Mucin Biosynthesis

Identification of the cis-Regulatory Elements of Human C2GnT-M Gene

Shuhua Tan and Pi-Wan Cheng

Department of Biochemistry and Molecular Biology, College of Medicine, and Eppley Institute for Research in Cancer and Allied Diseases, University of Nebraska Medical Center, Omaha, Nebraska

Mucin glycan is the primary determinant of mucin functions. These functions are expanded by three branch structures, including core 2, core 4, and blood group I, which are synthesized by core 2 β1,6 N-acetylglucosaminyltransferase-M (C2GnT-M). Alteration of C2GnT-M gene expression is expected to have a profound effect on mucin functions, which prompted us to study the regulation of this gene. Quantitative real-time PCR analysis of the expression of this gene in 24 human tissues and airway epithelial cells showed that this gene was expressed primarily in mucus-secretory tissues. 5′ Rapid amplification of cDNA ends analysis, coupled with sequence alignment with human genome database, revealed that this gene was comprised of three exons and two introns. Northern blotting using exon 1 probe showed the presence of this exon in all transcripts, suggesting the presence of cis-regulatory elements in the proximal region upstream of and/or near the transcription initiation site (+1). Analysis of this DNA region (−417/+187) by a promoter-reporter transient transfection assay, coupled with serial deletion and linker scanning mutagenesis, revealed two positive regulatory regions, including –291/–282, and –62/–43. Further, the promoter activity was enhanced by all-trans retinoic acid (ATRA) and IL-13. Thus, the promoter region is specific to hC2GnT-M gene and subject to regulation by ATRA and IL-13. These cis-regulatory elements may be useful for construction of a mucus cell–specific vector for therapy of mucus hypersecretory diseases.

Keywords: mucin; C2GnT-M; quantitative real-time PCR; cis-regulatory elements; promoter

CLINICAL RELEVANCE

Identification of the cis-regulatory elements of the hC2GnT-M gene would afford an opportunity for constructing a mucus cell-specific vector for gene therapy of mucus hypersecretory diseases.

Mucins are high–molecular weight glycoproteins heavily glycosylated with oligosaccharides linked O-glycosidically to ser/thr in the protein backbone (1). To date, 13 membrane-tethered and 8 secreted human mucins have been reported (2–4). Membrane-tethered mucins are involved in cell–cell recognition (4), immune responses (5), carcinogenesis, and tumor metastasis (5, 6). Secreted mucins are the major determinants of the rheologic properties of mucus, which protects the underlined epithelium by maintaining hydrated environment, and can serve as the first line of defense against airborne and ingested pathogens by trapping and then clearing them from the airways and gastrointestinal tract, respectively. The above-mentioned mucin functions reside primarily in the carbohydrate moieties, which make up 60–90% of the molecules by weight. For membrane-tethered mucins, some carbohydrates, such as sialyl Lewis X without or with sulfation at C6 of GlcNAc, can serve as the signal to direct leukocyte trafficking during inflammation and leukocyte recirculation, respectively (7–9). Carbohydrate structures in secreted mucins are more heterogeneous than membrane-tethered mucins (10–13) so that the secreted mucins can fulfill their functional obligation to bind more variety of pathogens. For example, hundreds of different mucin glycans have been identified from airway mucins of a single individual (10, 12), while only a fraction of them was detected in membrane-tethered mucins found in leukocytes (14). Understanding regulation of mucin glycan biosynthesis is important for understanding the functions of mucins.

Mucin glycans are synthesized in the Golgi apparatus and are determined primarily by the amounts and Golgi localization of glycosyltransferases. Biosynthesis of mucin glycans in the Golgi apparatus is initiated by the transfer of N-acetylgalactosamine (GalNAc) from UDP-GalNAc to the hydroxyl group of Ser/Thr in the peptide backbone. After the initial glycosylation step, mucin carbohydrate can be extended by adding either Gal or GlcNAc to GalNAc to generate core 1 (Galβ1,3GalNAc) or core 3 (GlcNAcβ1,3GalNAc), respectively. Further expansion of these two core structures involves extension from the β1,3Gal or GlcNAc arm, or C6 of GalNAc as shown in Figure 1. Addition of β1,6GlcNAc to core 1 and core 3 disaccharides at GalNAc would form core 2 and core 4 branch structures, respectively. Blood group I is another β1,6GlcNAc branch structure formed by adding β1,6GlcNAc to Gal of blood group i, GlcNAcβ1,3Gal. The β1,6GlcNAc branch structure on core 2, core 4, and blood group I allows generation of additional carbohydrate structures and thus expansion of functional potential of mucins. Changes of mucin glycan branch structures can affect immune function and be associated with certain diseases. For example, overexpression of core 2 branch reduces acquired immunity in an animal model (15, 16) and exhibits hematologic disorders and cancer in patients with Wiskott-Aldrich syndrome (17, 18). On the other hand, ablation of core 2 structure decreases inflammatory response in a mouse model (19).

To date, three enzymes that exhibit C2GnT activity have been reported (20). They are C2GnT-1 (or L, leukemia type) (21), C2GnT-2 (or M, mucus type) (22), and C2GnT-3 (or T, thymus type) (23), respectively. C2GnT-L is ubiquitously expressed (21), whereas C2GnT-M and C2GnT-3 are expressed primarily in mucus-secretory tissues (22) and thymus (23), respectively. C4GnT activity is only expressed by C2GnT-M (20), while IGnT activities are exhibited by IGnT (24, 25) and C2GnT-M (22, 23, 26). Therefore, C2GnT-M enzyme is unique in that it exhibits C2GnT, C4GnT, and blood group I activities.

(*Received in original form September 5, 2006 and in final form February 2, 2007*)

This work was supported by grants from NIH RO1 HL48282 (to P.-W.C.), the Cystic Fibrosis Foundation (CFF-NIH), and the Nebraska Research Initiative-Cancer Glycobiology Program.

Correspondence and requests for reprints should be addressed to Pi-Wan Cheng, Ph.D., Department of Biochemistry and Molecular Biology, University of Nebraska Medical Center, 984525 Nebraska Medical Center, Omaha, NE 68198–5870. E-mail: pcheng@unmc.edu

Am J Respir Cell Mol Biol Vol 36. pp 737–745, 2007
Originally Published in Press as DOI: 10.1165/rcmb.2006-0334OC on February 15, 2007
Internet address: www.atsjournals.org

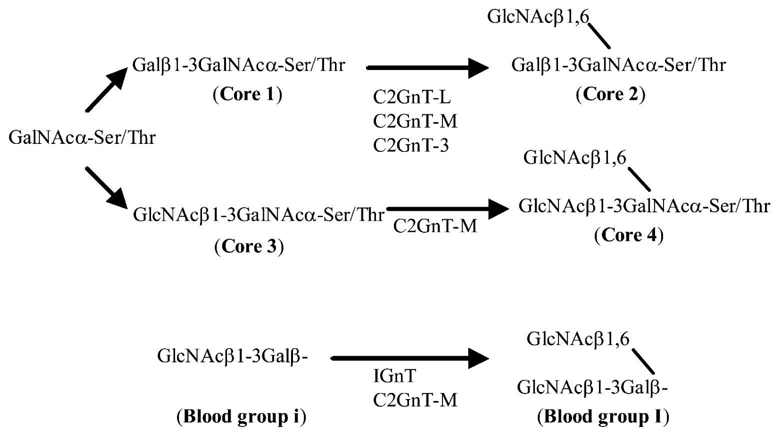

Figure 1. Biosynthesis pathway of mucin cores 1, 2, 3, and 4 as catalyzed by respective branching enzymes.

The multi-acceptor specificity of C2GnT-M was demonstrated by our group using enzyme purified from bovine tracheal epithelium (26). This multi-acceptor specificity was later confirmed with recombinant enzyme generated by cloned cDNA (15). The report also showed that this gene was expressed primarily in mucus secretory tissues (15). Given the capability of C2GnT-M enzyme to generate all known mucin β6GlcNAc branch structures, alteration of C2GnT-M activity is expected to have a profound effect on the branch structure of mucin glycans and thus mucin functions in mucus secretion. Understanding the regulation of the expression of this gene is thus important for understanding mucin functions.

To date, tissue-specific expression of mouse (27, 28) and human (29) C2GnT-L, post-translational mechanism for regulation of rat C2GnT-L (30), and characterization of 5′ untranslated region of human C2GnT-L (29) have been reported. However, very little is known about how C2GnT-M gene is regulated. Recently, we showed that expression of C2GnT-M gene was inhibited by epidermal growth factor (EGF) (31), but enhanced by Th2 cytokines and all-trans retinoic acid (ATRA) in airway epithelial cells (32). Also, it has been shown that C2GnT-M is down-regulated in colorectal cancer and its re-expression causes growth inhibition of colon cancer cells (33). However, the molecular mechanism at the level of transcription regulation is still not available. In this study, we characterize the transcripts of hC2GnT-M gene and identify its 5′-untranslated regions in various mucus-secretory tissues and airway epithelial cells. Furthermore, we identify cis-regulatory elements in the promoter regions, which are critical for promoter activity under basal and ATRA and Th2 cytokine–treated conditions.

MATERIALS AND METHODS

Materials

TaqManUniversal PCR Master Mix 200Reactions was purchased from Applied Biosystems (ABI, Foster City, CA). Human Sure-Race Multi-Tissue RACE Panels (10 ng/in 5× well) and human multiple-choice Northern blot premade membrane were from OriGene Technologies, Inc. (Rockville, MD). StrataScript First-Strand Synthesis System and QuickChange II XL site-directed mutagenesis kit were from Stratagene (La Jolla, CA). Total RNA extraction TRI reagent was from Molecular Research Center, Inc. (Cincinnati, OH). GeneRacerKit (for full-length, RNAligase-mediated rapid amplification of 5′ and 3′ cDNA ends [RLM-RACE]), DMRIE-C transfection reagent, as well as the cloning vectors pCRII, were from Invitrogen (Carlsbad, CA). PNA was from EY Laboratories, Inc. (San Mateo, CA). Oligonucleotide primers were synthesized by the Molecular Biology Core Lab, the University of Nebraska Medical Center. [α-^{32}P]dCTP (1 mCi, 6,000Ci/mmol in 50 μL H_2O) and [γ-^{32}P]ATP were purchased from MP Biomedical (formerly ICN Diagnostic, Inc., Solon, OH). RNA Millennium Markers (0.5–9.0 Kb) was obtained from Ambion (Austin, TX).

Cell Culture

A549 human lung carcinoma cells were grown in F-12 K medium containing 10% fetal bovine serum (FBS); HBE1 human bronchia epithelial cells were cultured in F-12 medium supplemented with insulin (5 μg/ml), transferrin (5 μg/ml), EGF (10 ng/ml), dexamethasone (0.1 μM), cholera toxin (20 ng/ml), and bovine hypothalamus extract (15–30 μg/ml); human pancreas duct epithelioid carcinoma cell (PANC-1) and human colon colorectal adenocarcinoma epithelial cell LS 180 were grown in MEM medium with 10% FBS. All media were supplemented with 50 U/ml penicillin and 50 μg/ml streptomycin.

Isolation of Total RNA and Polyadenylated RNA

Total RNA from the cultured cells were isolated using TRI reagent. DNA-free DNase treatment and removal reagents (Ambion Inc.) were used to remove the trace amount of DNA contaminant in RNA preparation. Polyadenylated RNA was isolated from the total RNA with Poly-ATract mRNA Isolation System III.

Quantitative Real-Time PCR Analysis of hC2GnT-M Gene Expression in Various Tissues and Cell Lines

The cDNAs from 24 human tissues for quantitative real-time PCR were purchased from Sure-Race Multi-Tissue RACE Panels (10 ng/in 5× well). Human trachea polyA$^+$ RNA from Clontech was reverse-transcribed into cDNA using random primer and StrataScript First-Strand Synthesis System. Gene-specific primer sets and probes were designed with Primer Express 2.0 software (ABI) according to the published gene sequences of hC2GnT-M (22) (Gene Bank Accession No. AF102542) and human β-actin (Gene Bank Accession No. BC013835). The primer set and probe for hC2GnT-M were as follow: forward primer, 5-AAAGACACTTATAGCCCAGATGAACA-3; reverse primer, 5-GA AGTCATGTCTGAGATGTCGTACTTG-3; probe, 6FAM-CTCTGTT CCCAACCAC-NFQ. The primer set and probe for human β-actin were as follows: forward primer, 5-GCCGTCTTCCCCTCCAT-3; reverse primer, 5-CTGACCCATGCCCACCAT-3; probe, 6FAM-ACGCCCTG GTGCCTG-NFQ. The 20× primers and probe mix (18 μM of each primer and 5 μM of probe), 2× TaqManUniversal PCR Master Mix, and cDNA were used to set up the reaction, which contained all reagents for PCR. The standard curves for hC2GnT-M and β-actin were generated by serial dilution of cDNA prepared from A549 total RNA. Triplicate reactions from each cDNA were performed for real-time PCR, which included 1 cycle at 50°C for 2 min; 1 cycle at 95°C for 10 min; and 40 cycles at 95°C for 10 s, and at 60°C for 1 min. PCR products were measured continuously with an ABI PRISM 7500 Sequence Detection System (Applied Biosystems). The relative amounts of the transcripts were normalized to the amount of human β-actin transcript in the same cDNA sample.

5′RACE and 3′RACE

For 5′RACE and 3′RACE in A549 and HBE cells, GeneRacerKit was used. This 5′RACE technique was based on RNA ligase-mediated (RLM-RACE) and oligo-capping amplification of cDNA ends (RACE) method, which produced the selective ligation of an RNA oligonucleotide to the 5′ ends of decapped mRNA using T4 RNA ligase. For 5′RACE, total RNA was treated with calf intestinal phosphatase to remove the 5′ phosphates, then treated with tobacco acid pyrophosphatase to remove the 5′ cap structure from intact, full-length mRNA, and leave a 5′phosphate to ligate with the 3′end of the GeneRacer RNA Oligo. Subsequently, the ligated mRNA was reverse-transcribed into RACE-ready first-strand cDNA using random primer, and amplified with Gene Racer 5′ Primer (5-CGACTGGAGCACGAGGACACTG A-3) as well as gene-specific primer (GSP) 1 (hC2GnT-M-Exon 2-GSP 1: 5-G CGGTGTTCCGTGGCAGCAAATGTGAA-3) for the first round PCR reaction; the condition was 94°C 2 min; 94°C 30 s, 72°C 3 min (5 cycles); 94°C 30 s, 70°C 3 min (5 cycles); 94°C 30 s, 68°C 30 s, 72°C 3 min (30 cycles); 72°C 10 min. The product generated was used as the template for the second round nested PCR reaction with Gene Racer 5′ nested primer (5-GGACACTGACATGGACTGAAGGAGTA-3) and gene-specific primer 2 (hC2GnT-M-Exon 2-GSP 2: 5-GAGGG CAGCAGGTGGCTCCATCACTCTC-3); the condition for the second round nested PCR was 94°C 5 s, 60°C 30 s, 72°C 3 min (30 cycles); 72°C 10 min. The PCR products were cloned into pCRII 2.1 vector for DNA sequencing.

For 3′RACE, hC2GnT-M-3RACE-GSP 1 outer primer (5-GCC TTGCAAATTGCTGCCTGGGTGA-3) and Gene Racer 3′ reverse primer (5-GCTGTCAACGATACGCTACGTAACG-3) were used for the first-round PCR; hC2GnT-M-3RACE-GSP 2 inner primer (5-TG CTGCCTGGGTGAATGCTGCTTGTTC-3) and Gene Racer 3′ nested primer (5-CGCTACGTAACGGCATGACAGTG-3) were used for the second-round PCR. The condition for two rounds of nested PCR was the same as that used for the 5′RACE experiment. For 5′RACE in 24 different human tissues, Sure-Race Multi-Tissue RACE Panel (OriGene Technologies, Inc.), which contained cDNAs prepared from 22 individual adult human tissues and 2 human fetal tissues (fetal brain and liver), were used as the templates for another round of 5′RACE analysis. The RACE was also performed by two rounds of touch-down PCR reaction to reduce the background and nonspecific reaction. The first round of PCR was performed in 20 μl of reaction system using a sealed 48-well plate containing 24 human RACE-ready cDNAs with outer adapter primer1 (ADP1, 5′ CGGAATTCGTCACTCAGCG 3′ 10 μM) and hC2GnT-M-Exon 2-GSP 1 primer to enrich the specific cDNAs; the condition was 94 °C 3 min; 10 cycles of 94°C 30 s, 68°C 30 s; 15 cycles of 94°C 30 s, 62°C 30 s, 72°C 3 min; then extension at 72°C for 6 min. The second round of nested PCR was performed by using 1 μl of 1:10 diluted reaction product as the template with inner adapter primer 2 (ADP2, 5′ AGCGCGTGAATCAGATCG 3′) and hC2GnT-M-Exon 2-GSP 2 primer to complete the 5′RACE reaction, the condition was 94 °C 3 min; 5 cycles of 94°C 30 s, 65°C 30 s, 72°C 3 min; 5 cycles of 94°C 30 s, 62°C 30 s, 72°C 3 min; 25 cycles of 94°C 30 s, 60°C 30 s, 72°C 3 min; then extension at 72°C for 6 min. The RACE products were analyzed on 1.2% agarose gel and the target bands were recovered and cloned into pCRII 2.1vector for DNA sequencing analysis.

Northern Blot Analysis

Total RNA (25–35 μg/lane) and RNA marker were fractionated on 2.2 M formaldehyde containing agarose gel (1%) (wt/vol) electrophoresis and transferred to a Nytran membrane using Turboblotter system (Schleicher and Schuell Bioscience, Keene, NH) and ultraviolet-crosslinked. The membrane was prehybridized in a solution containing 50% formamide, 5× standard saline citrate (SSC), 5× Denhardt's solution, 5 mM EDTA, 0.1% SDS, and 100 μg/ml denatured salmon sperm DNA at 42°C for 4 h. The membrane was hybridized with [α-^{32}P]-labeled probes in the prehybridization solution supplemented with 10% dextran sulfate (MW ~ 500,000 Da) at 42°C for 18 h. The membrane was washed with 0.1% SDS in 1× SSC at 55°C for 30 min once, then washed with 0.1% SDS in 0.25× SSC at 55°C for 30 min twice and exposed to Kodak BioMax MR film with intensifying screen at −80°C for 48–72 h.

In addition, Multiple-Choice Northern Blot premade membrane (OriGenes) was used for Northern blot analysis of 11 human tissues. For preparation of 120-bp exon 1 probe, DNA template was first amplified from A549 cDNA using the following primers: forward primer (5-CTGGGAAGCCCTGGGATTCTGCTAATACCTATCACTGTAGG TGC-3) and reverse primer (5-CTTTGCCAGGCGTCAGCTTGGTC-3). Subsequently, these probes were labeled with [α-^{32}P] dCTP by hot asymmetric PCR with the reverse primer, which only make the antisense strand alone. Fifty-microliter PCR labeling reaction was set up as follows: 10× PCR buffer (minus Mg^{2+}) 5 μl, dNTP (10 mM dATP, dGTP, and dTTP, and 60 μM dCTP) 1 μl, 50 mM $MgCl_2$ 1.5 μl, [α-^{32}P]dCTP (1 mCi, 6,000 Ci/mmol in 50 μl H_2O) 12 μl, 10 μM reverse primer 1 μl, 10 ng/μl DNA template 1 μl, Taq DNA polymerase 1 μl (2.5 units), and H_2O to 50 μl. PCR was performed at 94°C, 30 s; 50°C, 30 s; 70°C, 1 min for 35 cycles.

Creation of Promoter–Reporter Constructs

The pGL3-basic vector (Promega, Madison, WI), which was promoterless and enhancerless, was used to create the promoter–reporter constructs. Briefly, various lengths of nest-deleted DNA fragments upstream of the transcription initiation site were PCR-amplified from A549 genomic DNA with multiple Bgl II–tailed forward primers and one common Hind III–tailed reverse primer (Table 1). The PCR products were cloned into the PCR2.1 cloning vector and confirmed by DNA sequencing with T7 promoter primer. Subsequently, the DNA fragments were cut out and subcloned into luciferase reporter vector pGL3-Basic at Bgl II/Hind III restriction sites.

Generation of the Linker-Scanning Mutants

The QuickChange II XL site-directed mutagenesis kit (Stratagene, La Jolla, CA) was used to generate a series of linker-scanning mutants, and each mutant contained the 10-bp sequence 5′-GCAGATCCGC-3′ not known to cover any important cis-regulatory elements (34). All mutants were confirmed by sequence analysis. Seven linker-scanning mutant constructs were created from the template construct p508, and spanning a 70-bp region of the hC2GnT-M promoter from −321 to −252, relative to the transcription initiation site. Three linker-scanning mutant constructs were generated from the template construct p259, and spanning a 30-bp region of the hC2GnT-M promoter from −72 to −43. Primers used to create the linker-scanning mutants are listed in Table 2.

Transient Transfection and Luciferase Assay

A previously described transfection method (35) was modified and used in this experiment. The confluent A549 cells in 48-well culture plate were rinsed once with 500 μl of *N*-2-hydroxyethylpiperazine-*N*9–2-ethanesulfonic acid (HEPES)-buffered saline (HBS) (20 mM HEPES [pH 7.2] containing 100 mM NaCl). These cells were then exposed to 500 μl transfection solution that had been prepared as described below. In 12- × 75-mm polystyrene tubes (Becton Dickinson, Lincoln Park, NJ), the following reagents were sequentially added, gently mixed, and subsequently incubated for 15 min at room temperature after each addition: 100 μl HBS containing 5 μg PNA and 2.75 μl of 2 mg/ml DMRIE-C reagent,100 μl HBS containing 0.8 μg construct, and 0.2 μg pSV-β-galactosidase control plasmid. The 200-μl mixture was transferred to each well, which contained 300 μl plain medium (serum-free) without antibiotics, and was then gently mixed. The plate was incubated for 6–8 h at 37°C under a water-saturated environment in the presence of 5% CO_2. The conditioned media were then replaced with 1.0 ml of complete medium containing 10% FBS and antibiotics (50 U/ml penicillin and 50 mg/ml streptomycin). After incubation for 40 h, cells were lysed with 1× passive lysis buffer (100 μl/well for 48-well plate) and the supernatants were isolated after a brief centrifugation. A 20-μl aliquot of the cell lysates was mixed with 100 μl luciferase assay reagent, and then the light produced was measured with a luminometer. For transfection of LS180, Panc1 and HBE cells, 5 μg MAA, 0.05 μg HSA, and 10 μg transferrin instead of 5 μg PNA was used, respectively.

β-Galactosidase Assay

For quantitative β-galactosidase assay, 20 μl of the same cell lysate as that used for luciferase assay was added to each well of a 96-well plate. The sample volume was brought to 50 μl with H_2O, and 50 μl of H_2O was used as the blank. Next, 100 μl of β-galactosidase assay reagent (Pierce, Rockford, IL) was added to each well. After incubation at 37°C for 30 min, 100 μl of 0.5 M Na_2CO_3 was added to each well to stop the reaction. The

TABLE 1. SEQUENCES OF THE OLIGONUCLEOTIDES USED FOR PCR TO PRODUCE 5′- TO 3′ SERIAL TRUNCATION OF hC2GnT-M PROMOTER-LUCIFERASE CONSTRUCTS

Primer Sequences	Orientation	Product Regions
5-tagatctAATCTCTTTTCTCTTCCACTCCAG-3	F	−417/+187
5-tgaattagaTCTAGAGCAAGGGGGAACTTTGAG-3	F	−364/+187
5-tgaattagaTCTAAGGAGGAACCTACAGGCTGTC-3	F	−321/+187
5-tgaattagaTCTGATTAGAAACCAGTTTGATCAC-3	F	−249/+187
5- tgaattagaTCTAGAACATTTATTTAATGTTAGA-3	F	−158/+187
5-tgaattagaTCTAGCTCCAGCCTAATAAGTAACAC-3	F	−72/+187
5-tgaattagaTCTGGGTTTTCCTATTTATCTATCC-3	F	−42/+187
5- tgaattagaTCTCTCGCATTACTTCTCTGAGTCA-3	F	−19/+187
5- tgaattagaTCTAAGTCACGGGAACTGCCCTTG-3	F	+18/+187
5-tgaagtaagcttGTCAGCTTGGTCTTCTCATTGAC-3	R	

Definition of abbreviations: F, forward primer; R, reverse primer.

Bgl II (AGATCT) and Hind III (AAGCTT) restriction sites added at the 5′ end of the primers used for subcloning are underlined. Ranges of the products relative to the transcription initiation site are indicated.

plate was then read at 405 nm with an enzyme-linked immunosorbent assay plate reader.

Statistical Analysis of Data

The prism (GraphPad, San Diego, CA) software package was used for computing two-tailed P values using an unpaired t test.

RESULTS

Expression Levels of hC2GnT-M Gene in Various Tissues and A549 Cells

Previously, Yeh and coworkers (22) employed Northern blot analysis using ORF probe to show that hC2GnT-M gene was expressed primarily in mucus-secreting tissues. In the present study, quantitative real-time PCR was performed to evaluate the expression levels of hC2GnT-M gene in 24 human tissues and A549 lung carcinoma cells. The data after normalization with β-actin (Figure 2) showed that this gene was expressed in all mucus-secretory tissues, which confirmed previous report (22), although the expression levels varied. Specifically, hC2GnT-M was expressed at the highest level in colon; at the second highest level in testis, stomach, and small intestine; at a moderate level in adrenal glands, kidney, trachea, thyroid glands, uterus, ovary, pancreas, and A549 cells; at a low level in lung, pituitary glands, prostate, peripheral blood lymphocytes, and fetal liver; but at an undetectable level in brain, heart spleen, liver, muscle, placenta, fetal brain, fat, and mammary glands.

Identification and Cloning of the 5′ Flanking Sequence of hC2GnT-M Gene in Mucus-Secretory Tissues and Lung Epithelial Cells

To understand the regulation of C2GnT-M gene, we initiated studies to identify and clone the 5′ flanking sequence of hC2GnT-M gene. Alignment of the previously reported 2.1-kb hC2GnT-M transcript sequence (22) with the human genome database at NCBI website (http://www.ncbi.nlm.nih.gov/BLAST/) suggested that the reported hC2GnT-M transcript was a spliced product containing two exons. One of the exons was 1,799 bp in length and contained a 5′ untranslated region (UTR) (59 bp), an open reading frame (ORF, 1,314 bp), and a 3′UTR (426 bp). The second exon was 293 bp in length and located 1,297 bp upstream of the other exon. The 1,297-bp DNA fragment contained the splicing signature sequences, GT and AG, at the 5′ and 3′ ends, respectively (36).

TABLE 2. PRIMER PAIRS USED TO CONSTRUCT p508 AND p259 LINKER-SCANNING MUTANTS WHICH SPAN A 70-bp REGION OF hC2GnT-M PROMOTER FROM −321 TO −252 AND A 30-bp REGION FROM −72 TO −43

Primer Sequence	Orientation	Mutant	Mutated Region
5-GTGCTAGCCCGGGCTCGAGAGCAGATCCGCAACCTACAGGCTGTCCTG-3	F	LS1	−321/−312
5-CAGGGACAGCCTGTAGGTTGCGGATCTGCTCTCGAGCCCGGGCTAGCAC-3	R		
5-GCCCGGGCTCGAGATCTAAGGAGGGCAGATCCGCCTGTCCCTGTCTTTCAGCAAATTGC-3	F	LS2	−311/−302
5-GCAATTTGCTGAAAGACAGGGACAGGCGGATCTGCCCTCCTTAGATCTCGAGCCCGGGC-3	R		
5-CGAGATCTAAGGAGGAACCTACAGGGCAGATCCGCCTTTCAGCAAATTGCATCCAGTCTG-3	F	LS3	−301/−292
5-CAGACTGGATGCAATTTGCTGAAAGGCGGATCTGCCCTGTAGGTTCCTCCTTAGATCTCG-3	R		
5-GAGGAACCTACAGGCTGTCCCTGTGCAGATCCGCATTGCATCCAGTCTGAGCTTCAC-3	F	LS4	−291/−282
5-GTGAAGCTCAGACTGGATGCAATGCGGATCTGCACAGGGACAGCCTGTAGGTTCCTC-3	R		
5-CAGGCTGTCCCTGTCTTTCAGCAAGCAGATCCGCGTCTGAGCTTCACTTTGCAACCCTG-3	F	LS5	−281/−272
5-CAGGGTTGCAAAGTGAAGCTCAGACGCGGATCTGCTTGCTGAAAGACAGGGACAGCCTG-3	R		
5-CCTGTCTTTCAGCAAATTGCATCCAGCAGATCCGCCACTTTGCAACCCTGATTAGAAACC-3	F	LS6	−271/−262
5-GGTTTCTAATCAGGGTTGCAAAGTGGCGGATCTGCTGGATGCAATTTGCTGAAAGACAGG-3	R		
5-GCAAATTGCATCCAGTCTGAGCTTGCAGATCCGCCCCTGATTAGAAACCAGTTTGATC-3	F	LS7	−261/−252
5-GATCAAACTGGTTTCTAATCAGGGGCGGATCTGCAAGCTCAGACTGGATGCAATTTGC-3	R		
5-CGCGTGCTAGCCCGGGCTCGAGATGCAGATCCGCCCTAATAAGTAACACATTAACTGGG-3	F	LS8	−72/−63
5-CCCAGTTAATGTGTTACTTATTAGGGCGGATCTGCATCTCGAGCCCGGGCTAGCACGCG-3	R		
5-CTAGCCCGGGCTCGAGATCTAGCTCCAGGCAGATCCGCAACACATTAACTGGGTTTTCC-3	F	LS9	−62/−53
5-GGAAAACCCAGTTAATGTGTTGCGGATCTGCCTGGAGCTAGATCTCGAGCCCGGGCTAG-3	R		
5-CTAGCTCCAGCCTAATAAGTGCAGATCCGCCTGGGTTTTCCTATTTATCTATCCTCTCGC-3	F	LS10	−52/−43
5-GATAAATAGGAAAACCCAGGCGGATCTGCACTTATTAGGCTGGAGCTAGATCTCGAGCC-3	R		

Definition of abbreviations: F, forward primer and R, reverse primer.

Underlined sequences refer to mutated nucleotides. The p508 mutants include LS1–7; the p259 mutants include LS8–10.

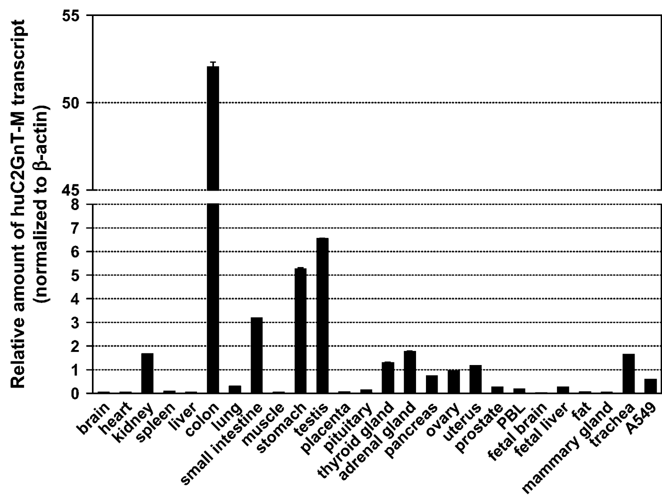

Figure 2. Expression level of hC2GnT-M gene in various tissues by quantitative real-time PCR. The expression level of the hC2GnT-M gene was normalized to that of β-actin measured in same cDNA preparation. Data obtained from triplicate experiments of same sample are shown as mean ± SD.

To identify the transcription initiation site, 5′RACE was performed. Commercial kits that were based on the presence of 5′cap structure of the transcripts were employed to ensure that only full-length transcripts were amplified. Further, nested PCR using two hC2GnT-M gene-specific primers was performed to increase the specificity of the RACE products. Blasting the human genomic DNA database with gene-specific sequences of the RACE products located these DNA sequences at the region upstream of hC2GnT-M gene at q21.3 on chromosome 15 (22). This result indicated that all the sequences obtained in this experiment were located in the region contiguously adjacent to the 5′ end of hC2GnT-M gene. The alignment also revealed a new exon and a new intron along with additional sequences that extended the 5′ end of the second exon of the previously identified transcript (22). The length of the second exon was extended to 333 and 401 bp. The 333-bp exon was found in all mucus secretory tissues and airway epithelial cells examined, while the 401-bp exon was detected in A549 cells only (Figures 3 and 4A). The 5′ ends of these exons were adjacent to an AG splicing signature sequence located at the 3′ end of the newly identified intron. Since there was only one exon identified beyond this intron, this intron was designated as intron 1. The newly identified exon was designated as exon 1, while the second exon was designated as exon 2. Intron 1 was ~ 4.5 kb in size and flanked by GT and AG splicing signature sequences at the 5′ and 3′ ends, respectively. Exon 1 was heterogeneous in size, which ranged from 69–198 bp depending on tissues and cells (Figures 3 and 4A). Since E1 obtained from intestine tissue was the longest, the 5′ end (which was located at 6,455 bp upstream of the translation start site ATG) of E1 from this tissue was designated as the transcription initiation site (+1).

In addition to 5′ RACE, we also performed 3′ RACE to identify 3′UTR. The 3′UTR sequence obtained from all tissues and cells examined was very similar to that reported previously (22) except for an additional 65 bp at the poly(A) tail region, which extended the length of 3′UTR to 491 bp. The longest transcript identified from small intestine tissue in this study has been submitted to the GenBank with an accession number of EF152283.

Detection of E1-Containing hC2GnT-M Transcripts in Different Tissues and A549 Cell by Northern Blotting

To confirm the presence of E1 in the three different size hC2GnT-M transcripts, Northern blot analysis of various tissues and A549 cells was performed using a 120-bp probe (+79/+198) located at the 3′ end of E1. As shown in Figure 4B, E1 was present in all transcripts with the following relative abundance: ~ 2.3–2.5 kb > > ~ 6.8–7.0 kb > ~3.6–3.8 kb. Similar Northern blots were obtained when exon 2 and exon 3 were used as the probes (data not shown).

Isolation and Functional Characterization of the Promoter of hC2GnT-M Gene

The DNA sequence around the transcription initiation site of hC2GnT-M gene is shown in Figure 5. To assess the functionality of the promoter region (−417/+187), we generated a series of 5′end truncated promoter fragments by PCR using genomic DNA of A549 cells as the template. The PCR products generated with multiple Bgl II–tailed forward primers and one common Hind III–tailed reverse primer were cloned into pGL3-Basic vector, which does not have promoter and enhancer. Oligonucleotides used for PCR are listed in Table 1. These truncated mutants were co-transfected with pSV-β-galactosidase plasmid into A549 cells to assess their promoter activities by transient transfection. pGL3-Basic was used as the negative control and pGL3-Control, which contains a SV40 promoter, as the positive control. As shown in Figure 6, there were two positive regulatory regions in the proximal core promoter: one was −321/−249 ($P < 0.001$, as compared with full-length promoter) and the other one was −72/−42 ($P < 0.001$, as compared with full-length promoter). This region contained the most critical sequence for the basal promoter activity. Similar results were obtained for another cell line, human colorectal adenocarcinoma epithelial cell LS180, which expressed high C2GnT-M activity, while human pancreas duct epithelioid carcinoma cell PANC-1, which did not express detectable C2GnT-M activity, showed a promoter activity close to the background (data not shown).

Identification of Important cis-Regulatory Elements

To further narrow down the critical cis-regulatory elements in the promoter region between −321 and −249, and between −72 and −42, a linker scanning analysis was performed by replacing consecutive

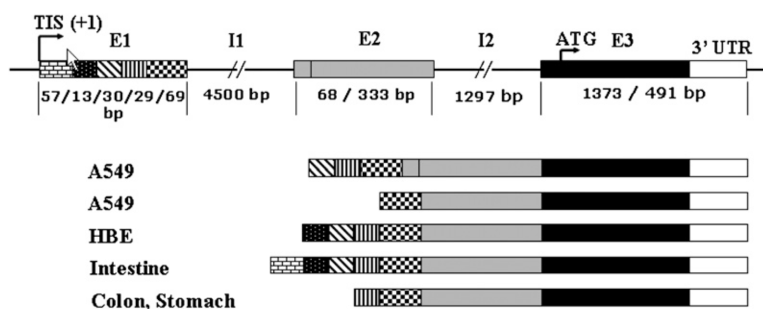

Figure 3. Schematic presentation of hC2GnT-M transcripts in major mucus-secretory tissues by alignment of the sequences of 5′RACE products with the human genome database. The *top diagram* represents the hC2GnT-M genomic structure at q21.3 on chromosome 15. The previously reported hC2GnT-M transcript (22) is ~ 2.1 kb in length, and contains 293-bp exon 2 and 1,799-bp exon 3, in which 426 bp is 3′UTR. Current study extends the sizes of these transcripts to 2,266–2,463 bp by identification of 69–198-bp exon 1, 333–401-bp exon 2, and 1,864-bp exon 3, including 491 bp of 3′UTR. TIS, transcription initiation site designated as +1; E, exon; I, intron; UTR, untranslated region.

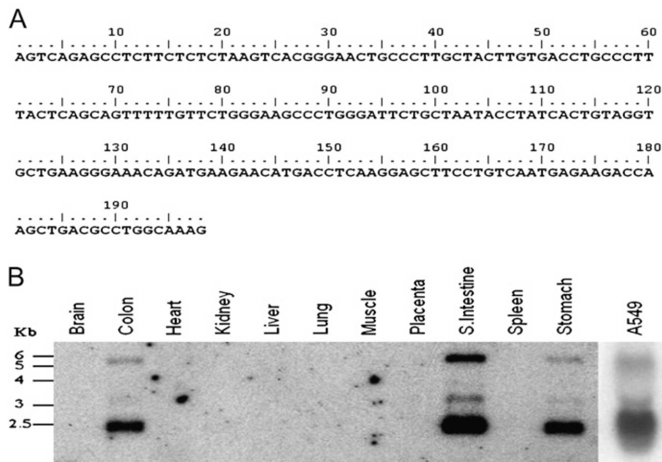

Figure 4. (*A*) The DNA sequence of exon 1 region of hC2GnT-M gene. The number +1 represents the transcription initiation site. (*B*) Northern blot of hC2GnT-M poly(A+) mRNA in eleven human tissues and A549 cells using E1 probe. The 12- bp probe was located near the 3′ end of exon 1, ranging from +79 to +198. Each *lane* contained ~ 2 μg poly(A+)RNA. The relative labeling intensity for these three bands was 2.3–2.5 kb > > 6.8–7.0 kb > 3.6–3.8 kb. The size estimate for these transcripts on Northern blot was within the margin of error (± 15%) of the sizes of the standards used for Northern blotting.

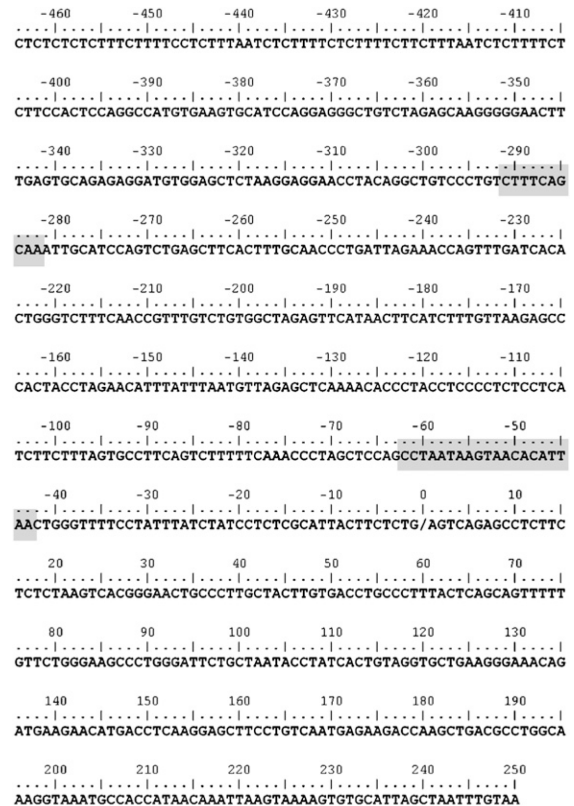

Figure 5. DNA sequence of proximal promoter of the hC2GnT-M gene. Numberings are relative to the transcription initiation site (+1) in intestinal C2GnT-M transcript identified by 5′ RACE. Two important positive regulatory domains (PRD1, −291/−282, and PRD2, −62/−43) (*see* Figure 6) are indicated by *gray shading*.

10-bp sequences with 5′-GCAGATCCGC-3′ in these two regions. This sequence is not known to contain any important cis-regulatory elements (34). As shown in Figure 7, a significant decrease in luciferase activity was observed for LS4 (−33%) ($P < 0.05$ as compared with p508), LS9 (−73%), and LS10 (−74%)($P < 0.0001$ as compared with p259) constructs, indicating that the two positive regulatory domains (PRD), PRD1 (−291/−282) (Figure 7A) and PRD2 (−62/−43) (Figure 7B), were important for the proximal basal promoter activity of hC2GnT-M gene.

ATRA and IL-13 Enhanced hC2GnT-M Promoter Activity in A549 Cells

It is known that vitamin A is involved in development and differentiation of airway epithelial mucosa, and our previous report demonstrated that ATRA up-regulated hC2GnT-M activity in human lung carcinoma cells (32). To examine whether ATRA could enhance hC2GnT-M promoter activity in A549 cells, these cells were transfected with two representative promoter-reporter constructs p604 and p259. As shown in Figure 8A, treatment with 100 nM ATRA for 96 h increased the promoter activities of p604 and p259 constructs by 118 and 150%, respectively. Also, treatment with 20 ng/ml IL13 for 24 h increased the promoter activities of p604 and p259 constructs by 37 and 63%, respectively (Figure 8B). These results were supported by increased C2GnT-M gene expression measured by quantitative real-time PCR after treatment of A549 cells with ATRA or IL13 (Figure 8C).

DISCUSSION

Human (22) and bovine (37) C2GnT-M cDNAs have been cloned and human C2GnT-M gene has been shown to be regulated by EGF (31), Th2 cytokines (32), and ATRA (32). However, very little is known about the regulatory mechanism of this gene. Given the importance of this enzyme in the formation of all three mucin glycan branch structures, we initiated current study to determine the structure of hC2GnT-M gene, identify the promoter region, and then characterize its cis-regulatory elements.

By performing quantitative real-time PCR, we not only confirm previous report that this gene is expressed primarily in mucus-secretory tissues (22), but also reveal the relative abundance of hC2GnT-M transcripts in these tissues. Further, we extend previous report (22) by showing that the hC2GnT-M gene is ~ 8.26 kb and contains three exons and two introns. The structure of the hC2GnT-M gene is similar to that of bovine C2GnT-M gene except that bovine gene is only 6.3 kb (37). In addition, there are differences in exon and intron utilization of the C2GnT-M gene between these two species. In the case of the bovine C2GnT-M gene, four different transcripts are distributed in two different sizes. Two of the transcripts are intronless and the other two contain intron 2. Within each group of similarly sized transcripts, one contains exon 1 and the other one does not. Furthermore, E1-containing transcripts are expressed only in trachea and testis. For the hC2GnT-M gene, there does not seem to have any tissue-specific expression of the transcripts. In addition, three different size transcripts, including 2.3–2.5, 3.6–3.8, and 6.8–7.0 kb, are identified. Exon 1 is present in all transcripts and had same 3′-end but different 5′-ends. Linking of exon 1 to GT at the 5′-end of intron 1 suggests that this is a potential splicing site. Heterogeneous 5′-ends of exon 1 suggest that they are multiple transcription start sites as reported for other genes (38, 39). We also observe heterogeneous sizes of exon 2s in the A549 cell line (Figure 3). These exons 2s share same 3′-end but exhibit different 5′-ends and are flanked by AG and GT at 5′ and 3′ ends in the hC2GnT-M gene, respectively, suggesting that they are products of alternative splicing events. Based on the genomic structure, sizes of these transcripts as estimated by Northern

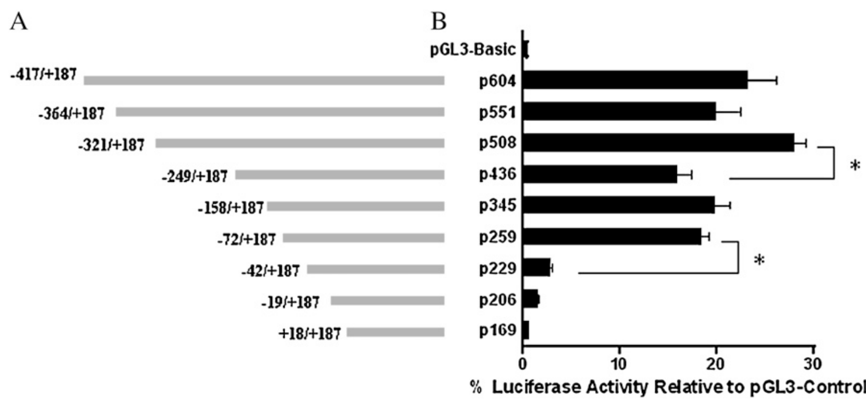

Figure 6. Functional characterization of proximal promoter of the hC2GnT-M gene in A549 cells. (*A*) Schematic representation of the 5′-flanking region of hC2GnT-M gene and serial truncation constructs. The *numbers* adjacent to truncation constructs represent the positions of 5′ and 3′ ends of the inserts, which are relative to the transcription initiation site (+1), and all the inserts share a common 3′ end at position +187. (*B*) The truncation constructs were co-transfected with pSV-β-galactosidase plasmid into A549 cells for transient expression, pGL3 basic and pGL3-control were used as the negative and positive control, respectively. Specific activity of luciferase in each construct was normalized to the specific activity of β-galactosidase and then plotted as a percentage of that of the positive control plasmid (pGL3-control). Data are shown as mean ± SEM; * $P < 0.001$; $n = 4$ independent samples.

analysis, and detection of exon 1, exon 2, and exon 3 in all transcripts, these three size transcripts are likely derived from variation of intron use. The smallest transcripts (∼ 2.3–2.5 kb) are likely intronless and contain exon 1 (69–198 bp), exon 2 (333–401 bp), and exon 3 (1,864 bp). The intermediate (∼ 3.6–3.8 kb)- and the largest (∼ 6.8–7.0 kb)-sized transcripts contain, in addition to all three exons, intron 2 (1.3 kb) and intron 1 (4.5 kb), respectively. These two size transcripts most likely represent incompletely processed messenger RNAs.

It has been reported that human (29) and mouse (27, 28) C2GnT-L genes, which span 53–60 kb and are distributed into 6 exons and 5 introns, use multiple transcription initiation sites in a tissue-specific manner. The ubiquitous nature of tissue expression of C2GnT-L gene supports the need for such mechanisms to control where, when, and how the C2GnT-L gene will be expressed to meet the specific need of each tissue/organ. Since C2GnT-M gene is expressed primarily in mucus-secretory tissues, it is not a surprise that the mechanism for regulating C2GnT-M gene expression is not as complex as that for C2GnT-L gene. Among the three C2GnT genes, C2GnT-3 gene (23) is the least known. The other two C2GnT genes are regulated differently. For example, expression of hC2GnT-L gene is only moderately enhanced by Th2 cytokines (32), inhibited by EGF (31), but not affected by ATRA (32). Expression of hC2GnT-M gene is inhibited by EGF (31), and greatly enhanced by Th2 cytokines and ATRA (32). These results further re-enforce the potential differences in biological functions between these two C2GnT isozymes as predicted from the difference in their tissue distribution.

Since E1 is present in all three mRNA transcript bands (Figures 3 and 4), we surmise that the promoter region of this gene would be located upstream of and/or near the 5′ end of E1. Therefore, we clone the promoter region (−417/+187) of hC2GnT-M gene and characterize the basal promoter activity by transient transfection assay of truncated promoter-reporter constructs. We identify two small regions, −62/−43 and −291/−282, that contain critical sequences for basal promoter activity. We find that the promoter activity is correlated with the level of C2GnT-M activity in different cells, supporting the notion that these promoter sequences are specific for hC2GnT-M gene. In addition, hC2GnT-M promoter activity is enhanced by ATRA and IL-13 (Figure 8). The presence of putative retinoic acid (RAREs) and STATx response elements at +40/+69 and +77/+85 may be responsible for the promoter activities of these reporter constructs in response to treatment with ATRA and IL-13, respectively. Detailed functional studies using mutated response element-reporter constructs and transcription factor expression plasmids are needed to confirm the prediction.

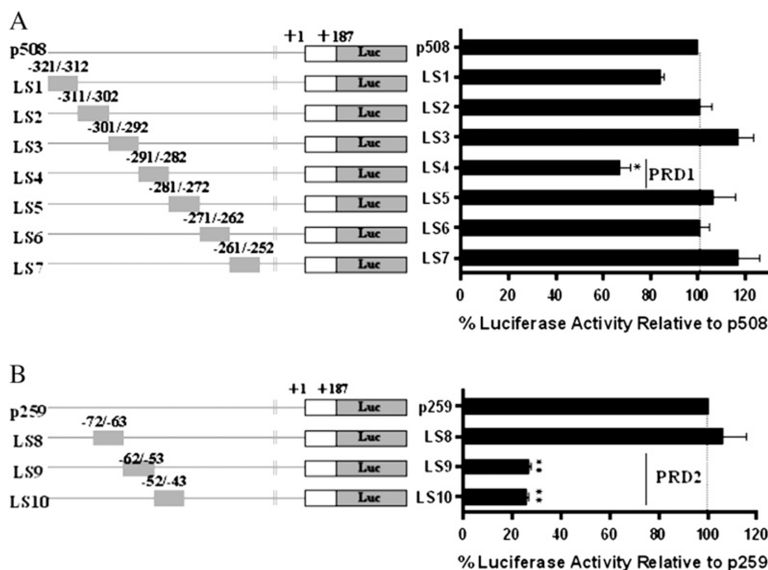

Figure 7. Linker-scanning mutational analysis of important regions in the proximal promoter of hC2GnT-M gene in A549 cell line. All mutant constructs were introduced with the 10-bp sequence of 5′-GCAGATCCGC-3′ not known to contain important cis-regulatory elements. Seven mutants (LS1 ∼ 7) spanning a 70-bp region of the hC2GnT-M promoter from −321 to −252, relative to the transcription initiation site were generated. Three mutants (LS8 ∼ 10) spanning a 30-bp region of the human C2GnT-M promoter from −72 to −43. Specific activity of luciferase in each construct was normalized to the specific activity of β-galactosidase and then plotted as a percentage of the wild-type plasmid. Two positive regulatory domains (PRD1, −291/−282, and PRD2, −62/−43) were identified. Data are shown as mean ± SEM; *$P < 0.05$; **$P < 0.01$; $n = 4$ independent samples.

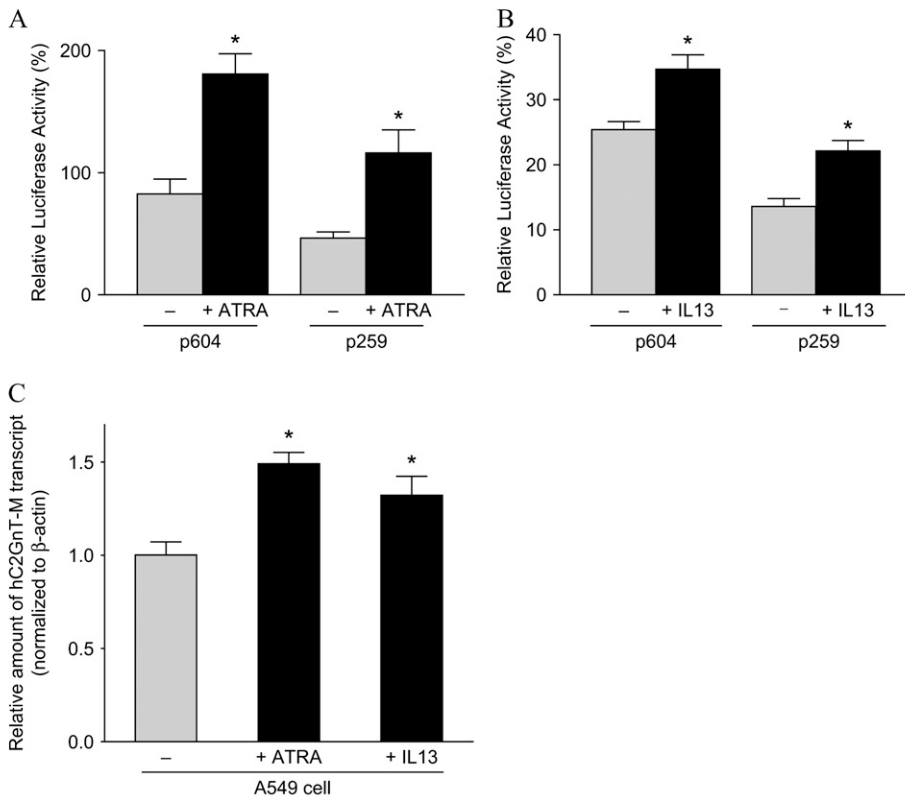

Figure 8. Effects of ATRA (*A*) and IL-13 (*B*) treatment on hC2GnT-M promoter activity and (*C*) expression of hC2GnT-M gene in A549 cells. Two representative hC2GnT-M gene promoter-reporter constructs, p604 and p259, were co-transfected with pSV-β-galactosidase plasmid into A549 cells for transient expression. After transfection, the cells were treated with 100 nM retinoic acid for 96 h or 20 ng/ml IL-13 for 24 h. The specific activity of luciferase was measured and normalized to that of β-galactosidase, then plotted as a percentage of that of the positive control plasmid (pGL3-control). In this assay, pGL3 basic and pGL3-control were used as the negative and positive control, respectively. For measurement of hC2GnT-M gene expression by quantitative real-time PCR, A549 cells were treated with ATRA for 72 h or IL-13 for 24 h first. After normalization to β-actin message, data are shown as means ± SEM; $^*P < 0.05$; $n = 4$ independent samples.

We believe that identification of the cis-regulatory elements of hC2GnT-M gene could lead to the development of mucus cell–specific gene therapy vectors. The therapeutic genes may include RNAi against secretory mucins and selective glycosyltransferase such as core 1 β1,3-galactosyltransferase (40) to suppress the production of mucins or fully glycosylated mucins, respectively. Also, other genes that can suppress mucus secretion may be delivered with same vector. The potency of this therapeutic strategy may be further enhanced by treatment with ATRA and/or Th2 cytokines.

***Conflict of Interest Statement*:** None of the authors has a financial relationship with a commercial entity that has an interest in the subject of this manuscript.

***Acknowledgments*:** The authors acknowledge Dr. Liping Liu's assistance in real-time PCR and Helen Cheng's assistance in cell culture and transient transfection.

References

1. Kim YS, Gum J Jr, Brockhausen I. Mucin glycoproteins in neoplasia. *Glycoconj J* 1996;13:693–707.
2. Rose MC, Voynow JA. Respiratory tract mucin genes and mucin glycoproteins in health and disease. *Physiol Rev* 2006;86:245–278.
3. Perez-Vilar J, Hill RL. The structure and assembly of secreted mucins. *J Biol Chem* 1999;274:31751–31754.
4. Gendler SJ, Spicer AP. Epithelial mucin genes. *Annu Rev Physiol* 1995; 57:607–634.
5. Kannagi R, Izawa M, Koike T, Miyazaki K, Kimura N. Carbohydrate-mediated cell adhesion in cancer metastasis and angiogenesis. *Cancer Sci* 2004;95:377–384.
6. Hollingsworth MA, Swanson BJ. Mucins in cancer: protection and control of the cell surface. *Nat Rev Cancer* 2004;4:45–60.
7. Uchimura K, Gauguet JM, Singer MS, Tsay D, Kannagi R, Muramatsu T, von Andrian UH, Rosen SD. A major class of L-selectin ligands is eliminated in mice deficient in two sulfotransferases expressed in high endothelial venules. *Nat Immunol* 2005;6:1105–1113.
8. Kawashima H, Petryniak B, Hiraoka N, Mitoma J, Huckaby V, Nakayama J, Uchimura K, Kadomatsu K, Muramatsu T, Lowe JB, *et al.* N-acetylglucosamine-6-O-sulfotransferases 1 and 2 cooperatively control lymphocyte homing through L-selectin ligand biosynthesis in high endothelial venules. *Nat Immunol* 2005;6:1096–1104.
9. Tsuboi S, Fukuda M. Roles of O-linked oligosaccharides in immune responses. *Bioessays* 2001;23:46–53.
10. Lamblin G, Lhermitte M, Klein A, Houdret N, Scharfman A, Ramphal R, Roussel P. The carbohydrate diversity of human respiratory mucins: a protection of the underlying mucosa? *Am Rev Respir Dis* 1991;144:S19–S24.
11. Hounsell EF, Lawson AM, Stoll MS, Kane DP, Cashmore GC, Carruthers RA, Feeney J, Feizi T. Characterisation by mass spectrometry and 500-mhz proton nuclear magnetic resonance spectroscopy of penta- and hexasaccharide chains of human foetal gastrointestinal mucins (meconium glycoproteins). *Eur J Biochem* 1989;186:597–610.
12. Xia B, Royall JA, Damera G, Sachdev GP, Cummings RD. Altered O-glycosylation and sulfation of airway mucins associated with cystic fibrosis. *Glycobiology* 2005;15:747–775.
13. Breg J, Van Halbeek H, Vliegenthart JF, Lamblin G, Houvenaghel MC, Roussel P. Structure of sialyl-oligosaccharides isolated from bronchial mucus glycoproteins of patients (blood group O) suffering from cystic fibrosis. *Eur J Biochem* 1987;168:57–68.
14. Piller F, Piller V, Fox RI, Fukuda M. Human T-lymphocyte activation is associated with changes in O-glycan biosynthesis. *J Biol Chem* 1988;263: 15146–15150.
15. Tsuboi S, Fukuda M. Overexpression of branched O-linked oligosaccharides on T cell surface glycoproteins impairs humoral immune responses in transgenic mice. *J Biol Chem* 1998;273:30680–30687.
16. Tsuboi S, Fukuda M. Branched O-linked oligosaccharides ectopically expressed in transgenic mice reduce primary T-cell immune responses. *EMBO J* 1997;16:6364–6373.
17. Hang HC, Bertozzi CR. The chemistry and biology of mucin-type O-linked glycosylation. *Bioorg Med Chem* 2005;13:5021–5034.
18. Machida E, Nakayama J, Amano J, Fukuda M. Clinicopathological significance of core 2 β1,6-N-acetylglucosaminyltransferase messenger RNA expressed in the pulmonary adenocarcinoma determined by in situ hybridization. *Cancer Res* 2001;61:2226–2231.
19. Ellies LG, Tsuboi S, Petryniak B, Lowe JB, Fukuda M, Marth JD. Core 2 oligosaccharide biosynthesis distinguishes between selectin ligands essential for leukocyte homing and inflammation. *Immunity* 1998;9:881–890.
20. Beum PV, Cheng PW. Biosynthesis and function of β 1,6 branched mucin-type glycans. *Adv Exp Med Biol* 2001;491:279–312.

21. Bierhuizen MF, Fukuda M. Expression cloning of a cDNA encoding UDP-GlcNAc:Gal β 1–3-GalNAc-R (GlcNAc to GalNAc) β 1–6GlcNAc transferase by gene transfer into CHO cells expressing polyoma large tumor antigen. *Proc Natl Acad Sci USA* 1992;89:9326–9330.
22. Yeh JC, Ong E, Fukuda M. Molecular cloning and expression of a novel β-1, 6-N-acetylglucosaminyltransferase that forms core 2, core 4, and I branches. *J Biol Chem* 1999;274:3215–3221.
23. Schwientek T, Yeh JC, Levery SB, Keck B, Merkx G, van Kessel AG, Fukuda M, Clausen H. Control of O-glycan branch formation: molecular cloning and characterization of a novel thymus-associated core 2 β1, 6-N-acetylglucosaminyltransferase. *J Biol Chem* 2000;275:11106–11113.
24. Bierhuizen MF, Mattei MG, Fukuda M. Expression of the developmental i antigen by a cloned human cDNA encoding a member of a β-1,6-N-acetylglucosaminyltransferase gene family. *Genes Dev* 1993;7:468–478.
25. Inaba N, Hiruma T, Togayachi A, Iwasaki H, Wang XH, Furukawa Y, Sumi R, Kudo T, Fujimura K, Iwai T, *et al.* A novel I-branching β-1,6-N-acetylglucosaminyltransferase involved in human blood group I antigen expression. *Blood* 2003;101:2870–2876.
26. Ropp PA, Little MR, Cheng PW. Mucin biosynthesis: Purification and characterization of a mucin β6 N-acetylglucosaminyltransferase. *J Biol Chem* 1991;266:23863–23871.
27. Sekine M, Nara K, Suzuki A. Tissue-specific regulation of mouse core 2 β-1,6-N-acetylglucosaminyltransferase. *J Biol Chem* 1997;272:27246–27252.
28. Sekine M, Taya C, Shitara H, Kikkawa Y, Akamatsu N, Kotani M, Miyazaki M, Suzuki A, Yonekawa H. The cis-regulatory element GSL5 is indispensable for proximal straight tubule cell-specific transcription of core 2 β-1,6-N-acetylglucosaminyltransferase in the mouse kidney. *J Biol Chem* 2006;281:1008–1015.
29. Falkenberg VR, Alvarez K, Roman C, Fregien N. Multiple transcription initiation and alternative splicing in the 5′ untranslated region of the core 2 β1–6 N-acetylglucosaminyltransferase i gene. *Glycobiology* 2003;13:411–418.
30. VanderElst IE, Datti A. β1,6 N-acetylglucosaminyltransferase (core 2 GlcNAc-T) expression in normal rat tissues and different cell lines: evidence for complex mechanisms of regulation. *Glycobiology* 1998;8:731–740.
31. Beum PV, Bastola DR, Cheng PW. Mucin biosynthesis: epidermal growth factor downregulates core 2 enzymes in a human airway adenocarcinoma cell line. *Am J Respir Cell Mol Biol* 2003;29:48–56.
32. Beum PV, Basma H, Bastola DR, Cheng PW. Mucin biosynthesis: upregulation of core 2 β 1,6 N-acetylglucosaminyltransferase by retinoic acid and Th2 cytokines in a human airway epithelial cell line. *Am J Physiol Lung Cell Mol Physiol* 2005;288:L116–L124.
33. Huang MC, Chen HY, Huang HC, Huang J, Liang JT, Shen TL, Lin NY, Ho CC, Cho IM, Hsu SM. C2GnT-M is downregulated in colorectal cancer and its re-expression causes growth inhibition of colon cancer cells. *Oncogene* 2006;25:3267–3276.
34. Karantzoulis-Fegaras F, Antoniou H, Lai SL, Kulkarni G, D'Abreo C, Wong GK, Miller TL, Chan Y, Atkins J, Wang Y, *et al.* Characterization of the human endothelial nitric-oxide synthase promoter. *J Biol Chem* 1999;274:3076–3093.
35. Cheng PW. Receptor ligand-facilitated gene transfer: enhancement of liposome-mediated gene transfer and expression by transferrin. *Hum Gene Ther* 1996;7:275–282.
36. Padgett RA, Grabowski PJ, Konarska MM, Seiler S, Sharp PA. Splicing of messenger rna precursors. *Annu Rev Biochem* 1986;55:1119–1150.
37. Choi KH, Osorio FA, Cheng PW. Mucin biosynthesis: Bovine C2GnT-M gene, tissue-specific expression, and herpes virus-4 homologue. *Am J Respir Cell Mol Biol* 2004;30:710–719.
38. Chen Z, Marsden PA, Gorczynski RM. Cloning and characterization of the human CD200 promoter region. *Mol Immunol* 2006;43:579–587.
39. Kawaji H, Frith MC, Katayama S, Sandelin A, Kai C, Kawai J, Carninci P, Hayashizaki Y. Dynamic usage of transcription start sites within core promoters. *Genome Biol* 2006;7:R118.
40. Ju T, Cummings RD, Canfield WM. Purification, characterization, and subunit structure of rat core 1 β1,3-galactosyltransferase. *J Biol Chem* 2002;277:169–177.