# *MUC8* Protein Structure and Transcription Regulation Mechanism by the Analysis of *MUC8* Gene and Protein

**Kyou-Nam Cho**

**Department of Medical Science**

**The Graduate School, Yonsei University**

# *MUC8* Protein Structure and Transcription Regulation Mechanism by the Analysis of *MUC8* Gene and Protein

**Directed by professor Joo-Heon Yoon**

**The Doctoral Dissertation Submitted to the Department of Medical Science, the Graduate School of Yonsei University in partial fulfillment of the requirements for the degree of Doctor of Philosophy**

**Kyou-Nam Cho**

**December 2004**

**This certifies that the Doctoral Dissertation of Kyou-Nam Cho is approved.**

---

Thesis Supervisor : Joo-Heon Yoon

---

Kwang-Chul Chung : Thesis Committee Member #1

---

Won-Jae Lee : Thesis Committee Member #2

---

Kwang-Hoon Lee : Thesis Committee Member #3

---

Dong-Seok Kim : Thesis Committee Member #4

**The Graduate School**

**Yonsei University**

**December 2004**

# ACKNOWLEDGEMENTS

또 하나의 과정이 끝났습니다. 늦게 시작한 공부가 힘들 때마다 항상 먼저 떠오르며, 가장 기뻐하실 어머님께 이 작은 결실을 드립니다.
석사과정을 지나 박사 과정을 마치기까지 짧지않은 시간을 보내면서 가장 좋았던 시간은 일을 처음 배웠던 석사과정 때였습니다. 그때는 다른 것에 신경 쓸 겨를도 없이 그냥 일이 좋아 열심히 했던 것 같습니다. 졸업을 하고, 어디서 일을 하던지 그 때의 초심을 잃지 않고, 사회의 한 구성원으로써 열심히 살아가는 제 모습을 그려봅니다.

박사과정 3년 동안 본 논문의 초안에서 완성까지 지도해주시며, 많은 관심과 배려를 베풀어 주신 윤주헌 교수님께 감사의 말씀을 드립니다. 바쁘신 중에도 본 논문을 검토해 주신 정광철 교수님, 이원재 교수님, 이광훈 교수님, 김동석 교수님께도 감사의 말씀을 드립니다.
실험실 식구들간에 화합할 수 있도록 잦은 유흥문화를 제공해주신 김창훈 선생님, 최재영 선생님, 김진국 선생님, 정유삼 선생님, 유종범 선생님께도 감사드리며, 하시는 연구에 많은 진전 있으시기를 바랍니다.
선배로써 잘 해준 것이 없어 늘 미안한 후배 기한과 지현이가 내년에는 원하는 학교에 진학할 수 있기를 바랍니다.
힘든 내색 한 번 내지않고, 성실한 모습 보여준, 후배 욱열, 수진, 영은, 영덕에게도 감사드립니다.
15년 넘게 좋은 친구가 되어준 용준, 해성, 경식, 창환, 정석에게도 고마운 마음을 전하며, 그 우정이 영원하길 바랍니다.
학생이라는 이유로 많은 부분에서 이해해주신 장인어른, 장모님, 누님들, 매형들께 죄송스런 마음과 사랑하는 마음을 전합니다.

끝으로, 제 곁에서 항상 든든한 동행자가 되어준 아내 보영이와 이제 걸음마를 시작하는 우리 아들 현승이에게도 사랑하는 마음을 전합니다.

# Table of Contents

# List of Figures

# List of Tables

**Abstract**

# *MUC8* Protein Structure and Transcription Regulation Mechanism by the Analysis of *MUC8* Gene and Protein

**Kyou-Nam Cho**

*Department of Medical Science*

*The Graduate School, Yonsei University*

(Directed by Professor Joo-Heon Yoon)

*MUC8* gene expression is overexpressed in nasal polyp epithelium and is also increased by treatment with inflammatory mediators in nasal epithelial cells. These data suggest that *MUC8* may be one of important mucin genes expressed in human airway. The MUC8 protein has been determined only to the extent of its short C-terminal sequence (313 amino acids), which can be inferred from 3′-end of its cDNA. This unique sequence appears to be composed of two types of consensus repeats, but this provides no clues into its function. Accordingly, we decided to clone a complete human *MUC8* cDNA sequence. We have sequenced the full-length *MUC8* cDNA

from normal human nasal epithelial cells. *MUC8* cDNA is 10292 bp long, and can be separated into 1733 bp of a 5′-untranslated region, 8097 bp of an open reading frame, and 459 bp of a 3′-untranslated region. *MUC8*′s open reading frame is derived from 9 exons. Deduced MUC8 peptide sequences reveal three unique tandem repeat regions, each composed of mucin-specific amino acids. The protein sequence of MUC8 contains both secreted- and membrane-bound domains. In addition, we cloned the – 1644 to + 87 region of genomic DNA upstream of the *MUC8* transcription start site, and we examined the mechanism by which prostaglandin E2 (PGE2), an arachidonic acid metabolite, increases *MUC8* gene expression levels. Here, we show that ERK MAP kinase is essential for PGE2-induced *MUC8* gene expression in normal human nasal epithelial cells and that p90 ribosomal S6 protein kinase 1 (RSK1) mediates the PGE2-induced phosphorylation of cAMP-response element binding protein (CREB). Our results also indicate that cAMP-response element (CRE) at – 803 region of the *MUC8* promoter is an important site of PGE2-induced *MUC8* gene expression. In conclusion, these studies give insights into the molecular mechanism of PGE2-induced *MUC8* gene expression and enhance our understanding on mucin hypersecretion during inflammation.

---

**Key words : *MUC8* gene, Mucus hypersecretion, Transcriptional regulation, PGE2**

# *MUC8* Protein Structure and Transcription Regulation Mechanism by the Analysis of *MUC8* Gene and Protein

**Kyou-Nam Cho**

***Department of Medical Science***

***The Graduate School, Yonsei University***

**(Directed by Professor Joo-Heon Yoon)**

## I. INTRODUCTION

Mucus hypersecretion is a common feature in respiratory tract infections, most notably, rhinosinusitis, chronic bronchitis, and cystic fibrosis. Mucins are the major components of mucus[1-4]. Although twenty mucin genes have been identified thus far, their functions remain unknown. The mucins are usually subdivided into two groups:, membrane-bound mucins and secreted mucins. MUC2[5], MUC5AC[6], MUC5B[7], MUC6[8], MUC7[9], MUC9[10] and MUC19[11] are secreted mucins, whereas MUC1[12], MUC3[13], MUC4[14], MUC11[15], MUC12[15], MUC13[16],

MUC17[17], MUC18[18] and MUC20[19] are membrane-bound mucins. However, the other mucin genes, including MUC8[20] have not been characterized as to whether they are secreted or membrane-bound forms.

Mucins are very densely *O*-glycosylated, and consequently, contain sequences rich in proline, threonine, and/or serine (PTS), which function as acceptors for the *O*-glycans during biosynthesis. PTS regions usually comprise relatively short, tandem repeated sequences. The peptide sequences of PTS regions exhibit limited amino acid composition and, consequently, limited codon usage at the DNA and mRNA levels. Each mucin contains a distinctive repetitive PTS-rich sequence, with respect to both amino acid sequence and the length of the repetitive unit[21].

The cDNA sizes of the secreted mucins tend toward the very large (15 – 40 kb cDNA). The cDNA sequences of secreted mucins exhibit multiple cysteine-rich domains in the flanking region of the mucin-like PTS regions, cysteine-knot, and von Willebrand Factor C and D (vWFC and D) domains. Both the number and the position of the cysteines play an essential role in the formation of disulfide-linked dimers and multimers[21]. In contrast, membrane-bound mucins characteristically have a transmembrane domain, a sea-urchin-sperm protein-enterokinase-agrin (SEA) domain, and one or two epidermal-growth factor (EGF)-like

domains. However, neither the SEA nor the EGF domains are specific to all mucins. MUC4 contains a nidogen domain, and a ′adhesion-associated domain in *MUC4* and other proteins′ (AMOP) domain[21].

Prostaglandins (PGs) are arachidonic acid metabolites with a wide range of biological actions. Moreover, it is known that cyclooxygenase (COX) converts arachidonic acid to prostaglandin H2 (PGH2), which is further metabolized to various PGs and thromboxanes[22]. These species are produced in a wide variety of tissues and function as lipid mediators. In particular, PGE2 mediates IL-1β-induced *MUC5AC* gene expression in human airway epithelium[23]. Because mucin hypersecretion is a hallmark of airway inflammation, it is important that we determine how PGE2 regulates airway mucin gene expression.

In a previous study, we found that the expression of *MUC8* mRNA increased as a function of differentiation in normal human nasal epithelial (NHNE) cell culture, and that *MUC8* mRNA levels were up-regulated in the nasal polyp epithelium, which is invariably stimulated by inflammatory mediators[2]. Interleukin-1β (IL-1β), tumor necrosis factor-α (TNF-α), and a mixture of inflammatory mediators were found to result in the up-regulation of *MUC8* mRNA *in vitro*[24]. In addition, we found that the IL-1β-induced expression of the *MUC8* gene was mediated by a sequential MAPK/RSK1/CREB cascade pathway in human airway

epithelial cells[25]. In that study, we were unable to use *MUC8*-specific promoter, as, at that time, only a portion of the *MUC8* cDNA had been identified. The aim of the present study was to characterize the complete cDNA and to clone the *MUC8* promoter, and also to determine whether the cAMP response element is essential for PGE2-induced *MUC8* gene expression in airway epithelial cells. *MUC8* cDNA is 10292 bp long, comprising 1733 bp of a 5′-untranslated region, 8097 bp of an open reading frame, and 459 bp of a 3′-untranslated region. *MUC8*′s open reading frame is derived from 9 exons. In addition, we cloned the – 1644 to + 87 region of genomic DNA upstream of the *MUC8* transcription start site and we examined the mechanism by which PGE2 increases *MUC8* gene expression levels. We found that extracellular signal-regulated kinase mitogen-activated protein kinase (ERK MAPK) is essential for PGE2-induced *MUC8* gene expression in normal human nasal epithelial (NHNE) cells. We also found that p90 ribosomal S6 protein kinase 1 (RSK1) mediates the PGE2-induced phosphorylation of cAMP response element binding protein (CREB). In addition, transcriptional activities of cloned *MUC8* promoter regions showed that CRE in *MUC8* promoter is an important site of PGE2-induced *MUC8* gene expression.

Taken together, these studies provide insights into the function of MUC8 and the mechanism of PGE2-induced *MUC8* gene expression, and extend our understanding of mucin

gene overexpression during airway mucosal inflammation.

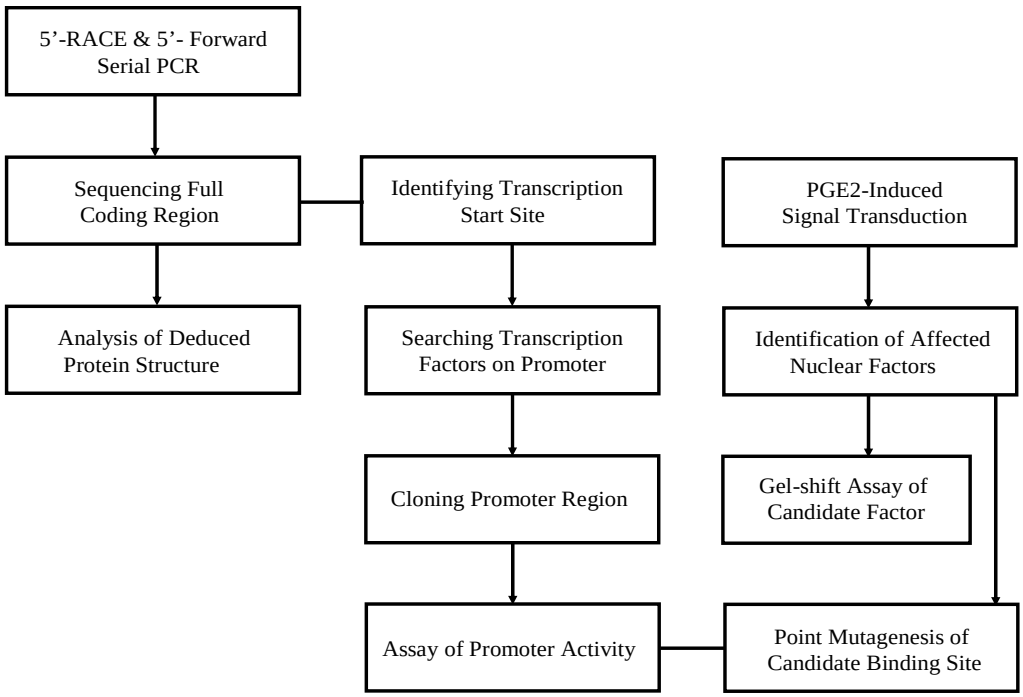


**Fig. 1. Experimental Flowchart.**

## II. MATERIALS AND METHODS

1. *MUC8* full cDNA sequence and protein structure

A. Cell cultures

Normal human nasal epithelial (NHNE) cells (passage 2) were seeded in 0.5 ml of culture medium onto 24.5 mm, 0.45 μm pore size, Transwell-clear (Costar) culture inserts. Cells were cultured with a 1:1 mixture of bronchial epithelial cell growth medium (BEGM, Clonetics):Dulbecco′s modified Eagle′s medium (DMEM, Invitrogen) containing all supplements, as previously described[26]. Cultures were grown submerged, during which time the culture medium was changed on the first day, and daily thereafter. The human lung mucoepidermoid carcinoma cell line, NCI-H292, was purchased from the American Type Culture Collection (CRL-1848, Manassas, VA), and was cultured in RPMI 1640 (Invitrogen) supplemented with 10% fetal bovine serum (FBS, Invitrogen) in the presence of penicillin-streptomycin, at 37 °C, in a humidified chamber at 5% $CO_2$.

B. RT-PCR

Total RNA was isolated using TRIzol (Invitrogen) from NHNE cells. cDNA was synthesized with random hexamers (PerkinElmer Life Sciences) using Moloney murine

leukemia virus-reverse transcriptase (PerkinElmer Life Sciences). In order to verify that the amplified products were from mRNA, and not from genomic DNA contamination, negative controls were constructed by omitting reverse transcriptase from the RT reaction. Oligonucleotide primers were designed based on GenBank™ accession number AC079031 of human chromosome 12q24.3, including the *MUC8* gene (Fig. 1 and Table 1). The established *MUC8* mRNA sequence (GenBank™ accession number U14383) location is from nucleotide 44861 to nucleotide 43472, in comparison to this genomic DNA sequence. Each PCR was performed from a cDNA template using 4 overlapped A1-A2, B1-B2, C1-C2, and D1-D2 pair-primers, and amplified fragments were joined to one another.

C. 5′-Rapid Amplification cDNA Ends PCR

The SMART™ RACE cDNA Amplification kit (Clontech) was used to synthesize first-strand cDNA from 1 μg of RNA using the oligo (dT) and SMARTII A oligo anchor primer. *MUC8* gene specific PCR was performed using an Advantage2 Polymerase kit (Clontech) with an RA59840 antisense primer (Table 1), and the anchor primer of SMART™ RACE kit was used as a sense primer. The PCR amplification reaction mixture (25 μl) contained

1.5 μl of cDNA, 10 mM dNTP mix, 10 μM of each primer, 2.5 μl of 10X Advantage 2 PCR buffer, and 0.5 μl of 50X Advantage 2 Polymerase mix. PCR was performed with a PerkinElmer Thermal Cycler Gene AMP® PCR System 2400. PCR parameters were: 94 °C for 3 minutes, followed by 5 cycles at 94 °C for 5 seconds and 72 °C for 3 minutes, 5 cycles at 94 °C for 5 seconds, 70 °C for 30 seconds, and 72 °C for 3 minutes, and 30 cycles at 94 °C for 5 seconds, 65 °C for 30 seconds, and 72 °C for 3 minutes, followed by a final elongation at 72 °C for 5 minutes. The amplified product was electrophoresed on 2% TBE agarose gel, and stained with ethidium bromide. The band was isolated, purified using QIA quick Gel extraction kit (Qiagen), and sequenced.

D. Primer Extension Analysis

When progressive 5′ RACE reactions could no longer amplify the additional upstream region from the 60114S primer site, we performed primer extension from the putative 5′-end. PE60033 primer (Table 1) was end-labeled with [$\gamma$-$^{32}$P]dATP and was annealed to 30 μg of total RNA prepared from NHNE cells at 42 °C for 30 minutes, followed by 10 minutes of incubation at room temperature. Extension was performed as described in the Primer extension kit (Promega) using 50 U of Avian Myeloblastosis Virus-Reverse

Transcriptase for 50 minutes at 42 °C. The reaction was halted by the addition of loading dye (98% formamide, 10 mM EDTA, 0.1% xylene xyanol, 0.1% bromophenol blue). ΦX174 DNA/*Hin*fI dephosphorylated marker was labeled with [$\gamma$-$^{32}$P]dATP. Manual sequencing of the fragment was performed using a T7 Sequenase Version 2.0 kit (Amersham Pharmacia Biotech). The sample was denatured for 10 minutes at 90 °C before loading onto 6% sequencing gel. The gel was then vacuum-dried and auto-radiographed for 4 days at –70 °C.

E. Ribonuclease Protection Analysis

In order to verify the transcription start site location as determined by 5' RACE-PCR and primer extension analysis, RNase protection analysis was performed. An RNA probe was amplified from Bacteria artificial chromosome (BAC) clone, RP11-503G7, including the human chromosome 12q24.3 locus, using RP60364 and RP59748 primers (Table 1). This 525 bp of PCR product was incorporated into the T7 promoter, and an [$\alpha$-$^{32}$P]UTP-labeled RNA probe was generated using the MAXIscript$^{TM}$ *in vitro* Transcription kit (Ambion). For analysis, 20 μg of total RNA was hybridized with this probe at 42 °C overnight. The RNA-RNA template was digested for 30 minutes at 37 °C with 0.5 U of

RNase A/T1, precipitated, and run on 6% sequencing gel with a [γ-$^{32}$P]dATP-labeled ΦX174 DNA/*Hinf*I dephosphorylated marker for purposes of size determination.

F. *In vitro* Transcription and Translation

The 60114S/57326AS RT-PCR product covers + 8/+ 2794 from the transcription start site, and was predicted to synthesize 37 kD of peptide from methionine in Kozak sequence at + 1773. *In vitro* transcription and translation experiments were performed with $T_NT$ Coupled Reticulocyte Lysate System (Promega) in accordance with manufacturer's instructions. Amino acid mixture lacking methionine, supplemented with [$^{35}$S]methionine, rabbit reticulocyte lysate, and SP6 RNA Polymerase was used, and incubated for 90 minutes at 37 °C. Translation products were separated by 15% SDS-PAGE, and gels were dried and subjected to autoradiography.

G. Northern Hybridization using new sequence probe

20 μg of total RNA was separated on 1.0% agarose gel containing 2.2 M formaldehyde, and was then transferred to a positively charged nylon Hybond-N+ membrane (Amersham Pharmacia Biotech). The cDNA probe fragment was amplified from NHNE mRNA using B1 and B2 primers (Table 1), and was labeled with [α-$^{32}$P]dCTP using a Random hexamer

labeling kit (Promega). The labeled probe was purified by filtration through a Sephadex G-50 (Amersham Pharmacia Biotech) spin column, and added to 20 ml of hybridization buffer containing 6X SSS, 5X Denhardt's solution, 1% SDS, 50% Formamide, and 2 mg of boiled salmon sperm DNA. The membrane was hybridized at 42 °C for 16 hours, and autoradiographed on X-ray film at –70 °C for 12 hours.

H. Zooblot analysis

A Zooblot was purchased from Seegene Inc., and probed with a $^{32}$P-labeled 55958S-55433AS fragment (Table 1). Hybridization and washing were performed according to the manufacturer's instructions.

I. Sequence Analysis

Nucleotide and predicted amino acid sequence homology searches with sequences in the GenBank database were conducted using NCBI′s BLATN and BLASTX programs. Domains in amino acid sequence were searched using using SMART program (http://smart.embl-heidelberg.de).

2. PGE2-induced *MUC8* gene expression

A. Inhibitor, Antibodies and dominant negative vector

PD98059 and anti-α-tubulin antibody were purchased from Calbiochem, and anti-phospho p44/42 MAP kinase (Thr202/Tyr204) antibody, anti-phospho-p38 MAP kinase (Thr180/Tyr182) antibody, anti-phospho-SAPK/JNK MAP kinase (Thr183/Tyr185) antibody, anti-phospho-RSK1 (Ser380) antibody, and anti-phospho-CREB (Ser133) antibody were purchased from Cell Signaling. Plasmid encoding kinase-deficient MEK1 mutant (*pcDNA5-MEK1DN*) was kindly provided by Dr. Jian-Dong Li (House Ear Institute, Los Angeles, CA, USA).

B. Real-Time Quantitative PCR of MUC8

Primers and probes were designed using PerkinElmer Life Sciences Primer Express® software, purchased from PE Biosystems. Commercial reagents (TaqMan PCR Universal PCR Master Mix, PE Biosystems) and conditions were applied according to the manufacturer's protocol. One microgram of cDNA (reverse transcription mixture), oligonucleotides at a final concentration of 800 nM of primers, and 200 nM TaqMan hybridization probe were used in a 25 μl volume. The probe of real-time PCR was labeled

with carboxy fluorescein (FAM) at the 5'-end and with the quencher carboxytetramethylrhodamine (TAMRA) at the 3'-end. The following primers and TaqMan probes were used: *MUC8*, forward (5'-TAACCCAATGCCACTCCTTC-3') and reverse (5'-GGAGTGTAACCTGGCTGCTC-3') and TaqMan probe (FAM-GGTTAGG GCTGACCACAGAA-TAMRA). *β2M*, forward (5'-CGCTCCGTGGCCTTAGC-3') and reverse (5'-GAGTACGCTGGATAGCCTCCA-3') and TaqMan probe (FAM-TGCTCGCGCTACTCTCTCTTTCTGGC-TAMRA). Real-time reverse transcription-PCR was performed on a PE Biosystems ABI PRISM® 7700 Sequence Detection System (Foster City, CA, USA). The thermocycler parameters were 50 °C for 2 min, 95 °C for 10 min, followed by 40 cycles of 95 °C for 15s and 60 °C for 1 min. All reaction were performed in triplicate. Relative quantity of *MUC8* mRNA was obtained using a comparative threshold method, and results were normalized against *β2M* as an endogenous control.

C. Western Blot Analyses

NCI-H292 cells were grown to confluence in 6 well plates. After treating with 10 nM PGE2, cells were lysed with 2X Lysis buffer [250 nM Tris-Cl (pH 6.5), 2% SDS, 4% β-

mercaptoethanol, 0.02% bromo phenol blue, and 10% glycerol]. Equal amounts of whole cell lysates were resolved by 10% SDS-PAGE, and then transferred to a polyvinyl difluoride membrane (PVDF; Millipore). Membranes were blocked with 5% skimmed milk in Tris-buffered saline [TBS; 50 mM Tris-Cl (pH 7.5) and 150 mM NaCl] for 2 hours at room temperature. And, blots were then incubated overnight with primary antibodies in TTBS (0.5% Tween 20 in TBS). After washing with TTBS, the blots were further incubated for 1 hour at room temperature with anti-rabbit or anti-mouse antibody (Cell Signaling) in TTBS, and then visualized by ECL (Amersham-Pharmacia).

D. *MUC8* Promoter Cloning

A BAC clone (clone number: RP11-503G7) was excised with *Kpn*I at – 1644 and *Sac*I at + 87 from the *MUC8* transcription start site, and ligated to pGL3-basic vector (Promega). The resultant pGL3-MUC8 (- 1644 / + 87) construct was excised with *Kpn*I at – 1644, and then, digested with *Bst*EII at – 1190, *Sma*I at – 973, and *Sma*I at –549. Each fragment was followed using the Klenow (Promega) expansion and ligation reactions. These constructs were designated pGL3-MUC8 (- 1190 / + 87), pGL3-MUC8 (- 973 / + 87), and pGL3-MUC8 (- 549 / + 87). For the point mutation of the

CRE site on the – 803 promoter region, the following primers were used: CREM1 sense, 5'-ATAACAAC**ACT**CGCACTTCCGCCC-3'; CREM1 antisense, 5'-GGCGGA AGTGCG**AGT**GTTGTTAT-3'; CREM2 sense, 5'-ATAACAACTG**GT**GCACTTCC GCCC-3'; and CREM2 antisense, 5'-GGGCGGAAGTGC**AC**CAGTTGTTAT-3'. Site-specific mutations (underlined) were corroborated by DNA sequencing.

E. Electro mobility shift analysis

PGE2-treated NCI- H292 cells were washed with ice-cold PBS and pelleted. Pellets were resuspended in cell homogenization buffer (10 mM HEPES (pH 7.9), 1.5 mM $MgCl_2$, 10 mM KCl, 0.5 mM DTT, and 0.5 mM phenylmethylsulfonylfluoride), incubated for 10 min on ice and centrifuged. Cells were then resuspended in cell homogenization buffer containing 0.05% (V/V) NP-40, and then homogenized. Nuclei were the pelleted, and resuspended in cell resuspension buffer (40 mM HEPES (pH 7.9), 0.4 M KCl, 1 mM DTT, 10% (V/V) glycerol, 0.1 mM phenylmethylsulfonylfluoride, 0.1% (W/V) aprotinin, and 0.3 M NaCl). This nuclear extract was then centrifuged at 24,000 RPM for 15 min at 4 °C, and the supernatant was aliquoted and stored at –70 °C. For EMSA, oligonucleotides corresponding to the consensus CRE sequence (5'-

AGAGATTGCCTGACGTCAGAGAGCTAG-3'), and the MUC8 CRE sequence (5'-AACAGATAACAACTGACGCACTTCCGCCCG-3') were synthesized, annealed, and end-labeled with [γ-$^{32}$P]ATP using T4-polynucleotidekinase (Promega). Nuclear extract was incubated at room temperature for 30 min with the $^{32}$P-labeled CRE probes in binding buffer (Promega). Oligo-nuclear protein complexes were separated from the probes by electrophoresis through 6% nondenaturing polyacrylamide gel in 0.5X Tris-borate-EDTA (TBE) buffer. Supershift experiment was conducted using 2 μl of anti-phospho-CREB antibody (Cell Signaling). The gel was dried and autoradiographed on Phosphor Imager.

F. Transient transfection and luciferase analysis

NCI-H292 cells were transiently transfected with pGL3-basic, pGL3-MUC8 (- 1644 / + 87), pGL3-MUC8 (- 1190 / + 87), pGL3-MUC8 (- 973 / + 87), pGL3-MUC8 (- 549 / + 87), pGL3-MUC8 CREM1 and pGL3-MUC8 CREM2 constructs using a FuGENE6 Transfection Reagent (Roche Applied Science) according to the manufacturer′s instructions, incubated for 48 hours, treated with 10 nM of prostaglandin E2 (PGE2) for 24 hours, harvested, and assayed for luciferase activity using a luciferase assay system (Promega), according to the manufacturer′s instructions. β-Galactosidase activity was also

assayed to standardize sample transfection efficiencies. To confirm that the luciferase activity of each construct was from PGE2, we assayed the activity of each construct in the absence of PGE2.

# III. RESULTS

## 1. *MUC8* full cDNA sequence and protein structure

### A. RT-PCR, 5′ RACE-PCR and Sequence Determination of *MUC8* Coding Region

It has been reported a partial cDNA sequence of *MUC8* in GenBank™ (accession number U14383). To amplify and confirm the entire coding sequences of *MUC8* cDNA, the overlapped 4 pair-primers of A1-A2, B1-B2, C1-C2, and D1-D2 were designed on the basis of GenBank™ accession number AC079031 (Fig. 2). Total RNAs were isolated from NHNE cells, and RT-PCR was performed using primer pairs shown in Table 1. Both B1-B2 and C1-C2 products were amplified and sequenced, composing of 5135 bp. In addition, we further confirmed upstream of 2622 bp (60114S-57335AS product) and downstream of 1137 bp (D1-E1 product) by serial PCRs. The latter part sequences were matched with a reported *MUC8* partial sequence (U14383). Taken together with our sequence and U14383 sequence, the Exon-Intron boundary and 6354 bp of intron were determined by comparing

the genomic and cDNA nucleotide sequences. Interestingly, we found that an intron spanning 6354 bp located between Exon1 and Exon2. We made four probes for Northern analysis corresponding to this region, and found no specific band, suggesting that this 6354 bp region is truly intronic sequences (Data not shown). As shown in Table 2, splice acceptor and donor sequences agree with the 'GT-AG' rule proposed by Mount[27]. However, we could not amplify the upstream sequences of AC079031 (nucleotide 60114) using RT-PCR. Therefore, 5′-RACE was employed using RA59840 primer and an adaptor primer, and we amplify about 200 bp PCR fragment. The overall cDNA size predicted by 5′ RACE was about 10.2 kb. Large fragment of about 9 kb might be expected exon 1 including tentative 5′-untranslated region (UTR). In this cDNA sequence, we found that there are three tandem repeat regions including 26 known 41 bp repeat unit[20]. The newly-found tandem repeats are 61-nucleotide, and 12-nucleotide. The tandem repeat component of 61-nucleotide is TGGGGTTCCCGACAGGCCCCGGC TTCAGGCAGGGCTGGATTCAGGCAGCAGCCCACACCCC, and is repeated 7 times, and its deduced amino acid sequence shows three types of analogous polypeptides (Fig. 3A). Also, the tandem repeat component of 12-nucleotide is TCCCGGTTATGT, and is repeated 42 times, and shows fifteen types of analogous polypeptide (Fig. 3B). These

tandem repeat sequences have not been observed in any *MUC* cDNAs known thus far, and the combined contents of *MUC* protein-specific Ser, Thr, Pro, Gly, and Ala are about 60% and 50% of the total amino acids.

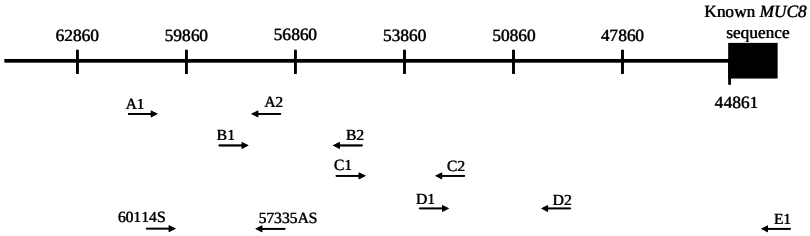


**Fig. 2. PCR strategy for sequencing *MUC8* full cDNA.**

**Table 1. Primers used for various application as described in the text**

| Name | Sequence | Location[a] |
|---|---|---|
| RT-PCR | | |
| A1 | GCCCCCAAGTCCCACAAGAAG | 60184/60163 |
| A2 | CTCAGCTGTGTGTGCTCACACCT | 57424/57447 |
| B1[b] | GAATCGGTGTGTGTGTCAATGGA | 57493/57470 |
| B2[b] | AAGGAGTGGCATTGGGTTAGG | 55938/55959 |
| C1[d] | CCTAACCCAATGCCACTCCTT | 55958/55937 |
| C2 | GGCCAGTGTTTGGATGTAGAGA | 52359/52381 |
| D1 | TGCGGCCGTGAGCGCTCACATTCTT | 53810/53785 |
| D2 | ACGCCACTCATGTGAGCAGCT | 49213/49234 |
| 60114S | CTGCTGCCAGACACGGAGCGCA | 60114/60092 |
| 57335AS | GTGGTGCATTCAAACCCATGGTT | 57335/57358 |
| E1 | TCAAGCACAGGGTGCCCTGCCAGGGAA | 43379/43406 |
| 5'RACE-PCR | | |
| RA59840 | AAGAGGGTAGTCTATCTCAGCCG | 59840/59863 |
| Primer extension | | |
| PE60033 | GCGCCTGCCTCGCAACCGCAGCCG | 60033/60057 |
| RNase protection | | |
| RP60364 | CTCCGGTCTCGCGGAACCGAAGT | 60364/60341 |
| RP59748 | CTCTCTGTGACACCTCCCCGTTCGATT | 59748/59775 |
| *In vitro* transcription/translation | | |
| 60114S | CTGCTGCCAGACACGGAGCGCA | + 8/+29[c] |
| 57326 | CCTGACTGTGGTGCATTCCAA | +2794/+2773[c] |
| Zooblot | | |
| 55433AS | GAGCACCTGTGCAAGCTTGTG | 55433/55412 |

[a] Base on GenBank™ accession number AC079031
[b] Also used to generate the northern probe
[c] Numbered from transcription start site
[d] Also Used to generate the zooblot probe

**Table 2. Exon-Intron Boundary in between Exon 1 and Exon 2**

| 5'-Splice donor (position on AC079031) | Intron size (bp) | 3'-Splice acceptor (position on AC079031) |
|---|---|---|
| - - - - -CACAGGGGAC<u>A</u>/gt (51336) | 6354 | ag/<u>C</u>AAAGGATGCC- - - - - (44981) |
| Exon 1 | | Exon 2 |

A WGSRQAPASGRAGFRQQPTPRGVPDRLRLQAGLDSGSTPYPVGFLTGPGIRQGWIQAAPRTR
WGSRHALASGRAGFRQHPVPAGVPDMPRLQAGLDSGSTPYPVGFLTGPGFRQGWIQAAPRTR
WGSRHALASAGLDSG

B PGYVPVMSGLR
PGYVQVMSRLC
SGYVLVRSRLC
PGQVRVRFRLF
LGYVWLG
PGYVRVMSQLG
PGYVQVVSGLG
PGYFQVMSRLG
PGYVRVTSRLR
PGYVQVRSRLG
PGYVALG
PGQVRVMSGLC
PGQVPVISGLC
PGQVLLCSGYF
FQVRSQLG

**Fig. 3. Deduced amino acids of novel 61-nt (A) and 12-nt (B) tandem repeats in *MUC8* coding region.**

B. Primer Extension and Ribonuclease Protection Analysis

To determine the transcription start sites for *MUC8* gene, primer extension and RPA were carried out using a radiolabeled probe and RNAs isolated from NHNE cells. For the Primer extension, we used PE60033 primer (Table 1), which is approximately 100 bp upstream of the 5′-end of our RACE-PCR product as estimated from agarose gel. The primer extension reaction yielded a product of 85 bp (Fig. 4A). Alternatively, RNase protection analysis was performed using PCR product probe amplified from pBeloBAC11 clone. The template was total RNA of NHNE cells. RNA protected the probe fragment, indicating two putative transcription start sites. One site is matched with the site identified by primer extension and the other site is at 2 bp upstream (Fig. 4B). Of these two sites, we designated the one that was in accordance with one of the RNase protection analysis results as the major transcription start site. However, we could not exclude the possibility that the other site may be involved as another transcription start site. Taken together, total *MUC8* transcript size is estimated as 10292 bp. This total *MUC8* cDNA sequence was deposited into GenBank™ accession number BK005559.

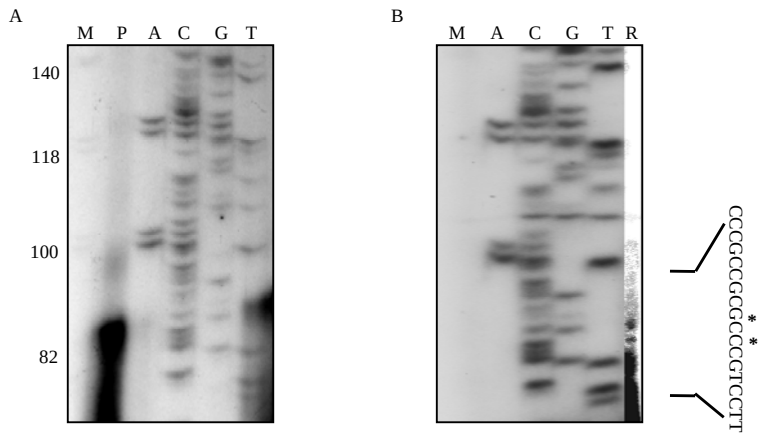


**Fig. 4. Primer extension (A) and RNase protection assay (B) of *MUC8* transcript**

C. Transcription and Translation in vitro

A methionine residue of ATG at nucleotide +1734 is contained within the optimal context for initiation of translation, GCCGCCGATGC, as described by Kozak[25]. To confirm that *MUC8* translation starts at this methionine, we synthesized PCR product (60114S-57326AS, Table 1) covering +8/+2794, and predicted about 37 kDa peptide product from this candidate translation start site. *In vitro* transcription and translation result showed expected 37 kDa band (Fig. 5). This result suggests that MUC8 protein starts at +1734 position. Consequently, 5′-UTR is from +1 to +1733. *MUC8* gene transcript is total 10292 bp, and is composed of 1733 bp of 5′-UTR, 8100 bp of open reading frame (ORF), and 459 bp of 3′-UTR. *MUC8* total transcript sequence and amino

acid sequence from above Met are seen in Fig. 6. *MUC8*′s ORF is derived from 9 exons and 8 introns. *MUC8* protein is composed of 2699 amino acids, and its molecular weight is 290 kDa according to Georgetown University′s Protein Information Source (http://www-nbrf.georgetown.edu). The combined content of Ser, Thr, Pro, Gly, and Ala is 44.9% of the total amino acids. Also, this deduced amino acid sequence doesn′t have homology to any *MUC* proteins as assessed by National Center for Biotechnology Information (NCBI)′s BLASTX program.

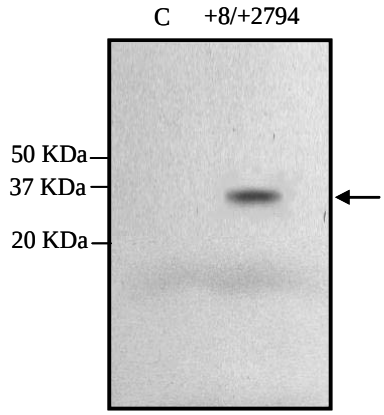


**Fig. 5. *In vitro* transcription and translation using 60114S / 57326AS product covering +8/+2794 from the transcription start site.**

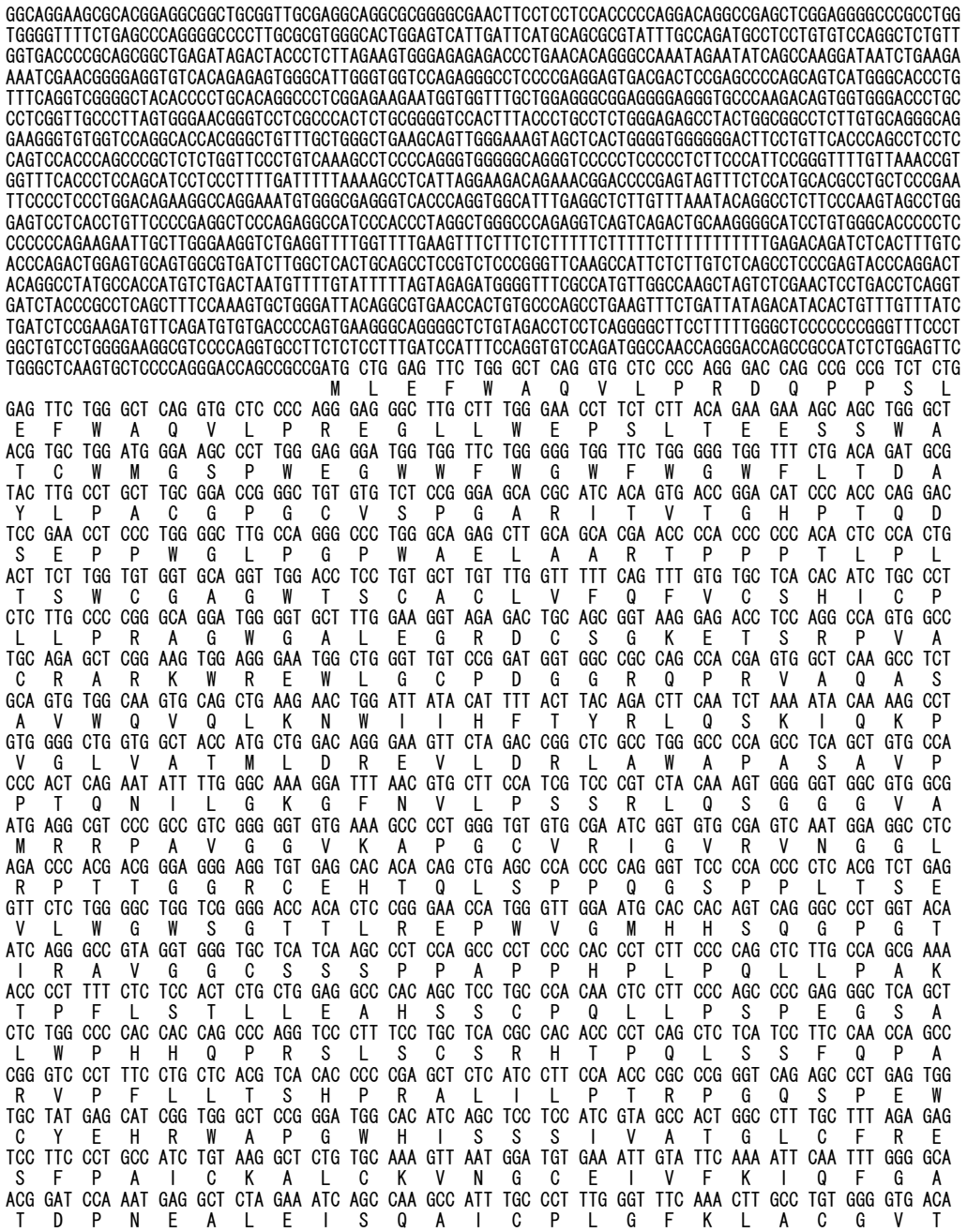


**Fig. 6. *MUC8* 's full cDNA sequence and deduced amino acid sequence**

```
GCC AAA AAG GAG AAA GTA TGT TCC ATT TCT TGC AGC CCC CAG CCC CAT CTG CGA GCC GGC TGT CAG CTC CAG CCC
 A   K   K   E   K   V   C   S   I   S   C   S   P   Q   P   H   L   R   A   G   C   Q   L   Q   P
TCC GCA TGC ACC CCT CTT CCT TCC CGT AAT GGC ACC TCA GTG TAT TTT GGA TGC ATA GTC TGC TCT TGT GGG TGC
 S   A   C   T   P   L   P   S   R   N   G   T   S   V   Y   F   G   C   I   V   C   S   C   G   C
CGT CTG GTG ACA CAG CAG GTC CTA AGG CCC CAT GCT TCC CTA GAA AAG GAA CAA AGG AAC GTC AGT CAG CTC TTG
 R   L   V   T   Q   Q   V   L   R   P   H   A   S   L   E   K   E   Q   R   N   V   S   Q   L   L
CTT GGA GAA CGC TGT GTA ACA AAC AGC CCC TGG GCC CTC AGT GAC AAT GCC TTG CTC ACA GGG TTG CAG GGT GAC
 L   G   E   R   C   V   T   N   S   P   W   A   L   S   D   N   A   L   L   T   G   L   Q   G   D
AAG AGG GGC TTT GCT TCT TCA TTT GGA CTT TGG TTA GGC CCA GCT ATA ATC CAC TGG CCT CTT TCT GTG GGA TGC
 K   R   G   F   A   S   S   F   G   L   W   L   G   P   A   I   I   H   W   P   L   S   V   G   C
AGG CGT TCA TTC TCC CTC AGT GGC TCA GGA GGC AGA GCA GAG CCA TAT AAA CCT AGG AGA AGC CCG TGC TTG AAG
 R   R   S   F   S   L   S   G   S   G   G   R   A   E   P   Y   K   P   R   R   S   P   C   L   K
CCT CAT GTT GTG TCT GTC AAG GAA GTT TCA AGG CTA GGA CCA GCC TCC ACG GGG CAG AGA AGT CGT GCT TTC TGC
 P   H   V   V   S   V   K   E   V   S   R   L   G   P   A   S   T   G   Q   R   S   R   A   F   C
TCT GGT GGG GTG AGA TGG CTC AGT TTG TCA TGC AGG CGA CCC AGG CGA CAC CAG TCA GGT GGC CTC TTC CTG GCA
 S   G   G   V   R   W   L   S   L   S   C   R   R   P   R   R   H   Q   S   G   G   L   F   L   A
TTG CAG TTA GAA TGT GCC TTG AGC CAC ATG TCA AGG CAA TTG AGT GTT TGG AGT CCT CAA CGT GCC CCC TTC CAG
 L   Q   L   E   C   A   L   S   H   M   S   R   Q   L   S   V   W   S   P   Q   R   A   P   F   Q
TCA TCC TGC TCC CGA GGA TGT GCT GTT GCC TGG TTC CGA GCC TGC TGC AGC TCC GCG GGC CGC CCC CTC CCT GTT
 S   S   C   S   R   G   C   A   V   A   W   F   R   A   C   C   S   S   A   G   R   P   L   P   V
CAC CCA GGG GAG CAG GCG TGT TCC CTC CGC AGG GGC TTG AGA CCT GCC GTC CTT TCC CCT GGA CCC TCC CTC TCC
 H   P   G   E   Q   A   C   S   L   R   R   G   L   R   P   A   V   L   S   P   G   P   S   L   S
CCC AAG CCC CAA CCC AAT GCC ACT CCT TCC CGA GGC CGA TGG TGG CTT TGC GTG AGG TGG GCC CTG CTG AGC AGC
 P   K   P   Q   P   N   A   T   P   S   R   G   R   W   W   L   C   V   R   W   A   L   L   S   S
AGA GAA TTT CTT AGA ATT TTC ATC GCC AGA TGG CTC TGG GTT AGG GCT GAC CAC AGA AAA CCA GCA GGC ACG TTG
 R   E   F   L   R   I   F   I   A   R   W   L   W   V   R   A   D   H   R   K   P   A   G   T   L
GAA AGC AGG AGT GGA ATG CAG AAG GCA GAG GGA GCC ACA CGG AGC AGC CAG GTT ACA CTC CTG CCG TCG TGT GAC
 E   S   R   S   G   M   Q   K   A   E   G   A   T   R   S   S   Q   V   T   L   L   P   S   C   D
TGC TGC AGA CCC CCT GTC CCC GCA CCT CAG CGG GCT TGG GCC GTT CCA GGC CTG CAG CAC CGT GGG CTC CAG CCA
 C   C   R   P   P   V   P   A   P   Q   R   A   W   A   V   P   G   L   Q   H   R   G   L   Q   P
CCT CCT GCC ATT GCC CCT CCC AGT CTG TGT CAG GAG CAC TGG AGT CTC CCA GTC AGT TCC TCA GGA CGC TAT GTC
 P   P   A   I   A   P   P   S   L   C   Q   E   H   W   S   L   P   V   S   S   S   G   R   Y   V
ACC GTC ATG ATG TCC GTG AGC TCC CAT GCG CCT GGC TCT CCC CGT TAC CAC ACT TAC CTT TCT CTC CTG GCA GCC
 T   V   M   M   S   V   S   S   H   A   P   G   S   P   R   Y   H   T   Y   L   S   L   L   A   A
ACT CCC TGG CCC TCC TTG GAG AGT GTT TCA GCT GGC CAT GCT CCG TAT GCA GTG AGA GAG TGC ACA AGC TTG CAC
 T   P   W   P   S   L   E   S   V   S   A   G   H   A   P   Y   A   V   R   E   C   T   S   L   H
AGG TGC TCA CAC CAG AAA GGG ACG TTG CAG GTT GGG AGG AGG GTC CTT GTG CCA GGC CCT GGC CCC CGA GTC TGC
 R   C   S   H   Q   K   G   T   L   Q   V   G   R   R   V   L   V   P   G   P   G   P   R   V   C
CCC TTG CAG TGG GGT GCC CAG CCT GGT GTG GGG TCA GGC TGT GGG AAG GAG GAG AGG CCA GGG ACA AGC AGC CGG
 P   L   Q   W   G   A   Q   P   G   V   G   S   G   C   G   K   E   E   R   P   G   T   S   S   R
GAT GAG CTT GGG CTC TTT CCC TTG AAA GTG GCA GAA ATG CAA ACC CAA AGC AGC CCG TAC CTG TGG GGT TCC CGA
 D   E   L   G   L   F   P   L   K   V   A   E   M   Q   T   Q   S   S   P   Y   L   W   G   S   R
CAG GCC CCG GCT TCA GGC AGG GCT GGA TTC AGG CAG CAG CCC ACA CCC CGT GGG GTT CCC GAC AGG CTC CGG CTT
 Q   A   P   A   S   G   R   A   G   F   R   Q   Q   P   T   P   R   G   V   P   D   R   L   R   L
CAG GCA GGG CTG GAT TCA GGC AGC ACC CCG TAC CCC GTG GGG TTC CTG ACA GGC CCC GGC TTC AGG CAG GGC TGG
 Q   A   G   L   D   S   G   S   T   P   Y   P   V   G   F   L   T   G   P   G   F   R   Q   G   W
ATT CAG GCA GCA CCC CGT ACC CGC TGG GGT TCC CGA CAC GCC CTG GCT TCA GGC AGG GCT GGA TTC AGG CAG CAC
 I   Q   A   A   P   R   T   R   W   G   S   R   H   A   L   A   S   G   R   A   G   F   R   Q   H
CCC GTA CCC GCT GGG GTT CCC GAC ATG CCC CGG CTT CAG GCA GGG CTG GAT TCA GGC AGC ACC CCG TAC CCC GTG
 P   V   P   A   G   V   P   D   M   P   R   L   Q   A   G   L   D   S   G   S   T   P   Y   P   V
GGG TTC CTG ACA GGC CCC GGC TTC AGG CAG GGC TGG ATT CAG GCA GCA CCC CGT ACC CGC TGG GGT TCC CGA CAC
 G   F   L   T   G   P   G   F   R   Q   G   W   I   Q   A   A   P   R   T   R   W   G   S   R   H
GCC CTG GCT TCA GCA GGG CTG GAT TCA GGG ACA GGG GTT CCA GCA GCC TCA GCT TTA AGT GGG GAG TTA AGG GGT
 A   L   A   S   A   G   L   D   S   G   T   G   V   P   A   A   S   A   L   S   G   E   L   R   G
TCA ATG TCG ACA GGA AAA ATA GAA TTT TAT GAA GCT CTG AGC AAA GTC CCG AGT CCT CTC CGA ATG GGA TCT TCA
 S   M   S   T   G   K   I   E   F   Y   E   A   L   S   K   V   P   S   P   L   R   M   G   S   S
GGC TGC CAA GTC CCT GCA GCT CGG GGG TGT GAA GCT GTG ACT GAC AGG CCT CGG TCT ACC GAG ACC ACA GAC CGA
 G   C   Q   V   P   A   A   R   G   C   E   A   V   T   D   R   P   R   S   T   E   T   T   D   R
GGA GGC AGA GGC GGG CCC CAA AAC CAG ATG GGA ACG GAC ACC GTA GGG GGC CAG GTC CCC ACC CGA GGC CCA CCT
 G   G   R   G   G   P   Q   N   Q   M   G   T   D   T   V   G   G   Q   V   P   T   R   G   P   P
```

**Fig. 6. Continued**

```
GCC TCA CTG GAC CAG CAC CGG GGC TCG AGT CTG CAC CCC ACC TAC CCA CTT CCC CGC AGT GTC CCG GCC ACC CCA
 A   S   L   D   Q   H   R   G   S   S   L   H   P   T   Y   P   L   P   R   S   V   P   A   T   P
CGC AGC TCT CTG GAC ATC GTG TCC CGG CTG CCC CAC ACA AGT CTC CGA GGC CCG GGC AGC TCT CTG GAC ACC TTT
 R   S   S   L   D   I   V   S   R   L   P   H   T   S   L   R   G   P   G   S   S   L   D   T   F
CTG CTA CTT GTG CCA CAC TGT TGG TTC TTG GGC CCG AGC TGT TCC CAC CTC CCA TGC CCA GCA GGT CCC GGT TAT
 L   L   L   V   P   H   C   W   F   L   G   P   S   C   S   H   L   P   C   P   A   G   P   G   Y
GTC CCG GTT ATG TCC GGT TTA CGT CCT GGT TAT GTC CAG GTT ATG TCC CGG TTG TGT TCC GGT TAT GTC CTG GTT
 V   P   V   M   S   G   L   R   P   G   Y   V   Q   V   M   S   R   L   C   S   G   Y   V   L   V
AGG TCC CGG TTA TGT CCG GGT CAG GTC CGG GTT AGG TTC CGG TTA TTT CTG GGT TAT GTC TGG TTA GGT CCC GGT
 R   S   R   L   C   P   G   Q   V   R   V   R   F   R   L   F   L   G   Y   V   W   L   G   P   G
TAT GTC CGG GTT ATG TCC CAG TTA GGT CCC GGT TAT GTC CAG GTT GTG TCC GGG TTA GGT CCC GGT TAT TTC CAG
 Y   V   R   V   M   S   Q   L   G   P   G   Y   V   Q   V   V   S   G   L   G   P   G   Y   F   Q
GTT ATG TCC AGG TTA GGT CCT GGT TAT GTC CGG GTT ACA TCC AGG TTA CGT CCT GGT TAC GTC CAG GTT AGG TCC
 V   M   S   R   L   G   P   G   Y   V   R   V   T   S   R   L   R   P   G   Y   V   Q   V   R   S
CGG TTA GGT CCG GGT TAT GTC GCG TTA GGT CCC GGT CAG GTC CGG GTT ATG TCC GGG TTA TGT CCG GGT CAG GTT
 R   L   G   P   G   Y   V   A   L   G   P   G   Q   V   R   V   M   S   G   L   C   P   G   Q   V
CCG GTT ATT TCT GGG TTA TGT CCG GGT CAG GTC CTG TTA TGT TCC GGT TAT TTC CAG GTT AGG TCC CAG TTA GGT
 P   V   I   S   G   L   C   P   G   Q   V   L   L   C   S   G   Y   F   Q   V   R   S   Q   L   G
CTG GGC GCA TCT GTC AGT CAT TGC TGT GTC CCA GCG ATG GCA AAT CTC AGT GTC TGC GGC CGT GAG CGC TCA CAT
 L   G   A   S   V   S   H   C   C   V   P   A   M   A   N   L   S   V   C   G   R   E   R   S   H
TCT TCT CAG CTC TGC AGG TCA GCT GCA GGG CTC CCA TCC ACC CCA GGG AGG AGT CAT GCC CGG CTG GAG CCT GGG
 S   S   Q   L   C   R   S   A   A   G   L   P   S   T   P   G   R   S   H   A   R   L   E   P   G
AGA CTG TGT TTC CCC ATG AGA GAC CAG GGG ACG GGT GAG CCA GGC CTG GGA GGT ATA AGG CAA TGT CTG CAA GGG
 R   L   C   F   P   M   R   D   Q   G   T   G   E   P   G   L   G   G   I   R   Q   C   L   Q   G
CGT CTG GGG AAG ATG CCC CTC TGT GAT GAA GAG AGA AGG ACA CAC AAA GAT GTC ATC CGT TTT TCA TCC TGC CCA
 R   L   G   K   M   P   L   C   D   E   E   R   R   T   H   K   D   V   I   R   F   S   S   C   P
CTC GCC CCA CTC CTG GCT TTG AAT GCT GCC CTG GGA GCA TAT GCT CCC GAG AGC TGT GGC AGC CTT TTT GCA CTC
 L   A   P   L   L   A   L   N   A   A   L   G   A   Y   A   P   E   S   C   G   S   L   F   A   L
ATG AGG CAG CGA GCT CCA GGA GGA AGA GCC AAC GTG CTG ACA CAG TGG GGC AGA AAG ACA GAG ACA ACA GAA TGG
 M   R   Q   R   A   P   G   G   R   A   N   V   L   T   Q   W   G   R   K   T   E   T   T   E   W
GGA GCT GGC GGC CTG GAC CAG CTA CCT CCA GAC ATT TCA TCC GGG AAG GAG CTG TCC GTA CAG TTT GAG CTG CTC
 G   A   G   G   L   D   Q   L   P   P   D   I   S   S   G   K   E   L   S   V   Q   F   E   L   L
TTT GTT GGA TGT TCT GTT ACT TGC ATC CTA TGG ATT TGG CCT CTG GCT GCA ACG GGA CAG CAT CAG TTT GGG GAA
 F   V   G   C   S   V   T   C   I   L   W   I   W   P   L   A   A   T   G   Q   H   Q   F   G   E
ACA AAA ACA TGT CTG TAT TCA GGC AGA GAG GCA TTT AAC GTG GGG ACA GGT ACA CAA ATG TCG GAA GGG CTG TGC
 T   K   T   C   L   Y   S   G   R   E   A   F   N   V   G   T   G   T   Q   M   S   E   G   L   C
GGC CAT CGC GGA CAC GAG GGG GCT GGG GAG CTC GGG ACA CTC TCA GTG CCA ACG GGA CTG GAC ACG AGG GGG CTG
 G   H   R   G   H   E   G   A   G   E   L   G   T   L   S   V   P   T   G   L   D   T   R   G   L
GGA TGC TCT CGT GCC AAT GGG ACC GGA CAT GAG GGG GCT GGG ACA CTC CCA GGG CCA ATG GGA CTG GCT CTG GTG
 G   C   S   R   A   N   G   T   G   H   E   G   A   G   T   L   P   G   P   M   G   L   A   L   V
ACC TGC AGA GCA GGA TTT TTG GCT GCC TCT GTG CAG GTG TTT CCA GAG GAG ACT GCA GTG TGT GTC CGA GTG GCG
 T   C   R   A   G   F   L   A   A   S   V   Q   V   F   P   E   E   T   A   V   C   V   R   V   A
GAG GTG AAG ATG ACC TGC CCT CAG TGT GGG TGG CAC CTT CCA GCT GGG GTC TGG AGA GAA CAG AAA CAG AAA AGG
 E   V   K   M   T   C   P   Q   C   G   W   H   L   P   A   G   V   W   R   E   Q   K   Q   K   R
GGA ATG TGT CTC TCC CCA CGG CAG CTG GGA TGC ACG CTT CCT CTC CTG CCC TCA GAC AGA ACT GCA GGC GCC CCA
 G   M   C   L   S   P   R   Q   L   G   C   T   L   P   L   L   P   S   D   R   T   A   G   A   P
GCC TTT GGG CTC AAG GAC TCG CAC TGG CAG CCC CGG GTC TCA CGC CTG TGG CCT CGG GCT GAG TCG CGC CAT CAG
 A   F   G   L   K   D   S   H   W   Q   P   R   V   S   R   L   W   P   R   A   E   S   R   H   Q
CTT CCC GGG TTC GGA TGC CTT CTG ACG TGG ACT GAG CCA CAC TAC CGG CAT CCG GTG TTG TCC AGC CTG CAG ACA
 L   P   G   F   G   C   L   L   T   W   T   E   P   H   Y   R   H   P   V   L   S   S   L   Q   T
GCC TGT TGT GGG ACT TCC CAG CCT CCA TCA TCT GGT TCC CCT AAC AGA CCC TGC TCA CTT CTC TCT CTG TGT CAT
 A   C   C   G   T   S   Q   P   P   S   S   G   S   P   N   R   P   C   S   L   L   S   L   C   H
GTT GGC TCT GTC TCT CTG GAG ACC CCT GAC AAA CAC AGT TGT GCT GGA GGT GAA GGA AGT GAG GAG GCT GTG TTA
 V   G   S   V   S   L   E   T   P   D   K   H   S   C   A   G   G   E   G   S   E   E   A   V   L
TGG GTT ACC GTG TGT CCT CAG TAT CAC TCA CTG GAG TCC CAA CCC CCA GAA CCC CAG AAT GTA ACA CTA TTG GAA
 W   V   T   V   C   P   Q   Y   H   S   L   E   S   Q   P   P   E   P   Q   N   V   T   L   L   E
ATT GGG TCT TTG CAG ATG AAA TTG CAA ATG TTA AAA GGA AGT TCC CAG ATT TGG CTG GTC TCT ACA TCC AAA CAC
 I   G   S   L   Q   M   K   L   Q   M   L   K   G   S   S   Q   I   W   L   V   S   T   S   K   H
TGG CCC CCT TGT GAA AAG ACA CAC AGG CCT AGG GGA GAC GGT GTG TGG GGA CGG GAG CAG AGG CTG CAG GGA CGC
 W   P   P   C   E   K   T   H   R   P   R   G   D   G   V   W   G   R   E   Q   R   L   Q   G   R
```

**Fig. 6. Continued**

```
AAC CAC ACA GGC GGG AAG GCC AAC ACC ACC AGG AAC AGG AAG AGA CAC GGG TCC TCC CCT GGA GCC TCC GGA GGC
 N   H   T   G   G   K   A   N   T   T   R   N   R   K   R   H   G   S   S   P   G   A   S   G   G
AGC GTG GCT CCA ATA ACA CTG GTT CTG GAC ATC AGG TCT CTA AAT CTC AAG GGA AGA CAT TTC TGT GCT TCA AGC
 S   V   A   P   I   T   L   V   L   D   I   R   S   L   N   L   K   G   R   H   F   C   A   S   S
TGT ACT TTT TTT TTC TTT TTG ACG GCG TCT TAC TCT GTC GCT CAG GCT GGA GTG CAG TGT CGC GAT CTC AGC TCA
 C   T   F   F   F   F   L   T   A   S   Y   S   V   A   Q   A   G   V   Q   C   R   D   L   S   S
CCA CCA TCT CTG CCT CCT GGG TTC AAG CGA TTC TCC TGC CTC AGC CTC CCA AGT AGC TGG GTT TAC AGG TGC TCA
 P   P   S   L   P   P   G   F   K   R   F   S   C   L   S   L   P   S   S   W   V   Y   R   C   S
CCA CTA TAC CTG GTT AAT TTT TGT ATT TTT AGT AAA GAC AGG GTT TCA CCA TGT TGG TCA GGC TGG TCT GGA ACT
 P   L   Y   L   V   N   F   C   I   F   S   K   D   R   V   S   P   C   W   S   G   W   S   G   T
CCT GAT GTC AAG GGA TCC ACC CGC CTC GGC CTC TCA AAT TGT GAG GAT TAC AGG CGT GAG CCA CTG CAC CCA ACC
 P   D   V   K   G   S   T   R   L   G   L   S   N   C   E   D   Y   R   R   E   P   L   H   P   T
TTC ACT GCA CTT TGC TAT GGA GCC TGG GAA ACG ACC ACA GGG TGC TCC AAG AAA GGG AGC TTC TAC CCC ACA CAG
 F   T   A   L   C   Y   G   A   W   E   T   T   T   G   C   S   K   K   G   S   F   Y   P   T   Q
ATG AGA GCC CAG CAG GAC AGC CTC AGG GAG GCC CAC AGT GCA CAG GGG CCA GGG CTG CTG ATG GGA GGG GGC CGC
 M   R   A   Q   Q   D   S   L   R   E   A   H   S   A   Q   G   P   G   L   L   M   G   G   G   R
TGG GTC CCT GGG GGT GCA GCC TCA CTG GTG TGG AGG CCG AAC AGC TTC CCT TCC AGT TGG CAG GGC CAC CTG GAA
 W   V   P   G   G   A   A   S   L   V   W   R   P   N   S   F   P   S   S   W   Q   G   H   L   E
GCC CAG CTG CAG GGA AGC CTG GGG CAC ACA GGG TGT GGA CGG CCG GCC CGC TCA TGC ACC GTG CCC CAG GGT GGG
 A   Q   L   Q   G   S   L   G   H   T   G   C   G   R   P   A   R   S   C   T   V   P   Q   G   G
GGT GCT GCC AGG AGG AAA TGT CAG GGG CCT GGG GCC TCG TGG ACA GTC TGG AAG TTT CTC TTT CCT GAA TGC ATA
 G   A   A   R   R   K   C   Q   G   P   G   A   S   W   T   V   W   K   F   L   F   P   E   C   I
ATC AGT CCT TCG AAG ATC CCA CAA ACC CAA ACC AAA CCC ACA GAA GCC GAC CAG CGT GTC GGA TGC AGC AGC TCC
 I   S   P   S   K   I   P   Q   T   Q   T   K   P   T   E   A   D   Q   R   V   G   C   S   S   S
GGT GGG AAC GGG CCC GGA AGC CAG TGT CTT TCC GGA GGA CAC AGG GGA CAC AAA GGA TGC CAG CCC GGA CGG AAC
 G   G   N   G   P   G   S   Q   C   L   S   G   G   H   R   G   H   K   G   C   Q   P   G   R   N
CAG TGC ACG TCC ACC ACG AGC TGC CCA CGT CCT CTC CAG GAA GGG ACC CGG GTC CAC GAG CTG CCC ACG TCC TCT
 Q   C   T   S   T   T   S   C   P   R   P   L   Q   E   G   T   R   V   H   E   L   P   T   S   S
CCA GGA AGG GAC CCG GGT CCA CGA GCT GCC CAC GTC CTC TCC AGG AAG GGA CCC GGG TCC ACG AGC TGC CCA CGT
 P   G   R   D   P   G   P   R   A   A   H   V   L   S   R   K   G   P   G   S   T   S   C   P   R
CCT CTC CAG GAG GGG ACA CCG GGT TCA CGA GCT GCC CAC GCC CTC TCC AGG AGG GGA CAC CGG GTT CAC GAG CTG
 P   L   Q   E   G   T   P   G   S   R   A   A   H   A   L   S   R   R   G   H   R   V   H   E   L
CCC ACG TCC TCT CCA GGA GGG GAC ACC GGG TTC ATG AGC TGC CCA CGC CCT TTC CAG GAA GGG ACC CCG GGT TCA
 P   T   S   S   P   G   G   D   T   G   F   M   S   C   P   R   P   F   Q   E   G   T   P   G   S
CGA GCT GCC CAC GTC CTC TCC AGG AAG GGA CCC CGG GTT CAC GAG CTG CCC ACG TCC TCT CCA GGA AGG GAC CCC
 R   A   A   H   V   L   S   R   K   G   P   R   V   H   E   L   P   T   S   S   P   G   R   D   P
GGG TTC ACG AGC TGC CCA CGT CCT CTC CAG GAA GGG ACC CGG GTC ACG AAC TGC CCA CGT CCT CTC CAG GAA GGG
 G   F   T   S   C   P   R   P   L   Q   E   G   T   R   V   T   N   C   P   R   P   L   Q   E   G
ACC CCG GGT TCA CGA GCT GCC CAC GTC CTC TCC AGG AGG GGA CAC CGG GTT CAC GAG CTG CCC ACG CCC TCT CCA
 T   P   G   S   R   A   A   H   V   L   S   R   R   G   H   R   V   H   E   L   P   T   P   S   P
GGA AGG GAC CCC GGG TTC ATG AGC TGC CCA CGT CCT CTC CAG GAA GGG ACC CGG GTC ACG AAC TGC CCA CGC CCT
 G   R   D   P   G   F   M   S   C   P   R   P   L   Q   E   G   T   R   V   T   N   C   P   R   P
CTC CAG GAG GGG ACC CGG GTC ACG AGC TGC CCA CGT CGT CTC CAG GAA GGG ACC CGG GTC ACG AGC TGC CCA CGT
 L   Q   E   G   T   R   V   T   S   C   P   R   R   L   Q   E   G   T   R   V   T   S   C   P   R
CCT CTC CAG GAA GGG ACC CGG GTC ACG AAC TGC CCA CGC GCT CTC CAG GAG GGG ACA CCG GGT TCA CGA GCT GCC
 P   L   Q   E   G   T   R   V   T   N   C   P   R   A   L   Q   E   G   T   P   G   S   R   A   A
CAC GCC CTC TCC AGG AAG GGA CCC CGG GTT CAC GAG CTG CCC ACG TCC TCT CCA GGA GGG GAC ACC GGG TTC ACG
 H   A   L   S   R   K   G   P   R   V   H   E   L   P   T   S   S   P   G   G   D   T   G   F   T
AGC TGC CCA CGT CCT CTC CAG GAG GGG ACA CCG GGT TCA CGA GCT GCC CAC GCC CTC TCC AGG AGG GGA CAC CGG
 S   C   P   R   P   L   Q   E   G   T   P   G   S   R   A   A   H   A   L   S   R   R   G   H   R
GTT CAC GAG CTG CCC ACG TCC TCT CCA GGA AGG GAC CCG GGT CAC GAG CTG CCC ACG TCC TCT CCA GGA GGG GAC
 V   H   E   L   P   T   S   S   P   G   R   D   P   G   H   E   L   P   T   S   S   P   G   G   D
ACC GGG TTC ACG AGC TGC CCA CGC ACT TTC CAG GAA GGG ACC CCG GGT TCA GGT CTC CTG CCG GCC CAC ATC GTG
 T   G   F   T   S   C   P   R   T   F   Q   E   G   T   P   G   S   G   L   L   P   A   H   I   V
CCT TTG TGT AAA TCA GAA GAA AGA TGAGGAACAGGCCCTCCTCTCTCTCCAGGCAGGCTTTGGTGGAGGGGCTGGATCTCCTGCCGCACCTT
 P   L   C   K   S   E   E   R   *
CCCTGGCAGGCACCCTGTCGTTGAGCCCCAGAACTGCAGGCGGCCGGCAGAGAAGGGGTCCATGATGGCGCCTCGGTGCGGCCTTGGACCTGCCCCCATG
GACCTGGAGACAGGGTTTCTCCTCATTGGCCAGGCTGGTCTCGAACTCCTGACCTCAGACGATCCACCTGCCTCAGCCTCCCGAAGTGTTGGGATTACAG
CACGAGCCACTGTGCCCGGCCATCATTCCTTTTTACTGCTGACTAATAGTCTGCTGTGTGAATCCACCGCTAGAAACCCACTCATCAGTTGATGGTCATG
TGGGTTGCTTCTCGTATTCGCTTATTATGAACAGTGCTGGAATAAACGTTCCTGTGCACTCTTGGCATATGCCTAGGAGTGGAACTGCTGGGTA
```

**Fig. 6. Continued**

D. Northern Blot Analysis

To confirm the size of full-length of *MUC8* cDNA, Northern Blot was performed. Further confirmation of the identity between our new sequence and *MUC8* was provided by Northern blot analysis in which we observed that a probe from our new sequence, B1-B2 (Table 1), showed intense polydispersed hybridization band typical of *MUC* probe (Fig. 8).

E. Zooblot Analysis

A zooblot experiment was performed using a 55958S/55433AS product (Table 1), *MUC8* cDNA as probe in an attempt to assess the extent of evolutionary conservation of the *MUC8* sequence. The hybridization was performed at low stringency and the membrane was subsequently washed at increasingly higher stringency conditions. As shown in Fig. 9, band was detected in only human genomic DNA. However, some of hybridization signal observed after the lower stringency washes appeared to correspond to non-specific stickiness of the probe to satellite bands, which all disappeared after the most stringent washes. These data suggest that *MUC8* may be expressed in only human.

However, a conclusion regarding this apparent lack of conservation of the mucin genes waits the identification of *MUC8* orthologues.

F. *MUC8* Domains Expected from Peptide Sequence

We searched functional domains from *MUC8* peptide sequence using SMART program (http://smart.embl-heidelberg.de), and depicted the domains in Fig. 10. The results showed EGF-like, Cysteine-rich or knot, AMOP, nidogen domain and transmembrane region found in other mucins. Amino acid sequences composing these domains are homologous to corresponding domains sequences of other mucins (Fig. 11). EGF-like, AMOP, nidogen domain and transmembrane region are found in mainly membrane-bound mucins, and Cysteine-rich or knot and vWFC domains are secreted mucin-specific. However, *MUC8* peptide sequence shows both membrane bound and secreted mucin domains. Computer-based domain search doesn't give a definite evidence about localization of *MUC8*.

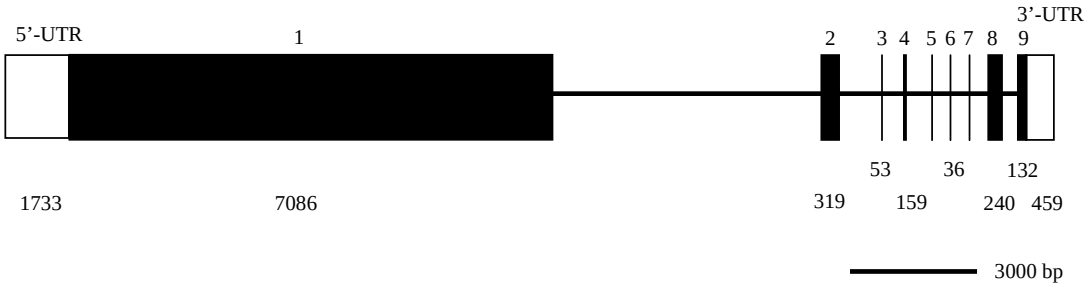


**Fig. 7. Diagram of *MUC8* gene. Closed boxes represent exons, which are shown with exon numbers and sizes, and open boxes represent untranslated regions.**

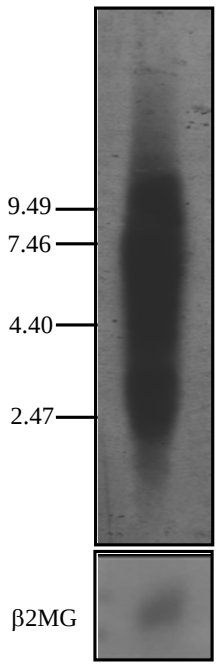


**Fig. 8. Northern hybridization using the novel cDNA sequence probe, B1-B2.**

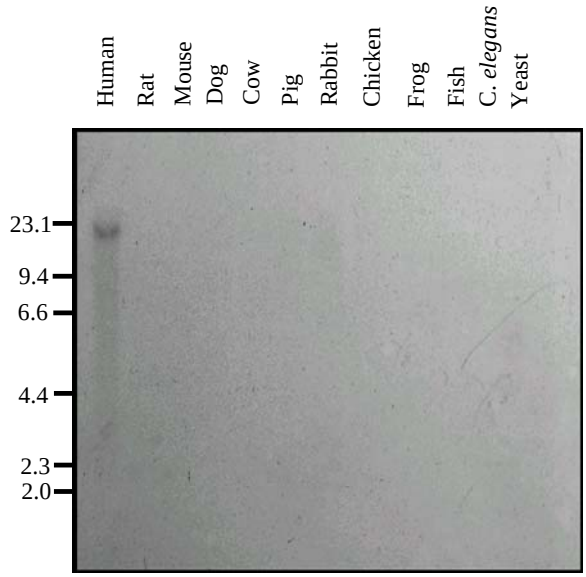


**Fig. 9. Zooblot of *MUC8* gene. This figure shows the probing of a zooblot containing immobilized *Eco*RI digested genomic DNA from various species, with a 52 bp $^{32}$P-labeled probe.**

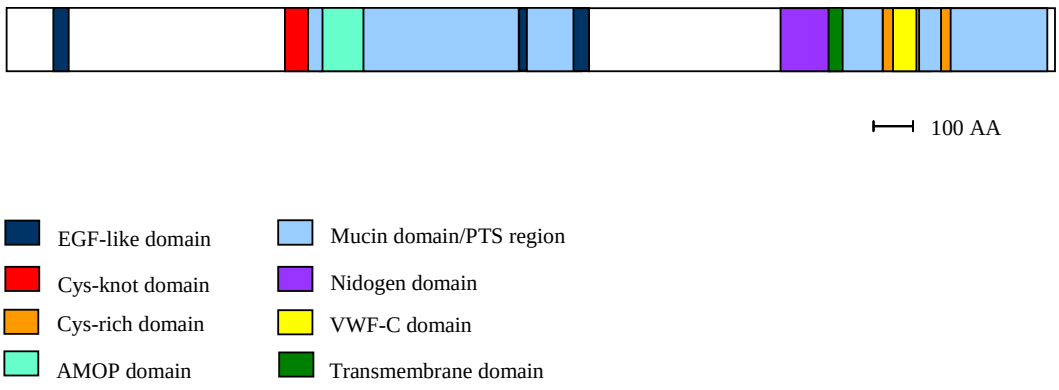


**Fig. 10. Domains predicted from the *MUC8* peptide sequence.**

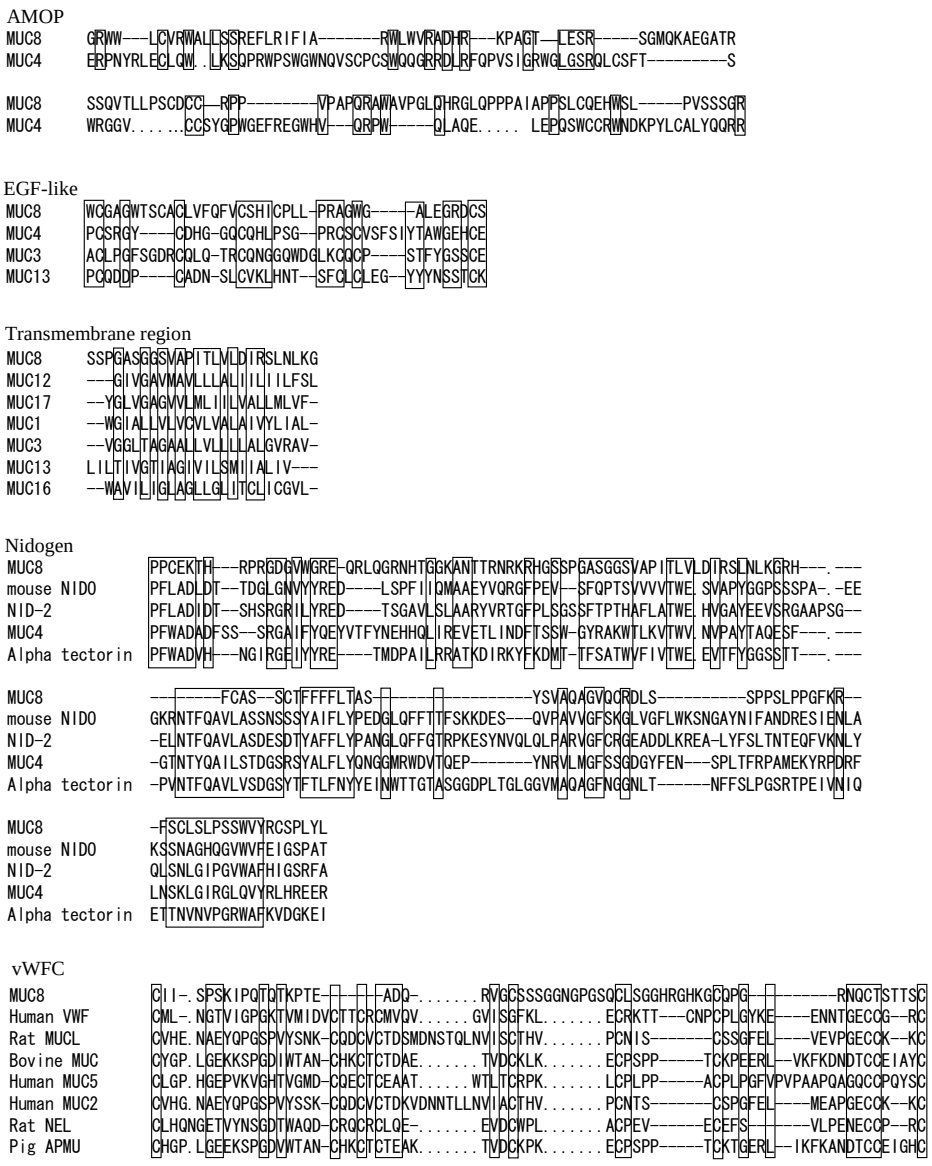


**Fig. 11. Alignments of the domains of MUC8, other proteins, and other mucins, using the SMART and ClustalW programs.** Open boxes indicate similar amino acids in each domain.

2. PGE2-induced *MUC8* gene expression

A. PGE2 can induce *MUC8* gene expression through ERK MAP kinase signaling in NHNE and NCI-H292 cells

To determine whether PGE2 can induce *MUC8* gene expression within NHNE cells, we carried out RT-PCR after treatment with varying concentrations of PGE2. As the dose of PGE2 was increased from 1 to 1000 nM, there was a gradual increase in *MUC8* gene expression from 10 nM. As shown in Fig. 12A, 10 nM PGE2 significantly induce *MUC8* gene expression. No corresponding change was found in the expression of internal control, and β2-Microglobulin. In order to determine whether PGE2 induced *MUC8* gene expression in a time-dependent manner, we examined the expression level of *MUC8* after various lengths of exposure to 10 nM PGE2 (Fig. 12B). *MUC8* gene expression was significantly increased at 24 hours of exposure to PGE2. The results showed that the *MUC8* gene expression was significantly increased after treatment with PGE2 in NHNE cells. We used 10 nM PGE2 for all the subsequent experiments. As a next step, to investigate which MAP kinase signal pathway is activated within NHNE cells stimulated by PGE2, we performed a Western blot analysis using phospho-specific

antibodies. ERK MAP kinase was maximally activated at 5 min, and this effect decreased after 20 min (Fig. 13). However no change was detected in the activation of p38 and JNK. NCI-H292 cells and A549 cells treated IL-1β were used as positive controls for p38 and JNK activation, respectively. It thus appeared that stimulation by PGE2 activates the ERK MAP kinase pathway in NHNE cells. Because PGE2 significantly increased both *MUC8* gene expression and ERK MAP kinase activity, we wanted to determine whether PGE2-induced *MUC8* gene expression involves the ERK MAP kinase pathway. As a next step, 20 μM PD98059, specific MEK1/2 inhibitor, was applied before treatment with PGE2. Pretreatment of NCI-H292 cells with PD98059 for 1 hr clearly inhibited ERK MAP kinase and significantly suppressed PGE2-induced *MUC8* gene expression in NHNE cells (Fig. 14A). When the same experiments (Figs. 12, 13, and 14) were performed using NCI-H292 cells, human lung mucoepidermoid carcinoma cell line, the results were the same in the NCI-H292 cells as in the NHNE cells (data not shown). To further confirm whether ERK MAP kinase was involved in PGE2-induced *MUC8* gene expression, cells were transiently transfected wih DNA construct encoding MEK1 dominant negative (DN). Overexpression of MEK1 DN suppressed the PGE2-induced ERK MAP kinase activity. Consistently, PGE2-induced

*MUC8* gene expression was significantly suppressed by MEK1 DN (Fig. 14B). These results show that the activation of ERK MAP kinase via MEK1 was essential for PGE2-induced *MUC8* gene expression in NCI-H292 cells.

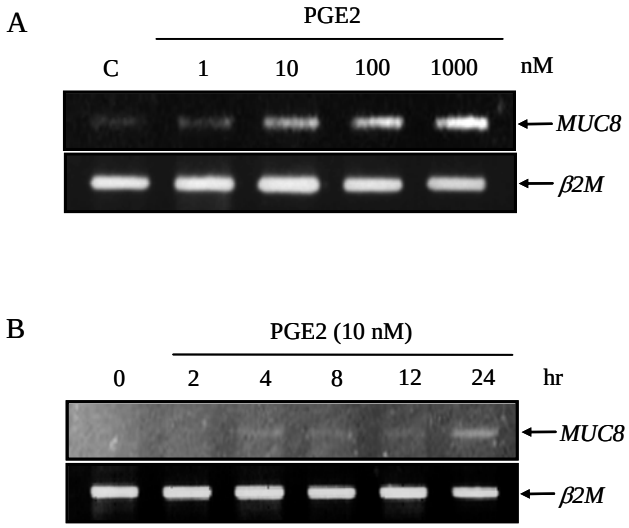


**Fig. 12. Effect of PGE2 on *MUC8* gene expression in NHNE cells.**
Confluent cells were treated with PGE2 (1, 10, 100, and 1000 nM) for 24 h (A). Confluent cells were treated with PGE2 (10 nnM) for 2, 4, 8, 12, and 24 h (B), and cell lysates were harvested for RT-PCR. C, control. β2-microglobulin (*β2M*) was employed as an internal control.

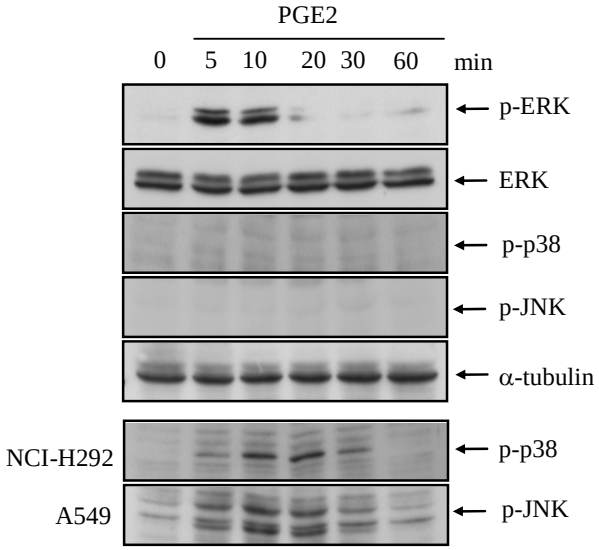


**Fig. 13. PGE2 induces *MUC8* gene expression via ERK MAP Kinase signaling.** Confluent cells were treated with PGE2 (10 nM) for 5, 10, 20, 30, and 60 min, and cell lysates were harvested for Western blot analysis. Representative Western blots using phospho-specific antibodies showed transient activation of ERK but not p38 and JNK, and the maximum effect is at 5 min.

`

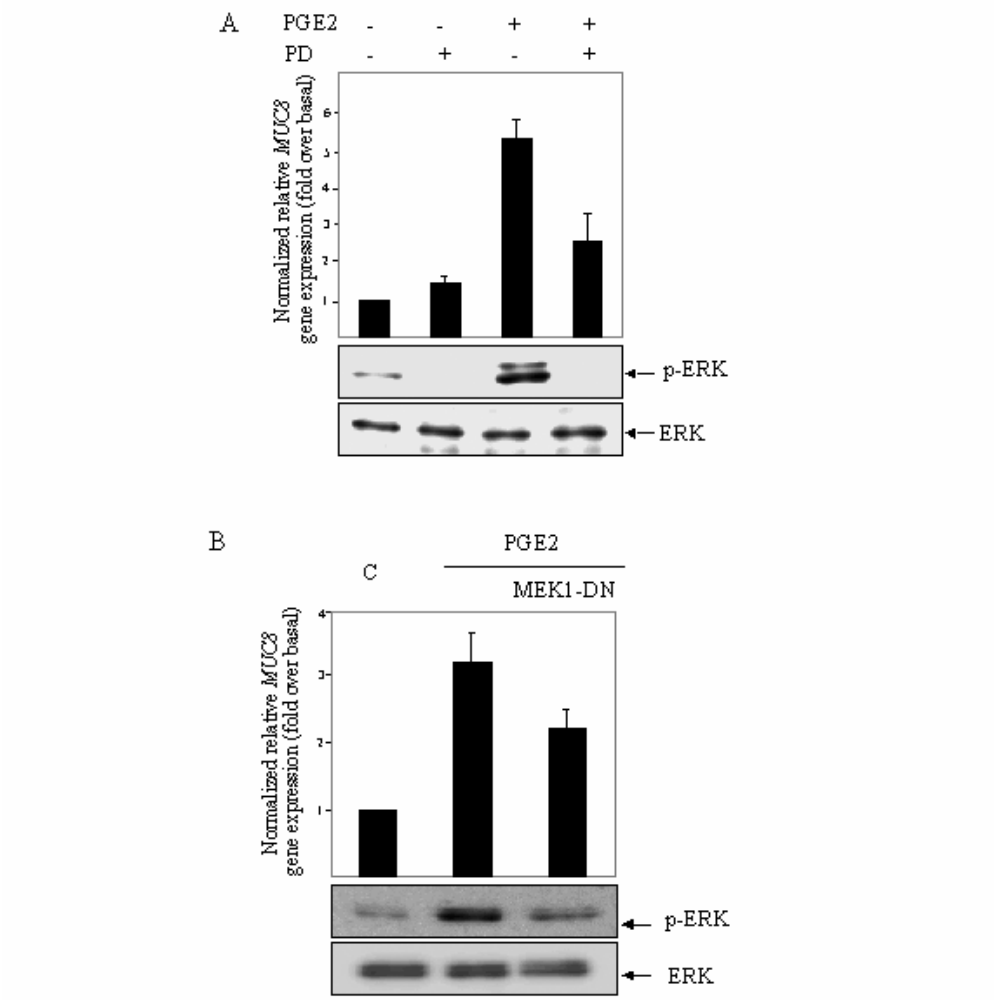


**Fig. 14. PGE2 induces *MUC8* gene expression via ERK MAP Kinase signaling.**
Confluent cells were pretreated for 1 h with 20 $\mu$M PD98059 and then stimulated for 24 h with PGE2 prior to collection of total RNA for Real-Time Quantitative PCR. Pretreated cells were stimulated for 5 min with PGE2 prior to collection of cell lysates for Western blot analysis (A). The cells were transiently transfected with MEK1-dominant negative (MEK1-DN) construct and stimulated with PGE2 for 24 h prior to Real-Time Quantitative PCR, and transfected cells was stimulated with PGE2 for 5 min prior to Western blot analysis (B). Real-time PCR figures shown are representatives of three independent experiments.

B. Effect of RSK1 on PGE2-induced *MUC8* gene expression

To determine which molecules are involved in the down stream signaling of ERK MAP kinase in PGE2-induced *MUC8* gene expression, we investigated RSK1. RSK1 has been reported to be activated by ERK MAP kinase. The phosphorylation of RSK1 by PGE2 peaked at 5 min and then decreased at 20 min after PGE2 stimulation (Fig. 15A). To determine whether RSK1 plays an important role in PGE2-induced *MUC8* gene expression, an RSK1 mutant study was performed. RSK1 is a currently candidate for the mediation of cytokine-induced CREB phosphorylation at $Ser^{133}$. Overexpression of RSK1 DN (D205N, a dominant-negative construct encoding RSK1 protein with $Asp^{205}$ phosphorylation site mutated to $Asn^{205}$) suppressed the PGE2-induced CREB and RSK1 phosphorylation. Consistently, PGE2-induced *MUC8* gene expression was significantly suppressed by RSK1 DN (D205N) (Fig. 15B). These results show that RSK1 is required for PGE2-induced *MUC8* gene expression.

C. Effect of CREB on the PGE2-induced *MUC8* gene expression

To determine whether CREB plays a role in *MUC8* gene expression, we performed a Western blot analysis using phospho-CREB antibody. The phosphorylation of CREB by PGE2 peaked at 10 min and then decreased at 60 min after PGE2 stimulation (Fig. 16A). Furthermore, PGE2-induced *MUC8* gene expression was significantly suppressed in cells transfected with plasmid encoding CREB DN (S133A, a dominant-negative construct encoding CREB protein with $Ser^{133}$ phosphorylation site mutated to $Ala^{133}$) (Fig. 16B). These findings suggest that the activation of CREB is essential for PGE2-induced *MUC`* gene expression via ERK MAP kinase and RSK1.

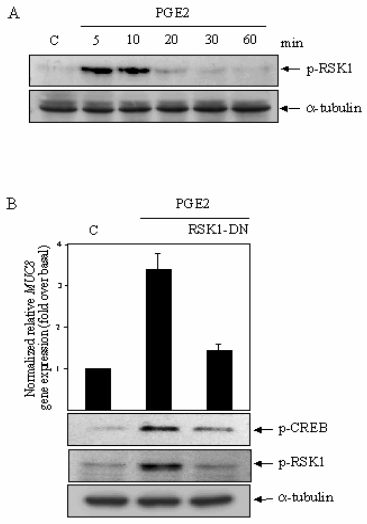


**Fig. 15. Effect of RSK1 on PGE2-induced *MUC8* gene expression.**
Confluent cells were stimulated for the indicated times with PGE2, and then total proteins were collected for Western blot (A). The cells were transiently transfected with RSK1-dominant negative (RSK1-DN) construct and stimulated with PGE2 for 24 h prior to Real-Time Quantitative PCR, and transfected were stimulated with PGE2 for 10 min prior to Western blot analysis (B). Real-time PCR figures shown are representatives of three independent experiments.

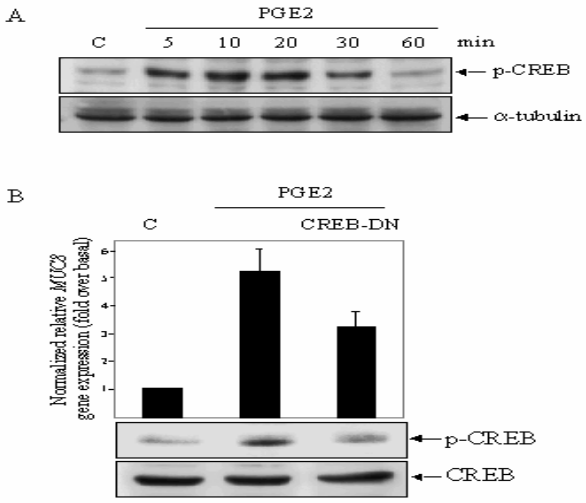


**Fig. 16. Effect of CREB on PGE2-induced *MUC8* gene expression.**
Confluent cells were stimulated for the indicated times with PGE2, and then total proteins were collected for Western blot (A). The cells were transiently transfected with CREB-dominant negative (CREB S133A) construct and stimulated with PGE2 for 24 h prior to Real-Time Quantitative PCR, and transfected were stimulated with PGE2 for 10 min prior to Western blot analysis (B). Real-time PCR figures shown are representatives of three independent experiments.

D. Identification of the binding complex between CREB and CRE in response to PGE2

From pBeloBAC11, we cloned *MUC8* promoter covering –1644 to +87 (Fig. 17). CRE site is located in –803 region of *MUC8* promoter. To analyze the DNA binding activity of PGE2-activated CREB, we performed EMSA using nuclear extracts from NCI-H292 cells after PGE2 treatment. As shown in Fig. 18, the activity of consensus CRE (CREc) and *MUC8* specific CRE (CREs) oligonucleotides remarkably increased in response to PGE2.

To distinguish any specific CRE-binding complexes, competition and supershift analysis were using 50-, and 100-fold excesses of non-radiolabeled (cold) CREs oligonucleotide and anti-phospho-CREB antibody, respectively. The specific band was found to be selectively inhibited by the specific CRE competitor and was supershifted by anti-phospho-CREB antibody (Fig. 18). These results indicated that activated CREB binds to a *cis*-acting element, CRE, in the *MUC8* promoter.

E. Identification of PGE2-Responsive Regions within *MUC8* Promoter

Various deletion clones (-1190 to +87, -973 to +87 and –549 to +87) were made on the basis of above –1644 to +87 clone. Cells were then transiently transfected with the various deletion mutants and treated with PGE2 (10 nM) for 24 h, respectively. As shown in Fig. 19A, PGE2 selectively increased luciferase activity of –1644/-973 region of *MUC8* promoter. However, the effect was decreased on fragments covering -549/+87 region, indicating that the -1644/-973 region of *MUC8* promoter may be necessary to observe a response to PGE2. We examined whether activation of CRE is required for PGE2- induced *MUC8* transcription by performing selective mutagenesis of the CREB-binding site at -803.

As a shown in Fig. 19B, mutant constructs MI and M2 decreased responsiveness of wild-type *MUC8* promoter construct. These results showed that CRE in the regulatory region of *MUC8* promoter was critical for the up-regulation of the transcriptional activity of *MUC8* induced by PGE2.

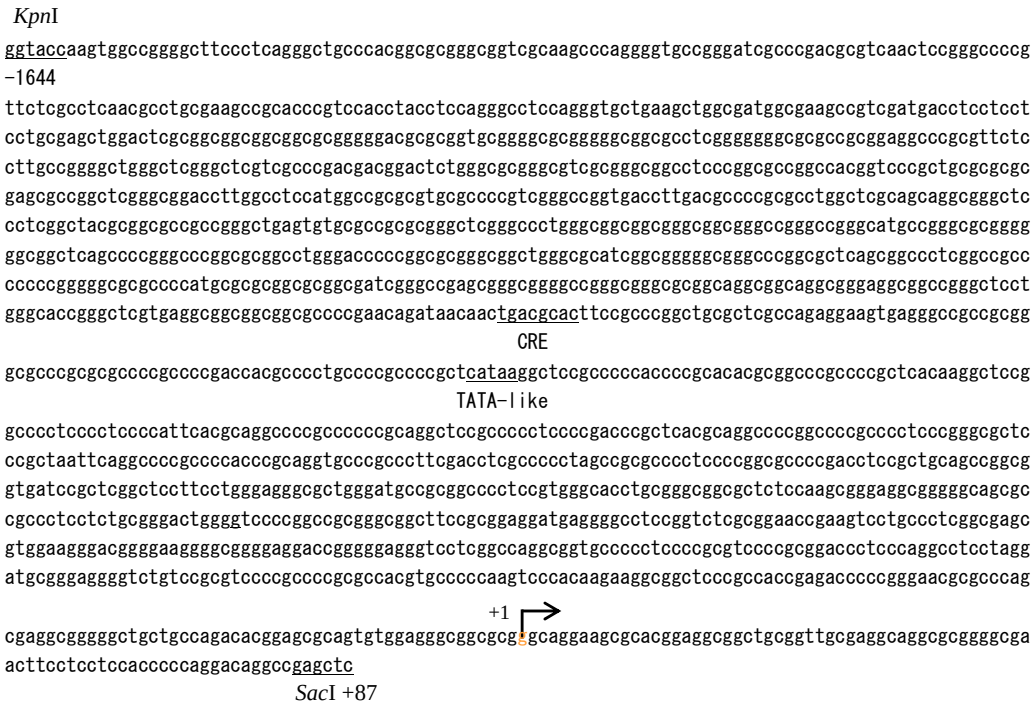


**Fig. 17. Nucleotide sequence of *MUC8* promoter.** The restriction enzyme sites used in the cloning of the *MUC8* promoter, and the position of the potential cAMP response element (CRE) site, are underlined. The transcription start site (+1) is indicated by a bent arrow.

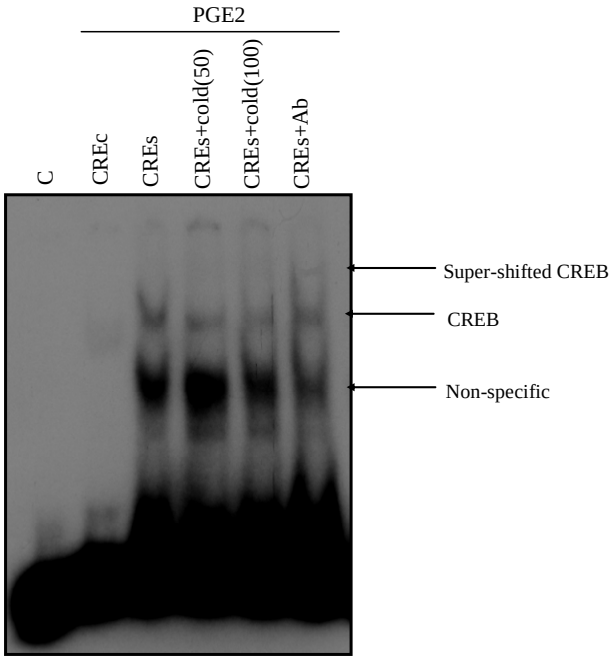


**Fig. 18. Identification of the CREB-CRE binding complex formed in response to PGE2.** Confluent cells were stimulated for 1 h with PGE2. Nuclear protein extract from PGE2-treated NCI-H292 cells were subjected to EMSA. Nuclear proteins were incubated with CREc, CREs, 50-, and 100-fold excesses of cold probe or anti-phospho-CREB antibody before EMSA. The labeled nuclear proteins were separated by electrophoresis on 6% polyacrylamide gels, and the gels were dried and exposed to autiradiography at - 70°C.

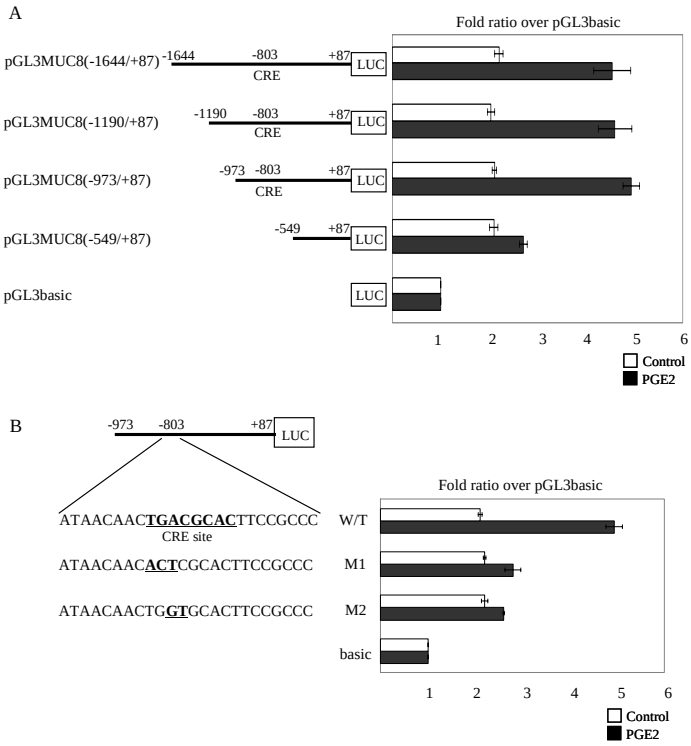


**Fig. 19. CRE is required for PGE2-induced *MUC8* transcription.** NCI-H292 cells were transiently transfected with various *MUC8* promoter luciferase reporter constructs and stimulated with PGE2 for 24 h. Luciferase activity was then assessed in PGE2-treated and –untreated cells (A). Cells were transfected with the *MUC8* promoter constructs containing mutated CRE sites as indicated (B). The luciferase activities were displayed after correction for transfection efficiency using the β-galactosidase activity of the cell lysates to standardize the values. The values shown are means ±standard deviations of experiments performed in triplicate.

# IV. DISCUSSION

## 1. *MUC8* cDNA sequencing and protein structure

Presently, full-length mucin cDNAs corresponding to seven human genes - namely *MUC1*, *MUC2*, *MUC7*, *MUC13*, *MUC15*, *MUC18*, and *MUC20* – have been characterized. The other *MUC* genes have been identified only as partial cDNA sequences. In this study, we

confirmed the 10292 bp of the complete *MUC8* cDNA sequence extended from the known partial cDNA sequence (GenBank™ accession number U14383), and deduced the amino acid sequence. The *MUC8* cDNA sequence does not exhibit homology with other mucin genes. Zooblot data suggest that *MUC8* may be expressed only in human (Fig. 9). However, a conclusion regarding this apparent lack of conservation of mucin genes awaits the identification of *MUC8* orthologues. MUC8 protein is composed of 2699 amino acids, and its expected size is about 290 kDa. Like other mucins, the *MUC8* cDNA sequence possesses tandem repetitive sequences in which there are rich hydroxyl amino acids, such as serine, threonine, proline, glycine, and alanine, are believed to be extensively glycosylated. Glycosylated domains form semi-rigid, extended structures. Many mucins exhibit length polymorphism of tandem repeats as the result of multiple alleles encoding different numbers of tandem repeats[29].

Cys-rich, Cys-knot and vWFC domains in MUC8 are found in secreted mucins, and are expected to function in polymerization. In addition, EGF-like, nidogen and AMOP domains can also be found, as well as the trans-membrane region characteristic of membrane-bound mucins. These suggest that the MUC8 protein sequence encompasses both secreted- and membrane-bound domains. As these functions were predicted by the SMART program, it remains uncertain as to whether MUC8 should be classified as a membrane-bound or secreted

mucin. Interestingly, three EGF-like domains were found in *MUC8*. The EGF-like domain in MUC1 activates the ras pathway via Grb2[30], and interacts with β-catenin, an important modulator of cell adhesion and growth[31, 32]. The EGF-like domain in MUC4 interacts with ErbB2, a member of the class-I EGF receptor tyrosine kinase[14]. This complex may well be involved in a signaling pathway required for the proliferation and differentiation of airway epithelial cells. In addition, considering that MUC8 is expressed in the ciliated cells of human nasal polyps, and not in goblet cells[33], it is possible that MUC8 is also involved in ciliated cell differentiation.

**2. PGE2-induced *MUC8* Gene Expression**

The molecular mechanism of *MUC8* gene expression up-regulation by inflammatory mediators, remains poorly understood. Recently, we described the signal transduction pathway by which IL-1β induces *MUC8* gene expression[25], and in the present study, we investigated the mechanism of *MUC8* gene expression up-regulation by PGE2 in normal human nasal epithelial cells. Our results show that only ERK MAP Kinase activation is required for PGE2-induced *MUC8* gene expression (Fig. 13), although several studies[34-36] have suggested that more than one MAP Kinase is necessary for the signal transduction of various inflammatory mediators.

Moreover, the activation of ERK MAP kinase by various stimulants mainly occurs through MEK1. Thus, in this study, we investigated whether ERK MAP kinase activation by PGE2 is MEK1-dependent or –independent. The overexpression of MEK1 mutant significantly reduced PGE2-induced ERK MAP kinase phosphorylation and *MUC8* gene expression (Fig. 14B), showing that the ERK MAP kinase activation required to induce *MUC8* gene expression in human airway epithelial cells occurs via an MEK1-dependent pathway.

To date, the signal molecules involved in the downstream signaling of ERK MAP kinase, for PGE2-induced *MUC8* gene expression, have not been yet identified. However, recently, it was reported that IL-1β-induced *MUC8* gene expression is mediated by sequential ERK MAPK/RSK1/CREB activation in airway epithelial cells[25]. In the present study, we found that RSK1 and CREB are also important downstream molecules of ERK MAP Kinase activation in PGE2-induced *MUC8* gene expression (Figs. 13, 14). MSK1 and RSK1 are known substrates of CREB in NHNE cells[37, 25]. However, MSK1 did not affect PGE2-induced *MUC8* gene expression (data not shown). On the other hand, RSK1 phosphorylates several transcription factors, e.g., CREB[38], c-Fos[39], CCAAT/enhancer binding protein[40], nuclear factor-κB[41], and the estrogen receptor[42], and interacts with transcriptional coactivator CREB-binding protein (also known as p300)[43]. Moreover, PGE2 receptors, such as EP1-4, are coupled to cAMP

upregulation[44], which implies CREB activation. And, CREB is a known potent regulator of the expression of mucin genes (*MUC2, MUC5AC, MUC5B,* and *MUC6*) in the p15 arm of chromosome 11 (11p15)[45]. In a previous study, we reported the possible involvement of CREB in IL-1β-induced *MUC8* gene expression[25], but at that time, since the *MUC8* promoter sequence was not known, we could not determine whether CREB binds to a *MUC8* specific promoter. However, we already cloned *MUC8* promoter region[20]. In the present study, our results show that the – 973 / - 549 region of the *MUC8* promoter is involved in response to PGE2, and that CRE in the – 803 region of the *MUC8* promoter is important for *MUC8* gene up-regulation by PGE2. Gerritsen et al. reported that p300 and CREB-binding protein (CBP) act as co-activators of p65 transactivation and may play an important role in the cytokine-induced expression of various immune and inflammation genes[46]. These findings suggest that CREB may interact directly or indirectly with other transcription factor(s) and that non-DNA binding transcriptional co-activators, such as p300 and CBP, which were believed to function as bridging proteins between DNA-binding transcription factors and basal transcription factors, play a role as integrators of diverse signaling pathways leading to *MUC8* gene expression.

Thus, to induce *MUC8* gene expression, IL-1β transduces through the IL-1β receptor/Ras/Raf/ERK/RSK1CREB cascade pathway[25], and PGE2 through the EP1-

4/MEK1/ERK/RSK1/CREB cascade pathway. Although these two substances stimulate their own membrane receptors, they seem to share common signaling molecules, downstream of ERK MAP kinase.

Shimamoto et al reported that the $Ca^{2+}$-regulation of exocytic events and PGE2 release are activated in acetylcholine-stimulated antral mucous cells, and that the PGE2 released induces cAMP accumulation, which enhances $Ca^{2+}$-regulated exocytosis[47]. Gray et al. reported that the induction of *MUC5AC* gene by IL-1β involves COX2-generated PGE2[23]. From these reports, we cannot exclude the possibility that IL-1β stimulates PGE2 secretion, and that secreted PGE2 induces *MUC8* gene expression in an autocrine manner.

In summary, our results demonstrate that ERK MAP Kinase is essential for PGE2-induced *MUC8* gene expression, and that the activations of RSK1 and CREB are crucial required for the intracellular mechanisms that mediate *MUC8* gene expression. This study also demonstrated that CRE in the *MUC8* promoter may play a role in these processes by binding CREB. Further analysis of the signal pathways activated by various stimulators may yield deeper insights into the signaling mechanism of *MUC8* gene expression.

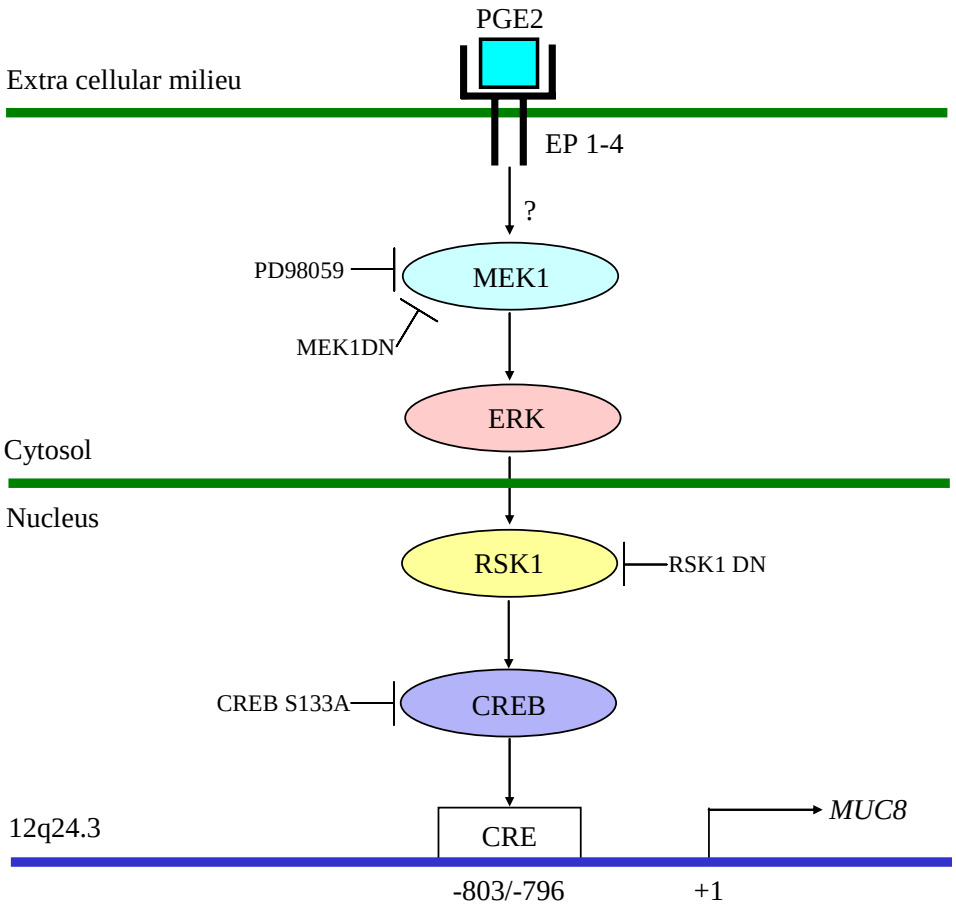


**Fig. 20. The signaling pathway of PGE2-induced *MUC8* gene expression in airway epithelial cell.**

# V. CONCLUSION

In conclusion, *MUC8* full cDNA sequence is possible to explain the function of *MUC8* protein, and *MUC8* promoter region is used as tool to analyze the *MUC8* gene transcription mechanisms induced by various stimulators.

## REFERENCES


1. Yuta A, Ali M, Sabol M, Gaumond E, Baraniuk JN. Mucoglycoprotein hypersecretion in allergic rhinitis and cystic fibrosis. Am J Physiol (Lung Cell Mol Physiol) 1997; 273: L1203-1207.

2. Kim SS, Kim KS, Lee JG, Park IY, Koo JS, Yoon JH. Levels of intracellular protein and messenger RNA of mucin and Lysozyme in normal human nasal and polyp epithelium. Laryngoscope 2000; 110: 276-280.

3. Chung MH, Choi JY, Lee WS, Kim HN, Yoon JH. Compositional difference in middle ear effusion: mucus versus serous. Laryngoscope 2002; 112: 152-155.

4. Nadel JA. Role of epidermal growth factor receptor activation in regulating mucin synthesis. Respir Res 2001; 2: 85-89.

5. Gum JR Jr, Hicks JW, Toribara NW, Siddiki B, Kim YS. Molecular cloning of human intestinal mucin (*MUC2*) cDNA. Identification of the amino terminus and overall sequence similarity to prepro-von Willebrand factor. J Biol Chem 1994; 269: 2440-2446.

6. Van de Bovenkamp JH, Hau CM, Strous GJ, Buller HA, Dekker J, Einerhand AW. Molecular cloning of human gastric mucin *MUC5AC* reveals conserved cysteine-rich D-domains and a putative leucine zipper motif. Biochem Biophys Res Commun 1998;

245: 853-859.

7. Keates AC, Nunes DP, Afdhal NH, Troxler RF, Offner GD. Molecular cloning of a major human gall bladder mucin: complete C-terminal sequence and genomic organization of *MUC5B*. Biochem J 1997; 324: 295-303.

8. Ho SB, Roberton AM, Shekels LL, Lyftogt CT, Niehans GA, Toribara NW. Expression cloning of gastric mucin complementary DNA and localization of mucin gene expression. Gastroenterology 1995; 109: 735-747.

9. Bobek LA, Tsai H, Biesbrock AR, Levine MJ. Molecular cloning, sequence, and specificity of expression of the gene encoding the low molecular weight human salivary mucin (*MUC7*). J Biol Chem 1993; 268: 20563-20569.

10. Lapensee L, Paquette Y, Bleau G. Allelic polymorphism and chromosomal localization of the human oviduct gene (*MUC9*). Fertil Sterile 1997; 68: 702-708.

11. Chen Y, Zhao YH, Kalaslavadi TB, Hamati E, Nehrke K, and Le AD et al. Genome-wide search and identification of a novel gel-forming mucin MUC19/Muc19 in glandular tissues. Am J Respir Cell Mol Biol. 2004; 30:155-165.

12. Spicer AP, Parry G, Patton S, Gendler SJ. Molecular cloning and analysis of the mouse homologue of the tumor-associated mucin, *MUC1*, reveals conservation of potential O-glycosylation sites, transmembrane, and cytoplasmic domains and a loss of

minisatellite-like polymorphism. J Bio Chem 1991; 266: 15099-15109.

13. Van Klinken BJ, Van Dijken TC, Oussoren E, Buller HA, Dekker J, Einerhand AW. Molecular cloning of human *MUC3* cDNA reveals a novel 59 amino acid tandem repeat region. Biochem Biophys Res Commun 1997; 238: 143-148.

14. Moniaux N, Nollet S, Porchet N, Degand P, Laine A, Aubert JP. Complete sequence of the human mucin *MUC4*: a putative cell membrane-associated mucin. Biochem J 1999; 338: 325-333.

15. Williams SJ, McGuckin MA, Gotley DC, Eyre HJ, Sutherland GR, Antalis TM. Two novel mucin genes down-regulated in colorectal cancer identified by differential display. Cancer Res 1999; 59: 4083-4089.

16. Williams SJ, Wreschner DH, Tran M, Eyre HJ, Sutherland GR, McGuckin MA. *Muc13*, a novel human cell surface mucin expressed by epithelial and hemopoietic cells. J Biol Chem 2001; 276: 18327-18336.

17. Gum JR, Crawley SC, Hicks JW, Szymkowski DE, Kim YS. *MUC17*, a novel membrane-tethered mucin. Biochem Biophys Res Commun 2002; 291: 466-475.

18. Sers C, Kirsch K, Rothbacher U, Riethmuller G, Johnson JP. Genomic organization of the melanoma-associated glycoprotein *MUC18*: implications for the evolution of the immunoglobulin domains. Proc Natl Acad Sci USA 1993; 90: 8514-8518.

19. Higuchi T, Orita T, Nakanishi S, Katsuya K, Watanabe H, and Yamasaki Y et al. Molecular cloning, genomic structure, and expression analysis of MUC20, a novel mucin protein, up-regulated in injured kidney. J Biol Chem 2004 ; 279: 1968-1979.

20. Shankar V, Gilmore MS, Elkins RC, Sachdev GP. A novel human airway mucin cDNA encodes a protein with unique tandem repeat organization. Biochem J 1994; 300: 295-298.

21. Jan D, John WAR, Hans AB, Alexandra WCE. The MUC family: an obituary. Trends in Biochemical Science 2002; 27: 126-132.

22. Narumiya S, Sugimoto Y and Ushikubi F. Prostanoid receptors: structures, properties, and functions. Physiol Rev 1999; 79: 1193-1226.

23. Yoon JH, Moon HJ, Seong JK, Kim CH, Lee JJ, and Choi JY et al. Mucociliary differentiation according to time in human nasal epithelial cell culture. Differentiation 2002; 70: 77-83.

24. Yoon JH, Kim KS, Kim HU, Linton JA, Lee JG. Effects of TNF-alpha and IL-1 beta on mucin, lysozyme, IL-6, and IL-8 in passage-2 normal human nasal epithelial cells. Acta Otolaryngol 1999; 119: 905-910.

25. Song KS, Seong JK, Chung KC, Lee WJ, Kim CH, and Cho KN et al. Induction of *MUC8* gene expression by interleukin-1 beta is mediated by a sequential ERK

MAPK/RSK1/CREB cascade pathway in human airway epithelial cells. J Biol Chem 2003; 278: 34890-6.

26. Yoon JH, Kim KS, Kim SS, Lee JG, Park IY. Secretory differentiation of serially passaged normal human nasal epithelial cells by retinoic acid: expression of mucin and Lysozyme. Ann Otol Rhinol Laryngol 2000; 109: 594-601.

27. Mount SM. A catalogue of splice junction sequences. Nucleic Acids Res 1982; 10: 459-472.

28. Kozak M. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res 1987; 15: 8125-8148.

29. Vinall LE, Hill AS, Pigny P, Pratt WS, Toribara N, and Gum JR et al. Variable number tandem repeat polymorphism of the mucin genes located in the complex on 11p15.5. Hum Genet 1998; 102: 357-366.

30. Pandey P, Kharbanda S, Kufe D. Association of the DF3/MUC1 breast cancer antigen with Grb2 and the Sos/Ras exchange protein. Cancer Res 1995; 55: 4000-4003.

31. Yamamoto M, Bharti A, Li Y, Kufe D. Interaction of the DF3/MUC1 breast carcinoma-associated antigen and beta-catenin in cell adhesion. J Biol Chem 1997; 272: 12492-12494.

32. Li Y, Bharti A, Chen D, Gong J, Kufe D. Interaction of glycogen synthase kinase 3beta

with the DF3/MUC1 carcinoma-associated antigen and beta-catenin. Mol Cell Biol 1998; 18: 7216-7224.

33. Gerritsen ME, Williams AJ, Neish AS, Moore S, Shi Y, Collins T. MUC8 as a ciliated cell marker in human nasal epithelium, Proc Natl Acad Sci USA. 1997; 94: 2927-2932.

34. Bernatchez PN, Allen BG, Gelinas DS, Guillemette G, Sirois MG. Regulation of VEGF-induced endothelial cell PAF synthesis: role of p42/44 MAPK, p38 MAPK and PI3K pathways. Br J Pharmacol 2001; 134: 1253-1262

35. Waetzig GH, Seegert D, Rosenstiel P, Nikolaus S, Schreiber S. p38 mitogen-activated protein kinase is activated and linked and to TNF-alpha signaling in inflammatory bowel disease. J Immunol 2002; 168: 5342-5351.

36. Plath KE, Grabbe J, Gibbs BF. Calcineurin antagonists differentially affect mediator secretion, p38 mitogen-activated protein kinase and extracellular signal-regulated kinase from immunologically activated human basophils. Clin Exp Allergy 2003; 33: 342-350

37. Song KS, Lee WJ, Chung KC, Koo JS, Yang EJ, and Choi JY et al. Interleukin-1 beta and tumor necrosis factor-alpha induce MUC5AC overexpression through a mechanism involving ERK/p38 mitogen-activated protein kinases-MSK1-CREB activation in human airway epithelial cells. J. Biol. Chem. 2003; 278: 23243-23250.

38. Xing J, Ginty DD, Greenberg ME. Coupling of the RAS-MAPK pathway to gene activation by RSK2, a growth factor-regulated CREB kinase. Science 1996; 273: 959-963

39. Chen RH, Abate C, Blenis J. Phosphorylation of the c-Fos transrepression domain by mitogen-activated protein kinase and 90-kDa ribosomal S6 kinase. Proc Natl Acad Sci USA. 1993; 90: 10952-10956.

40. Buck M, Poli V, van der Geer P, Chojkier M, Hunter T. Phosphorylation of rat serine 105 or mouse threonine 217 in C/EBP beta is required for hepatocyte proliferation induced by TGF alpha. Mol Cell 1999; 4: 1087-1092.

41. Kim KW, Kim SH, Lee EY, Kim ND, Kang HS, Kim HD and et al. Extracellular signal regulated kinase/90-KDA ribosomal S6 kinase/nuclear factor kappa B pathway mediates phorbol 12-myristate 13-acetate-inducedmegakaryocytic differentiation of K562 cells. J Biol Chem 2001; 276: 13186-13191.

42. Joel PB, Smith J, Sturgill TW, Fisher TL, Blenis J, Lannigan DA. p90rsk1 regulates estrogen receptor-mediated transcription through phosphorylation of Ser-167. Mol Cell Biol 1998; 18: 1978-1984.

43. Nakajima T, Fukamizu A, Takahashi J, Gage FH, Fisher T and Blenis J et al. The signal dependent coactivator CBP is a nuclear target for pp90RSK. Cell 1996; 86: 465-474.

44. Bukhave K and Rask-Madsen J. Saturation kinetics applied to in vitro effects of low prostaglandin E2 and F 2 alpha concentrations on ion transport across human jejunal mucosa. *Gastroenterology* 1980; 78: 32-42.

45. Van Seuningen I, Pigny P, Perrais M, Porchet N, Aubert JP. Transcriptional regulation of the 11p15 mucin genes. Towards new biological tools in human therapy, in inflammatory disese and cancer? Front Biosci 2001; 6: D1216-D1234.

46. Gerritsen ME, Williams AJ, Neish AS, Moore S, and Shi Y et al. CREB-binding protein/p300 are transcriptional coactivators of p65. Proc Natl Acad Sci USA. 1997; 94: 2927-2932.

47. Shimamoto C, Fujiwara S, Kato M, Ito S, Katsu K-I, and Mori H et al. Inhibition of ACh-stimulated exocytosis by NSAIDs in guinea pig antral mucous cells: autocrine regulation of mucin secretion by PGE2. Am J Physiol Gastrointest Liver Physiol Sep. 2 2004.

국문요약

# *MUC8* 유전자 및 단백 분석을 통한 *MUC8* 단백 구조와 전사 조절 기전 연구

( 지도교수 윤 주 헌 )

연세대학교 대학원 의과학과

조 규 남

우리의 이전 연구에서는 MUC8 유전자가 만성 부비동염의 비용 상피에서 과발현되며, 배양 조건에서는 염증 매개체에 의해 발현이 유도되는 것을 보여주었다. 그러나, 지금까지 MUC8 단백의 기능은 잘 알려져 있지 못하다. 우리는 이미 알려진 MUC8 단백 서열이 극히 제한적이기 때문이라고 생각하여, MUC8 유전자를 포함하고 있는 염색체 12q24.3 locus의 염기 서열을 바탕으로, 이미 알려진 서열의 상위 서열에 대하여 RT-PCR과 5′-RACE-PCR을 수행하였다. Primer extension과 RPA를 통해 전사 개시 부위를 확인할 수 있었고, Kozak sequence를 통해 번역 개시

부위를 유추할 수 있었다. MUC8 cDNA는 전체 10292 bp로, 1733 bp의 5′-UTR, 8100 bp의 ORF, 그리고 459 bp의 3′-UTR로 이루어진 구조이다. 8100 bp의 ORF는 mucin 특이적인 tandem repeat 영역이 3군데 있으며, 그 영역을 구성하는 아미노산들은 glycosylation 가능성이 있는 아미노산 들이 주를 이루고 있다. Computer program에 기초하여 분석된 MUC8 domain은 막 결합형 mucin domain과 분비형 mucin domain을 모두 보여 주고 있다. 그러나, 실재로 우리의 이전 결과에 따르면, MUC8 단백이 ciliated cell에서는 발현되고, 분비형태를 나타내지 않아 막 결합 mucin임을 보여주었다. MUC8의 합성 후 위치 또는 기능을 분석하는 데 있어서 computer에 기초한 서열 분석은 한계가 있으며, 확인된 서열을 토대로 특이적인 MUC8 항체를 이용한 추가 실험이 필요할 것으로 본다.

앞서 확인된 MUC8 전사 개시 부위를 참조하여, 상위영역 –1644부터 +87 영역을 BAC clone에서 cloning할 수 있었다. 그리고, MUC8의 발현을 일으키는 신호 전달 경로와 관련된 전사 조절 인자 확인을 위해, 염증 유발 물질인 PGE2를 이용하였다. PGE2 신호는 ERK/RSK1/CREB을 통해 전달되며, MUC8 promoter 영역을 결손, 그리고 point mutation 결과로 –803의 CRE site가 MUC8 발현의 기본적인 activity를 유발함을 알 수 있었다. Cloning된 MUC8 promoter는 다양한 stimulator들이 MUC8 발현을 유도하는 과정을 이해하는 tool로써 사용될 수 있을

것으로 본다.

핵심되는 말 : MUC8 유전자, 점액과분비, 전사조절, PGE2