Glycoconjugate Journal (1996) **13:** 797–807

Promoter of the canine tracheobronchial mucin gene

MUKESH VERMA[1], VANDAVALLI V. S. MURTHY[2], SUSAN MATHEW[2], DEBABRATA BANERJI[2], RABINDER N. KURL[3], MATTHEW J. OLNES[3], JAMES R. YANKASKAS[4], CLAUDIA BLASS[1] and EUGENE A. DAVIDSON[1]*

[1]*Department of Biochemistry and Molecular Biology, Basic Science Building, Georgetown University Medical Center, 3900 Reservoir Road, NW, Washington, DC 20007, USA*
[2]*Memorial Sloan-Kettering Cancer Center, New York 10021, USA*
[3]*Graduate program in Pathobiology, Brown University, Providence, RI 02903-4923, USA*
[4]*Department of Medicine, University of North Carolina at Chappel Hill, NC 27599-7020, USA*

Received 2 October 1995; revised 2 January 1996

The mucin gene is up-regulated in diseases such as cystic fibrosis (CF) and asthma. To understand the mechanisms involved in transcriptional regulation of mucin gene expression we have characterized the region of the mucin gene up-stream of the transcriptional start site and analysed the *cis*-acting elements required for mucin promoter activity. We isolated clones from a dog genomic library containing the promoter region for the tracheobronchial mucin gene (*TBM*). The authenticity of the promoter was tested by nucleotide sequencing, primer extension analysis, electrophoretic mobility shift assay (EMSA) and reporter gene expression analysis. The canine *TBM* promoter is different from housekeeping gene promoters (as it is not rich in GC content and contains TATA- and CAAT-like sequences) and different from that of regulatory genes (because it contains many TATA- and CAAT-like sequences and multiple transcriptional initiation sites). Reporter gene analysis using canine *TBM* promoter-chloramphenicol acetyltransferase (CAT) fusion plasmids established the regions responsible for promoter activity and verified the positions of the major mucin transcriptional initiation sites. Reporter gene analysis also established that a region of the canine *TBM* promoter and first exon containing all of the transcriptional initiation sites is more active in mucin expressing cells (e.g. CT1 cells-immortalized canine tracheal epithelial cells, human CFT1 cells-immortalized tracheal epithelial cells from a CF subject, or HBE1 cells-immortalized tracheal epithelial cells from non-CF subject) than in mucin non-expressing cells (COS7, 3T3), suggesting cell specificity. The promoter region contained cAMP response element (CRE) sequences, and the *TBM* gene transcription was enhanced when cAMP analogs were added to transfected cells. EMSA indicated the presence of at least two DNA binding proteins in CT1 cells. This is the first report describing the characterization of a *TBM* gene promoter. The information obtained in the present studies will be valuable in understanding mucin gene regulation in normal and pathological conditions.

Keywords: cyclic AMP response element, cystic fibrosis, epithelial cells, glycoprotein, immortalization, mucin, papilloma virus vectors, promoter, transcription, transcription factors

Abbreviations: CF, cystic fibrosis; CRE, cAMP response element; CREB, cAMP response element binding protein; CT, cholera toxin; CTM, canine tracheal mucin; EGF, epidermal growth factor; EMSA, electrophoretic mobility shift assay; GRE, glucocorticoid response element; INS, insulin; PBS, phosphate buffered saline; RA, retinoic acid; RT-PCR, reverse transcriptase polymerase chain reaction; TF, transferrin; TRE, thyroid response element; VDRE, vitamin D response element.

Accession number: L40433.

Introduction

Mucin is a class of glycoproteins synthesized and secreted by epithelial goblet and mucous gland cells of the respiratory and gastrointestinal tracts. Abnormalities in mucin production contribute to morbidity and mortality in diseases of both organ systems. Efforts to understand the pathophysiology of these diseases have led to the development of animal models and the analysis of such models has suggested that the tracheobronchial mucin gene (members of *MUC5* gene family) is up-regulated in

*To whom correspondence should be addressed. Phone (202)-687-1501; Fax (202)-687-7186.

disease processes such as CF and asthma [1–3]. Our laboratory has focused considerable attention upon mucin of the airways [4–6]. Previously, we cloned and sequenced the cDNA of the canine *TBM* gene [7–10] and characterized its gene product [5]. The promoter of the *TBM* gene has not been characterized. Most models of gene regulation implicate regions at the 5′ end of the gene. For this reason, we have isolated canine genomic clones coding for the amino terminal portion of the canine *TBM*. In this report we describe the characterization of genomic clones representing the promoter region of the canine *TBM* gene.

Promoter regions of *MUC1* (breast cancer mucin gene) and *MUC2* (intestinal mucin gene) have been characterized to some extent [11, 12]. In the case of *MUC1*, only 750 bp of 5′ sequence are required to obtain epithelial specific expression, and positive and negative regulatory elements are found in the immediate proximal sequences which can act with up-stream regulatory elements. Shirotani *et al.* [13] found a responsive mucin element (RME) in the *MUC1* promoter located between –531 and –520 which binds to a nuclear protein isolated from colon cancer cells. This sequence does not show any homology with previously known *cis*-elements. Southwestern analysis identified this nuclear protein to be of 70 kDa. Such studies have not been performed with the *TBM* gene.

Mucin gene expression can be induced by certain secretagogues [8, 10]. Results of our studies on the expression of canine *TBM* mRNA by nuclear run-on and RT-PCR analysis suggest that this induction is regulated at the transcriptional level (Verma *et al.*, unpublished). It was, therefore, intriguing to characterize the canine *TBM* promoter and study regulation of expression of the mucin gene at the transcription level.

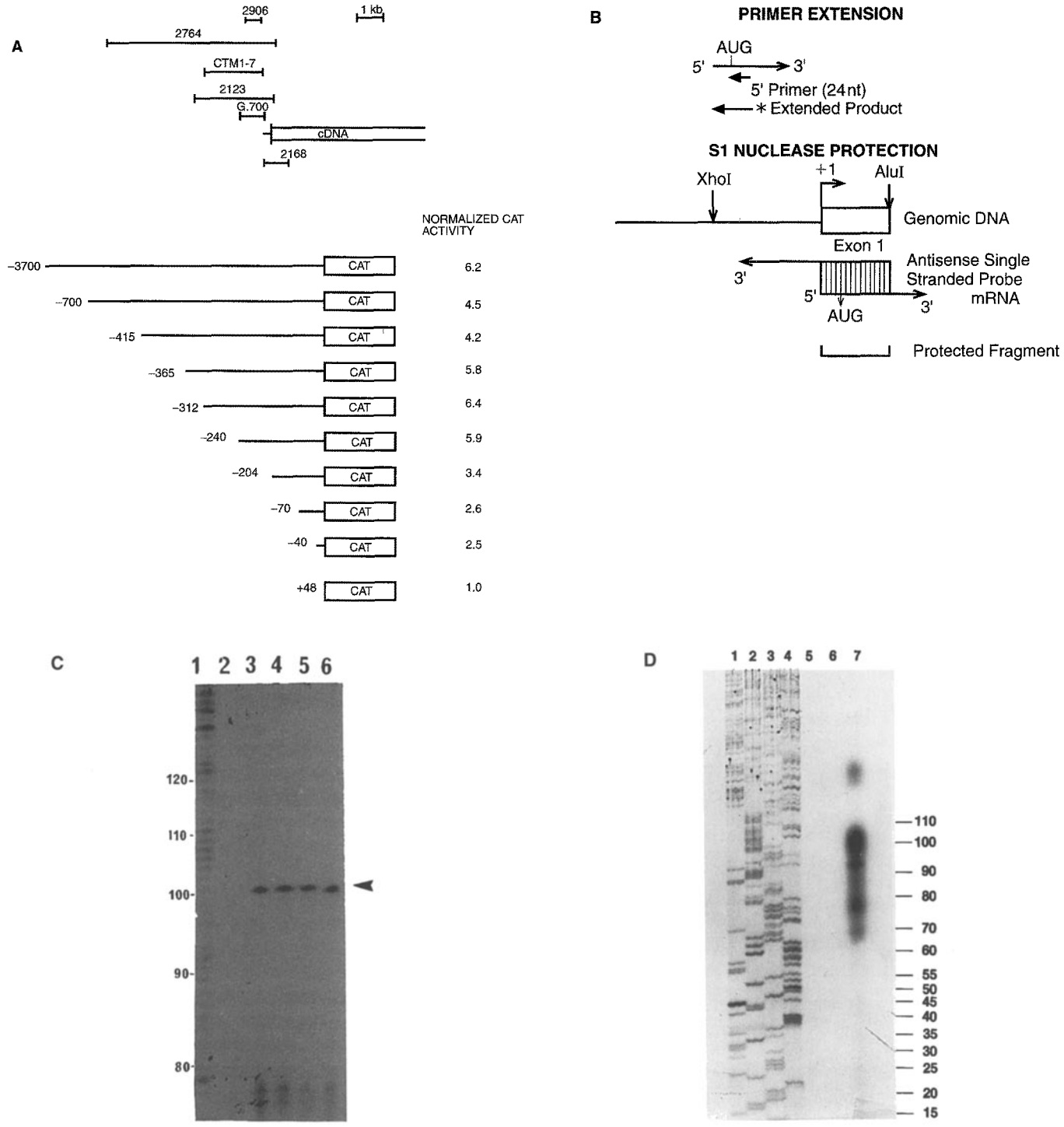
Materials and methods

The HPV18 vector (CCB/18) containing the *E6* and *E7* genes was provided by Dr Richard Schlegel (Georgetown University, Washington, DC). Oligonucleotides were purchased from Oligos etc, (Cleveland, OH) and Midland, Inc. (Midland, TX). RT-PCR assay system and other PCR-related reagents were from Perkin Elmer Cetus, Inc. (Branchburg, NJ). Media for culture, Rad-Prime labeling kit, mRNA isolation kit, molecular size markers (1 kb ladder, 123 bp ladder, RNA ladder, and lambda DNA digested with *Hind* III), double stranded oligonucleotides used for EMSA and restriction enzymes were obtained from GIBCO/BRL Life Technologies, (Gaithersburg, MD). Retinoic acid (RA), hydrocortisone (HC), transferrin (TF), insulin (bovine) (INS), cholera toxin (CT), epidermal growth factor (EGF), trypsin-EDTA-saline (100×), penicillin, streptomycin, and neomycin (100×), and protease type XVI were purchased from Sigma Chemical Co., (St Louis, MO). Cellagen and Gene Screen nylon membrane discs were obtained from ICN Biochemicals (Cleveland, OH). Radioisotopes were purchased from Amersham and Dupont NEN (Boston, MA). All other reagents were ultrapure (molecular biology grade) and obtained from Sigma, Gibco or Fisher. Centricon 30 tubes were purchased from Amicon (Beverly, MA). Antibodies directed against AP1 and CREB were the products of Oncogene Science (Uniondale, NY). Poly (dldC) was obtained from Boehringer Mannheim Biochemicals (Indianapolis, IN). Texas Red conjugated, anti mouse IgG (H + L) was obtained from Vector Laboratories, (Burlingame, CA). Control CAT constructs and dog genomic library were purchased from Stratagene (San Diego, CA). For making nested deletions Erase-a-base system from Promega was used. Constructs containing SV40 promoter, SV40 enhancer, or both were the products of Promega. CF (CFT1, CFT1–4, homozygous for phenylalanine deletion at 508) and non-CF (HBE1, clone 9) cell lines were generated with HPV 18 *E6* and *E7* genes as described [14]. CT1 and CT3 cells were isolated in our laboratory [10, 15].

Isolation of genomic clones containing the promoter region

A canine genomic library was screened using, as probe, fragments from clone GMV2168 (with 72 nt in the 5′ untranslated region and 500 nt in the coding region of the cDNA, see [7]) located at the 5′ end of the canine *TBM* cDNA followed by Southern hybridization. One clone GMV2764 (with about a 6 kb insert, spanning the promoter region and part of the 5′ end of the cDNA) and a few other clones with smaller inserts were isolated (Fig. 1). Further restriction analysis, PCR, and hybridiza-

Figure 1. (A) Nested-deletion analysis of the promoter of the canine *TBM* gene and relative CAT activity. The location of some of the genomic clones is shown on the top part of the figure. DNA fragments from the promoter region were subcloned into the pSVoCAT plasmid [27, 36]. These constructs were transfected into CT1 and CAT activity was determined 48 h later. The transfection efficiency was determined by co-transfection with beta galactosidase gene containing construct [22]. All the constructs were made by deletion of the clone GMV2764 (with about 6 kb insert which covers the promoter region and about 500 nt in the 5′ end region of the canine *TBM* cDNA). The numbers represent the length (bp) of the insert isolated from GMV2764. The CAT activity results are representative of three independent experiments. (B) The schematic diagram to perform primer extension analysis and S1 mapping. The details are described in the text. (C) S1 nuclease mapping. One genomic fragment (0.23 kbp long) located upstream of the 5′ end of the cDNA was isolated from clone GMV 2764, labeled at the *Alu* I site (10^6 dpm), was restricted with *Xho* I and precipitated with 10 μg of tRNA or poly(A^+) RNA from CT1 cells; dissolved in 80% formamide, 0.4 M NaCl, 40 mM PIPES, pH 6.5, 1 mM EDTA; denatured at 90 °C for 10 min; and hybridized at 52 °C.



After 3 h, the sample was diluted 10-fold in ice cold S1 buffer (0.25 M NaCl, 30 mM potassium acetate, pH 6.5, 1 mM $ZnCl_2$, 500 U S1 nuclease). After incubation at 37 °C for 30 min, the digestion was stopped by phenol extraction, and the DNA fragments were precipitated. The product was analysed on a 6% sequencing gel (the size of the protected fragment was about 100 bp). 1, Molecular size standard; 2, hybridization done in the presence of tRNA; 3, hybridization done in the presence of CT1 mRNA; 4, hybridization done in the presence of HBE1 mRNA; 5, hybridization done in the presence of CFT1 mRNA; 6, hybridization done in the presence of mRNA isolated from dog tracheal epithelial cells (primary cells). (D) Primer extension analysis. RNA was isolated according to Bailey and Verma [37] and primer extension analysis was performed as described [16, 21]. The potential transcription initiation sites are shown in the figure. The major transcription initiation site is numbered as +1 (see Fig. 2). Lanes 1–4 are sequencing reaction for G, A, T, C respectively (the ladder does not belong to the canine *TBM* template and serves as a size marker only); lane 5, primer extension using tRNA (control), lane 6, primer extension without any RNA (control); lane 7, primer extension using tracheal epithelial cell mRNA.

tion analysis were used to identify clones CTM1–7 (2.1 kb insert, located at the 5′ end of the cDNA which also covers the promoter region), GMV23 (2.9 kb insert, located at the 5′ end of the cDNA and which contains the promoter region), GMV22 (2.8 kb insert, located at the 5′ end of the cDNA, it contains the promoter region and is similar to GMV23 except that it lacks about 100 bases at the 5′ end), CTM 1–7 (2.1 kb insert, located at the 5′ end of the cDNA and contains the promoter region, see Fig. 1), GMV2906 (0.5 kb insert, located in the promoter region at –519/+40, numbered according to Fig. 2), GMV23–700 (subclone of GMV23, –649/+51), GMV23–400 (subclone of GMV23, –351/+51), GMV-300 (subclone of GMV23, –247/+51), GMV23–200 (subclone of GMV23, –146/+51), GMV23–150 (subclone of GMV23, –97/+51), GMV23–70 (subclone of GMV23, –18/+51), GMV22–60 (subclone of GMV22, –9/+51), GMV23–45 (subclone of GMV23, +2/+51), GMV23–40 (subclone of GMV23, +7/+51) spanning the 5′ flanking region of the canine *TBM* gene. Subclones of the above clones were made in the pBSKII$^+$ plasmid (for nucleotide sequencing and in determining the transcription initiation site of the canine *TBM* gene) and in pSVoCAT (for CAT assay) [16]. The junctions of the constructed subclones were sequenced to verify the orientation of the insert. Nested deletions of the promoter region were made using Exonuclease III and mung bean

```
                HSE-like Sequence
-1517  AAATTGCCATTTGAAATCTTTATTAGTTAGCAAAAAAATCTAACATAATCATCAGAGTTTTAAGTAG
                        p53 Binding Site
-1450  TAGTGACAGAAGATAAAAATGCCTCCATATCCCAAATTAACATCTCCGTATACCAAGTAGCTAACCT
         HNF3
-1383  TTATTGTTTTACAACTAAAGTCAGGGTTTTTATTTTCAGAGACAGATTCTGCAAACTTGACATTAAG
                                                              AP4
-1316  CGAGGTTGGCATGGTTAGATGTGGAGCTAGTGTAGTTGCAGAAGGTGTGGCAGCACAGCTGTGGTCT
                  AP5
-1249  CTGGTGGAGCTGAGGAAGACAGTGGAGCTAGTGTAGTTGCAGAAGGTGTGGTCTCTGGTGGAGCTGA

-1182  GGAAGATGGTGGAGCTCCTGTAGTTGCAGAAGGTGTGGGTGGAGCCAGCTGTGGTCTCTGATATACT
         fos BLE-1 Site                                  Elk1 Site
-1115  ATTGCGCCACCATGTCGTCAATAATTCTGTTTAGTAGTTCATATATATAAGAAATCAGGAAATTTTA

-1048  TTTTGGCCAGGAGCTGAAGTTGGTTGTTTAGTAGTAGTAGTTTGAGTGGTGACCGAAGTAGTAGTCT

-981   GGTAGTGTGAGTGGCAGCTGAAGTTTCAGAAGTGTCAGGTGCAAGAGTAGTTGGGGAAGAATTAGGT
       AP3          TCE
-914   GTGGGAATTGGGCAGCTGCGTGGGATGGAGGCTAGTGTAGTTGGCAGAAGGTATAAATAAACCCATT
         E2aEcb Site                                              PEA3
-847   GTCTCACATCCTTTTTTAAGAACAGTTTGTAACATGGGCTTATTAAACCCTGGGAACCGCAAGTTGG

-780   GCAAATAAAGGTCAACATAATAACTCATACCATAAGTTTTGCAGAATATGTCTATCAGTCAGCAACC
                                        TRE/VDRE
-713   TCTGAGAACTACCAAGGCTCTGTCTAGGTCACATAGGTCATGACTAGCCCACCCACACCTCCTGCAA
                Delta Site                          CACCC Factor
-646   CTACACCAGCTCCACCATCTTCCTAGCTCCACCAGAGACCACAGCTGCCCCACCCACACTCTGCAAC

-579   TACACCAGCTCACATTTCTAGTCCACAGCCACAGTGCAATAAATCAAAGAGTTATTTTGAACCCCAG
         ANF Site                                          ATBF1
-512   ACCTGCATCTGGATTGTGGGATAGAGGCATTTTTCTTGTTAATAGGTCTGTTGTGGTGAATAATTAG

-445   ATTGAAAGATTAATGAAGTGCTTAATCGGTCAGGCGTTAGTCCGCGTTCTTACCACTTATTTCCGGC
         Oct AP1   CRE
-378   CTATTTGAGTCATATTAAAAAGAAATGCCAGAAATTTTCCGGCATTATAGGTCAGACTTGCCAGACA
           CRE   c-Myc Binding Site
-311   ATATCCATGCGATCACGTTGACTGACTTTACGGAGTTTACAAGAAATGCTACGGTAACCTATATAGT

-244   TGCCACCTATGCGTCACTAAAAAAAGAGGTAAATCGAAGGAATCGAGGACTTTATAGGTACATTGAG

-177   AACTGAGCTGGAAATGCCCTGAGCTGGAAATGCCGAAAATGTTCTAGTTACGATGCCATTGGATATA
                                                                        ▼
-110   TCAAGGTGGTATATCCAGTGATTTTTTCTCCATTTTAGCTTCCTTAGCTCCTGAAAATCTCTGCCAA
                                          POB Site +1       ▼
-43    GCTCCTCGAGAAACCACTTGTCTCACATCCTTGTTTTAAGAACAGTTTGTAACCATAAACTTATTTA
       ▼                ▼                                          TIS
25     AAACCCTGGGAACGCAAGTGGGCAAACAAATAAAGAACCCCGGAACGGGTGCGGGGTGTTCAAGAAT

92     GTCCATT
```

Figure 2. Nucleotide sequence of the promoter region of the canine *TBM* gene. DNA fragments from clone 2168 (Fig. 1A, located towards the 5′ end of the gene, from 1 to 1258 nt in [7]) were chosen as probes to screen the dog genomic library. Genomic clones containing the promoter region were sequenced. The major transcription initiation site is marked as +1 and other numbers represent the location of the bases with respect to the major transcription initiation site. Other transcription initiation sites are marked as arrow heads. The translation initiation site ATG is shown as TIS. Some response elements (in bold) which were identified by computer analysis are also shown.

nuclease [17]. All genomic clones described in this report were isolated from the Stratagene canine genomic library.

Transfection and reporter assay

Protocols described by Virmani *et al.* [18] and Freshney [19] were followed for the isolation and maintenance of the primary culture of canine tracheal epithelial cells. The NIH 3T3 feeder cells were prepared according to Freshney [19] and used as described previously [10, 14]. Electroporation was with a Gene Pulser Electroporation apparatus (Bio Rad, Richmond, CA) with a 0.35 cm electrode gap in the sample cuvette. Briefly, 5×10^6 cells in late log phase were trypsinized, harvested ($600 \times \mathbf{g}$ for 5 min at 4 °C), and suspended in PBS (without Ca^{2+} and Mg^{2+}) in a final volume of 0.4 ml. DNA (10 μg) was mixed with cells by holding the cuvettes on the two 'window sides' and flicking the bottom. After 5 min at 4 °C, the cuvette was placed in the holder of the electroporation apparatus (kept at room temperature) and shocked once at 250 V and 960 mF capacitance. After electroporation, cells were placed at 4 °C for 10 min, diluted 20-fold in culture medium and grown in 100 mm petri dishes for 12 h; medium was changed and the CAT activity measured after 48 h [20]. Data were normalized as described [20–22]. All assays were performed a minimum of three times and representative data are shown. In some of the experiments, cells were treated with 8BrcAMP 24 h following transfection and the CAT activity was determined after 48 h. The transfection efficiency was determined by co-transfection with beta-galactosidase gene containing construct and necessary corrections in the transient expression were made [22].

Preparation of nuclear extracts

Nuclear extracts were made according to Parker and Topol [23]. Protein concentration was determined by a colorimetric assay [24].

EMSA (electrophoretic mobility shift assay)

The binding assay was a slight modification of a published protocol [25]. The fragments to be tested or the double stranded oligonucleotides were end labeled by standard protocols using gamma-^{32}P-ATP and T4 polynucleotide kinase [21]. Unincorporated radioactivity was removed by Sephadex G-50 chromatography or the labeled probes were purified on native acrylamide gels [22]. The reaction mixture contained 20 mM Hepes, pH 7.6, 50 mM NaCl, 10% (w/v) glycerol, 1 mM DTT, 1 mM EDTA, 0.05% Nonidet P40, 3 μg poly d(I-C), 15 000 cpm 5′-end-labeled probe, 10 μg of extract. The reaction mixture was incubated at 22 °C for 20 min and electrophoresed at 8 V cm^{-1} through a 6% polyacrylamide native gel in 1 × TBE buffer (0.089 M Tris, 0.089 M boric acid, 1 mM EDTA); the gel was subsequently dried and autoradiograph developed. Tests for specificity of binding included the ability to inhibit binding with 'self'-competition, but not with unrelated sequences.

Primer extension analysis

The 5′ boundary of the exon was determined by S1 nuclease mapping performed according to Killen *et al.* [26]. One genomic fragment (0.23 kbp long) located upstream of the 5′ end of the cDNA was isolated from clone GMV 2764, labeled at the *Alu* I site (10^6 dpm), was digested with *Xho* I and precipitated with 10 μg of tRNA or poly(A^+) RNA from CT1 cells; dissolved in 80% formamide, 0.4 M NaCl, 40 mM PIPES, pH 6.5, 1 mM EDTA; denatured at 90 °C for 10 min; and hybridized at 52 °C. After 3 h, the sample was diluted 10-fold in ice cold S1 buffer (0.25 M NaCl, 30 mM potassium acetate, pH 6.5, 1 mM $ZnCl_2$, 500 U S1 nuclease). After incubation at 37 °C for 30 min, the digestion was stopped by phenol extraction, and the DNA fragments were precipitated. The product was analysed on a 6% sequencing gel (the size of the protected fragment was about 100 bp). It corresponded to the major transcription initiation site determined by primer extension analysis.

To identify the 5′ end of the canine *TBM* transcript, an analytical primer extension experiment was performed using an end labeled synthetic oligonucleotide (5′TTCACGGGCATGGATCTTGAGAATGGACATTCTTGA 3′) corresponding to the 5′ 31 bp of the 5′ end of the cDNA (at nucleotide position 35-60 [7]). This oligonucleotide was annealed with 5 μg of poly (A^+) RNA from CT1 cells, and the primer was extended with 2 U of avian myeloblastosis virus reverse transcriptase at 44 °C in 60 mM Tris-HCl, pH 8.3, 80 mM NaCl, 40 mM KCl, 5 mM $MgCl_2$, 10 mM dithiothreitol. After 45 min, the reaction was stopped by phenol extraction, and the reaction products were precipitated with ethanol. The products of S1 nuclease and primer extension analysis were analysed on 6% sequencing gels [20].

Chimeric gene construction

To subclone the 5′-flanking sequences of the canine *TBM* transcript, pSV_0CAT [26, 27] was linearized at the *Hind* III site, the site was filled using the Klenow fragment of DNA polymerase, and the product ligated with *Xho* I linkers. All subsequent constructs were obtained in this derivative of pSV_0CAT. To obtain canine *TBM* promoter CAT constructs, the insert of clone CTM1–7 was treated with *Bal*31 for varying intervals and the deleted fragments were obtained [22]. The ends of these fragments were repaired with the Klenow DNA polymerase and after adding *Xho* I linkers, cloning was done in pSV_0CAT [27]. Thus the basic construct used for reporter gene analysis was pSVoCAT. Some constructs with their insert sizes are shown in Fig. 1. For some experiments, CAT constructs with SV40 promoter but no enhancer, or SV40 enhancer but no promoter (Promega, Madison, WI) were used.

These constructs were used to test whether mucin gene promoter contained some enhancer elements.

Results

Identification of first exon and nucleotide sequence of the canine TBM *promoter*

We previously isolated a 3773 bp cDNA clone for the canine *TBM* from a library constructed by specific primer extension utilizing poly(A^+) RNA from canine tracheal epithelial cells [7]. This cDNA encodes 72 bp of the 5′-untranslated region (UTR) and the complete coding region. We used 72 bp of the 5′-UTR of this cDNA as a probe to obtain the genomic clone CTM1–7. The size of the insert of this clone was about 2.1 kb. This clone was then used as a probe to isolate other genomic clones (Fig. 1A). Several sub-clones and deletion constructs of these clones were generated using standard methods. The nucleotide sequence of the promoter region is shown in Fig. 2.

S1 nuclease protection and primer extension were utilized to identify the 5′ boundary of the exon and the transcription initiation site(s) of the gene (Fig. 1B–D). The major and minor transcription initiation sites are shown in Fig. 1(B–D) and 2. Primer extension analysis indicated multiple transcription initiation sites (Fig. 1D), and revealed *TBM* messages of various lengths, the predominant ones starting 55, 81, 89, 110 and 129 bases from the translation start site. Based on the intensity of the signal seen on the autoradiogram, the major transcription start site is shown as +1 (Fig. 2). Similar results were obtained with human cells CFT1 and HBE1 [14] (results not presented). Multiple transcription initiation sites have also been reported for other mucin genes [11, 12].

The promoter region of the canine *TBM* gene was identified by reporter gene expression analysis and its nucleotide sequence was determined. Genomic clones identified here represent the putative first exon. No splice acceptor site was present in this region. The promoter region was not rich in G + C content and did not exhibit homology with either the *MUC1* or *MUC2* promoters [11, 12]. There were many TATA-like (e.g. nt starting at +18, +7, –7, –101, –114, –133, –192, –215, –227, –237, –252, –310, –334, –366, –376, –388, –424, –436, –452, –474, –491, –528, and –540) and CAAT-like (e.g. nt starting at +39, +11, –20, –29, –80, –95, –109, –121, –184, –230, –306, –313, –336, –350, –368, –392, –485, –535, and –543) sequences. Several other response elements were also present (Fig. 2). CRE-like sequences were present at positions –228/–239, –289/–306, and –367/–373 and AP1 like sequences were present at –286/–294 and at –363/–374. TREp (thyroid response element) and VDRE-like (vitamin D response element-like) sequences were located at –677/–671. They included the palindrome GTCATGAC (–677/–671). The TREp is an optimized palindromic rat growth hormone TRE which also functions as an efficient RARE (retinoic acid response element) [28].

To demonstrate the functionality of the promoter, reporter gene (CAT) expression analyses were performed by making nested deletions in the promoter region fused to the CAT gene [17]. These experiments indicated that the promoter identified in the present investigation was functional (Fig. 1A). The CAT activity was determined in mucin-expressing and non-expressing cell lines. Results indicated that the CAT activity was many fold higher in mucin expressing tracheal epithelial cells compared to fibroblasts or other mucin non-expressing cells (Fig. 3) indicating that some cell specificity exists for mucin gene expression.

As shown in Fig. 1(A) the CAT activity increased about six-fold (compared to the promoterless plasmid pSVo-CAT) when GMV23–3700 with the very large insert (about 3.7 kb) was used. This clone spans the complete promoter region and 5′ end of the cDNA of the *TBM* gene. The activity was reduced to some extent when the insert was shorter (e.g., an increase of four-fold was observed when GMV23–415 with about a 0.4 kb insert was used). Surprisingly, constructs with 0.24–0.32 kb inserts showed five- to six-fold more CAT activity. Results of further deletion indicated that a minimum of 40 bases up-stream of the major transcription initiation site (indicated as +1 in Fig. 2) was needed for promoter activity. In all experiments the transfection efficiency was the same (determined according to [22]).

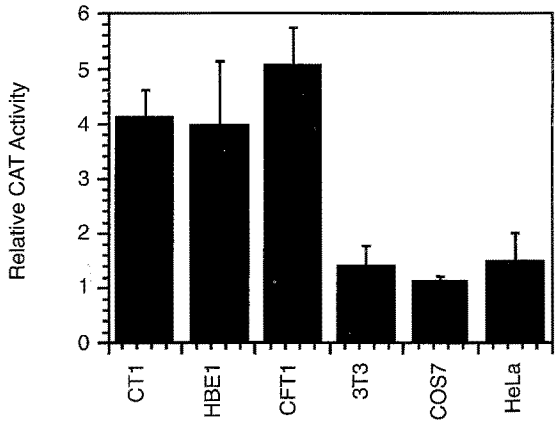


Figure 3. CAT activity in different cell lines. Mucin expressing (CT1, HBE1 and CFT1) and non-expressing (3T3, HeLa and COS7) cells were transfected with GMV2906 (–519/+40; the numbering is according to Fig. 2). The relative CAT activity was determined according to Sambrook *et al.* [20]. The transfection efficiency was determined by co-transfection with beta galactosidase gene containing construct [22].

Secondary structure of the 5′-end of the canine TBM *gene*

The hypothetical secondary structure of the 5′-end of the canine *TBM* gene is shown in Fig. 4. Hairpin structures containing 6, 3, 7, 9, and 6 unpaired bases in the indicated loops were observed between nucleotides +7 and +103. We found that the 5′-untranslated region of the gene presented a stem-and-loop structure. Such secondary structures have been shown to modulate gene expression in higher eukaryotes [29, 30]. The inhibitory effect of the loop depends both upon its thermodynamic stability and its position relative to the transcription start site. According to the model proposed by Jbilo *et al.* [31] the hairpin structure in the 5′-untranslated region of the canine *TBM* is expected to have only a slight effect on expression for two reasons: it has a low thermodynamic stability of $-21.7\ \text{kcal mol}^{-1}$, and it is located 7 bases downstream of the cap site. Hairpins that are 5–20 nucleotides downstream of the cap site do not inhibit ribosome binding. Thus, it is likely that the translation machinery melts this secondary structure. All the above presumptions are based on the computer model only and a functional analysis is yet to be done.

Canine TBM *promoter contains binding sites for regulatory proteins*

Initially, DNA fragments located in the promoter region (from 40 to 500 nt up-stream of the transcription site) were used for EMSA (Fig. 5). Nuclear extracts from CT1 and HeLa cells were used for this analysis. The region identified as the promoter of the canine *TBM* gene indeed contained binding sites for nuclear proteins (based on their distinct mobility in EMSA) (Fig. 5). Clones containing shorter than a 45 nt insert did not show any protein binding activity. Results of these experiments indicated that the promoter and flanking regions provide binding sites for nuclear proteins.

Previously we demonstrated that cAMP and its analogs enhanced mucin gene expression in CT1 cells, a response which required *de novo* RNA and protein synthesis [10]. Therefore we examined the promoter region for any

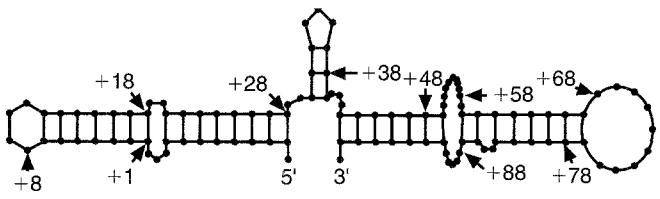


Figure 4. Schematic representation of secondary structure in the untranslated 5′-end region of the canine *TBM* gene. Hairpin structures containing 6, 3, 7, 9, and 6 unpaired bases in the indicated loops were observed between nucleotides +7 and +103. The minimum destabilizing energy of this secondary structure was estimated to be $-21.7\ \text{kcal mol}^{-1}$ using the program FoldRNA [38–40]. A +1 indicates the transcription initiation site.

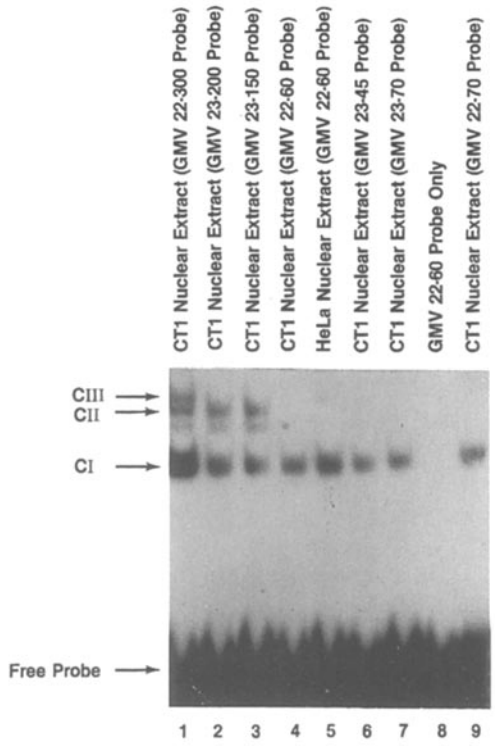


Figure 5. The promoter region of canine *TBM* gene contains binding sites for nuclear proteins. EMSA of the nuclear proteins, from canine tracheal epithelial cells (CT1 cells) as well as from HeLa cells, was performed. The product (protein-DNA complex) was analysed on a native 6% polyacrylamide gel. All the probes mentioned above are DNA fragments isolated from the canine genomic DNA. The location of the probes is as follows: GMV22–300, (–247/+51); GMV23–200, (–146/+51); GMV23–150, (–97/+51); GMV 22–60, (–9/+51); GMV23–45, (+2/+51); GMV23–70, (–18/+51). The numbering of the bases is based on the numbering shown in Fig. 2. None of the probes show any complex formation in the absence of the nuclear extract (not shown in the figure). CI, CII and CIII designate three proteins that bind to the *TBM* promoter.

CRE-like sequences; two such sequences were found (Fig. 2). EMSA indicated that CRE-like nucleotide sequences CATGCGATCACGTTGACTGACTT (–289/–306) and ATTTGAGTCATATTAAA (–360/–376) contained binding sites for CT1 nuclear proteins (Fig. 6); another CRE-like sequence is also present at location –228/–234. EMSA with CT1 nuclear extract indicated that these sequences provide binding sites for different nuclear proteins.

Competition with CRE and mutant oligonucleotides

The motif CATGCGATCACGTTGACTGACTT (–289/–306) contained a nucleotide sequence similar to a CRE whereas the other motif ATTTGAGTCATATTAAA (–376/–360) contained a nucleotide sequence similar to the AP1 site (the consensus binding site for CREB is

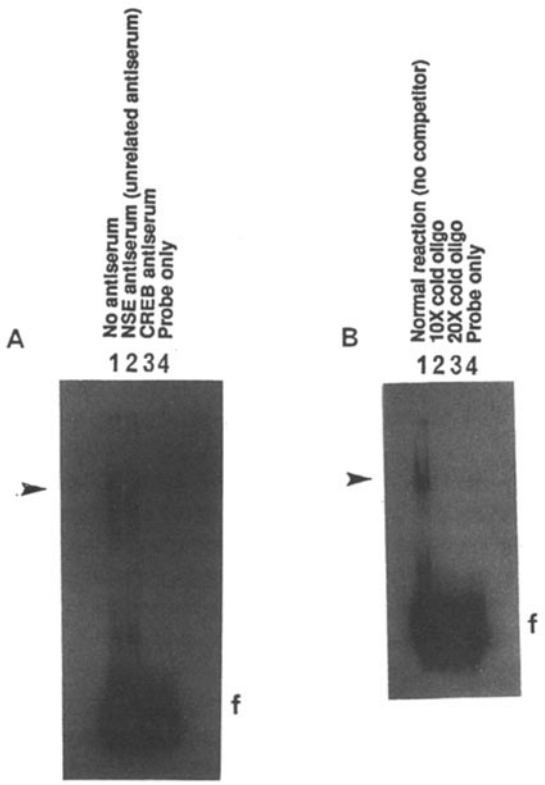


Figure 6. (A) Treatment of the nuclear extract with anti-CREB antiserum. The probe used for these experiments was a fragment (–296/–373) isolated from clone GMV2906 (–519/+40, the numbering is according to Fig. 2). Nuclear extract made from CT1 cells was incubated with anti-CREB antiserum at 37 °C for 30 min followed by EMSA. f, free probe. (B) Competition assay. The probe used for this experiment was an oligonucleotide (–360/–376). The binding of nuclear proteins and oligonucleotide probe was competed out with molar excess of cold oligonucleotide (–360/–376) as shown.

GATTGGCTGACGTCAGAGAGCT whereas for AP1 it is CTAGTGATGAGTCAGCCGGATC). These sequences contained binding sites for different nuclear proteins because the mobility of the complex was different when the above CRE-like (CATGCGATCACGTTGACTGACTT) and AP1-like (ATTTGAGTCATATTAAA) nucleotides were used. Mutated oligonucleotides of CRE-like and AP1-like oligonucleotides did not form DNA-protein complexes and the retardation band disappeared when the respective cold oligonucleotides were included in the binding assay (Fig. 6B). Mutants in the non-conserved CRE-sequence did not affect the protein binding activity. These results indicate that the complexes generated were due to specific binding. The retarded bands disappeared when the nuclear extract was pre-incubated with CREB antiserum (lane 3, Fig. 6A) but not with unrelated neuron specific enolase antiserum (lane 2, Fig. 6A). This indicates that the binding protein complex contained CREB.

Up-regulation of the canine TBM *expression by 8BrcAMP*

CT1 cells grown in selective medium (1–2 × 10^6 cells/ml^{-1}) were exposed to secretagogues, such as cAMP, 8BrcAMP, 2-butyryl cAMP, catechol, PGE1, PGE2, arachidonic acid and dexamethasone and their effects evaluated on the canine *TBM* stable transcript levels using quantitative RT-PCR [10]. These experiments and their results have been described previously [10]. It was of interest to know what secretagogues might induce mucin synthesis so that mucin purified from these induced cells could be compared with the mucin synthesized by normal primary (without immortalization) tracheal epithelial cells. The results indicated enhanced canine *TBM* message levels due to the treatment with the above agents. Furthermore, cAMP alters ion transport resulting in the elevation of CFTR expression. Therefore, we determined whether cAMP (and its analogs) altered expression of the mucin gene. Results suggested that cAMP analogs induced higher *TBM* transcript levels and this induction required *de novo* RNA and protein synthesis (as determined by evaluating RNA and protein synthesis in the presence of actinomycin D and cycloheximide which are transcription and translation inhibitors, respectively) [10].

To investigate. whether induction of canine *TBM* by cAMP is regulated at the transcription level, two assays were performed, *viz.*, nuclear run-on and CAT. Nuclear run-on data indicated that regulation occured at the transcription level (Fig. 7). To perform these experiments, CT1, HBE1, and CFT1 cells were grown in growth factor and hormone supplemented medium. At approximately 80% confluence, the medium was replaced with medium containing 8BrcAMP and incubation was continued for 2 h at 37 °C. Nuclei were isolated and elongation of nascent transcripts was performed in a reaction buffer containing ^{32}P-UTP. For detection of specific transcripts, newly synthesized transcripts were isolated and hybridized to filters containing an excess of plasmid DNA (MVM 428, located in the 3′ end region of the cDNA, see [7]). The level of transcription was determined by autoradiography and quantified by scanning densitometry or direct counting of the hybridized area of the nylon membrane. The transcription was found to be maximum in CFT1 cells after treatment with 8BrcAMP. Similar results were obtained when CFT1–4 cells were used (CFT1–4 is another immortalized epithelial cell line isolated from a CF patient homozygous for phenyl alanine at 508). In control cells (3T3), no signal was obtained. When a promoter-CAT construct (GMV2906-CAT, –519/+40) was used to determine the level of transcription of the canine *TBM* gene in the presence and absence of the cAMP analogs, regulation was also found to be at the transcription level (Fig. 8). Twenty-four hours following transfection, cells were treated with 8BrcAMP in the presence or absence of the transcription inhibitor

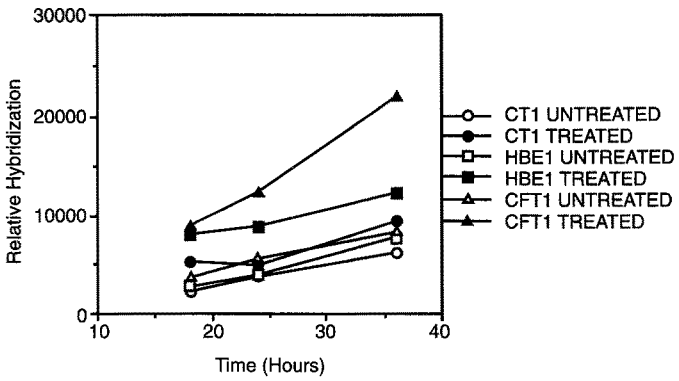


Figure 7. Nuclear run-on assay. CT1, HBE1, and CFT1 cells were grown in growth factor and hormone supplemented medium. At approximately 80% confluence, the medium was replaced with medium containing 8BrcAMP (1 μM) and incubation was continued for 2 h at 37 °C. Nuclei were isolated and elongation of nascent transcripts was assayed in a reaction buffer containing ^{32}P-UTP. For detection of specific transcripts, newly synthesized transcripts were isolated and hybridized to filters containing an excess of plasmid DNA MVM428 (cDNA of the canine *TBM* gene covering the 3′ end of the cDNA, see [7]). The level of transcription was determined by autoradiography and quantified by scanning densitometry or direct counting of the hybridized area of the nylon membrane. Results are the representative of at least three independent experiments. For control, either plasmids with no insert, or insert containing glyceraldehyde phosphate dehydrogenase (*GAPDH*) gene were used. Levels of GAPDH did not change as a result of treatment with 8BrcAMP.

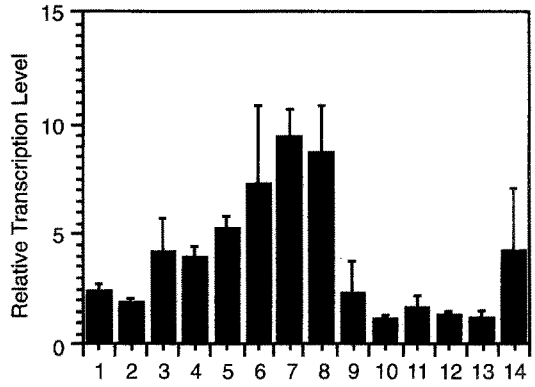


Figure 8. CAT assay in CT1, CFT1, HBE1 and CFT1 cells after treatment with 8BrcAMP. For the treatment of cells 1 μM 8BrcAMP was used. Cells were treated with 8BrcAMP 24 h following transfection and the CAT activity was determined after 48 h. Lane identification: 1–4, and 13, 14, untreated cells; 5–12, treated cells; 5–8, cells treated with 8BrcAMP; 9–12, cells treated with 8BrcAMP (1 μM) and actinomycin D (10 μg ml^{-1}) added simultaneously; 1, 5, and 9, CT1 cells; 2, 7, and 10, HBE1 cells; 3, 4, 6, 8, 11 and 12, CFT1 cells; 1–12, cells transfected with GMV2906–CAT (–519/+40, the numbering is according to Fig. 2) containing the canine *TBM* promoter; 13, CFT1 cells transfected with pSVoCAT; 14, CFT1 cells transfected with pSV2CAT. Transfection efficiency was determined according to [22]. Results are average of three independent experiments and bars represent the standard deviation.

actinomycin D (10 μg ml^{-1}) and CAT activity determined after 12 h. Results indicated much higher CAT activity in CFT1 cells compared to other cells and this increased transcription was not observed in the presence of actinomycin D. The transcription of the mucin gene was thus enhanced by 8BrcAMP (Fig. 8). The half-life of the canine *TBM* mRNA was not altered as a result of treatment with 8BrcAMP supporting the notion that regulation of the canine *TBM* by 8BrcAMP occurred at the transcription level, as opposed to the post-transcription level (Fig. 9). Northern blot analysis of CAT mRNA produced in the transfected CT1 cells demonstrated that CAT activity from these constructs was representative of the RNA expressed in response to 8BrcAMP. When RNA was isolated from CT1 cells pre-treated with 8BrcAMP and used for primer extension experiments, the transcription started at the major transcription initiation site (results not shown). Induction of the canine *TBM* gene transcription by 8BrcAMP involved the major transcription start site (marked +1 in Fig. 2). When other epithelial cells, such as HBE1, CFT1, CFT1-4, were treated with 8BrcAMP, induction in the expression of the mucin gene was observed. However, the maximum levels of the transcript were found in CFT1. Under similar conditions the expression of GAPDH did not change.

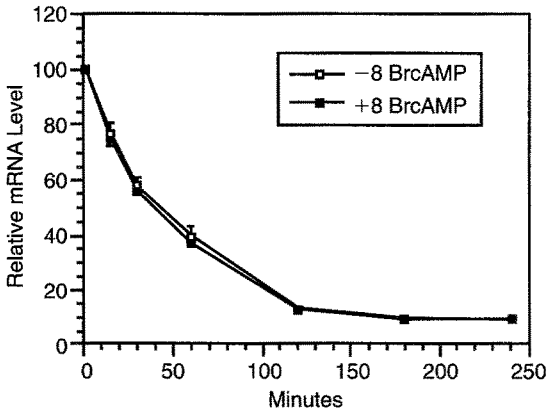


Figure 9. Half-life of the canine *TBM* mRNA in 8BrcAMP treated and untreated cells. CT1 cells were treated with 1 μM 8BrcAMP and half-life of the *TBM* mRNA was determined as described [10, 16]. Briefly, CT1 cells grown to 80% confluency were treated with 1 μM 8BrcAMP either in the presence or absence of actinomycin D (10 μg ml^{-1}). At different time intervals cells treated with actinomycin D were analysed for the *TBM* mRNA expression and the half-life of the *TBM* mRNA was compared with those of untreated cells [10, 16]. The half-life of the *TBM* mRNA was 45–60 min in 8BrcAMP-treated or untreated cells.

Discussion

The canine *TBM* promoter showed interesting features such as multiple transcription initiation sites, more than one TATA- and CAAT-like sequence, and binding sites for CREB and other transcription factors. The intestinal mucin gene *MUC2* contains a single TATA-like sequence and also has multiple transcription initiation sites [12] although all of the detectable sites occur within a narrow range of nucleotides. The *MUC1* (gene encoding lactating mammary gland mucin) promoter contains positive and negative response elements in its promoter region. The promoter regions of other mucin genes have not been characterized.

The secondary structure of the canine *TBM* 5′ non-translated region exhibited a free energy of $-21.7\ \text{kcal mol}^{-1}$ as calculated by the FOLD program of the GCG analysis package and $-21.2\ \text{kcal mol}^{-1}$ as calculated by the SQUIGGLE program of the same package. This region should be efficient for transcription since less than $-50\ \text{kcal mol}^{-1}$ free energy has been reported not to cause inhibition in transcription of a gene. The parameters combined for analysis included hairpin loops, internal loops,. and bulges, mismatches, impaired terminal nucleotides and other structures. There is no homology in the promoter region of *MUC1*, *MUC2*, and canine *TBM*.

Results presented in this report indicate the following: expression of canine *TBM* appears to be tightly controlled. A certain level of mucin remains in the cell all the time which can be up-regulated by certain secretagogues (such as dexamethasone, prostaglandins, cAMP etc.). The promoter region of the canine *TBM* contains two CREs that bind to epithelial cell nuclear proteins. Previously we demonstrated the induction of canine *TBM* mRNA levels by 8BrcAMP, a process which required *de novo* RNA and protein synthesis [10]. To delineate the sequences required for response to 8BrcAMP, we constructed a series of promoter deletion mutants. Our studies defined a region between –289 and –376 base pairs upstream of the transcriptional initiation site that conferred cAMP responsiveness when placed adjacent to the promoter. Within the promoter region (–289/–376) of the mucin gene, two consensus CREs were present (one at –289/–306 and another at –367/–373 nt). Deletion analysis and reporter gene assays strongly suggested that the CRE motifs were essential for the up-regulation of the mucin gene. The CRE is present in many other genes, such as *PEPCK*, *VIP* (vasoactive intestinal peptide), parathyroid hormone, proenkephalin, insulin, α-chorionic gonadotrophin, c-fos, cytomegalovirus enhancer and *BLV LTR* [31]. This indicates that the CREB (cAMP response element binding protein) is involved in regulation of a number of genes.

Results of EMSA suggested that CREs provide sites for binding of nuclear proteins. To test whether this protein is CREB, the nuclear extract was pre-incubated with anti-CREB antiserum followed by EMSA. Since the retardation band disappeared as a result of treatment with CREB antiserum, it indicated that CREB is involved in the binding reaction.

CREB is a transcription factor that is phosphorylated in response to extracellular stimuli which elevate intracellular cAMP and activate the cAMP-regulated protein kinase (cAPK) [33]. CREB is constitutively expressed, present under both basal and induced conditions [34] which activate transcription after phosphorylation on residue Ser^{133} [34]. The protein kinases that have been shown to phosphorylate CREB on Ser^{133} are protein kinase A and calmodulin dependent protein kinases.

We have observed that expression of mucin gene is altered by treatment with a variety of agents, e.g. dexamethasone, carbachol, prostaglandins, arachidonic acid, atropine, benzo-α-pyrene, retinoic acid. It is not clear why these agents show different levels of induction of mucin transcript and what is the underlying mechanism for this effect. Characterization of mucin synthesized from primary culture and immortalized cells suggests that mucin synthesized by these cells is similar to that synthesized by primary tracheal epithelial cells in terms of amino acid and carbohydrate composition (Verma *et al.*, unpublished results). Basbaum's group have reported an alteration (induction) in the levels of mucin genes as a result of exposure to irritants (such as nitric oxide and SO_2) and infection [35]. It is quite possible that the regulation of mucin gene(s) under external stimuli is also mediated by the cAMP pathway.

Acknowledgements

This work was supported by grants from the Cystic Fibrosis Foundation (CF 6147 to M.V.) and the National Institutes of Health (HL28650 to E.A.D. and ES05303 to R.N.K.). We are grateful to Drs Raziuddin, James Gum, Young Kim, Dean Rosenthal, J. Shaper, Ram Shukla, Sandra Gendler, Alvin Berger and D. Lyn for their critical comments.

References

1. Chambers JA, Harris A (1993) *J Cell Sc* **105**: 417–22.
2. Collins FS (1992) *Science* **256**: 774–79.
3. Robinson CB, Martin WR, Ratcliff JL, Holland PV, Wu R, Cross CF (1993) *Am Rev Resp Dis* **148**: 385–89.
4. Feldhoff PA, Bhavanandan VP, Davidson EA (1979) *Biochemistry* **18**: 2430–36.
5. Ringler NJ, Selvakumar R, Woodward HD, Simet IM, Bhavanandan VP, Davidson EA (1987) *Biochemistry* **26**: 5322–28.

6. Ringler NJ, Selvakumar R, Woodward, HD, Bhavanandan VP, Davidson EA (1988) *Biochemistry* **27**: 8056–63.
7. Verma M, Davidson EA (1993) *Proc Natl Acad Sc USA* **90**: 7144–48.
8. Verma M, Davidson EA (1994) *Glycoconj J* **11**: 172–79.
9. Verma M, Madhu M, Marrota C, Lakshmi CV, Davidson EA (1993) *Cancer Biochem Biophys* **14**: 41–51.
10. Verma M, Sanadai A, Davidson EA (1994) *Glycobiology* **4**: 825–36.
11. Kovarik A, Peat N, Wilson D, Gendler SJ, Taylor-Papadimitriou J (1993) *J Biol Chem* **268**: 9917–26.
12. Gum JR, Hicks JW, Toribara NW, Siddiki B, Kim Y (1994) *J Biol Chem* **269**: 2440–46.
13. Shirotani K, Taylor-Papadimitriou J, Gendler S, Irimura T (1994) *J Biol Chem* **269**: 15030–35.
14. Yankaskas JR, Cotton CU, Knowles MR, Gatzy JT, Boucher RC (1985) *Am Rev Resp Dis* **132**: 1281–87.
15. Verma M, Davidson EA (1994) *Cancer Biochem Biophys* **14**: 123–31.
16. Hames BD, Higgins SJ (1993) (eds) *Gene Transcription: A Practical Approach*. New York: IRL press.
17. Koh J, Sferra TJ, Collins FS (1993) *J Biol Chem* **268**: 15912–21.
18. Virmani AK, Naziruddin B, Desai VC, Lowry JP, Graves DC, Sachdev GP (1992) *In Vitro Cell Dev Biol* **28A**: 120–27.
19. Freshney RI (1992) (ed.) *Culture of Epithelial Cells*. pp. 25–223. New York: Wiley-Liss, Inc.
20. Sambrook J, Fritisch E, Maniatis T (1989) *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor Press, NY: Cold Spring Harbor Laboratory.
21. Ausubel FM, Brent R, Kingston RE, Moore DD, Smith JA, Siedman JG, Struhl K (1987) *Current Protocols in Molecular Biology*. New York: John Wiley and Sons.
22. Murray EJ (1993) (ed.) *Gene Transfer and Expression Protocols* pp. 17–330. Clifton, NJ: Humana Press.
23. Parker CS, Topol J (1984) *Cell* **36**: 357–66.
24. Verma M, Kurl RN (1993) *Biochem Molec Biol Int* **30**: 293–303.
25. Jones KA, Kadonaga JT, Rosenfeld PJ, Kelley TJ, Tjian R (1987) *Cell* **48**: 79–89.
26. Killen PD, Burbelo P, Martin GR, Yamada Y (1988) *J Biol Chem* **263**: 12310–14.
27. Burbelo PD, Bruggeman LA, Gabriel GC, Klotman PE, Yamada Y (1991) *J Biol Chem* **266**: 22297–302.
28. De-Luca LM (1991) *FASEB J* **5**: 2924–30.
29. Kozak M (1986) *Proc Natl Acad Sci USA* **83**: 2850–54.
30. Kozak M (1991) *J Biol Chem* **266**: 19867–70.
31. Jbilo O, Toutant J, Vatsis KP, Chatonnet A, Lockridge O (1994) *J Biol Chem* **269**: 20829–37.
32. Lalli E, Sassone-Corsi P (1994) *J Biol Chem* **269**: 17359–62.
33. Alberts AS, Montminy M, Shenolikar S, Feramisco JR (1994) *Mol Cell Biol* **14**: 4398–407.
34. Borsook D, Konradi C, Falkowski O, Comb M, Hynam SEH (1994) *Mol Endocrinol* **8**: 240–48.
35. Jany B, Gallup M, Tsuda T, Basbaum C (1991) *Biochem Biophys Res Commun* **181**: 1–8.
36. Young MF, Findlay DM, Dominguez P, Burbelo P, McQuillan C, Kopp JB, Robey PG, Termine JD (1989) *J Biol Chem* **264**: 450–56.
37. Bailey JM, Verma M (1991) *Analyt Biochem* **196**: 11–18.
38. Zuker M, Stiegler P (1981) *Nucleic Acids Res* **9**: 133–48.
39. Freier SM, Kierzek R, Jaeger JA, Sugimoto N, Caruthers MH, Neilson T, Turner DH (1986) *Proc Natl Acad Sci USA* **83**: 9373–77.
40. Verma M, Olnes MJ, Kurl RN, Davidson EA (1995) *Gene* **154**: 255–57.