J Biochem 127, 845–854 (2000)

Molecular Cloning and Genomic Analysis of Mouse GalNAc α2,6-Sialyltransferase (ST6GalNAc I)[1]

Nobuyuki Kurosawa,[*,2] Shou Takashima,[*,3] Mari Kono,[*,4] Yuzuru Ikehara,[†,5] Mio Inoue,[*,6] Yuriko Tachida,[*,6] Hisashi Narimatsu,[†] and Shuichi Tsuji[*,7]

[*]Molecular Glycobiology, Frontier Research Program, The Institute of Physical and Chemical Research (RIKEN), Wako, Saitama 351-0198; and [†]Division of Cell Biology, Institute of Life Science, Soka University, 1-236 Tangi-cho, Hachioji, Tokyo 192-8577

Received January 26, 2000, accepted February 18, 2000

cDNA clones encoding mouse GalNAc α2,6-sialyltransferase (ST6GalNAc I) were isolated from a mouse submaxillary gland cDNA library. The deduced amino acid sequence of cDNA clones is 526 amino acids in length and has highly conserved motifs among sialyl transferases, sialyl motifs L, S, and VS. The expressed recombinant enzyme exhibited similar substrate specificity to chicken ST6GalNAc I. The mouse ST6GalNAc I gene was expressed in submaxillary gland, mammary gland, colon, and spleen. The mouse ST6GalNAc I gene was also cloned from a mouse genomic library, which was divided into 9 exons spanning over 8 kilobases of genomic DNA. The genomic structure of the mouse ST6GalNAc I gene was similar to that of the mouse ST6GalNAc II gene. Unlike the ST6GalNAc II gene, however, which has a housekeeping gene–like promoter with GC-rich sequences, the ST6GalNAc I gene has two promoters and they do not contain GC-rich sequences but contain putative binding sites for tumor-associated transcription factors such as c-Myb, c-Myc/Max, and c-Ets. Analysis of the 5′-RACE PCR products suggested that the mouse ST6GalNAc I gene expression is regulated by these two promoters in tissue-specific manners.

Key words: genomic organization, sialyl-Tn, sialyltransferase, ST6GalNAc I.

The NeuAcα2,6GalNAc-structure is known to be a common feature of *O*-linked oligosaccharides and α-series gangliosides. The transfer of sialic acid to *N*-acetylgalactosamine (GalNAc) from CMP-sialic acid is catalyzed by a family of sialyltransferases, GalNAc α2,6-sialyltransferases (ST6GalNAc-family).

So far, the cDNAs of six members of the GalNAc α2,6-sialyltransferase family (ST6GalNAc I–VI) have been cloned from chicken (cST6GalNAc I and II) (*2*, *3*), mouse (mST6GalNAc II–VI) (*4–7*, footnote 8),[8] rat (rST6GalNAc III) (*8*), and human (hST6GalNAc I, II, and VI) (*9*, footnote 9).[9] ST6GalNAc I exhibits the broadest substrate specificity, transferring sialic acid with an α2,6-linkage to the GalNAc residues of the following structures: GalNAc-*O*-Ser/Thr, Galβ1,3GalNAc-*O*-Ser/Thr, and NeuAcα2,3Galβ1,3GalNAc-*O*-Ser/Thr. ST6GalNAc II has been considered to have narrower substrate specificity. But we recently found that mouse ST6GalNAc II has similar substrate specificity to ST6GalNAc I. Both genes are expressed in secretory organs, such as the submaxillary and mammary glands, so these enzymes are considered to be involved in the biosynthesis of *O*-glycans of mucin. On the other hand, ST6GalNAc III, IV, and V exhibit restricted substrate specificity, utilizing only the NeuAcα2,3Galβ1,3GalNAc-sequence as an acceptor, although they show some differences in their substrate preferences. ST6GalNAc III can transfer sialic acid to both NeuAcα2,3Galβ1,3GalNAc-*O*-Ser/Thr and ganglioside GM1b. ST6GalNAc IV exhibits strong activity toward NeuAcα2,3Galβ1,3GalNAc and *O*-glycans. ST6GalNAc V has been considered to be the most probable candidate for GD1α synthase. The substrate specificity of ST6GalNAc VI has not been reported.

The sialyl-Tn antigen (STn; NeuAcα2,6GalNAc-*O*-Thr/Ser) is a mucin-associated carbohydrate antigen and the

[1] This work was supported by the following grants: Grants-in-Aid for Scientific Research on Priority Areas, Nos 10152263, 10178104, and 10178105, and for Scientific Research, B11480184, from the Ministry of Education, Science, Sports and Culture of Japan. The nucleotide sequences reported in this paper have been submitted to the GenBank™/EBI Data Bank with accession numbers Y11274 (mouse full length ST6GalNAc I cDNA), AJ271036 (5′-flanking sequence and exon 1a of the mouse ST6GalNAc I gene), Y10294 (5′-flanking sequence and exon 1b of the mouse ST6GalNAc I gene), and Y10295 (exons 2–9 of the mouse ST6GalNAc I gene)

Present addresses [2] Department of Biochemistry, Nagoya University School of Medicine, 65 Tsurumai-cho, Showa-ku, Nagoya 466-8550, [3] Laboratory for Neural Regeneration, Brain Science Institute, RIKEN, Wako, Saitama 351-0198, [4] Section on Biochemical Genetics, Genetics and Biochemistry Branch, NIDDK, National Institutes of Health, Bethesda, Maryland 20892, USA, [5] Division of Pathology, Aichi Cancer Research Institute, Chikusa-ku, Nagoya, [6] Laboratory for Protein Research, Genomic Science Center, RIKEN, Wako, Saitama 351-0198.

[7] To whom correspondence should be addressed. Tel: +81-463-85-5301, Fax: +81-463-85-5301, E-mail: stsuji02@yhb.att.ne.jp

Abbreviations: BSM, bovine submaxillary mucin; kb, kilobase(s); nt, nucleotide(s), PCR, polymerase chain reaction, RACE, rapid amplification of cDNA ends, STn, sialyl-Tn antigen. The abbreviated nomenclature for cloned sialyltransferases follows the system of Tsuji *et al.* (*1*).

[8] In addition to the references, the following data also support it: EMBL accession No. AB035174

[9] In addition to the references, the following data also support it: EMBL accession Nos. AJ251053 and AB035173

[10] In addition to the references, the following data also support it: EMBL accession No. AJ251053

epitopes can be expressed in a tumor-associated fashion in different gastrointestinal organs (*10*). The addition of a sialic acid through an α2,6-linkage to GalNAc-*O*-Thr/Ser prevents the elongation of *O*-linked oligosaccharides (*11*). Thus, the enzyme which catalyzes this reaction is a key enzyme for *O*-linked oligosaccharide biosynthesis. Among sialyltransferases so far characterized, ST6GalNAc I and II have been shown to have STn-antigen synthesis activity (*2*, *9*, our unpublished results for ST6GalNAc II). Previously, we have reported the genomic organization of mouse ST6GalNAc II gene and revealed its housekeeping gene-like expression mechanisms driven by GC-rich promoter (*4*). But we cannot explain the tumor-associated expression of STn-antigen by the ST6GalNAc II promoter. Understanding the mechanisms underlying differential expression of ST6GalNAc genes requires knowledge of the promoter function of these genes. In this study, we will report cloning of mouse ST6GalNAc I cDNA and genomic DNA, including promoter regions, and characterization of the recombinant enzyme.

MATERIALS AND METHODS

Materials—The materials used in this study were essentially the same as in previous studies (*4*). Glycoproteins (fetuin, asialofetuin, α1 acid glycoprotein, and bovine submaxillary mucin (BSM)) were purchased from Sigma. Asialo α1 acid glycoprotein was prepared by incubation at 50°C in 0.1 N H_2SO_4. NANase I, III, and Newcastle disease virus sialidase were from Oxford Glycosystems.

Polymerase Chain Reaction (PCR), Cloning, and Sequencing—Degenerate primers were designed to clone mouse ST6GalNAc I cDNA using sequence information of sialyl motifs L and S of chicken ST6GalNAc I cDNA (*2*). The primers used were 5′-primer, 5′-TGCACT(AC)GCTGTGC(CT)GT(GT)GT(GT)GG(GC)-3′, and 3′-primer, 5′-TG(AT)A(GC)CC(AG)TAGGCACT(GC)ACC(CT)GGTC(AG)CA-3′. PCR experiments were performed essentially according to the procedure described previously (*4*). A mouse submaxillary gland cDNA library was constructed and screened (*2*, *4*). DNA sequences were determined by the deoxynucleotide chain termination method (*12*) using a Thermo Sequenase cycle sequence kit (Amersham, USA).

Sialyltransferase Assay and Linkage Analysis—A truncated form of mouse ST6GalNAc I, lacking the first 33 amino acids of the open reading frame, was prepared by PCR amplification. The primers used were 5′-primer, 5′-TAAGGACTCGAGGGCAAAAGATTCCAG-3′ (nt. 195-221, synthetic *Xho*I site underlined), and 3′-primer, 5′-AGGGGCTCGAGCCTGATCTGGGACCTGGTCA-3′ (complementary nt. 1684-1713, synthetic *Xho*I site underlined). The amplified and *Xho*I-digested 1.5 kb fragment was inserted into the *Xho*I site of a pcDSA vector (*13*). The single insertion in the correct orientation was confirmed by restriction analysis and DNA sequencing, and the resulting plasmid was designated as pcDSA-B4X, which encodes a fusion protein of the IgM signal peptide sequence, a protein A IgG binding domain, and a truncated form of mouse ST6GalNAc I.

COS-7 cells on a 100-mm plate (5×10^6) were transiently transfected with 10 μg of pcDSA-B4X using the DEAE-dextran procedure, then cultured as described previously (*4*). After 48 h, the culture medium was collected, and the protein A-ST6GalNAc I fusion protein secreted into the medium was adsorbed to IgG-Sepharose (Pharmacia). The adsorbed fusion protein was used as an enzyme source as described previously (*4*). Each reaction mixture comprised 50 mM MES buffer (pH 6.0), 10 mM $MgCl_2$, 2 mM $CaCl_2$, 100 μM CMP-[^{14}C]NeuAc (9.1 kBq), 0.15 mM acceptor substrate, and enzyme preparation, in total volume of 20 μl. After 4 h of incubation at 37°C, the reaction was terminated by the addition of SDS-PAGE loading buffer, then the mixtures were directly subjected to SDS-PAGE (for glycoprotein acceptors). For glycolipid acceptors, the incubation mixtures were applied on a C-18 column (Sep-Pak Vac, 100 mg; Waters, Milford, MA), as described previously (*4*). The radioactive materials were visualized with a BAS2000 radio image analyzer (Fuji Film, Tokyo).

For linkage analysis of sialic acids, [^{14}C]NeuAc-incorporated fetuin was synthesized with mouse ST6GalNAc I, ST3Gal I (*13*), ST8Sia II (*14*), and ST6GalNAc III (*5*). The sialylated fetuin was treated with a linkage-specific sialidase, NANase I (specific for α2,3-linked sialic acids), NANase III (specific for α2,3-, α2,6-, and α2,8-linked sialic acids), or Newcastle Disease Virus sialidase (specific for α2,3- and α2,8-linked sialic acids) (*15*). After the sialidase treatment, the desialylated glycoprotein was subjected to SDS-PAGE (gradient gel, 5–20%).

Northern Blot Analysis—Total RNAs were prepared from tissues of adult ICR mice by the guanidium thiocyanate-CsCl method. The poly(A)$^+$ RNAs were purified by Oligotex-dT30 Super (Takara, Kyoto). Five micrograms of poly(A)$^+$ RNA was fractionated on a denaturing formaldehyde-agarose gel (1%) and transferred onto a nylon membrane, and then Northern hybridization was performed.

Isolation of Mouse ST6GalNAc I Genomic Clones—A mouse genomic cosmid library was constructed and screened as described previously (*4*). Cosmid DNA from the positive bacterial colonies was isolated and subjected to restriction analysis. The locations of the exons of the ST6GalNAc I gene were determined by PCR (GeneAmp XL PCR kit, Perkin Elmer, USA) with specific oligonucleotide primers or by hybridization.

Rapid Amplification of the 5′ cDNA Ends (RACE) by PCR—Amplification of the 5′-ends of mouse submaxillary gland ST6GalNAc I cDNA was performed as described previously (*4*). cDNA was synthesized by reverse transcription (Superscript II, GIBCO BRL) of mouse submaxillary gland poly(A)$^+$ RNA (5 μg) using primer RT-380, 5′-CCTGTTTCTGCAGGTCCTTTGAACTGA-3′ (complementary to nt. 338–364). The cDNA was A-tailed with 0.6 units of terminal deoxynucleotidyltransferase (Boehringer Mannheim, Germany) using 0.05 mM dATP. Two consecutive PCRs were performed with two nested sets of primers. In pair 1, the forward primer was *Not*I-$(dT)_{18}$ (Pharmacia, Sweden) and the reverse primer was RT-380. In pair 2, the forward primer was as above but without the T-tail, 5′-AACTGGAAGAATTCGCGGCCGCAGGAA-3′, and the reverse primer was PX-2, 5′-TGGATGGCTGGTCCCTAACAG-3′ (complementary to nt. 83–103). The cDNA was amplified through 35 cycles of a step program (94°C, 30 s; 55°C, 30 s; 72°C, 40 s). The amplification products were digested with *Eco*RI and *Hin*dIII, then subcloned into pUC119 and sequenced. Amplification of the 5′-ends of mouse colon and mammary gland ST6GalNAc I cDNA was performed using a 5′-full RACE core set (Takara) and 1 μg each of mouse colon and

```
A
   1                                                                                               AAAGG

   6 GCAGGGCTCCCTCGCTCTCACCTGAAGGAGCAGGCGATCTGACTGCTGCTTTAAGAAGCTTGCTCTCATCATACATGCTGTTAGGGACCAGCCATCCACC

 106 ATGACGAGATATTGCAGAGGACTTTCCCAGAGGCAAGCTTTCCTGCTTCTGACTGTGTTGGCATTGCTCTTTATTCTGCTCTTCGTTGTTAAGGATCCAA
   1  M  T  R  Y  C  R  G  L  S  Q  R  Q  A  F  L  L  L  T  V  L  A  L  L  F  I  L  L  F  V  V  K  D  P
           Exon 1  ↓  Exon 2
 206 GGGCAAAAGATTCCAGGCGACAGTTCATACTGAACAATGACTCAAGTGCCCAGGAGATTCTGCAAAAGGCTGAGCCCCAAGGACCCATCATGACACTGTC
  34 R  A  K  D  S  R  R  Q  F  I  L  N  N  D  S  S  A  Q  E  I  L  Q  K  A  E  P  Q  G  P  I  M  T  L  S

 306 ACCCAGAGTTCACAACAAAGAGGCAACCTCTGTCAGTTCAAAGGACCTGAAGAAACAGGAGAGAGAGGCAGTCCAAGGAGAGCAAGCTGAGGGGAAGGAG
  68   P  R  V  H  N  K  E  A  T  S  V  S  S  K  D  L  K  K  Q  E  R  E  A  V  Q  G  E  Q  A  E  G  K  E

 406 AAGAGGAAGTTAGAGACCATAAGGCCAGCACCAGAGAATCCTCAGAGCAAGGCAGAGCCTGCTGCAAAGACGCCTGTGTCAGAACACCTGGACAAAGTAC
 101  K  R  K  L  E  T  I  R  P  A  P  E  N  P  Q  S  K  A  E  P  A  A  K  T  P  V  S  E  H  L  D  K  V

 506 CCAGAACTCCAGGAGCACTGTCAACAAGGAAGACACCAATGGCCACAGGAGCTGTCCCAGCTAAGAAGAAAGTGGTCCAGGCTACCAAATCCCCTGCCTC
 134 P  R  T  P  G  A  L  S  T  R  K  T  P  M  A  T  G  A  V  P  A  K  K  K  V  V  Q  A  T  K  S  P  A  S

 606 TTCCCCACACCCCACCACACGAAGAAGGCAAAGGCTGAAGGCCTCTGAGTTCAAGTCTGAGCCTCGGTGGGATTTTGAGGAGGAATATAGCTTGGATATG
 168   S  P  H  P  T  T  R  R  R  Q  R  L  K  A  S  E  F  K  S  E  P  R  W  D  F  E  E  E  Y  S  L  D  M
                   ↓  Exon 3
 706 AGCAGCCTGCAGACGAACTGCTCTGCTTCTGTGAAGATCAAGGCCTCCAAGTCACCATGGCTACAGAATATCTTTCTGCCCAACATCACTCTGTTCCTGG
 201  S  S  L  Q  T  N  C  S  A  S  V  K  I  K  A  S  K  S  P  W  L  Q  N  I  F  L  P  N  I  T  L  F  L
                                                                                          ↓  Exon 4
 806 ACTCTGGACGCTTCACCCAGAGCGAGTGGAACCGTCTAGAGCACTTCGCTCCTCCCTTTGGCTTCATGGAACTCAATCAGTCCCTGGTACAGAAGGTGGT
 234 D  S  G  R  F  T  Q  S  E  W  N  R  L  E  H  F  A  P  P  F  G  F  M  E  L  N  Q  S  L  V  Q  K  V  V

 906 GACCCGCTTCCCTCCAGTTCGCCAGCAGCAGCTCCTCCTGGCCAGCCTCCCTACTGGGTACTCCAAGTGTATCACCTGTGCTGTAGTGGGCAACGGGGGC
 268   T  R  F  P  P  V  R  Q  Q  Q  L  L  L  A  S  L  P  T  G  Y  S  K  C  I [T  C  A  V  V  G  N  G  G]
                                                                ↓  Exon 5
1006 ATCCTGAATGATTCACGTGTTGGCCGGGAGATAGACAGCCATGACTATGTTTTCCGACTGAGTGGAGCCGTTATTAAAGGATATGAACAGGATGTGGGGA
 301 [I  L  N  D  S  R  V  G  R  E  I  D  S  H  D  Y  V  F  R  L  S  G  A  V  I  K  G  Y  E  Q  D  V  G]
                                                                                               ↓  Exon 6
1106 CCCGGACATCCTTCTATGGCTTCACTGCTTTCTCTCTGACCCAGTCTATCCTCATTTTGGGTAGACGAGGCTTCCAGCATGTGCCTCTGGGAAAGGACGT
 334 [T  R  T] S  F  Y  G  F  T  A  F  S  L  T  Q  S  I  L  I  L  G  R  R  G  F  Q  H  V  P  L  G  K  D  V
                                                                                                  ↓  Exon 7
1206 CCGATATCTACACTTCCTGGAAGGCACCCGGAACTATGAGTGGCTGGAAGCTATGTTTTTGAATCAGACCTTGGCAAAAACCCACCTTTCCTGGTTCAGG
 368   R  Y  L  H  F  L  E  G  T  R  N  Y  E  W  L  E  A  M  F  L  N  Q  T  L  A  K  T  H  L  S  W  F  R
                                                                                         ↓  Exon 8
1306 CATAGGCCTCAGGAAGCCTTCCGGAATGCCTTGGACTTGGATCGATACTTGCTGCTGCACCCAGACTTTCTCCGTTACATGAAGAACAGGTTTCTGAGGT
 401  H  R  P  Q  E  A  F  R  N  A  L  D  L  D  R  Y  L  L  L  H  P  D  F  L  R  Y  M  K  N  R  F  L  R
                                                                                         ↓  Exon 9
1406 CAAAGACCCTGGACACTGCCCATTGGAGAATATACCGCCCCACCACTGGTGCCCTCCTGCTGCTCACAGCCCTTCATCTCTGTGACAAGGTCAGCGCCTA
 434 S  K  T  L  D  T  A  H  W  R  I  Y  R |P  T  T  G  A  L  L  L  L  T  A  L  H  L  C  D  K  V  S  A  Y

1506 TGGCTTCATCACCGAGGGCCATGAGCGCTTCTCTGACCACTACTATGATACATCGTGGAAAAGGTTGATCTTTTATATCAACCATGACTTCAGGTTAGAG
 468   G |F  I  T  E  G  H  E  R  F  S  D  H  Y  Y  D  T  S  W  K  R  L  I  F  Y  I  N  H  D  F  R  L  E

1606 AGAATGGTGTGGAAGCGGTTGCATGATGAAGGCATCATCTGGCTCTACCAGCGTCCACAAAGTGACAAAGCAAAGAACTGACCAGGGCCCAGATCAGGGT
 501  R  M  V  W  K  R  L  H  D  E  G  I  I  W  L  Y  Q  R  P  Q  S  D  K  A  K  N  *
1706 GTTCCCCTCGACCACTTACAACACAGAATGCAAGGGCTCCTCAAATGCTCCTACCTGTGTCTCTGGAATCAGACCTGCTCCACACCTTTCAAGAGTCTAG
1806 ATCTACAGAGTAAGTTCACTGTCTGACAGAGCTACACAGAGAAACCCTCTCCAAAACAAAACAAAAGAAAACAACAACAAAAGCAAACAGGAAAATCACT
1906 GAAAAGTAACATTGACAGTAATGTTTTACTTTAAAATAAATAAAATCAGTCAAGTTACTTCCCCCCAGGAATAGGGAGAGATCCCCAAAGTGTCAGCACC
2006 ACCCACAAGGCATTGACTTCAACCTCACAGTGTGCTCACCTGACATTTTGGGTGAGTAGACTGAACACATTCAAGGTACGGCAGGATTTCAGAGTGAGCC
2106 CTCTGTGCCACTGAATGCTTGTAATCTTATTGAGAGAGAGAGAGGGAGAGAGAGAGAGAGAGAGAGAGTGTGTGTATGTGTGTGTGTACATCTGTGTA
2206 AATGTACAGGTTTACACATTTTCAGGAGTTCAAGGCTAGCATTGGCTACACAGCAAGTTCAAGGTCAGCCTGGGTTACAGGAAGACTCTGTCTCACATCA
2306 AAGAACTTCCCAGAGTGGAATTTTGTGTTAATTCAGAGCCATGATGGGGCACCATGAGCTGCCATTGAATAAAGGGACTACCCTCTGA
```

B

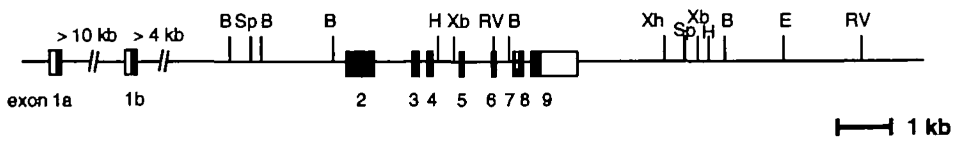

Fig. 1. **Nucleotide sequence, genomic structure, and restriction map of the mouse ST6GalNAc I gene.** A: The nucleotide and amino acid sequences are numbered from the most probable transcription initiation site and initiation methionine, respectively. The transmembrane domain is double underlined. Sialyl motifs L, S, and VS are boxed by solid line, dashed line, and bold solid line, respectively. The polyadenylation signal is underlined. The nucleotide differences between sequences from ICR mice and NIH3T3 cells are nt. 223, C-T; nt 225, A-C; nt. 235, C-T; nt. 236, T-G, nt. 240, C-G, nt 247, T-G; nt 263, T-A; nt 266, T-G; nt. 267, G-A; nt. 527, T-C; nt 1528, G-C; nt. 1692, G-T; nt. 1960, T-G, and sequences, AAGAGAAAGAGAGAGAGAGAGAGAGA and GT, are inserted between nt. 2156-2157 and nt 2182-2183, respectively. B The genomic organization of the mouse ST6GalNAc I gene is shown. The restriction enzyme sites are B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; RV, *Eco*RV; Sp, *Spe*I; Xb, *Xba*I; and Xh, *Xho*I. Closed boxes represent coding regions, and open boxes represent 5′- and 3′-untranslated regions.

mammary gland poly(A)$^+$ RNA according to the manufacturer's instructions. For cDNA synthesis, the mouse ST6GalNAc I cDNA specific primer, B4-RT, 5′-ACTCTGGGTGACAG-3′ (complementary to nt. 301–314) was used, and the cDNA was self-ligated. Then two consecutive PCRs were performed with the self-ligated cDNA as a template. The first PCR was performed using the first primer set, 5′-GAGGCAAGCTTTCCTGCTTC-3′ (B4-S1, nt. 135 to 154) and 5′-CCTCTGCAATATCTCGTCATG-3′ (B4-A1, complementary to nt. 105 to 125). The second PCR was performed with 1 μl of the above reaction mixture as a template using the second primer set, 5′-GTTGTTAAGGATCCAAGGGC-3′ (B4-S2, nt. 190 to 209) and 5′-GAGAGCAAGCTTCTTAAAGC-3′ (B4-A2, complementary nt. 53 to 72). The amplified products were blunt-ended and cloned into the *Eco*RV site of pBluescript II SK(+), and sequenced.

RESULTS

Cloning of Mouse ST6GalNAc I cDNA—To isolate the cDNA encoding mouse ST6GalNAc I, we performed PCR using two degenerate primers based on the sequences of two highly conserved regions, sialyl motifs L and S of chicken ST6GalNAc I (*2*), with mouse submaxillary gland cDNA as a template. The amplified 0.5-kb PCR fragment was sequenced and found to encode a part of protein homologous to chicken ST6GalNAc I. Thus the amplified PCR fragment was used as a probe to screen mouse submaxillary gland cDNA library. Several overlapping clones were obtained and sequenced, and one of these cDNA clones, named B4, encoded 526 amino acids and exhibited a type II transmembrane topology (Fig. 1A). The predicted amino acid sequence of B4 revealed domain structures such as an N-terminal cytoplasmic tail, a transmembrane domain, a stem region, and a C-terminal active domain including highly conserved regions, sialyl motifs L, S, and VS. These features are common to all sialyltransferases cloned to date (*2–9*, *16–18*). The stem region of B4 is very long, like chicken and human ST6GalNAc I (*2*, *9*). In addition, B4 contains a Kurosawa motif [Cys-Xaa$^{76-82}$-Cys-Xaa$^{1-2}$-Cys-Ala-Xaa-Val-Xaa$^{150-160}$-Cys (Xaa denotes any amino acid residue)], which has been shown in the ST3Gal-family (*13*, *19–21*), ST6GalNAc I from chicken and human (*2*, *9*) and ST6GalNAc II from mouse, chicken, and human (*3*, *4*, footnote 10).[10] So far, six members of the ST6GalNAc family (GalNAc α2,6-sialyltransferase) have been cloned from various sources. We performed a multiple alignment of amino acid sequences of these ST6GalNAc enzymes (chicken ST6GalNAc I and II, human ST6GalNAc I, II, and VI, rat ST6GalNAc III, mouse ST6GalNAc II, III, IV, V, VI, and newly cloned B4) using sequences between highly conserved sialyl motifs L and S, and the result was represented schematically (Fig. 2). In Fig. 2, the branch length shows similarity among these enzymes. ST6GalNAc I and II families form one cluster in this schema, suggesting that ST6GalNAc I and II are derived from a common ancestral gene. The newly cloned B4 is included in ST6GalNAc I subfamily. Comparison of the overall predicted amino acid sequence of B4 with those of chicken and human ST6GalNAc I showed 49.8 and 70.5% identity, respectively. These results strongly suggest that B4 is a mouse orthologue of chicken and human ST6GalNAc I.

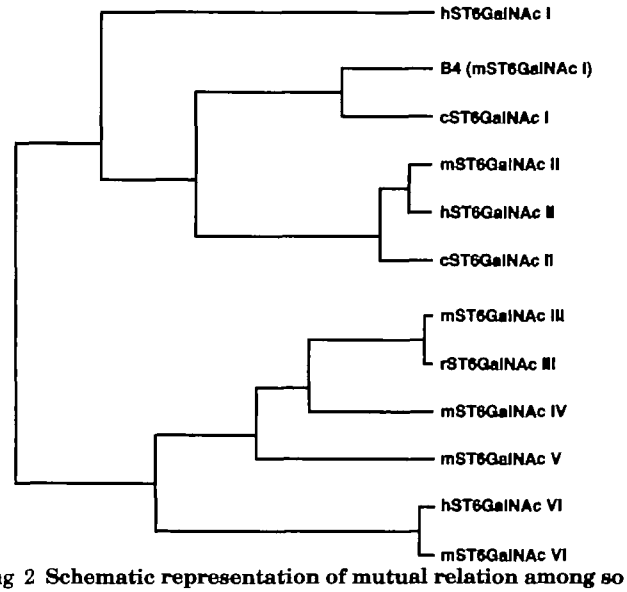

Fig 2 **Schematic representation of mutual relation among so far cloned GalNAc α2,6-sialyltransferases.** The dendrogram was constructed according to Higgins and Sharp (*38*). Prefixes of c, h, m, and r mean chicken, human, mouse, and rat, respectively.

TABLE I. **Comparison of acceptor specificities of mouse B4 (mST6GalNAc I) and chicken ST6GalNAc I (cST6GalNAc I).** The table shows the relative activities as to the incorporation of sialic acids into asialo-fetuin as a substrate. Each acceptor substrate was used at the concentration of 0.15 mM. A value of 0 indicates less than 0.1% activity. R represents the remainder of the *N*-linked oligosaccharide chain. NT means not tested.

Acceptor	Representative structures of carbohydrates	B4 (mST6GalNAc I) (%)	cST6GalNAc I (*2*) (%)
Fetuin	NeuAcα2,3Galβ1,3GalNAc-Ser/Thr NeuAcα2,3Galβ1,3(NeuAcα2,6)GalNAc-Ser/Thr NeuAcα2,6(3)Galβ1,4GlcNAc-R	120	148
Asialo-fetuin		100	100
Asialo-agalacto-fetuin		91.0	95.0
BSM	NeuAcα2,3Galβ1,3GalNAc-Ser/Thr NeuAcα2,6GalNAc-Ser/Thr	19.0	16.0
Asialo-BSM		169	194
α1 acid glycoprotein	NeuAcα2,6(3)Galβ1,4GlcNAc-R	0	6.30
Asialo-α1 acid glycoprotein		0	4.20
Ovomucoid		2.50	7.30
NeuAcα2,3Galβ1,3GalNAc-benzyl		0	NT
Asialo-GM1	Galβ1,3GalNAcβ1,4Galβ1,4Glcβ1,1Cer	0	0
GM1b	NeuAcα2,3Galβ1,3GalNAcβ1,4Galβ1,4Glcβ1,1Cer	0	NT
Paragloboside	Galβ1,4GlcNAcβ1,3Galβ1,4Glcβ1,1Cer	0	NT

Acceptor Substrate Specificity of Mouse ST6GalNAc I—To verify that B4 encodes a mouse orthologue of chicken and human ST6GalNAc I, we constructed an expression vector, pcDSA-B4X, which expresses the fusion gene encoding the secretable form of B4 fused to the IgG-binding domain of *Staphylococcus aureus* protein A. The vector was transfected into COS-7 cells, which secreted the protein A-fused B4 into the medium. Then the secreted enzyme was adsorbed to IgG-Sepharose gel, which was used as an enzyme source. Table I shows the substrate acceptor specificity of B4. Both B4 and chicken ST6GalNAc I efficiently transferred sialic acids to asialo fetuin and asialo BSM. No significant activity was observed toward α1 acid glycoprotein, which has *N*-linked oligosaccharides. In addition, oligosaccharides and glycosphingolipids did not serve as acceptors for B4. The acceptor specificity of B4 was almost identical to that of chicken ST6GalNAc I.

For linkage analysis, [^{14}C]NeuAc-incorporated fetuins were synthesized by B4, mouse ST3Gal I (*13*), ST6GalNAc III (*14*), and ST8Sia II (*5*). Each sialylated fetuin was treated with linkage-specific sialidases, *i.e.*, NANase I (specific for α2,3-linked sialic acids), NANase III (specific for α2,3-, α2,6-, and α2,8-linked sialic acids), and Newcastle disease virus sialidase (specific for α2,3- and α2,8-linked sialic acids) (*15*). The [^{14}C]NeuAc residue of fetuin sialylated by B4 was removed by the NANase III treatment but not by the NANase I or Newcastle disease virus sialidase treatments (Fig 3). The digestion pattern was virtually identical to that of ST6GalNAc III but differed from those of ST3Gal I and ST8Sia II, suggesting that the incorporated sialic acids bound through α2,6-linkage. B4 was shown to have α2,6-sialyltransferase activity toward GalNAc-*O*-Ser/Thr and almost identical substrate specificity to that of chicken ST6GalNAc I. Therefore, we concluded that B4 is the mouse orthologue of chicken and human ST6GalNAc I, and we named it mouse ST6GalNAc I.

Northern Blot Analysis of the ST6GalNAc I Gene—In our previous study, the expression levels and patterns of mouse sialyltransferase genes in various tissues were examined by competitive PCR technique, and the ST6GalNAc I gene was found to express in submaxillary gland, colon, and lactating mammary gland (*22*). In this study, we performed Northern blot analysis to evaluate the transcript size and the expression levels and patterns of the mouse ST6GalNAc I gene in various tissues (Fig. 4). As we expected, the ST6GalNAc I transcripts were observed in submaxillary gland, lactating mammary gland, and colon, and additionally in spleen. The expression levels in submaxillary gland, lactating mammary gland, and spleen were higher than in colon. But the expression levels in lactating mammary gland, colon, and spleen seemed to vary according to animal conditions, because we also observed different expression patterns. In our previous study, the expression level in colon was relatively high but that in spleen was very low (*22*), and sometimes the expression level in lactating mammary gland was also relatively low. However, we could not determine the reason for these differential expression patterns. The size of the main transcript in these tissues was 2.5-kb, but minor transcripts ranging from 2.0 to 4.0-kb were also detected. The expression pattern of the mouse ST6GalNAc I gene seems to be more restricted than that of the mouse ST6GalNAc II gene (*4*), which is expressed ubiquitously. In addition, the expression level of the ST6GalNAc I gene was lower than that of the ST6GalNAc II gene in these tissues (*22*). The size of the obtained ST6GalNAc I cDNA clone from the mouse submaxillary gland cDNA library was about 2.5-kb. In this study, we will focus on this 2.5 kb transcript.

Isolation of the Mouse ST6GalNAc I Gene—Screening of an NIH3T3 cell genomic cosmid library with mouse ST6GalNAc I cDNA probe resulted in the isolation of two independent overlapping clones, COS-B410 and COS-B41. COS-B41 was found to include the entire mouse ST6GalNAc I gene. The mouse ST6GalNAc I gene was divided into 9 exons with introns ranging from 84 bp to over 4 kb, and spanning over 8 kb (Fig. 1). Exon 1 contained the entire 5′-

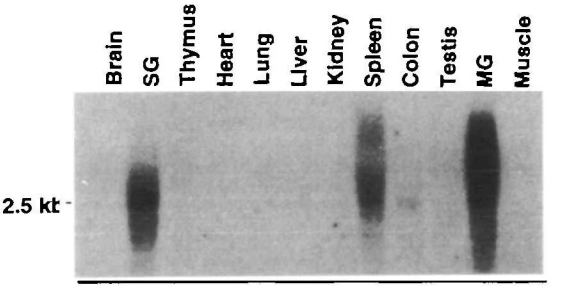

Fig. 4 **Northern blot analysis of the mouse ST6GalNAc I gene expression.** Poly(A)$^+$ RNA (5 μg) was electrophoresed and hybridized with ST6GalNAc I cDNA probe. SG, submaxillary gland; MG, lactating mammary gland

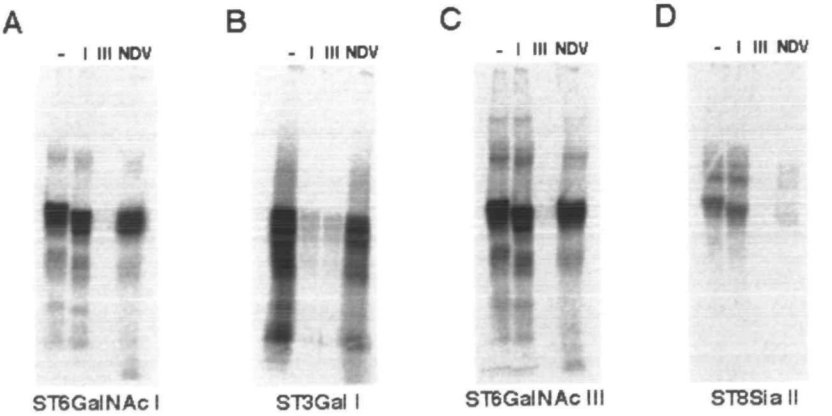

Fig 3. **Linkage analysis of incorporated sialic acids by mouse ST6GalNAc I.** [^{14}C]Sialylfetuin sialylated with B4 (ST6GalNAc I) (A), mouse ST3Gal I (B), mouse ST6GalNAc III (C), or mouse ST8Sia II (D) was digested with linkage-specific sialidases, NANase I (specific for α2,3-linked sialic acids), NANase III (specific for α2,3-, α2,6-, and α2,8-linked sialic acids), and Newcastle disease virus sialidase (specific for α2,3- and α2,8-linked sialic acids) After digestion, desialylated fetuin was subjected to SDS-PAGE (gradient gel of 5–20%), and radioactivity was detected with a BAS2000 radio image analyzer (Fuji Film, Tokyo). –, without sialidase treatment, I, NANase I treatment, III, NANase III treatment, NDV, Newcastle disease virus sialidase treatment.

untranslated region, the first 39 amino acids of a coding region containing a cytoplasmic domain (11 a.a.), a hydrophobic signal anchor sequence (19 a.a.), and a part of stem region (Fig. 5). Exon 2 encoded a stem region. Exons 3–9 encoded the putative active domain of the enzyme. Exon 9 contained a translation termination codon and a long 3′-untranslated region. Poly(A)$^+$ addition occurred 18 nucleotides downstream from the T residue of the polyadenylation signal (AATAAA). Introns of the mouse ST6GalNAc I gene were located in the protein coding regions and all exons were spliced at corresponding positions to those of the mouse ST6GalNAc II gene (Table II) (*4*). The splice junction in exon 3–4 occurred after the first nucleotide of the amino acid codon, those in exons 1–2, 4–5, and 6–7 occurred after the second nucleotide of the amino acid codon, and those in exons 2–3, 5–6, and 8–9 occurred between codons (Table II). Exons 4, 5, 8, and 9 of the mouse ST6GalNAc I and II genes were more highly conserved than exons 1, 2, 3, 6, and 7 of those genes (Fig. 5).

Mapping of the Transcription Initiation Sites and Identification of Two Different Promoters—5′-RACE PCR was performed in order to determine the transcription initiation sites of the mouse ST6GalNAc I gene. When submaxillary gland poly(A)$^+$ RNA was used for the reaction, three transcription initiation sites were identified (Fig. 6B). But most of the transcription initiation sites were mapped at the adenosine 105 nucleotides upstream from the initiation codon. We also performed RNase protection assay and determined the transcription initiation sites (data not shown). Some of these coincided with the transcription initiation sites determined by 5′-RACE PCR, but several extra transcription initiation sites were also identified, indicating that submaxillary gland ST6GalNAc I mRNA is transcribed from multiple transcription initiation sites.

When colon and mammary gland poly(A)$^+$ RNAs were used for the reaction, three transcription initiation sites were also identified (Fig. 6A). But these transcription initiation sites were different from those identified for submaxillary gland ST6GalNAc I mRNA. The sequences of the RACE products contained the same sequences as the mouse submaxillary gland ST6GalNAc I mRNA, but they also contained different sequences in their 5′-ends. In these cases, most of the transcription initiation sites were also mapped at the adenosine 105 nucleotides upstream from the initiation codon.

The results of 5′-RACE PCR suggest the existence of two different tissue-specific promoters. Exon 1 (exon 1b in Fig. 1B) and its 5′-flanking region, which we first found, seemed

Exon	1	2	3	4	5	6	7	8	9
ST6GalNAc I	39	166	56	58	46	35	30	33	63
ST6GalNAc II	41	20	58	57	46	35	28	33	55
Homology (%)	17	—	15	47	56	20	29	57	60

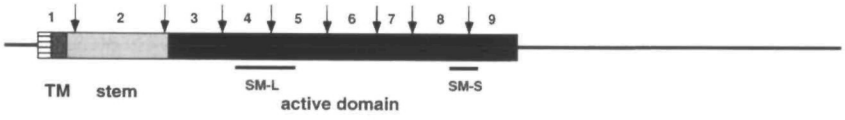

Fig. 5 **Comparison of exon structure of the mouse ST6GalNAc I and II genes.** Numbers of amino acids in each exon and the amino acid sequence homology between mouse ST6GalNAc I and II are shown. The domain structure of mouse ST6GalNAc I is shown schematically. The boxes indicate translated sequences and horizontal bars indicate untranslated sequences. sialyl motifs L and S are denoted as SM-L and SM-S, respectively. TM indicates a transmembrane domain. Exons are marked and numbered (1–9), and splicing sites are indicated by vertical arrows.

TABLE II. **Exon/intron junctions of the mouse ST6GalNAc I gene.** The nucleotide sequences at the intron (lowercase letters) and exon (uppercase letters) junctions are shown. The derived amino acid sequence is displayed below the nucleotide sequence. Exons are numbered from the 5′ end. The exon and intron sizes are indicated in bp. The numbering starts at the initiator methionine as position +1.

Exon		Splice donor	Intron	Splice acceptor
1	>214	AAA GAT TCC (38) Lys Asp Ser	AGgtaagaagatgcc---	>4–10 kb	---tattctctgtgcagG	CGA CAG TTC (40) Arg Gln Phe
2	499	CTG CAG ACG (205) Leu Gln Thr	gtgagtctggacctt—	757	---tcttcctgcacatag	AAC TGC TCT (206) Asn Cys Ser
3	169	AAT CAG TCC (261) Asn Gln Ser	Cgtgagtgcttactt---	92	---tgctgtgttccagTG	GTA CAG AAG (263) Val Gln Lys
4	172	TAT GTT TTC (318) Tyr Val Phe	CGgtatgttctccct---	614	—atctttctacacagA	CTG AGT GGA (320) Leu Ser Gly
5	499	CTG GGA AAG (365) Leu Gly Lys	gtgagaagggaggag—	336	---gtccttggtcttcag	GAC GTC CGA (366) Asp Val Arg
6	104	TCC TGG TTC (399) Ser Trp Phe	AGgtacccctccct—	195	---atctctggccacagG	CAT AGG CCT (401) His Arg Pro
7	90	ATG AAG AAC (429) Met Lys Asn	AGgtaagggtaggaa---	86	—ggtattttgaacagG	TTT CTG AGG (431) Phe Leu Arg
8	499	TGT GAC AAG (463) Cys Asp Lys	gtgagcctgcaaacc—	150	---tccattttcctgcag	GTC AGC GCC (464) Val Ser Ala

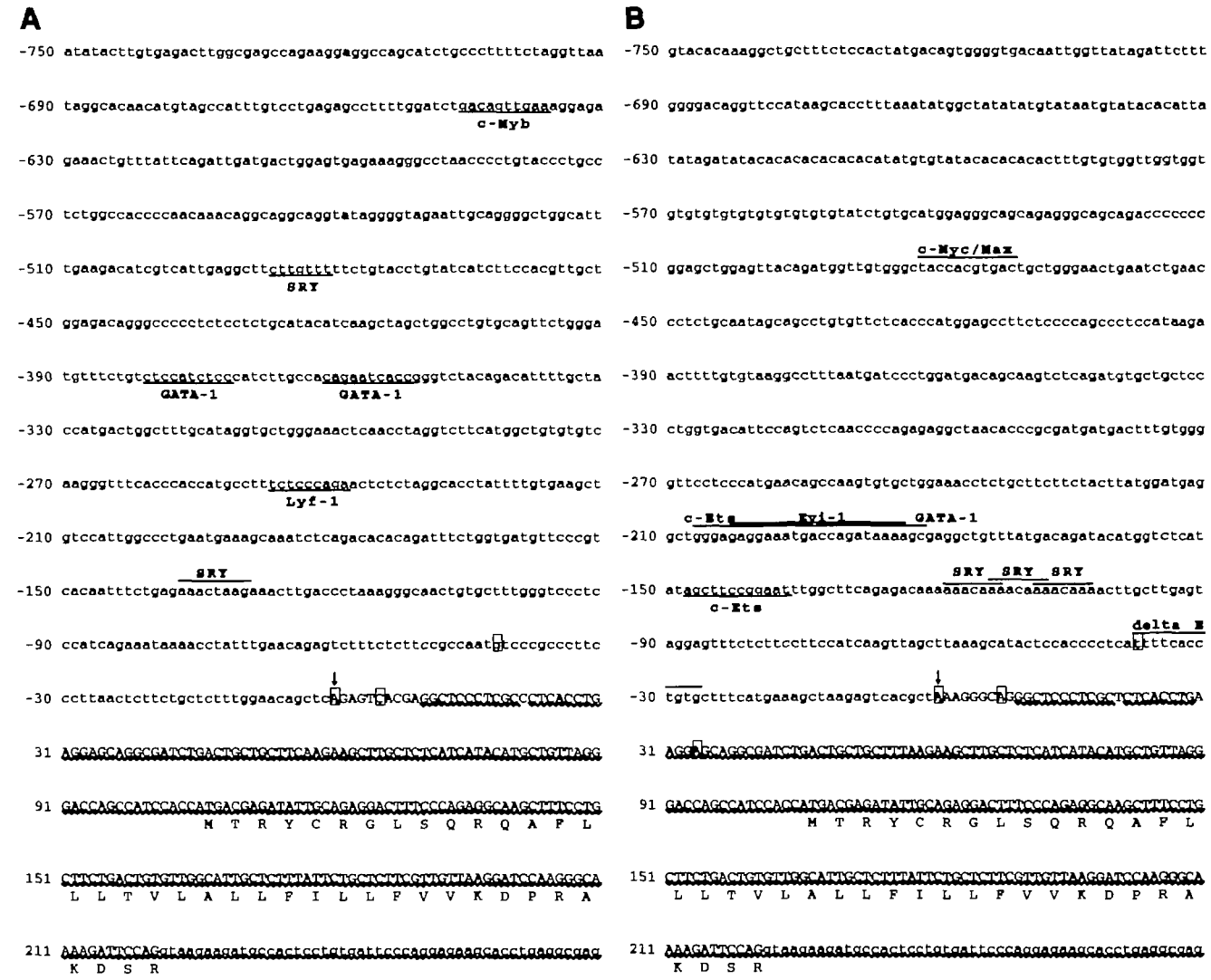

Fig. 6 **5′-Flanking regions of the mouse ST6GalNAc I gene.** The sequence of the 5′-flanking region, exon 1a (colon and mammary gland specific), and part of the intron are shown in A; and the sequence of the 5′-flanking region, exon 1b (submaxillary gland specific), and part of the intron are shown in B. Transcription initiation sites determined by 5′-RACE PCR are boxed. The most probable transcription initiation start site is numbered as +1 for convenience and marked by a vertical arrow. The first exon is typed in upper case letters and 5′-untranslated sequences and introns are typed in lower case letters. Putative binding sequences of transcription factors are overlined (forward orientation) or underlined (reverse orientation). Identical sequences in A and B are shown by the waved underline.

to be submaxillary gland-specific and were located more than 4 kb upstream from exon 2. On further analysis of the COS-B41 clones, we found another promoter and exon 1 (exon 1a in Fig. 1B), which seemed to be colon- and mammary gland–specific and were located more than 14 kb upstream from exon 2. Interestingly, exons 1a and 1b encoded the same amino acid sequence but had different promoters. The 5′-flanking region of the submaxillary gland–specific exon 1b contains no canonical TATA and CCAAT box (Fig. 6B). Sp1-binding sites are often observed in the immediate upstream region of the transcription initiation sites of the sialyltransferase genes (*4*, *23*, *24*), but there is no Sp1-binding site in this 5′-flanking region. Instead, several putative transcription factor–binding sites, *i.e.*, two c-Ets–binding sites (*25*), a GATA-1–binding site (*26*), an Evi-1–binding site (*27*), a c-Myc/Max–binding site (*28*, *29*), a delta E–binding site (*30*), and three SRY-binding sites (*31*) were observed in this 5′-flanking region (Fig. 6B). On the other hand, there is also no canonical TATA and CCAAT box or Sp1-binding site in the immediate upstream region of the transcription initiation sites of exon 1a (Fig. 6A). However, several putative transcription factor–binding sites, *i.e.*, a c-Myb–binding site (*32*), a Lyf-1–binding site (*33*), two GATA-1–binding sites, and two SRY-binding sites are observed in this 5′-flanking region. Comparison of the 5′-flanking regions of exon 1a and 1b revealed no significant similarity, suggesting that the tissue-specific expression of the mouse ST6GalNAc I gene is regulated in quite different manners.

DISCUSSION

In this study, we cloned mouse ST6GalNAc I cDNA and examined the acceptor substrate specificity of the enzyme. We also determined the genomic structure of the mouse ST6GalNAc I gene.

The evaluation of substrate specificities revealed that ST6GalNAc I exhibited high α2,6-sialyltransferase activity toward GalNAc-*O*-Ser/Thr. Recently, we found that mouse ST6GalNAc II also exhibits the same activity (data not shown). Siaα2,6GalNAc-*O*-Ser/Thr structures occur in a cancer-associated carbohydrate antigen, STn. It is well known that elevation of STn antigen correlates with a poor prognosis in gastric cancer patients. To elucidate the relation between STn antigen expression and prognosis in gastric cancers, it is necessary to identify the substance of a STn synthase and examine its expression in carcinogenesis. Among sialyltransferases so far characterized, only ST6GalNAc I and II can synthesize STn antigen.

Although ST6GalNAc I and II exhibit similar substrate specificity, the expression patterns of these genes are quite different. The expression of the ST6GalNAc II gene was ubiquitously observed in mouse tissues (*4*), whereas the expression of the mouse ST6GalNAc I gene was limited to submaxillary gland, mammary gland, colon, and spleen. In addition, the expression level of the ST6GalNAc II gene was much higher than that of the ST6GalNAc I gene (*22*).

The differences between the expression patterns of these genes can be explained by their promoter structures. The promoter region of the mouse ST6GalNAc II gene is embedded in GC-rich sequences and contains no canonical TATA and CAAT box but has putative binding sites for a transcription factor, Sp1 (*4*). So this promoter has features of a housekeeping gene promoter. The promoters of the mouse ST6GalNAc III (*34*), ST6GalNAc IV (*34*), ST8Sia II (*23*), and ST8Sia IV (*24*) genes also have similar structures. On the other hand, the mouse ST6GalNAc I gene has two promoters, and both of them contain no canonical TATA and CAAT box and no GC-rich sequence. Significantly, however, they have putative binding sites for tumor-associated transcription factors, such as c-Myb, c-Myc, and c-Ets. These features are quite different from those of housekeeping gene-like promoters of other sialyltransferase genes. The results of 5′-RACE PCR suggest that tissue-specific expression of the ST6GalNAc I gene is regulated by these two promoters. So far, tissue specific promoters of a sialyltransferase gene have been reported for the rat ST6Gal I gene (*35*). We cannot yet properly evaluate the ST6GalNAc I promoter activity because of unavailability of ST6GalNAc I–expressing mouse cell lines. But the colon-mammary gland–specific promoter worked well in a mouse embryonal carcinoma cell line, P19, although this does not express the ST6GalNAc I gene, while the submaxillary gland specific promoter worked poorly in this cell line (data not shown). These results may also support the idea of tissue-specific promoters of the mouse ST6GalNAc I gene.

Recently, the tumor-associated expression of the human ST6GalNAc I gene was reported (*9*). Considering above results, the ability of mouse ST6GalNAc I to synthesize STn antigen, and the tissue specific promoters of the gene, this ST6GalNAc I appears to be the strongest candidate for a mouse STn synthase in carcinogenesis. Both promoters of the mouse ST6GalNAc I gene have some elements for binding tumor-associated transcription factors. We are now analyzing which promoter is used and which transcription factors are involved in the tumor-associated expression of STn antigen. We will also examine the function of STn antigen in carcinogenesis by using ST6GalNAc I knockout mice, generation of which is now in progress.

Genomic structural analyses of the mouse sialyltransferase genes have progressed (*4*, *23*, *24*, *34*, *36*). The genomic structures of the mouse ST8Sia II and IV genes are similar to each other (Fig. 7). ST8Sia II and IV exhibit similar substrate specificity and they can synthesize polysialic acid. Recently, we found that the genomic structures of the mouse ST6GalNAc III and IV genes are also similar to each other (Fig. 7). ST6GalNAc III and IV also exhibit similar substrate specificity, *i.e.*, they utilize the NeuAcα2,3Galβ1,3GalNAc sequence as an acceptor. In this study, we have shown that the genomic structures of the mouse ST6GalNAc I and II genes are similar to each other, and that ST6GalNAc I and II exhibit similar substrate specificity. Therefore, we hypothesize that these sialyltransferases with similar substrate specificity arose from a com-

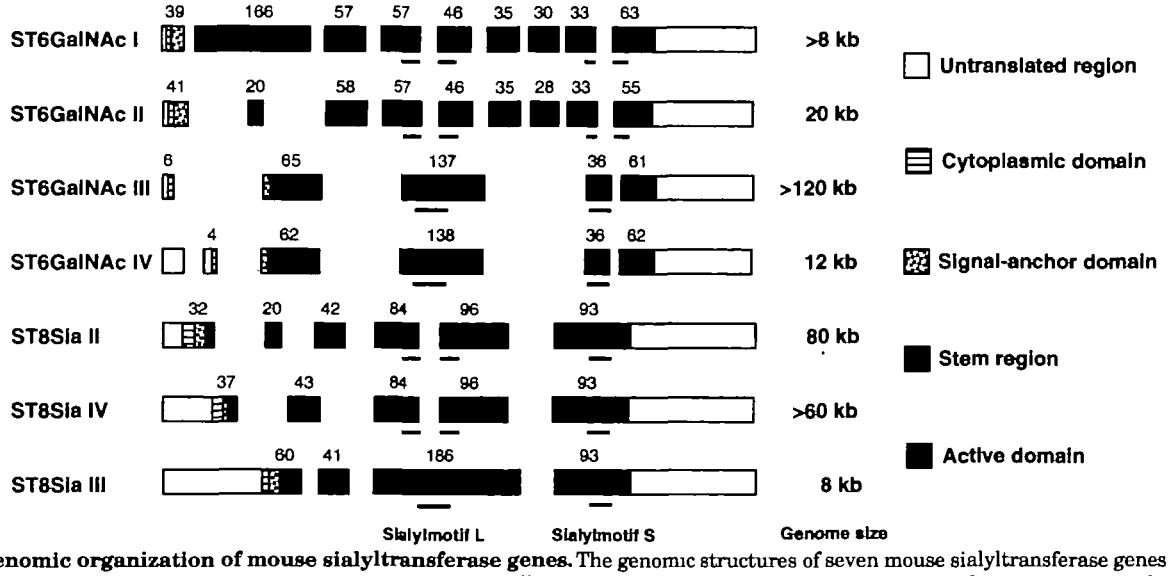

Fig. 7 **Genomic organization of mouse sialyltransferase genes.** The genomic structures of seven mouse sialyltransferase genes are presented. The protein domain structure is represented schematically by a rectangle, which is subdivided to show the major structural elements of the protein. Sialyl motifs L and S are underlined

mon ancestral gene. All mouse sialyltransferases characterized to date have sialyl motifs L, S, and VS (*37*). Thus, the prototype of sialyltransferase gene probably had sialyl motifs L, S, and VS, and it evolved into several subfamilies of sialyltransferase genes. The biological significance of these multiple gene families is unclear at this stage. One interpretation is that they may be important for fine control of the expression of sialylglycoconjugates, resulting in stage- and tissue-specific variety of glycosylation patterns. Study on the transcription regulation of ST6GalNAc genes will certainly help us to understand the regulation mechanisms of sialylglycoconjugate expression patterns and their biological functions.

We wish to thank Dr. Yoshitaka Nagai, Director of the Glycobiology Research Group, and Dr. Tomoya Ogawa, Coordinator of the Group, Frontier Research Program of the Institute of Physical and Chemical Research (RIKEN), for their support in this work

REFERENCES

1 Tsuji, S., Datta, A.K., and Paulson, J.C. (1996) Systematic nomenclature for sialyltransferases. *Glycobiology* **6 (7)**, v–vii
2 Kurosawa, N., Hamamoto, T., Lee, Y.-C., Nakaoka, T., Kojima, N., and Tsuji, S. (1994) Molecular cloning and expression of GalNAc α2,6-sialyltransferase. *J. Biol. Chem* **269**, 1402–1409
3. Kurosawa, N., Kojima, N., Inoue, M., Hamamoto, T., and Tsuji, S. (1994) Cloning and expression of Galβ1,3GalNAc-specific GalNAc α2,6-sialyltransferase. *J. Biol. Chem* **269**, 19048–19053
4. Kurosawa, N., Inoue, M., Yoshida, Y., and Tsuji, S. (1996) Molecular cloning and genomic analysis of mouse Galβ1,3GalNAc-specific GalNAc α2,6-sialyltransferase *J. Biol. Chem* **271**, 15109–15116
5 Lee, Y.-C., Kaufmann, M., Kitazume-Kawaguchi, S., Kono, M., Takashima, S., Liu, H., Kurosawa, N., Pircher, H., and Tsuji, S. (1999) Molecular cloning and functional expression of two members of mouse NeuAcα2,3Galβ1,3GalNAc GalNAcα2,6-sialyltransferase family, ST6GalNAc III and IV *J. Biol. Chem* **274**, 11958–11967
6 Okajima, T., Fukumoto, S., Ito, H., Kiso, M., Hirabayashi, Y., Urano, T., Furukawa, K., and Furukawa, K. (1999) Molecular cloning of brain-specific GD1α synthase (ST6GalNAc V) containing CAG/glutamine repeats. *J. Biol. Chem* **274**, 30557–30562
7. Ikehara, Y., Shimizu, N., Kono, M., Nishihara, S., Nakanishi, H., Kitamura, T., Narimatsu, H., Tsuji, S., and Tatematsu, M. (1999) A novel glycosyltransferase with a polyglutamine repeat; a new candidate for GD1α synthase (ST6GalNAc V) *FEBS Lett.* **463**, 92–96
8. Sjoberg, E.R., Kitagawa, H., Glushka, J., van Halbeek H., and Paulson, J.C. (1996) Molecular cloning of a developmentally regulated *N*-acetylgalactosamine α2,6-sialyltransferase specific for sialylated glycoconjugates. *J. Biol. Chem* **271**, 7450–7459
9. Ikehara, Y., Kojima, N., Kurosawa, N., Kudo, T., Kono, M., Nishihara, S., Issiki, S., Morozumi, K., Itzkowitz, S., Tsuda, T., Nishimura, S., Tsuji, S., and Narimatsu, H. (1999) Cloning and expression of a human gene encoding an *N*-acetylgalactosamine-α2,6-sialyltransferase (ST6GalNAc I). a candidate for synthesis of cancer-associated sialyl-Tn antigens. *Glycobiology* **9**, 1213–1224
10 Itzkowitz, S.H., Yuan, M., Montgomery, C.K., Kjeldsen, T., Takahashi, H.K., Bigbee, W.L., and Kim, Y. (1989) Expression of Tn, sialosyl-Tn and T antigens in human colon cancer. *Cancer Res.* **49**, 197–204
11. Saitoh, O., Gallagher, R.E., and Fukuda, M. (1991) Expression of aberrant *O*-glycans attached to leukosialin in differentiation-deficient HL-60 cells. *Cancer Res.* **51**, 2854–2862
12. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) DNA sequencing with chain-termination inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467
13 Lee, Y.-C., Kurosawa, N., Hamamoto, T., Nakaoka, T., and Tsuji, S. (1994) Molecular cloning and expression of Galβ1,3GalNAc-α2,3-sialyltransferase from mouse brain *Eur. J. Biochem* **216**, 377–385
14. Kojima, N., Yoshida, Y., Kurosawa, N., Lee, Y.-C., and Tsuji, S. (1995) Enzymatic activity of a developmentally regulated member of the sialyltransferase family (STX). evidence for α2,8-sialyltransferase activity toward *N*-linked oligosaccharides *FEBS Lett* **360**, 1–4
15. Paulson, J.C., Weinstein, J., Dorland, L., van Halbeek, H., and Vliegenthart, J.F. (1982) Newcastle disease virus contains a linkage-specific glycoprotein sialidase. Application to the localization of sialic acid residues in *N*-linked oligosaccharides of alpha 1-acid glycoprotein *J Biol. Chem* **257**, 12734–12738
16. Drickamer, K. (1993) A conserved disulphide bond in sialyltransferases. *Glycobiology* **3(1)**, 2–3
17 Livingston, B.D. and Paulson, J.C. (1993) Polymerase chain reaction cloning of a developmentally regulated member of the sialyltransferase gene family *J Biol. Chem.* **268**, 11504–11507
18. Geremia, R.A., Harduin-Lepers, A., and Delannoy, P. (1997) Identification of two novel conserved amino acid residues in eukaryotic sialyltransferases implications for their mechanism of action *Glycobiology* **7(2)**, v–vii
19 Lee, Y.-C., Kojima, N., Wada, E., Kurosawa, N., Nakaoka, T., Hamamoto, T., and Tsuji, S. (1994) Cloning and expression of cDNA for a new type of Galβ1,3GalNAc α2,3-sialyltransferase *J Biol Chem.* **269**, 10028–10033
20 Kono, M., Ohyama, Y., Lee, Y.-C., Hamamoto, T., Kojima, N., and Tsuji, S. (1996) Mouse β-galactoside α2,3-sialyltransferases comparison of *in vitro* substrate specificities and tissue specific expression *Glycobiology*, **7**, 469–479
21. Kono, M., Takashima, S., Liu, H., Inoue, M., Kojima, N., Lee, Y.-C., Hamamoto, T., and Tsuji, S. (1998) Molecular cloning and functional expression of a fifth type α2,3-sialyltransferase (mST3Gal V: GM3 synthase) *Biochem. Biophys. Res. Commun* **253**, 170–175
22. Takashima, S., Tachida, Y., Nakagawa, T., Hamamoto, T., and Tsuji, S. (1999) Quantitative analysis of expression of mouse sialyltransferase genes by competitive PCR. *Biochem Biophys. Res. Commun* **260**, 23–27
23 Yoshida, Y., Kurosawa, N., Kanematsu, T., Kojima, N., and Tsuji, S. (1996) Genomic structure and promoter activity of the mouse polysialic acid synthase gene (mST8Sia II). *J Biol. Chem* **271**, 30167–30173
24 Takashima, S., Yoshida, Y., Kanematsu, T., Kojima, N., and Tsuji, S. (1998) Genomic structure and promoter activity of the mouse polysialic acid synthase (mST8Sia IV/PST) gene. *J. Biol. Chem.* **273**, 7675–7683
25. Woods, D.B., Ghysdael, J., and Owen, M.J. (1992) Identification of nucleotide preferences in DNA sequences recognized specifically by c-Ets-1 protein. *Nucleic Acids Res.* **20**, 699–704
26 Whyatt, D.J., deBoer, E., and Grosveld, F. (1993) The two zinc finger-like domains of GATA-1 have different DNA binding specificities *EMBO J* **12**, 4993–5005
27. Perkins, A.S., Fishel, R., Jenkins, N., and Copeland, N.G. (1991) Evi-1, a murine zinc finger proto-oncogene, encodes a sequence-specific DNA-binding protein *Mol. Cell Biol* **11**, 2665–2674
28 Blackwood, E.M. and Eisenman, R.N. (1991) Max a helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with Myc. *Science* **251**, 1211–1217
29. Blackwell, T.K., Kretzner, L., Blackwood, E.M., Eisenman, R.N., and Weintraub, H. (1990) Sequence-specific DNA binding by the c-Myc protein *Science* **250**, 1149–1151
30 Sekido, R., Murai, K., Funahashi, J., Kamachi, Y., Fujisawa-Sehara, A., Nabeshima, Y., and Kondoh, H. (1994) The delta-crystallin enhancer-binding protein delta EF1 is a repressor of E2-box-mediated gene activation. *Mol. Cell Biol* **14**, 5692–5700
31. Pontiggia, A., Rimini, R., Harley, V.R., Goodfellow, P.N., Lovell-Badge, R., and Bianchi, M.E. (1994) Sex-reversing mutations

affect the architecture of SRY-DNA complexes. *EMBO J.* **13**, 6115–6124
32 Majello, B., Kenyon, L.C., and Dalla-Favera, R. (1986) Human c-myb protooncogene: nucleotide sequence of cDNA and organization of the genomic locus. *Proc. Natl. Acad. Sci. USA* **83**, 9636–9640
33 Hahm, K., Ernst, P., Lo, K., Kim, G.S., Turck, C., and Smale, S.T. (1994) The lymphoid transcription factor LyF-1 is encoded by specific, alternatively spliced mRNAs derived from the *Ikaros* gene *Mol. Cell Biol.* **14**, 7111–7123
34 Takashima, S., Kurosawa, N., Tachida, Y., Inoue, M., and Tsuji, S. (2000) Comparative analysis of the genomic structures and promoter activities of mouse Siaα2,3Galβ1,3GalNAc GalNAc α2,6-sialyltransferase genes (ST6GalNAc III and IV): characterization of their Sp1 binding sites. *J. Biochem.* **127**, 399–409
35. Wang, X., O'Hanlon, T.P., Young, R.F., and Lau, J.T. (1990) Rat β-galactoside α2,6-sialyltransferase genomic organization: alternate promoters direct the synthesis of liver and kidney transcript. *Glycobiology* **1**, 25–31
36. Yoshida, Y., Kurosawa, N., Kanematsu, T., Taguchi, A., Arita, M., Kojima, N., and Tsuji, S. (1996) Unique genomic structure and expression of the mouse α2,8-sialyltransferase (ST8Sia III) gene *Glycobiology* **6**, 573–580
37. Tsuji, S. (1999) Molecular cloning and characterization of sialyltransferases in *Sialobiology and Other Novel Forms of Glycosylation* (Inoue, Y., Lee, Y.C., and Troy, F.A., eds. pp. 145–154, Gakushin Publishing Company, Osaka
38. Higgins, D.G. and Sharp., P.H. (1988) Clustal: A package for performing multiple sequence alignment on a micro computer *Gene* **73**, 237–244