhl, A. E., Waldon, D. J., Kawabe, T. T., Rea, T. J., Groppi, V., and Vogeli, G. Hair-specific expression of chloramphenicol acetyltransferase in transgenic mice under the control of an ultra-high-sulfur keratin promoter. Proc. Natl. Acad. Sci. USA, *87:* 6848–6852, 1990.
36. Liou, G. I., Geng, L., Al-Ubaidi, M. R., Matragoon, S., Hanten, G., Baehrt, W., and Overbeek, P. A. Tissue-specific expression in transgenic mice directed by the 5′-flanking sequences of the human gene encoding interphotoreceptor retinoid-binding protein. J. Biol. Chem., *265:* 8373–8376, 1990.
37. Lem, J., Applebury, M. L., Falk, J. D., Flannery, J. G., and Simon, M. I. Tissue specific and developmental regulation of rod opsin chimeric genes in transgenic mice. Neuron, *6:* 201–210, 1991.
38. Arklie, J., Taylor-Papadimitriou, J., Bodmer, W. F., Egan, M., and Millis, R. Differentiation antigens expressed by epithelial cells in the lactating breast are also detectable in breast cancers. Int. J. Cancer, *28:* 23–29, 1981.