Promoters of Human Cosmc and T-synthase Genes Are Similar in Structure, Yet Different in Epigenetic Regulation*

Received for publication, March 24, 2015, and in revised form, May 29, 2015 Published, JBC Papers in Press, June 10, 2015, DOI 10.1074/jbc.M115.654244

Junwei Zeng1, Rongjuan Mi1,2, Yingchun Wang‡, Yujing Li3, Li Lin4, Bing Yao5, Lina Song1,3, Irma van Die4, Arlene B. Chapman1, Richard D. Cummings3, Peng Jin5, and Tongzhong Ju1,4,‡

From the Departments of1 Biochemistry, 3Human Genetics, and 4Medicine, Emory University School of Medicine, Atlanta, Georgia 30322 and the 5Department of Molecular Cell Biology and Immunology, VU University Medical Center, Amsterdam, The Netherlands

Background: Transcriptional regulation of human T-synthase and Cosmc is unknown.
Results: Minimal promoters for Cosmc and T-synthase are located in CpG islands, and the promoter of Cosmc, but not T-synthase, was hypermethylated in Tn4 B cells.
Conclusion: Promoters of Cosmc and T-synthase are similarly structured, yet different in epigenetic regulation.
Significance: These data aid in understanding the role of these genes in biology and diseases.

The T-synthase (core 1 β3-galactosyltransferase) and its molecular chaperone Cosmc regulate the biosynthesis of mucin type O-glycans on glycoproteins, and evidence suggests that both T-synthase and Cosmc are transcriptionally suppressed in several human diseases, although the transcriptional regulation of these two genes is not understood. Here, we characterized the promoters essential for human Cosmc and T-synthase transcription. The upstream regions of the genes lack a conventional TATA box but contain CpG islands, cCpG-I and cCpG-II for Cosmc and tCpG for T-synthase. Using luciferase reporter essays, site-directed mutagenesis, ChIP assays, and mithramycin A treatment, we identified the core promoters within cCpG-II and tCpG, which contain two binding sites for Krüppel-like transcription factors, including SP1/SP3, respectively. Methylome analysis of Tn4 B cells, which harbor a silenced Cosmc, confirmed the hypermethylation of the Cosmc core promoter but not for T-synthase. These results demonstrate that Cosmc and T-synthase are transcriptionally regulated at a basal level by the specificity protein/Krüppel-like transcription factor interactions (3–7), vascular biology (8), angiogenesis (9), and lymphangiogenesis (10). Interestingly, the endoplasmic reticulum-localized molecular chaperone Cosmc (core 1 β3-galactosyltransferase-specific molecular chaperone or C1GalT1C1) is required for active T-synthase formation in vivo by preventing its aggregation and subsequent proteasomal degradation (11, 12). Disruption of either T-synthase or Cosmc is embryonic lethal due to hemorrhaging, suggesting the important role of O-glycosylation in vascular biology (9, 13).

The structures of O-glycans change during cell differentiation and animal development, indicating that expression of glycosyltransferases is both temporally and spatially regulated. Yet the details on the mechanisms of regulation of these enzymes, especially at the transcriptional level, are poorly understood. Unusual expression of Tn and its sialylated version, sialyl-Tn (STn), due at least in part to compromised expression of either T-synthase and/or Cosmc, are found in several human diseases, such as Tn syndrome (14), IgA nephropathy (IgAN) (15), and human tumors (16). For example, somatic mutations in Cosmc were identified in the blood cells from patients with Tn syndrome (17, 18). Tn/STn antigens, which are highly expressed in tumor cell lines and human primary cervical cancer tissue, also arise from point mutations in the Cosmc coding region or the deletion of Cosmc (19, 20). Furthermore, the Tn4 cells derived

Mucin type O-glycosylation is one of the most abundant protein post-translational modifications and is initiated and completed by a set of glycosyltransferases in the Golgi apparatus. Among them is the core 1 β3-galactosyltransferase or C1GalT1 (T-synthase), which is the key enzyme that transfers galactose from UDP-galactose to GalNAcα1-Ser/Thr (Tn antigen) in glycoproteins to form the core 1 disaccharide Galβ1–3GalNAcα1-Ser/Thr. Core 1 serves as a precursor for a variety of complex O-glycans (1). These core 1-based O-glycans play wide ranging roles in many pathways, including immunity (2–4), cell-cell interactions (3–7), vascular biology (8), angiogenesis (9), and lymphangiogenesis (10). Interestingly, the endoplasmic reticulum-localized molecular chaperone Cosmc (core 1 β3-galactosyltransferase-specific molecular chaperone or C1GalT1C1) is required for active T-synthase formation in vivo by preventing its aggregation and subsequent proteasomal degradation (11, 12). Disruption of either T-synthase or Cosmc is embryonic lethal due to hemorrhaging, suggesting the important role of O-glycosylation in vascular biology (9, 13).

The structures of O-glycans change during cell differentiation and animal development, indicating that expression of glycosyltransferases is both temporally and spatially regulated. Yet the details on the mechanisms of regulation of these enzymes, especially at the transcriptional level, are poorly understood. Unusual expression of Tn and its sialylated version, sialyl-Tn (STn), due at least in part to compromised expression of either T-synthase and/or Cosmc, are found in several human diseases, such as Tn syndrome (14), IgA nephropathy (IgAN) (15), and human tumors (16). For example, somatic mutations in Cosmc were identified in the blood cells from patients with Tn syndrome (17, 18). Tn/STn antigens, which are highly expressed in tumor cell lines and human primary cervical cancer tissue, also arise from point mutations in the Cosmc coding region or the deletion of Cosmc (19, 20). Furthermore, the Tn4 cells derived

This work was supported, in whole or in part, by National Institutes of Health Grant RO1DK80876 (to T.J.). The authors declare that they have no conflicts of interest with the contents of this article.

1 Both authors contributed equally to this work.
2 Present address: Dept. of Molecular and Cell Biology, Henry M. Goldman School of Dental Medicine, Boston University, MA 02118.
3 Present address: Dept. of Biological Sciences, Carnegie Mellon University, Pittsburgh, PA 15213.
4 To whom correspondence should be addressed: Emory University School of Medicine, 1510 Clifton Rd. NE, Rm. 4029, Atlanta, GA 30322, Tel.: 404-727-6166; Fax: 404-727-2738; E-mail: tju@emory.edu.
5 The abbreviations used are: T-synthase, core 1 β3-galactosyltransferase (UDP-Gal:1-N-acetylgalactosaminyl-α1-Ser/Thr β3-galactosyltransferase); Cosmc, core 1 β3-galactosyltransferase specific molecular chaperone; TF, transcription factor; TSS, transcription start site; Tn antigen, GalNAcα1-Ser/Thr; KLF, Krüppel-like transcription factor; SP, specificity protein; OGT, O-linked N-acetylgalactosaminyltransferase; IP, immunoprecipitation; GT, glycosyltransferase; IgAN, IgA nephropathy; STn, sialyl-Tn; MeDIP, methylated DNA immunoprecipitation.
from an individual whose leukocytes carry Tn antigen harbor a silenced Cosmc but an active T-synthase due to hypermethylation of the putative promoter region of Cosmc (21). Galactose-deficient O-glycans or Tn/STn antigens are commonly found in the hinge region of immunoglobulin A1 (IgA1) molecules from sera of patients with IgAN, and they also occur in serum IgA1 from normal individuals in our study (22). Although the mechanism underlying a fraction of IgA1 from both healthy controls and IgAN patients carrying Tn/STn antigens is not fully understood, several studies have shown that under-galactosylated glycos or Tn/STn antigens are commonly found in the hinge region of immunoglobulin A1 (IgA1) molecules from patients with IgAN (23–26). Hypermethylation of Cosmc in peripheral B cells or immortalized B cells from patients with IgAN is associated with a lower transcription level of Cosmc and T-synthase in peripheral B cells or immortalized B cells from patients with IgAN (23–26). Hypermethylation of Cosmc occurs in many human pancreatic cancer samples and is associated with Tn antigen expression and malignant behavior of the tumor cells (27). Therefore, a major mechanism for the expression of Tn antigen on cells is the loss of functional Cosmc, or T-synthase, including both genetic mutations and epigenetic dysregulation of Cosmc (19, 21). Knock-out of either T-synthase or Cosmc in mice causes embryonic lethality and Tn antigen expression (9, 13), and specific deletion of T-synthase in endothelial and hematopoietic cells in mice causes defective lymphatic vessel formation (10) and loss of high endothelial venule integrity (28); specific disruption of either Cosmc (29) or T-synthase (30) in hematopoietic cells confers megathrombocytopenia and bleeding in mice.

Human T-synthase is located on 7p21 (earlier was designated 7p14–p13) with three exons (31). By contrast, human Cosmc is encoded by a single exon gene located on Xq24 (11). Interestingly, human Cosmc and T-synthase are ubiquitously and coordinately expressed, yet their expression level varies substantially from tissue to tissue or cell type to cell type (12, 31). Here, we report the identification and characterization of the promoters for human Cosmc and T-synthase. Our results show that the CpG islands at 5′-flanking regions of human Cosmc and T-synthase each containing two SP1/3-binding sites are the gene promoters. ChIP assays and site-directed mutagenesis analyses of either SP1/3 site confirmed the essential role of the SP1/3 sequence in the regulation of these two genes. Importantly, the methylome analysis of the Tn4 B cells harboring a silenced Cosmc revealed that the promoter, including the minimum promoter of Cosmc but not T-synthase, was hypermethylated.

**Experimental Procedures**

**Reagents—**All routine chemical reagents and mithramycin A were purchased from Sigma, Fisher, or VWR (Suwanee, GA). Deoxyoligonucleotides were ordered from Integrated DNA Technologies, Inc. (Coralville, IA).

**Cell Culture and Mithramycin A Treatment—**Two human cell lines, FEMX-I (kindly provided by the group of Dr. Oystein Fodstad, Norwegian Radium Hospital Research Foundation, Norway) and HEK293T cells (ATCC 293tsA1609neo), which have high T-synthase activities, were chosen for the promoter characterization and grown in DMEM (Invitrogen) containing 10% heat-inactivated fetal bovine serum (FBS, Invitrogen), penicillin (100 IU/ml), and streptomycin (100 IU/ml) at 37 °C, 5% CO₂. Mithramycin A was dissolved in sterile Milli-Q water to make a 1 mM stock solution. For mithramycin A treatment, FEMX-I cells and 293T cells were seeded with 5 × 10⁴ cells per T75 flask with the indicated concentration of mithramycin A for the desired duration of treatment. After drug exposure, cells were collected by centrifugation.

**Western Blot and Quantification—**Cells were lysed in an appropriate volume of 20 mM Tris-HCl buffer (pH 7.2) containing 300 mM NaCl, 0.5% Triton X-100, and proteinase inhibitor mixture (Roche Applied Science) for 30 min on ice. Lysates were then centrifuged at 15,000 rpm for 15 min at 4 °C. Protein
Promoters of Human Cosmc and T-synthase

cellular concentration in supernatants was determined by bicinchoninic acid (BCA) protein assay (Pierce). For each lysate, 10 μg of total reduced proteins were separated on Mini-PROTEAN TGX gels (Bio-Rad) and transferred electrophoretically onto a nitrocellulose membrane and immunoblotted with the indicated antibodies. The signals of immunoreactive proteins were visualized by enhanced chemiluminescence (Pierce).

Pictures were inverted and subtracted from background value, and band densitometric intensities were then measured by FluorChem and expressed as a percentage relative to β-actin in the same sample, and the integrated density value of T-synthase/β-actin at day 0 was set as 100%. The values were averaged from three different gels from three experiments.

Cloning and Characterization of the 5′-Flanking Region—A list of all oligonucleotide primers used in this study is provided in Table 1. To generate pGluc-CosmcP(−2066/537) (transcription starting site (TSS) as +1), the Cosmc 5′-flanking region was amplified by PCR Phusion high fidelity polymerase (New England Biolabs) using genomic DNA from normal human leukaemia as template and the oligonucleotides C-P-F1 and C-P-R1 covering −2161 to +687 bp as primers. The PCR product was cloned into PCR-TOPO vector (Invitrogen), and then partially digested with EcoRI and subcloned into the multiple cloning sites of the pBluescript SK(+) vector. The insert was finally subcloned to pGluc-Basic vector (New England Biolabs) using the engineered restriction sites for HindIII and BamHI. The nucleotide sequences of the plasmids were confirmed by sequencing.

Other deletion constructs were generated by PCR with pGluc-CosmcP(−2066/537) as template, namely pGluc-CosmcP(−550/90), pGluc-CosmcP(−400/90), pGluc-CosmcP(−260/90), pGluc-CosmcP(−160/90), pGluc-CosmcP(−100/90), and pGluc-CosmcP(−60/90). The primers used for the above constructs were the reverse primer C+90R, combined with one of the forward primers C−550F, C−400F, C−260F, C−160F, C−100F, and C−60F, respectively. pGluc-CosmcP-spl1(−42/−45) and pGluc-CosmcP-spl1(−17−15) were constructed by introducing three point mutations into pGluc-CosmcP(−260/90) with the QuikChange kit (Stratagene), respectively. The primers used for generating the mutations were C−42−45−SP1-F/R and C−17−15−SP1-F/R (Table 1). The PCR product was purified with the QIAquick gel extraction kit (Qiagen, Valencia, CA). The purified PCR product and plasmid pGluc-Basic were digested with Ncol and BamHI, purified with QIAquick gel extraction kit, and ligated according to the manufacturer’s instructions. The ligation mixture was transformed into Escherichia coli strain Top10 competent cells. The sequences of the putative promoter region in the resulting plasmids were confirmed by DNA sequencing. The CpG island for T-synthase (tCpG) was cloned by PCR using primers T-P-F5/T-P-R5 covering −666 to +747 bp and human genomic DNA as the template. Similar methods were used for T-synthase promoter construction, and the primers for constructing plasmids for T-synthase promoter activity analysis are also listed in Table 1.

Transient Transfection and Luciferase Assays—FEMX-I and 293T cells were transfected 1 day after seeding in 6-well plates with up to 1 × 10⁶ cells per well using FuGENE 6 and 9 (Roche Applied Science). For each construct, triplicates were performed, and 48 h after transfection, supernatants were collected, and luciferase activity was measured using the Gaussia luciferase assay kit (New England Biolabs). Briefly, luciferase activity was immediately measured by addition of 50 μl of assay reagents mixed with 20 μl of medium. The photon produced was measured by a Top Count NXT microplate scintillation and luminescence counter. Luciferase activities were calculated using fold difference as the activity of the control constructs pGluc-CCpGI-IIR (pGluc-CosmcP(−2066/537)-R) for Cosmc promoter constructs or pGluc-666/+747-R (pGluc-TsynF5R5-R) for T-synthase promoter constructs. The data were normalized with luciferase gene containing vector copy number. Unpaired Student’s t tests were performed to determine statistical significance with *, p < 0.05; **, p < 0.01, and ***, p < 0.001.

Chromatin Immunoprecipitation (ChIP) and PCR Analysis—Chromatin was isolated from formaldehyde-treated FEMX-I cells and was further treated with sonication. Antibodies used for immunoprecipitation were anti-SP1 and normal rabbit IgG (Millipore Upstate, Temecula, CA). PCR was carried out with the C−60F and C+90R primer sets to amplify SP1-binding sites containing regions within the Cosmc promoter. F−89 and R+63 primer sets were used for T-synthase ChIP analysis. To search for possible known cis-regulatory elements, the programs MacVector and Promoter Scan (PROSCAN, www-bimas.cit.nih.gov) were used.

DNA Methylation Analysis—DNA methylation profiling was performed as described (32). Briefly, genomic DNA of Tn4 and control C12 cells were isolated using a DNeasy Blood and Tissue kit (Qiagen). Methylated DNA immunoprecipitation (MeDIP) using 5mC-specific antibody was performed. Immunoprecipitated DNA along with input was used for library preparation with NEBNext ChIP-Seq Library Prep Reagent Set for Illumina, and 38-cycle single-end sequencing was performed using Version 4 Cluster Generation and Sequencing kits (parts 15005236 and 15002739) and Version 7.0 recipes. We use the standard Illumina Pipeline to process image and extract sequence.

FASTQ sequence files were aligned to the reference genome using Bowtie, keeping only unique nonduplicate genomic matches with no more than two mismatches within the first 25 bp. Unique and nonduplicate reads from nonenriched input genomic DNA and 5mC-enriched sequence set will be counted in 100-, 1000-, and 10,000-bp bins genome-wide and subsequently normalized to the total number of nonduplicate reads in millions. Input-normalized values will then be subtracted from 5mC-enriched values per bin to generate normalized 5mC signals. To determine genomic regions that present differential 5mC profiles, we identified true 5mC-enriched regions or “peaks” using the model-based analysis of ChIP-Seq (MACS) algorithm.

Quantitative RT-PCR—Total RNA of drug-treated 293T and FEMX-I cells was extracted with RNeasy mini kit (Qiagen). RT-PCR was carried out with 1 μg of mRNA as template followed by SuperScript First-strand Synthesis System for RT-PCR (Invitrogen) according to the manufacturer’s protocol. The quantitative real time PCR was performed in triplicate in a 20-μl reaction mixture containing gene-specific primers and
Power SYBR Green PCR Master Mix (Applied Biosystems, Grand Island, NY) using the 7500 Standard Real Time PCR systems (Applied Biosystems). Primers for human Cosmc, T-synthase transcripts, and internal control 18S rRNA were designed using Primer Express 3.0 software (Applied Biosystems) and are listed in Table 2. The relative quantifications were calculated by \( \Delta \Delta C_T \) method and normalized to the control samples.

**Results**

CpG Island-II at 5'-Flanking Region of Human Cosmic Has Gene Promoter Activity—Human Cosmic is a single exon gene originally designated within a 20-kb DNA fragment on Xq24, although it is now annotated as 4.47 kb in the GRCh38 Primary Assembly (www.ncbi.nlm.nih.gov). Analysis of the 5'-flanking region of Cosmic with the on-line program CpG Island Searcher identified two GC-rich regions or CpG islands (CpG-I) upstream of the Cosmic exon, designated as cCpG-I and cCpG-II. cCpG-I localizes between −4980’ and −4237’ bp (designated the translation initiation site as +1’ bp) with a size of 744 bp, and cCpG-II is between −3478’ and −2377’ bp with 1102 bp. Alignment between the 5'-untranslated region (5'-UTR) of the Cosmic cDNA revealed that the TSS falls in cCpG-II, suggesting that cCpG-II is the putative promoter of human Cosmic. However, this promoter region does not have a typical TATA box, consistent with early studies that CpG island promoters may often lack TATA boxes (33).

To experimentally confirm the promoter activity of the cCpG-II for Cosmic, we performed luciferase reporter assays using Gaussia luciferase-containing vector, pGluc (New England Biolabs). The constructs pGluc-cCpG-I-II, containing both CpG islands I and II, and pGluc-cCpG-I and pGluc-cCpG-II, containing cCpG island I and cCpG island II, respectively, as well as the reverse-oriented cCpG-I-II (pGluc-cCpG-I-II-R) and cCpG-II (pGluc-cCpG-II-R) are listed in Fig. 1A. All of the constructs were transfected into a human melanoma cell line, FEMX-I, and at 48 h post-transfection the luciferase activity in the cell culture media was measured. Compared with the activity of luciferase from pGluc-cCpG-I-II-R and the basic vector pGluc-Basic, −48-, 1-, and −562-fold increases of the luciferase activity were detected in the media of cells transfected with pGluc-cCpG-I-II, pGluc-cCpG-I, and pGluc-cCpG-II, respectively (Fig. 1B), indicating that human Cosmic CpG-II contains

**TABLE 2**

| Gene name | Sequence |
|-----------|----------|
| Cosmc     | Forward, 5’–CCTTGTAAACCAAAAGATGTGAGT–3’ |
|           | Reverse, 5’–TTCGACTAGTGGTGGTCAAGTC–3’ |
| T-synthase| Forward, 5’–GGACAGAATCATCTAGAAGGACAAATG–3’ |
|           | Reverse, 5’–CAATGTCTGTGTTCTACTTTATGTG–3’ |
| 18S rRNA  | Forward, 5’–CGGTCTACACCACAAGGA–3’ |
|           | Reverse, 5’–GCCGTATGTTATTTATCGACTACTC–3’ |

**A. Constructs of Reporter Assay for Cosmic Promoter**

**B. Luciferase Activity of Constructs for Cosmic Promoter in FEMX-I Cells**

**C. Luciferase Activity of Constructs for Cosmic Promoter in 293T Cells**

**FIGURE 1.** CpG islands at 5’-flanking regions of human Cosmic have gene promoter activity. A, schematic structure of human Cosmic promoter showing putative CpG islands at 5’-flanking regions predicted by bioinformatic tools. The name of each construct was assigned according to the numbers of CpG islands (shaded box, cCpG-I; open box, cCpG-II) located upstream of the initiation codon ATG. B and C, luciferase activity of FEMX-I (B) and 293T (C) cells transfected with human Cosmic promoter constructs containing CpG islands in pGluc-basic reporter plasmid. The activity of each construct was normalized to the copy number of the promoter region on the chromosome or plasmid with semi-quantitative PCR. Error bars represent mean ± 1 S.D. of three biological replicates. The luciferase activity of each construct is calculated as fold-change relative to the background activity of pGluc-Cosmc-CpG-I-II-R vector.
strong promoter activity. Importantly, pGluc-cCpG-II-R had no detectable activity, which is consistent with the importance of the promoter orientation. Unexpectedly, the activity of pGluc-cCpG-I was close to that of background, indicating that the CpG-I island itself does not contain a Cosmc promoter. Interestingly, the DNA fragment containing both cCpG-I and -II, pGluc-Cosmc-CpG-I-II, had much lower activity than pGluc-cCpG-II, 48-versus 562-fold, suggesting that CpG-I island may contain regulatory elements that could suppress the transcription of Cosmc in certain cell types.

We also tested the promoter activity of the constructs in another human cell line, HEK293T (human embryonic kidney 293T, 293T). We chose 293T and human melanoma FEMX-I for the reporter assay because they are developmentally distant cell types, yet they all have robust activity of T-synthase. The trend of reporter luciferase activity of constructs 2–5 in 293T cells was similar to that in FEMX-I cells confirming that cCpG-II is the promoter for human Cosmc (Fig. 1C). Interestingly, construct 1, having both cCpG-I and cCpG-II (pGluc-cCpG-I-II), had a similar activity of pGluc-cCpG-II in 293T cells (Fig. 1C). Combining this result with the result from FEMX-I cells, the data strongly suggest that there is TF regulating the activity of cCpG-I as a suppressor, and furthermore, this TF either does not exist in 293T cells or the pathway causing this TF to be released is not activated in 293T cells, in comparison with FEMX-I cells.

To fairly compare the promoter activities between the constructs, the relative activity or fold-change for each construct in each experiment was normalized by the apparent activity of luciferase divided by the copy number of the plasmid. The copy number of the plasmids in the transected cells was analyzed by semi-quantitative PCR using the primers listed in Table 1 for amplifying the vector-insert chimeras and genomic DNA isolated from the cells at the time when the media were harvested. The copy number of the plasmids was amplified in a linear range, and the DNA intensity was carefully measured using a densitometer. The copy number of the Cosmc coding region amplified by C-F9/C-R10 was used as an internal control. In all experiments with luciferase reporter assays, the activity was normalized and represented the same way.

**Identification of the Minimal Promoter of Cosmc**—To better understand the transcriptional regulation of Cosmc, the information on the minimal DNA region required for promoter activity was investigated. To identify the minimal promoter for Cosmc, DNA fragments with serial truncations at the 5′-end of the Cosmc promoter within the region −550 to +90 bp (TSS as +1) were generated by PCR and subcloned into a pGluc-Basic vector (Fig. 2A). These constructs were transfected into FEMX-I, and the medium was collected at 48 h post-transfection to measure luciferase activity. As shown in Fig. 2B, all of the constructs had some promoter activity, although variations between different constructs were observed. The smallest DNA fragment −60/+90 (pGluc-cCpG-II-60 +90) had substantial activity, indicating this 150-bp DNA in the cCpG-II is the minimal promoter for human Cosmc. Interestingly, the fragment of −260/+90 (pGluc-cCpG-II-260 +90) had the highest promoter activity, in contrast to the fragments of −160/+90 (pGluc-CpG-II-160 +90) with the lowest activity. These data suggested that the 70 bp within −260/−190 bp of Cosmc might have an enhancer element for FEMX-I cells. In 293T cells (Fig. 2C), a different pattern of luciferase activity from different constructs was observed: construct 1 (−550/+90) had the highest activity, and construct 3 (−260/+90) had relatively low activity in contrast to that in FEMX-I cells, indicating that the TF profiles of FEMX-I and HEK293 cells are very different. Neverthe-
Promoters of Human Cosmc and T-synthase

A. Constructs with Truncations at 5′-End of T-synthase Promoter

B. Luciferase Activity of Constructs in FEMX-I cells

C. Luciferase Activity of Constructs in 293T cells

less, the DNA fragment of −60/+90 in the cCpG-II in both cells gave a similar robust luciferase activity, suggesting the 150-bp DNA fragment (−60/+90) in cCpG-II is the minimal promoter for human Cosmc.

CpG Island at 5′-Flanking Region of Human T-synthase Has Gene Promoter Activity—Early studies showed that human Cosmc and T-synthase are coordinately expressed (12). To understand this phenomenon, we also analyzed the promoter for human T-synthase. Human T-synthase is composed of three exons on 7p14–13 stretching over 115 kb, which is much larger than Cosmc. Using a similar strategy for identification of the gene promoter for Cosmc, one GC-rich DNA fragment (CpG island) with 1.4 kb at the 5′-flanking region was identified, and the 5′-UTR of T-synthase cDNA fell in this region. Similar to the human Cosmc gene promoter, no typical TATA box was found in this putative promoter region for T-synthase. With a similar strategy to the promoter reporter assay for Cosmc, the 1.4-kb DNA fragment containing the CpG island (tCpG) of human T-synthase was cloned into a pGluc-Basic vector, designated as pGluc-tCpG (TSS as +1) (Fig. 3A), and its promoter activity was measured at 48 h post-transfection into both FEMX-I cells and 293T cells. As shown in Fig. 3, B and C, the luciferase activity of pGluc-tCpG (construct 1 −666/+747) was low but consistently greater than that of the reverse-oriented tCpG vector, pGluc-tCpG-R (construct 2 −666/+747-R), indicating that the tCpG has the promoter activity for human T-synthase.

To identify the minimal DNA fragment with its T-synthase promoter activity, serial constructs with truncations at the 5′-end of tCpG within −334 to +82 bp (Fig. 3A, constructs 3–5) were also analyzed for their promoter activity in both FEMX-I and 293T cells. In FEMX-I cells, the promoter activity increased along with the 5′-truncations of the tCpG (Fig. 3B), with construct 4 representing −169/+82 bp and construct 5 representing −89/+82 bp, having similar activity to each other, suggesting that 171 bp from −89/+82 bp was the minimal promoter for human T-synthase. These data also suggested that the 497-bp fragment from −166 to −169 bp contains a negative regulatory element, because the construct containing this region had little activity. In 293T cells (Fig. 3C), constructs 4 and 5 had the highest activity as seen in FEMX-I cells, although construct 3 showed much higher activity than that from FEMX-I cells. The differences in the activity observed in different cell lines might also arise from the differences in TF profiles between these two different types of cells. Nevertheless, these results from two different cell lines indicated that the minimal promoter for human T-synthase is the −89/+82-bp DNA fragment.

Two Binding Sites for SP1/3 in the 150-bp Minimal Promoter of Cosmc and 171-bp Minimal Promoter of T-synthase Are Essential for Basal Promoter Activity—Knowing the minimal DNA sequences that were required for the promoter activity of Cosmc and T-synthase, we investigated the potential TFs for transactivating the expression of these two genes. We used two programs, Promoter Scan program (www-bimas.cit.nih.gov) and MacVector Nucleotide Sequence Analysis-Mammalian TF, to search the possible TFs that bind to the Cosmc promoter cCpG-II and T-synthase promoter tCpG. No TATA or CCAAT box was found in either gene (Figs. 4 and 5); however, many putative TF-binding sequences were predicted in these promoters. Significantly, close to the TSS of mRNA for Cosmc and T-synthase, two putative binding sites for SP1/3, 5′-ACCCCGCCCC-3′ and 5′-GCTCCGCCCA-3′, were located on the minus strain in the cCpG-II and two SP1/3-binding sequences, 5′-ATAGGGCGGCA-3′ and 5′-GGGGCGGAGC-3′, were also found in tCpG. Because their binding sites are located upstream proximal to the TSS, SP1/3 are general TFs and Cosmc and T-synthase are ubiquitously expressed, we considered these
SP1/3-binding sequences in Cosmc and T-synthase are good candidates for regulating the minimal promoter activity.

To examine whether the promoter activity is controlled by these KLF/SP TF SP1/3-binding sequences within the Cosmc minimal promoter region, each of the SP1/3-binding sequences in the constructs was mutated individually, designated as pGluc-Cosmc-M1 and -M2 or doubly designated as pGluc-Cosmc-M1/2 (starred constructs in Fig. 6A). These constructs were transfected into FEMX-I cells, and the media were collected at 48 h post-transfection and measured for luciferase activities. Either mutation of the two SP1/3 sequences or double mutant abolished the promoter activity (Fig. 6B), demonstrating that the basal promoter activity for Cosmc is controlled by these two SP1/3-binding sites. Interestingly, in 293T cells, mutations in the SP1/3-binding sequences in the hCosmc minimal promoter only caused a modest decrease in promoter activity (Fig. 6C). These results are not totally surprising, because the profile of TFs in melanoma FEMX-I cells is likely very different from that in embryonic kidney 293T cells; this also indicates that the mutated sequences could be transactivated by some TFs existing in 293T cells that are expressed at low levels in FEMX-I cells.

Similar to the Cosmc promoter, the T-synthase minimal promoter region (−89/+82 bp) also contains two SP1/3 sites. Mutations were introduced into either SP1 sites, pGluc-Tsyn-M1 and pGluc-Tsyn-M2, or both sites, pGluc-Tsyn-M1/2 (starred constructs in Fig. 6D), which were then transfected into FEMX-I cells and 293T cells. As shown in Fig. 6E in FEMX-I cells, the luciferase activity in either one or both SP1/3 mutations was abrogated, demonstrating that these two SP1/3-binding sequences in Cosmc and T-synthase are good candidates for regulating the minimal promoter activity.

FIGURE 4. Promoter sequences of human Cosmc. The full sequence and potential transcription factor-binding sites predicted by MacVector 12 of promoter (cCpG-II) for human Cosmc is shown. Putative transcription factor consensus binding sites are indicated by bold, italic, underlines, or strikethroughs. Transcription start sites are indicated. The TSS for human Cosmc predicted by DBTSS (chromosome X:120625226–120631269) is −64 bp upstream. The binding sites and sequences by two critical TFs, SP1/3, upstream and close to the TSS are in bold and highlighted in yellow.
binding sites are critical for T-synthase basal promoter activity. Similarly to the Cosmc promoter mutants (Fig. 6C), the mutations in the SP1/3 sequences of the hT-synthase minimal promoter did not cause much loss of luciferase activity either (Fig. 6F), perhaps due to the difference in profile of TFs in melanoma FEMX-I cells from that in embryonic kidney 293T cells.

It should be noted that both Cosmc and T-synthase are endogenously expressed. The results of Fig. 6, B and E, demonstrated that the GC-rich sequences in the minimal promoters for Cosmc and T-synthase were essential for the promoter activity.

To further confirm the importance of the TSS sites in the minimal promoter for Cosmc, the constructs with deletions at the 3'-end of the pGluC-160 + 90 bp (Fig. 7A) were also generated, and their promoter activities were analyzed. As shown in Fig. 7B, all of the constructs pGluC-160 to −100, −160 to −60, and −260 to −150 bp lacked promoter activity, further demonstrating that the sequences around the TSS are essential to maintain Cosmc promoter activity.

**SP1 and SP3 Are Associated With the Core Promoters of Cosmc and T-synthase**—To confirm that SP1/3 sequence in the Cosmc core promoter was actually associated with KLF TF SP1/3 in vivo, a chromatin immunoprecipitation (ChIP) assay was performed. After the sheared chromatin from FEMX-I and 293T cells was immunoprecipitated (IP) with anti-SP1 and -SP3 antibody, the PCR of the promoter from 60 to 90 bp covering these two essential SP1/3 sequences in Cosmc was carried out. A product...
with the expected size of 150 bp was generated from the PCR after IP with anti-SP1 and anti-SP3 antibodies, while no detectable product was seen in the IP with the isotype control antibody and with blank material (Fig. 8A). These results demonstrate that SP1 and SP3 are associated with the Cosmc minimal promoter.

Similarly, ChIP assay for the human T-synthase promoter was also performed using primers covering /H11002 to /H11001 bp, including the two critical SP1 and SP3 sites in the minimal promoter. An expected product with a size of 171 bp corresponding to the minimal promoter for T-synthase was amplified (Fig. 8B) in the specific anti-SP1- and anti-SP3 antibody pulldown materials but not in the isotype control antibody and blank material experiment, confirming that SP1/3 sequences of human T-synthase minimal promoter were bound by SP1 and SP3 TFs.

Through these combined approaches, the promoters of human Cosmc and T-synthase are identified and characterized, and the full sequences are shown in Fig. 4 for Cosmc and Fig. 5 for T-synthase. It is noteworthy that both Cosmc and T-synthase are coordinately and ubiquitously expressed, and the identification of the important KLFs, SP1/3, clearly explains this phenomenon. Furthermore, these two genes are also differentially expressed in different tissues or cell types (12, 31), suggesting that the tissue-specific TFs are also involved in their transcriptional regulation by playing the enhancer or suppressor role. Indeed, there are many consensus sequences being recognized by putative TFs, such as AP2, /H9253-IRE, and other KLFs for both Cosmc and T-synthase. Taken together, our results suggest that the basal levels of human Cosmc and T-synthase are similarly transcriptionally regulated by KLF SP1/3 TFs and also enhanced or suppressed by some cell- or tissue-specific TFs.

Treatment with Mithramycin A Caused Reduction of Cosmc and T-synthase mRNAs and T-synthase Protein and Activity—

The above results indicated that KLFs, including SP1/3, are the important TFs regulating Cosmc and T-synthase transcription.
**Promoters of Human Cosmc and T-synthase**

**A. Constructs of Reporter Assay for Cosmc Promoter**

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| -60/90 | -160/-10 | -100/-10 | -160/-60 | -260/-150 |

**B. Luciferase Activity of Constructs in FEMX-I Cells**

![Graph showing relative luciferase activity for constructs 1 to 5](image)

**FIGURE 7. Characterization of Cosmc 3‘-truncated promoter constructs.**

---

Some studies (34, 35) reported that mithramycin A generally binds to GC-rich DNA sequences and consequently prevents the binding of TFs leading to compromised transcription. Thus, we explored whether mithramycin A could affect the binding of KLFs to the DNA sequence leading to alteration of Cosmc and T-synthase transcription.

To determine the concentration of mithramycin A for treatment, we cultured 293T and FEMX-I cells in the presence of a range of concentrations of mithramycin A for 8 days. Mithramycin A did not significantly alter the viability of 293T cells at concentrations \( \leq 100 \text{ nM} \) and FEMX-I cells at concentrations \( \leq 10 \text{ nM} \). Therefore, 293T and FEMX-I cells were treated with two different concentrations (50 and 100 nM for 293T cells and 5 and 10 nM for FEMX-I cells) of mithramycin A for different periods of time (0–8 days). To examine whether mithramycin A affected the transcription of Cosmc and T-synthase, quantitative RT-PCR to measure their mRNA levels was performed. As shown in Fig. 9A, the transcriptional levels of Cosmc and T-synthase were decreased in a dose-dependent manner in both cell lines tested. As expected, Western blots demonstrated that the protein levels of Cosmc and T-synthase were reduced in the mithramycin A-treated cells in a similar manner (Fig. 9B). To examine whether the mithramycin A treatment effect is relatively specific on the genes, including Cosmc and T-synthase who have the GC-rich promoters, a control glycosyltransferase, β-O-acetylglucosaminyltransferase (OGT), which catalyzes the addition of O-GlcNAc to cytoplasmic and nuclear proteins, was also analyzed by Western blotting. OGT was not significantly affected in cells by treatment with mithramycin A (Fig. 9B). Consistent with the decline of the protein levels, the activity of T-synthase in 293T and FEMX-I cells also decreased in a dose- and time-dependent manner upon treatment of mithramycin A as follows: moderate decrease at a low concentration (50 and 5 nM, respectively), and a significant decrease 80 and 60% at a high concentration (100 and 10 nM) for 293T cells and FEMX-I cells, respectively, after 8 days of treatment (Fig. 9C). Taken together, these results confirm the important role of the GC-rich sequence of CpG islands in the regulation of transcription and expression of Cosmc and T-synthase.

Methylome Analysis Revealed Hyper-methylation of Cosmc Promoter in Tn4 B Cells—DNA hypermethylation is generally found in CpG islands in gene promoter regions, and the degree of methylation reversibly correlates with the gene transcription level. Normal DNA methylation is essential for cell development and differentiation through regulation of gene transcription, chromatin structure, and accessibility (36). Dysregulation of DNA methylation is found to be associated with human diseases, such as carcinogenesis (37). Our previous study (21) using bisulfite sequencing showed that Tn4 B cells originally derived from a male individual with Tn-positive leukocytes have a silenced Cosmc due to core promoter hypermethylation but have a normally expressed T-synthase. To fully assess the DNA methylation in the Tn4 cells, we profiled the genome-wide DNA methylation by methylated DNA immunoprecipitation sequencing (MeDIP-seq). As shown in Fig. 10, A and B, in the control B cells (C12) immortalized from a healthy male individual both cCpG-II and tCpG were not methylated; the cCpG-II but not tCpG in Tn4 cells was hypermethylated. Examining the DNA sequences more closely, the hypermethylated region revealed by Methylome analysis falls in the central cCpG-II but not tCpG in Tn4 cells was hypermethylated.

**Discussion**

In this report, we have identified the promoters for both human Cosmc and T-synthase, two functionally related genes whose expression is essential for animal development and important for many biological processes through controlling the normal mucin type O-glycan biosynthesis. In addition, we have shown that these two genes are similarly regulated by GC-rich binding TFs, including the general TFs SP1 and SP3. Structurally, human Cosmc and T-synthase promoters have similarities as follows: CpG island with two GC-rich sequences proximal to the TSS, yet lacking a typical TATA box. Importantly, promoter activities of both Cosmc and T-synthase were reduced when SP1/3-binding sites were mutated, demonstrating that both genes are regulated by SP1/3 TFs. Furthermore, treatment with mithramycin A, a competitor for binding GC-rich DNA, results in a reduction in the endogenous transcription level of both Cosmc and T-synthase and a decrease in their activity levels, further confirming the essential role of SP1/3, and possibly other SP/KLF family members, in the regulation of the transcription and translation of these genes. Our data explain the nature of the coordinate expression of human
Promoters of Human Cosmc and T-synthase

Cosmc and T-synthase, as they have similar cis-elements controlling their transcription. A depiction of controlling elements within both Cosmc and T-synthase are shown in Fig. 11.

Many changes in the glycome are seen in the course of animal development and cell differentiation and transformation, indicating that glycosyltransferases (GTs) are temporally and spatially regulated. These changes often correlate with transcriptional regulation of GTs. Several examples of this type of regulation are known (38–44), although the underlying mechanistic explanations are poorly understood. We have been particularly interested in the regulation of the O-glycosylation pathway. Aberrant O-glycans Tn and STn antigens are commonly found in humans (16), such as breast (45), colon (46), and cervical (47) tumors. Yet, the status of transcription for Cosmc and T-synthase in these tumors has not yet been investigated. Significantly, epigenetic dysregulation, e.g. hypermethylation of Cosmc, was found in Tn4 B cells originally derived from a male individual with Tn-positive leukocytes in his peripheral blood (21). Epigenetic silencing of Cosmc is the third major molecular mechanism, in addition to the point mutations in the coding region, and deletion of the entire Cosmc gene (17, 19, 20, 48) to confer the expression of Tn and STn in pathological status, including tumors. Because of the hypermethylation of the core, minimal promoter of Cosmc in Tn4 cells, no transcripts of Cosmc were detected unless the cells were treated with 5'-aza-2-deoxycytidine, a DNA methylation inhibitor (21). Our methylome analysis confirmed the hypermethylation status of the Cosmc gene promoter in Tn4 cells but not the control B cells. This is of great interest in defining another molecular mechanism for silencing of Cosmc in tumor cells expressing these aberrant O-glycans. Indeed, Radhakrishnan et al. (27) reported that hypermethylation of Cosmc was found in human pancreatic cancer and associated with the expression of Tn antigen in the cancer. Interestingly, the overall structure of the T-synthase promoter is similar to Cosmc, yet T-synthase is not hypermethylated. The mechanism(s) for initiating and propagating methylation of Cosmc remains elusive. It is worth noting that Cosmc is on the X chromosome, and one X chromosome in female cells will undergo imprinting through hypermethylation, whereas T-synthase is on 7p21.3 and has never been shown to be hypermethylated in normal cells. The hypermethylation of Cosmc in Tn4 cells from a male individual is most likely through a different mechanism(s) unrelated to X chromosome inactivation, e.g. beyond imprinting, because other X-linked glyco-genes, including OGT and PIGA, are not silenced in these cells (21). It is possible that unknown DNA-binding proteins or transcription factors expressed in Tn4 B cells recognize and bind to the core promoter region of Cosmc, but not the T-synthase, and recruit the de novo methyltransferases DNMT-3a or -3b, which initiates the methylation. Exploring the mechanism(s) underlying the aberrant hypermethylation of Cosmc may aid in understanding the hypermethylation of Cosmc in human tumors (27).

Generally, GTs are transcriptionally regulated by two categories of TFs, depending on their roles in the glycan biosynthetic pathway. GTs involved in the core structure are ubiquitously transcribed under control of general TFs, although others involving the terminal and branching structure of glycans are strictly transcribed in a tissue-specific or differentiation stage-specific manner under control of specific TFs. For instance, human xylosyltransferase I, the first enzyme for synthesizing glycosaminoglycans (GAGs), including heparan, keratan sulfate, and chondroitin sulfate, which are important components of extracellular matrix, requires the activator protein 1 (AP-1) and specificity protein 1 (SP1) for transcriptional regulation (49). Moreover, although the ST6Gal-I gene is expressed in almost all human tissues regulated by AP-1, AP-2, and tissue-specific TFs (50), ST6Gal-2 shows a restricted tissue-specific pattern of expression, mostly expressed in embryonic and adult brain and controlled by neuron-related develop-
ment factors Sox5, Puralpha, and Olf1 (51). In the mucin type O-glycan biosynthetic pathway, the transcriptional regulation of a few GTs has been reported. The UDP-N-acetyl-α-/n-galactosamine:polypeptide N-acetylgalactosaminyltransferase-3 (ppGalNAc-T3) gene is restricted to epithelial glands and is regulated by the interaction of p300/CBP-associated factor with nuclear respiratory factor-1 (52). Core 2, 6-N-acetylgalactosaminyltransferase-I, which catalyzes the synthesis of core 2 structures in O-glycans, is a gene with a promoter that is putatively recognized by multiple tissue-specific TFs (53) and is regulated by SP1 in lymphocytes and epithelial cells (54).

Most normal O-glycans are synthesized from the common precursor core 1 disaccharide, the product of T-synthase. Thus, it is not surprising that T-synthase is regulated by KLF, a family of general TFs, including SP1 and SP3. Naturally, its molecular chaperone Cosmc is also under control in a similar fashion. Our findings explain how both Cosmc and T-synthase are coordinately and ubiquitously expressed. Krüppel-like factors, including SP1/SP3, are a family of ubiquitously expressed C2H2-type transcription factors.
zinc finger-containing DNA-binding proteins (55, 56). This family of TFs regulates expression of numerous genes in the metabolic pathway of nucleic acids (e.g. thymidylate synthase, adenine deaminase, and DNA polymerase), cell cycle, and proliferation, such as cyclin D1, E2F, c-Fos, and TGF-β (57–60).

Several GTs involved in O-glycan biosynthesis are also under control of SP1. In P19 cells, the ST6GalNAc-III gene is transcriptionally regulated by two of the three SP1-binding sites in its promoter, whereas in NIH3T3 cells, the ST6GalNAc-IV gene is transcriptionally under control of all three SP1-binding sites (61). Sato and Furukawa (62) found that the activation of the human 1,4-galactosyltransferase-V gene is through the sequential action of Ets-1 and SP1, and altered expression was involved in abnormal glycosylation characteristic of cancer cells.

We found that the cis-elements of both Cosmc and T-synthase lack the TATA box but have the GC-rich region (CpG island) containing two independent yet nearby SP1/3-binding sequences, which proved to be essential for their transcription. Generally the promoters of SP1-regulated genes lack a bona fide TATA-box that binds to the TATA-binding protein and nucleates the assembly of TATA-binding protein-associated factors (63). In a TATA-less element, SP1 facilitates a sequential recruitment of TFIIID and RNA polymerase II; SP1 binds to TFIIA via TAFII130 and TAFII55 subunits (64–66). Interestingly, the promoter activity of Cosmc/T-synthase requires both SP1 sites for their basal transcription, as seen in many GTs that are controlled by multiple TFs. Indeed, compelling evidence indicates that SP1-dependent activation and repression of target genes are modulated by its interactions with a repertoire of heterogeneous transcriptional factors, such as c-Myc, GATAs, p53, MEF2C, Smad2–4, Rb, NF-xB, NF-YA, VHL, MyoD, E2F, NFAT-1, YY1, and others (67). This might be the mechanism for the regulation of Cosmc and T-synthase in a tissue-specific manner, through the interaction of SP1 with the tissue or cell-specific transcription factor(s). These multiple TFs have to be
recruited by their corresponding DNA sequences in concert or sequential to the cis-regulatory element to form functional transcriptional complexes.

Mithramycin A is an antineoplastic antibiotic produced by Streptomyces plicatus because it has an inhibition activity on DNA/RNA polymerase. Many studies showed that it can interact with the GC-rich DNA, competing with TF SP1 binding, thereby inhibiting gene transcription (35). The promoters of human Cosmc and T-synthase are composed of GC-rich islands. Because mithramycin A inhibition of SP1 affects transcription of many genes, including some GTs (49), we tested whether mithramycin A could down-regulate their transcription. Both Cosmc and T-synthase transcripts, as well as T-synthase activity, declined upon long term treatment of mithramycin A. However, the protein level of another glycosyltransferase (OGT), which was determined as a control, was not significantly changed, suggesting that the overall translation machinery is not affected by mithramycin A. These data indicate that the decrease of functional transcripts of Cosmc and T-synthase was caused by the mithramycin A treatment.

The dysregulated expression of Cosmc and/or T-synthase may be important in IgAN. This is the most common primary glomerulonephritis (68–70) worldwide and is characterized by deposition of IgA1 in the mesangium. Studies suggest that a deficiency of galactose in the O-glycans and concomitant expression of the Tn antigen or STn in the hinge region of IgA1 might be responsible for initiating the pathogenesis (15, 23, 69). Several studies showed that the transcript levels of Cosmc and/or T-synthase are reduced in the B cells of patients with IgAN (24–26, 76–78). Suzuki et al. (75) reported that IgA1-secreting cell lines from patients with IgAN produce aberrantly glycosylated IgA1 due to lower transcripts of both Cosmc and T-synthase and up-regulated ST6GaNAc-II. Yet the mechanisms by which Cosmc and T-synthase are suppressed is unknown. Our results show that Cosmc and T-synthase are transcriptionally regulated by the general TF SP1 for their basal activity, and our data also suggest that other TFs may play enhancer or suppressor roles in overall control of gene transcription. It is also worth noting that the cCpG-I upstream of the promoter cCpG-II in Cosmc seems to play a suppressor role in the transcription of Cosmc in some types of cells. There are many potential transcription factor-binding sequences in cCpG-I, including NF-κB. It will be of great interest to define the role of these in plasma cells, which may be crucial to reveal the role of Cosmc and/or T-synthase in IgAN.

In summary, we have characterized the promoters for both human Cosmc and T-synthase, two functionally related genes that are similarly transcriptionally regulated by a general TF SP1/3 through two proximal SP1/3-binding sequences. Our data help to explain the observation that human Cosmc and T-synthase are ubiquitously expressed together, because their promoters are similar to each other. We also observed that the expression levels of human Cosmc and T-synthase in different tissues varies substantially (12, 31), indicating that tissue-specific TFs may also play important roles in the regulation of these two genes. Indeed, as shown in our results, the upstream DNA sequences of the minimal promoters harbor many putative TF recognition sites (Figs. 4 and 5), which could also affect their activities, indicating that additional TFs other than SP1/3 could bind to the sequences in this region to enhance or repress the expression of Cosmc and T-synthase in a tissue- or cell-specific manner. To fully understand the transcriptional regulation of these two genes, the tissue-specific TFs, for example the human B cell-specific or plasma cell-specific TFs and the pathways associated with these TFs, also need to be further characterized. A better understanding of the regulation of Cosmc and T-synthase will certainly aid in elucidating the mechanisms for the expression of abnormal O-glycans, Tn and STn, in IgA1 from patients with IgAN as well as in transformed cells from patients with tumors, such as colorectal carcinoma and breast cancer.
vessel biology and humoral immunity. Mol. Cell. Biol. 27, 8783–8796.
9. Xia, L., Ju, T., Westmucket, A., An, G., Ivanciu, L., McDaniel, J. M., Lupu, F., Cummings, R. D., and McEver, R. P. (2004) Defective angiogenesis and fetal embryonic hemorrhage in mice lacking core 1-derived O-glycans. J. Cell Biol. 164, 451–459.
10. Fu, J., Gerhardt, H., McDaniel, J. M., Xia, B., Liu, X., Ivanciu, L., Ny, A., Herrmans, K., Silansi-Mansat, R., McGee, S., Nye, E., Ju, T., Ramirez, M. I., Carmeliet, P., Cummings, R. D., et al. (2008) Endothelial cell O-glycan deficiency causes blood/lymphatic misconnections and consequent fatty liver disease in mice. J. Clin. Invest. 118, 3725–3737.
11. Ju, T., and Cummings, R. D. (2002) A unique molecular chaperone Cosmc required for activity of the mammalian core 1 beta 3-galactosyltransferase. Proc. Natl. Acad. Sci. U.S.A. 99, 16613–16618.
12. Ju, T., Aryan, R. P., Stowell, C. J., and Cummings, R. D. (2008) Regulation of protein O-glycosylation by the endoplasmic reticulum-localized molecular chaperone Cosmc. J. Cell Biol. 182, 531–542.
13. Wang, Y., Ju, T., Ding, X., Xia, B., Wang, X., Xia, L., He, M., and Cummings, R. D. (2010) Cosmos is an essential chaperone for correct protein O-glycosylation. Proc. Natl. Acad. Sci. U.S.A. 107, 9228–9233.
14. Berger, E. G. (1999) Tn-syndrome. Biochim. Biophys. Acta 1455, 255–268.
15. Hiki, Y. (2009) O-Linked oligosaccharides of the IgA1 hinge region: roles of its aberrant structure in the aberration and/or progression of IgA nephropathy. Clin. Exp. Nephrol. 13, 415–423.
16. Springer, G. F. (1984) Tn and Tn, general carcinoma autoantigens. Science 224, 1198–1206.
17. Crew, V. K., Singleton, B. K., Green, C., Parsons, S. F., Daniels, G., and Anstee, D. J. (2008) New mutations in C1GALT1 C1 in individuals with Tn positive phenotype. Br. J. Haematol. 142, 657–667.
18. Ju, T., and Cummings, R. D. (2005) Protein glycosylation: chaperone mutation in Tn syndrome. Nature 437, 1252.
19. Ju, T., Lanneau, G. S., Stowell, S. R., Willard, A. H., Li, X., Xia, B., Stowell, S. R., Willard, M. T., Wang, W., Xia, J. Y., Zhang, W., Chen, M., Li, C., Xiong, X., Xia, L., He, M., and Cummings, R. D. (2008) Human tumor antigen Tn is a neopiety antigen at Tn is a neopiety antigen at Tn and sialyl Tn arise from mutations in Cosmc. Protein. Proc. Natl. Acad. Sci. U.S.A. 105, 16613–16618.
20. Schietinger, A., Philp, M., Yoshida, B. A., Azadi, P., Liu, H., Meredith, S. C., and Schreiber, H. (2006) A mutant chaperone converts a wild-type protein into a tumor-specific antigen. Science 314, 304–308.
21. Mi, R. P., Wang, J., Liu, X., Yuen, K. V., Stowell, C. J., Chapman, A. B., Cummings, R. D., and Ju, T. (2012) Epigenetic silencing of the chaperone Cosmos in human leukocytes expressing Tn antigen. J. Biol. Chem. 287, 41523–41533.
22. Lehoux, S., Mi, R. P., Wang, J., Schjoldager, T. K., Clausen, H., Van Die, I., Han, Y., Chapman, A. B., Cummings, R. D., and Ju, T. (2014) Identification of distinct glycoforms of IgA1 in plasma from patients with IgA nephropathy and healthy individuals. Mol. Cell. Proteomics 13, 3097–3113.
23. Mestecky, J., Tomana, M., Moldoveanu, Z., Julian, B. A., Suzuki, H., Maeshima, Y., Fukushima, K., Nishizaki, K., Akagi, H., Narimatsu, Y., Nari matsumi, N., Takahashi, S. (2013) C1Galt1-deficient mice exhibit thrombocytopenia due to abnormal terminal differentiation of megakaryocytes. Blood 122, 1649–1657.
24. Ju, T., Brewer, K., D'Souza, A., Cummings, R. D., and Canfield, W. M. (2002) Cloning and expression of human core 1 #beta;3-galactosyltransferase. J. Biol. Chem. 277, 178–186.
25. Yao, B., Lin, L., Street, R. C., Zalewski, Z. A., Galloway, J. I., Wu, H., Nelson, D. L., and Jin, P. (2014) Genome-wide alteration of 5-hydroxy-methylcytosine in a mouse model of fragile X-associated tremor/ataxia syndrome. Hum. Mol. Genet. 23, 1095–1107.
26. Deaton, A. M., and Bird, A. (2011) CpG islands and the regulation of transcription. Genes Dev. 25, 1010–1022.
27. Lombò, F., Menéndez, N., Salas, J. A., and Méndez, C. (2006) The aureolic acid family of antitumor compounds: structure, mode of action, biosynthesis, and novel derivatives. Appl. Microbiol. Biotechnol. 73, 1–14.
28. Van Dyke, M. W., and Dervan, P. B. (1983) Chromomycin, mithramycin, and oligomycin binding sites on heterogeneous deoxyribonucleic acid. Footprinting with (methidiumpropyl-EDTA)iron(II). Biochemistry 22, 2373–2377.
29. Herrmann, A., Gowher, H., and Jeltsch, A. (2004) Biochemistry and biology of mammalian DNA methyltransferases. Cell. Mol. Life Sci. 61, 2571–2587.
30. Schmitt, F. C., Figueiredo, P., and Lacerda, M. (1995) Simple mucin-type O-glycan and its implications in the biology of breast cancer. Proc. Natl. Acad. Sci. U.S.A. 92, 9228–9233.
31. Liang, Y. J., Kuo, H. H., Lin, C. H., Chen, Y. Y., Yang, B. C., Cheng, Y. Y., Yu, A. L., Khoo, K. H., and Yu, J. (2010) Switching of the core structures of glycosiphoglycans from globo- to lacto-co-ganglio-series upon human embryonic stem cell differentiation. Proc. Natl. Acad. Sci. U.S.A. 107, 22564–22569.
32. Jung, E. M., Kwon, O., Kwon, K. S., Cho, Y. S., Rhee, S. K., Min, J. K., and Oh, D. B. (2011) Evidences for correlation between the reduced VCAM-1 expression and hyaluronan synthesis during cellular senescence of human mesenchymal stem cells. Biochem. Biophys. Res. Commun. 404, 463–469.
33. Chen, L., Zhang, W., Fregien, N., and Pierce, M. (1998) The her-2/neu oncogene stimulates the transcription of N-acetylgalactosaminyltransferase V and expression of its cell surface oligosaccharide products. Oncogene 17, 2087–2093.
34. Igaki, H., Ishihara, S., and Matsumoto, K. (2006) NFκB-p65 dependent transcriptional regulation of glycosyltransferases in human colon adenocarcinoma HT-29 by stimulation with tumor necrosis factor alpha. Biol. Pharm. Bull. 29, 2372–2377.
35. Kwon, H. Y., Kim, S. J., Kim, C. H., Son, S. W., Kim, K. S., Lee, J. H., Do, S. I., and Lee, Y. C. (2010) Tryptophol degrades and regulates human G3D3 synthase (hST8Sia I) gene expression in SK-MEL-2 human melanoma cells. Exp. Mol. Med. 42, 849–855.
36. Sengupta, K. P., Bouchie, M. P., and Kukuruzinska, M. A. (2010) N-Glycosylation gene DPAGT1 is a target of the Wnt/#beta;-catenin signaling pathway. J. Biol. Chem. 285, 31164–31173.
37. Xu, Q., Akama, R., Isaji, T., Lu, Y., Hashimoto, K., Kariya, Y., Fukuda, T., Du, Y., and Gu, J. (2011) Wnt/#beta;-catenin signaling down-regulates N-acetylgalactosaminyltransferase III expression: the implications of two mutually exclusive pathways for regulation. J. Biol. Chem. 286, 4310–4318.
38. Schmitt, F. C., Figueiredo, P., and Lacerda, M. (1995) Simple mucin-type carbohydrate antigens (T, sialosyl-Tn and sialosyl-Tn) in breast carcinoma. Virchows Arch. 427, 251–258.
39. Itozakowits, S. (1992) Carbohydrate changes in colon carcinoma. APMS Supplement 27, 173–180.
47. Terasawa, K., Furumoto, H., Kamada, M., and Aono, T. (1996) Expression of Tn and sialyl-Tn antigens in the neoplastic transformation of uterine cervical epithelial cells. *Cancer Res.* **56**, 2229–2232

48. Ju, T., Otto, V. L., and Cummings, R. D. (2011) The Tn antigen-structural simplicity and biological complexity. *Angewandte Chemie* **50**, 1770–1791

49. Muller, B., Prante, C., Klesieck, K., and Gotting, C. (2009) Identification and characterization of the human xylosyltransferase I gene promoter region. *J. Biol. Chem.* **284**, 30775–30782

50. Svensson, E. C., Soreghan, B., and Paulson, J. C. (1990) Organization of the β-galactoside a2,6-sialyltransferase gene. Evidence for the transcriptional regulation of terminal glycosylation. *J. Biol. Chem.* **265**, 20863–20868

51. Lehoux, S., Groux-Degroote, S., Cazet, A., Dhaenens, C. M., Maurage, C. A., Caillet-Boudin, M. L., Delannoy, P., and Krzewinski-Recchi, M. A. (2010) Transcriptional regulation of the human ST6GAL2 gene in cerebral cortex and neuronal cells. *Glycoconjug. J.* **27**, 99–114

52. Falkenberg, V. R., Alvarez, K., Roman, C., and Fregien, N. (2007) Control of core 2β1,6-N-acetylgalactosaminyltransferase-I transcription by Sp1 in lymphocytes and epithelial cells. *Glycoconjug. J.* **24**, 511–519

53. Kaczynski, J., Cook, T., and Urrutia, R. (2003) Sp1- and Kruppel-like factor family of transcription factors in cell growth regulation and transcription initiation and alternative splicing in the 5′-untranslated region of the core 2β1,6-N-acetylgalactosaminyltransferase-I gene. *Glyobiology* **13**, 411–418

54. Black, A. R., Black, J. D., and Azizkhan-Clifford, J. (2001) Sp1 and kruppel-like factor family of transcription factors in cell growth regulation and cancer. *J. Cell. Physiol.* **188**, 143–160

55. Levy, M., and Berger, J. (1988) Worldwide perspective of IgA nephropathy. *J. Am. Soc. Nephrol.* **18**, 1315–1316

56. Gill, G., Pascal, E., Tseng, Z. H., and Tjian, R. (1994) A glutamine-rich hydrophobic patch in transcription factor Sp1 contacts the dTAFI110 component of the Drosophila TFIID complex and mediates transcriptional activation. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 192–196

57. Yu, H. H., Chu, K. H., Yang, Y. H., Lee, J. H., Wang, L. C., Lin, Y. T., and Chiang, B. L. (2011) Genetics and immunopathogenesis of IgA nephropathy. *Clin. Rev. Allergy Immunol.* **41**, 198–213

58. Levy, M., and Berger, J. (1988) Worldwide perspective of IgA nephropathy. *Am. J. Kidney Dis.* **12**, 340–347

59. Kiernan, S., Feriozzi, S., Perez, M., Faraggiana, T., and Muda, A. O. (2007) Aberrantly glycosylated IgA1 in glomerular immune deposits of IgA nephropathy. *J. Am. Soc. Nephrol.* **18**, 3139–3146

60. Novak, J., Julian, B. A., Tomana, M., and Mestecky, J. (2008) IgA glycosylation is inherited in familial and sporadic IgA nephropathy. *Clin. Invest. Med.* **31**, E5–E13

61. Qin, W., Zhong, X., Fan, J. M., Zhang, Y. J., Liu, X. R., and Ma, X. Y. (2008) Studies of IgA nephropathy: past, present, and future. *Pediatr. Nephrol.* **25**, 2257–2268

62. Novak, J., Julian, B. A., Edberg, J. C., Alarcón, G. S., Kimberly, S. Y., Lifton, R. P., Tomino, Y., Mestecky, J., and Novak, B. A. (2008) Aberrant IgA1 glycosylation is inherited in familial and sporadic IgA nephropathy. *J. Am. Soc. Nephrol.* **19**, 1008–1014

63. Suzuki, Y., Suzuki, H., and Tomino, Y. (2008) Pathogenesis of IgA nephropathy: current understanding of mucosa-bone marrow axis in IgA nephropathy. *Nihon Jinzo Gakkai Shi* **50**, 456–460

64. Novak, J., Julian, B. A., Tomana, M., and Mestecky, J. (2008) IgA glycosylation and IgA immune complexes in the pathogenesis of IgA nephropathy. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 20863–20868