J. Biochem. Mol. Biol. Vol. 30, No. 2, pp. 95~100 March 31, 1997

Molecular Cloning and Expression of cDNAs Encoding Mouse Galβ1,3(4)GlcNAc α2,3-Sialyltransferase (mST3Gal III) and Galβ1,4(3)GlcNAc α2,3-Sialyltransferase (mST3Gal IV)

Kyoung-Sook Kim, Cheorl-Ho Kim[1], Deug-Yong Shin[2] and Young-Choon Lee*

Division of Molecular Glycobiology, and [2]Division of Protein Function,
Korea Research Institute of Bioscience and Biotechnology, Taejon 305-600
[1]Department of Biochemistry and Molecular Biology, College of Oriental Medicine,
Dongguk University, Kyung-Pook 780-350, Korea
(Received December 2, 1996)

Abstract : Two kinds of cDNA encoding mouse Galβ1,3(4)GlcNAc α2,3-sialyltransferase (mST3Gal III) and Galβ1,4(3)GlcNAc α2,3-sialyltransferase (mST3Gal IV) were isolated from mouse brain cDNA library by means of a PCR-based approach. The cDNA sequences included an open reading frame coding for proteins of 374 and 333 amino acids, respectively, and the primary structure of these enzymes suggested a putative domain structure consisting of four regions, like that in other glycosyltransferases. The deduced amino acid sequences of mST3Gal III and IV showed a 98% and 89% identity with rat ST3Gal III and human ST3Gal IV, respectively. Northern analysis indicated that the expression of mST3Gal III mRNA was abundant in heart, liver and adult brain, while that of mST3Gal IV mRNA was detected in all tissues tested except for testis, but the level was the highest in liver. Soluble forms of mST3Gal III and IV transiently expressed in COS cells exhibited enzyme activity toward acceptor substrates containing the terminal either Galβ1,3GlcNAc or Galβ1, 4GlcNAc sequences. The substrate preferences of both enzymes were stronger for tetrasaccharides than for disaccharides.
Key words : cDNA cloning, expression, sialyltransferase.

Sialic acids (NeuAc or Sia) are key determinants of carbohydrate structures invloved in a variety of biological functions, such as cell-cell interaction, adhesion, cell differentiation and clearance of asialoglycoproteins (Paulson, 1989; Brandely *et al.*, 1990; Varki, 1992). Sialic acid levels on cell surface change in a regulated manner during development and differentiation and oncogenic transformation (Roos, 1984; Kimber, 1989) The transfer of sialic acid from CMP-NeuAc to the terminal position of the carbohydrate group of glycoproteins and glycolipids is catalyzed by a family of glycosyltransferases called sialyltransferases (Kornfeld, 1985). Since each sialyltransferase exhibits strict specificity for acceptor substrate and linkages, more than 12 different sialyltransferases are required for the synthesis of all known sialyl-oligosaccharide structures (Broquet *et al.*, 1991). To date, thirteen members of the sialyltransferase gene family have been cloned (Kono *et al.*, 1996; Tsuji, 1996), and they have a domain structure similar to that of other glycosyltransferases: a short NH_2-terminal cytoplasmic tail, a hydrophobic transmembrane domain, and an extended stem region followed by a large COOH-terminal active domain. Comparison of these sequences has revealed highly conserved regions named sialylmotif in the active domain of these enzymes, which are not found in other glycosyltransferases. The sialylmotif has recently been used to clone new members of the sialyltransferase gene family using a polymerase chain reaction (PCR) homology approach employing degenerate primers to conserved regions (Tsuji, 1996).

There are sequences for three major types in the reaction pattern of sialyltransferase so far cloned: α2,3-, α2, 6- and α2,8-linkage. The α2,3-sialyltransferases act on N- and O- linked glycans of glycoproteins as well as gangliosides, while the α2,6- and α2,8-sialyltransferases are restricted to glycoproteins, and gangliosides and formation of polysialic acids, respectively.

Recently we cloned two kinds of cDNA encoding mouse brain Galβ1,3GalNAc α2,3-sialyltransferase (mST3GalI and mST3GalII) by a PCR method directed to sialylmotif (Lee *et al.*, 1993; Lee *et al.*, 1994). They

*To whom correspondence should be addressed.
Tel : 82-42-860-4260, Fax : 82-42-860-4597
E-mail : yclee@geri4680.geri.re.kr

exhibit the same acceptor substrate specificities and are able to synthesize only the NeuAc α2,3Galβ1,3GalNAc sequence in gangliosides and glycoproteins as well as oligosaccharides, but show a different acceptor substrate preference (Kojima *et al.*, 1994).

To compare gene expression and substrate specificities of the four β-galactoside α2,3-sialyltransferases from the same animal, the isolation of cDNAs of two other kinds of sialyltransferase was attempted from mouse brain cDNA library using sequence information obtained from enzymes previously cloned. Here, we report the cloning of cDNAs encoding mouse brain Galβ1, 3(4)GlcNAc α2,3-Sialyltransferase (mST3Gal III) and Galβ1,4(3)GlcNAc α2,3-Sialyltransferase (mST3Gal IV) and the expression of their active domain in COS cells.

Materials and Methods

Materials

Unless otherwise indicated, the materials used in this study were essentially the same as in the previous studies (Lee *et al.*, 1993; Lee *et al.*, 1994). Lacto-*N*-tetraose, lacto-*N*-neotetraose, Galβ1,3GalNAc, Galβ1, 3GlcNAc and Galβ1,4GlcNAc were purchased from Sigma (St. Louis, USA). CMP-[^{14}C]NeuAc (11 GBq/mmol) was from Amersham Corp. IgG-Sepharose was from Pharmacia (Uppsala, Sweden). Restriction and modification enzymes were from Takara Co.

Polymerase chain reaction (PCR)

PCR fragments used as probes for screening of the mouse brain cDNA library were amplified using T3 primer existing in Uni-ZAP XR vector as sense primer, and primers 101 and 102 deduced from the sialylmotif L of the previously cloned rat liver ST3Gal III (Wen *et al.*, 1992) and human placenta ST3Gal IV (Kitagawa and Paulson, 1994) as antisense primers, respectively. Primers 101: 5'-CTTCTCAAAGCCCTTCACAGG-3'and 102: 5'-CTCT(A)GGGGC(A)ATTGTTCAATCT-GAT-3' were synthesized with an Applied Biosystem 394 DNA Synthesizer, and cDNA purified from mouse brain cDNA library was used as a template for PCR. PCR amplification was carried out using a DNA thermal cycler (Perkin-Elmer Cetus), with 30 cycles consisting of denaturation at 94°C for 30 s, annnealing at 55°C for 30 s, and extension at 72°C for 1 min. The PCR products were blunt-ended, phosphorylated, then subcloned into the *Sma*I site on plasmids pUC118 or pUC119, and sequenced.

Screening of the mouse brain cDNA library and DNA sequence analysis

About 10^6 plaques of the mouse brain Uni-ZAP XR cDNA library (Lee *et al.*, 1993; Lee *et al.*, 1994) were screened on duplicate filters by plaque hybridization with the ^{32}P-labelled PCR product as probes. Positive clones were purified and finally selected as single colonies through *in vivo* excision of the pBluescript plasmids. The cDNA fragments obtained on *Eco*RI and *Xho*I digestion were subcloned into pUC118 and 119 vector plasmids, and single-stranded DNAs produced on infection with a helper phage (M13KO7) were used as templates for DNA sequencing. DNA sequences were determined by the dideoxynucleotide chain-termination method (Sanger *et al.*, 1977) using an Autocycle DNA sequencing kit and A. L. F. DNA sequencer (Pharmacia).

Nothern blot analysis

Five μg of total RNA was fractionated on a denaturing formaldehyde-agarose gel (1%) and then transferred onto a nylon membrane (Nytran, Schleicher & Schuell). Northern blotting was carried out by the same method as described previously (Lee *et al.*, 1993) using cDNA fragments labeled with [α-^{32}P]dCTP as probes.

Construction and expression of soluble forms of mST3Gal III and IV

To derive expression of soluble forms of mST3Gal III and IV, truncated forms of mST3Gal III and IV, which lacks 59 and 42 amino acids from the NH_2 terminus of the open reading frame, respectively, were prepared by PCR amplification using 5'-primers containing an in-frame *Eco*RI site (mST3Gal III; 5'-CTGTGAATTCGGC-TTCCTCCTGAAGCTGGA-3' and mST3Gal IV; 5'-CTGTGAATTCAAAGAGCCATGCTTCCAGGG-3') and 3'-primers located 50 bp downstream from the stop codon with an *Xho*I site (mST3Gal III; 5'-TTCC-ATCTCGAGTCCTCTCGAGTGCAGCTGCT-3' and mST3Gal IV; 5'-TTCCATCTCGAGTCAGATGG-CAAAGTAGGGAA-3'). PCR reaction was carried out with Taq polymerase by 94°C for 40 s, 55°C for 30 s and 72°C for 2 min. The amplified PCR fragments (1.0 kb and 926 bp) were subcloned into *Eco*RI/*Xho*I of pcDSA (Lee *et al.*, 1996) to yield expression vectors, pcDSA-mST3Gal III and IV, resulting in fusion of mST 3Gal III and IV to the IgM signal sequence followed by protein A IgG binding domain in vector. Each 3 μg of pcDSA-mST3Gal III and IV was transfected into COS-7 cells using the Lipofectin (BRL). After 48 h, the cell culture media was collected and the protein A fusion proteins expressed in the medium were absorbed into IgG-Sepharose (15 μl of resin per 10 ml of culture medium) at 4°C for 16 h. The resins were collected by centrifugation, washed three times with phosphate-buffered saline (PBS), suspended in 50 μl (final volume) of Dulbecco's modified Eagle medium (DMEM) without fe-

tal bovine serum, and used as soluble enzymes.

Sialyltransferase assay

The cloned enzyme assays were performed by essentially the same method as described previously (Lee *et al.*, 1994). In a total volume of 10 μl, the assay mixture containing 50 mM sodium cacodylate buffer (pH 6.0), 10 mM $MgCl_2$, 50 μM CMP-[^{14}C]NeuAc (0.9 Mbq/pmol), 0.5% Triton CF-54, 1 mM oligosaccharide acceptors, and 1 μl of enzyme preparation was incubated at 37℃ for 1 h. For separation of the product from CMP-NeuAc, the assay mixture was applied on a silica gel 60 HPTLC plate (Merck, Germany) and then developed with ethanol/pyridine/1-butanol/acetic acid/water (100:10:10:3:30). The radioactivity on each plate was detected and counted with the BAS2000 radioimage analyzer (Fuji Film, Japan).

Results

Cloning and nucleotide sequence of mST3Gal III cDNA

T3 primer, existing in Uni-ZAP XR vector and primer 101, complementary to the nucleotide sequence 568-589 of the sialylmotif L of the previously cloned rat liver ST3Gal III (Wen *et al.*, 1992) were used to clone mST3Gal III cDNA. When the amplified PCR fragments were subcloned and sequenced, one (0.7 kb) of them was found to correspond to the N-terminal part of mST3Gal III. This fragment was used as a hybridization probe to screen the mouse brain cDNA library. A single clone with 2.2 kb insert cDNA was obtained, subcloned and sequenced. The 2.2 kb cDNA contained a single open reading frame (1,125 bp) encoding a protein of 374 amino acids, including 150 bp of a 5' and 928 bp of a 3' non-coding region (Fig. 1). The identities between the open reading frames of the mouse and rat cDNAs were 98% and 99% at the nucleotide and amino acid sequence levels, respectively. As observed for other cloned sialyltransferases with a type II transmembrane topology (Fig. 3), the mST3Gal III is predicted to have an NH_2-terminal cytoplasmic tail (8 residues), a hydrophobic transmembrane domain (20 residues), an extended stem region (31 residues), and a large COOH-terminal active domain (315 residues). The mST3Gal III has two potential N-linked glycosylation sites (Fig. 1).

Cloning and nucleotide sequence of mST3Gal IV cDNA

Cloning of mST3Gal IV cDNA was carried out by the same strategy as that of mST3Gal III. T3 primer, existing in Uni-ZAP XR vector and primer 102, complementary to the nucleotide sequence 429-453 of the sialylmotif L of the previously cloned human placenta ST3Gal IV (Kitagawa and Paulson, 1994), were used to clone mST3Gal IV cDNA. When the amplified PCR fragments were subcloned and sequenced, one (0.6 kb) of them was found to correspond to the N-terminal part of mST3Gal IV. This fragment was used as a hybridization probe to screen the mouse brain cDNA library. A single clone with 1.5 kb insert cDNA was obtained, subcloned and sequenced. The 1.5 kb cDNA contained a single open reading frame (1,002 bp) encoding a protein of 333 amino acids, including 170 bp of a 5' and 323 bp

```
ATG GGA CTC TTG GTA TTT GTG CGC AAT CTG CTG CTA GCC CTC TGC CTC TTC CTG    54
MET Gly Leu Leu Val Phe Val Arg Asn Leu Leu Leu Ala Leu Cys Leu Phe Leu    18

GTC CTG GGA TTT TTG TAT TAT TCT GCC TGG AAG CTA CAC TTA CTC CAA TGG GAA   108
Val Leu Gly Phe Leu Tyr Tyr Ser Ala Trp Lys Leu His Leu Leu Gln Trp Glu    36

GAC TCC AAT TCA TTG CTT CTT TCC CTT GAC TCC GCT GGA CAA ACC CTA GGC ACA   162
Asp Ser Asn Ser Leu Leu Leu Ser Leu Asp Ser Ala Gly Gln Thr Leu Gly Thr    54

GAG TAT GAT AGG CTG GGC TTC CTC CTG AAG CTG GAC TCT AAA CTG CCT GCC GAG   216
Glu Tyr Asp Arg Leu Gly Phe Leu Leu Lys Leu Asp Ser Lys Leu Pro Ala Glu    72

CTA GCC ACC AAG TAC CCA AAC TTT TCC GAG GGA GCT TGC AAA CCC GGC TAC GCT   270
Leu Ala Thr Lys Tyr Pro Asn Phe Ser Glu Gly Ala Cys Lys Pro Gly Tyr Ala    90

TCA GCC ATG ATG ACT GCC ATC TTC CCC AGG TTC TCC AAG CCA GCA CCC ATG TTC   324
Ser Ala Met Met Thr Ala Ile Phe Pro Arg Phe Ser Lys Pro Ala Pro Met Phe   108

CTG GAT GAC TCC TTT CGC AAG TGG GCT AGG ATC CGG GAG TTT GTG CCA CCT TTT   378
Leu Asp Asp Ser Phe Arg Lys Trp Ala Arg Ile Arg Glu Phe Val Pro Pro Phe   126

GGG ATC AAA GGT CAA GAC AAT CTG ATC AAA GCC ATC TTG TCA GTC ACC AAA GAA   432
Gly Ile Lys Gly Gln Asp Asn Leu Ile Lys Ala Ile Leu Ser Val Thr Lys Glu   144

TAC CGC CTG ACC CCT GCC TTG GAC AGC CTC CAC TGC CGC CGC TGC ATC ATT GTA   486
Tyr Arg Leu Thr Pro Ala Leu Asp Ser Leu His Cys Arg Arg Cys Ile Ile Val   162

GGC AAT GGA GGG GTC CTC GCC AAC AAG TCT CTG GGG TCA CGG ATT GAC GAC TAT   540
Gly Asn Gly Gly Val Leu Ala Asn Lys Ser Leu Gly Ser Arg Ile Asp Asp Tyr   180

GAC ATT GTG ATC AGA CTG AAC TCA GCA CCT GTG AAG GGC TTT GAG AAG GAC GTG   594
Asp Ile Val Ile Arg Leu Asn Ser Ala Pro Val Lys Gly Phe Glu Lys Asp Val   198
                                  Sialyl motif L
GGC AGC AAG ACC ACC CTA CGC ATC ACC TAC CCT GAG GGC GCC ATG CAG CGA CCT   648
Gly Ser Lys Thr Thr Leu Arg Ile Yhr Tyr Pro Glu Gly Ala Met Gln Arg Pro   216

GAG CAA TAT GAA CGA GAC TCT CTC TTT GTC CTC GCT GGC TTC AAG TGG CAG GAC   702
Glu Gln Tyr Glu Arg Asp Ser Leu Phe Val Leu Ala Gly Phe Lys Trp Gln Asp   234

TTC AAG TGG CTG AAG TAC ATC GTC TAC AAG GAG AGA GTC AGT GCA TCT GAT GGC   756
Phe Lys Trp Leu Lys Tyr Ile Val Tyr Lys Glu Arg Val Ser Ala Ser Asp Gly   252

TTC TGG AAG TCC GTG GCC ACC CGA GTG CCC AAG GAG CCC CCT GAG ATC CGC ATC   810
Phe Trp Lys Ser Val Ala Thr Arg Val Pro Lys Glu Pro Pro Glu Ile Arg Ile   270

CTC AAC CCA TAC TTC ATC CAG GAG GCT GCC TTC ACT CTC ATT GGA CTG CCC TTC   864
Leu Asn Pro Tyr Phe Ile Gln Glu Ala Ala Phe Thr Leu Ile Gly Leu Pro Phe   288

AAC AAC GGC CTC ATG GGC AGA GGG AAC ATC CCA ACC CTT GGC AGT GTG GCA GTG   918
Asn Asn Gly Leu Met Gly Arg Gly Asn Ile Pro Thr Leu Gly Ser Val Ala Val   306

ACC ATG GCA CTA CAC GGC TGT GAT GAA GTG GCA GTC GCG GGC TTT GGC TAT GAC   972
Thr Met ala Leu His Gly Cys Asp Glu Val Ala Val Ala Gly Phe Gly Tyr Asp   324
                  Sialyl motif S
ATG AAC ACA CCC AAT GCA CCC CTG CAC TAC TAT GAA ACT GTG CGC ATG GCA GCC  1026
Met Asn Thr Pro Asn Ala Pro Leu His Tyr Tyr Glu Thr Val Arg Met Ala Ala   342

ATC AAA GAG TCC TGG ACA CAC AGC ATC CAG CGA GAG AAA GAG TTT CTG CGG AAG  1080
Ile Lys Glu Ser Trp Thr His Ser Ile Gln Arg Glu Lys Glu Phe Leu Arg Lys   360

CTA GTG AAG GCA CGT GTC ATC ACT GAC TTA AGC AGC GGT ATC TGA              1125
Leu Val Lys Ala Arg Val Ile Thr Asp Leu Ser Ser Gly Ile ---               374
```

Fig. 1. Nucleotide and deduced amino acid sequences of mST 3Gal III. The nucleotide and deduced amino acid sequences are numbered from the presumed start codon and initiation methionine, respectively. The *boxed* amino acid residues correspond to a putative hydrophobic transmembrane domain. Sialyl motif L and S are *underlined*. The *asterisks* indicate potential *N*-glycosylation sites (Asn-X-Ser/Thr). The N-terminal amino acid residue used to construct the fusion protein is indicated by the *arrow*. The nucleotide sequence of mST3Gal III has been submitted to the GeneBank™/EMBL Data Bank with accession number X84234.

of a 3' non-coding region (Fig. 2). The identities between the open reading frames of the mouse and human cDNAs were 85% and 89% at the nucleotide and amino acid sequence levels, respectively. As observed for other cloned sialyltransferases as well as the mST 3Gal III with a type II transmembrane topology (Fig. 3), the mST3Gal IV is predicted to have an NH_2-terminal cytoplasmic tail (8 residues), a hydrophobic transmembrane domain (18 residues), an extended stem region (16 residues), and a large COOH-terminal active domain (291 residues). The mST3Gal IV has three potential N-linked glycosylation sites (Fig. 2).

Northern blot analysis of mST3Gal III and IV

```
ATG ACC AGC AAA TCT CAC TGG AAG CTC CTG GCC CTG GCT CTG GTC CTT GTT GTT   54
MET Thr Ser Lys Ser His Trp Lys Leu Leu Ala Leu Ala Leu Val Leu Val Val   18

GTC ATG GTG TGG TAT TCC ATC TCC CGA GAA GAT AGG TAC ATT GAG TTC TTT TAT  108
Val Met Val Trp Tyr Ser Ile Ser Arg Glu Asp Arg Tyr Ile Glu Phe Phe Tyr   36

TTT CCC ATC TCA GAG AAG AAA GAG CCA TGC TTC CAG GGT GAG GCA GAG AGA CAG  162
Phe Pro Ile Ser Glu Lys Lys Glu Pro Cys Phe Gln Gly Glu Ala Glu Arg Gln   54

GCC TCT AAG ATT TTC GGC AAC CGT TCT AGG GAA CAG CCC ATC TTT CTG CAG CTT  216
Ala Ser Lys Ile Phe Gly Asn Arg Ser Arg Glu Gln Pro Ile Phe Leu Gln Leu   72

AAG GAT TAT TTC TGG GTA AAG ACA CCA TCC ACC TAT GAG CTG CCC TTT GGG ACT  270
Lys Asp Tyr Phe Trp Val Lys Thr Pro Ser Thr Tyr Glu Leu Pro Phe Gly Thr   90

AAA GGA AGT GAA GAC CTT CTT CTC CGG GTG CTG GCC ATC ACT AGC TAT TCT ATA  324
Lys Gly Ser Glu Asp Leu Leu Leu Arg Val Leu Ala Ile Thr Ser Tyr Ser Ile  108

CCT GAG AGC ATA AAG AGC CTC GAG TGT CGT CGC TGT GTT GTG GTG GGA AAT GGG  378
Pro Glu Ser Ile Lys Ser Leu Glu Cys Arg Arg Cys Val Val Val Gly Asn Gly  126

CAC CGG TTG CGG AAC AGC TCG CTG GGC GGT GTC ATC AAC AAG TAC GAC GTG GTC  432
His Arg Leu Arg Asn Ser Ser Leu Gly Gly Val Ile Asn Lys Tyr Asp Val Val  144
                                  Sialyl motif L
ATC AGA TTG AAC AAT GCT CCT GTG GCT GGC TAC GAG GGA GAT GTG GGC TCC AAG  486
Ile Arg Leu Asn Asn Ala Pro Val Ala Gly Tyr Glu Gly Asp Val Gly Ser Lys  162

ACC ACC ATA CGT CTC TTC TAT CCT GAG TCG GCC CAC TTT GAC CCT AAA ATA GAA  540
Thr Thr Ile Arg Leu Phe Tyr Pro Glu Ser Ala His Phe Asp Pro Lys Ile Glu  180

AAC AAC CCA GAC ACG CTC TTG GTC CTG GTA GCT TTC AAG GCG ATG GAC TTC CAC  594
Asn Asn Pro Asp Thr Leu Leu Val Leu Val Ala Phe Lys Ala Met Asp Phe His  198

TGG ATT GAG ACC ATC TTG AGT GAT AAG AAG CGG GTG CGA AAA GGC TTC TGG AAA  648
Trp Ile Glu Thr Ile Leu Ser Asp Lys Lys Arg Val Arg Lys Gly Phe Trp Lys  216

CAG CCT CCC CTC ATC TGG GAT GTC AAC CCC AAA CAG GTC CGG ATT CTA AAC CTC  702
Gln Pro Pro Leu Ile Trp Asp Val Asn Pro Lys Gln Val Arg Ile Leu Asn Leu  234

TTC TTT ATG GAG ATT GCA GCA GAC AAG CTC CTG AGC CTG CCC ATA CAA CAG CCT  756
Phe Phe Met Glu Ile Ala Ala Asp Lys Leu Leu Ser Leu Pro Ile Gln Gln Pro  252

CGA AAG ATC AAG CAG AAG CCA ACC ACG GGT CTG CTA GCC ATC ACC TTG GCT CTA  810
Arg Lys Ile Lys Gln Lys Pro Thr Thr Gly Leu Leu Ala Ile Thr Leu Ala Leu  270
                                                  Sialyl motif S
CAC CTC TGC GAC TTA GTG CAC ATT GCT GGT TTT GGC TAT CCA GAT GCC TCC AAC  864
His Leu Cys Asp Leu Val His Ile Ala Gly Phe Gly Tyr Pro Asp Ala Ser Asn  288

AAG AAG CAG ACC ATC CAC TAC TAT GAA CAG ATC ACA CTT AAG TCT ATG GCG GGA  918
Lys Lys Gln Thr Ile His Tyr Tyr Glu Gln Ile Thr Leu Lys Ser Met Ala Gly  306

TCA GGC CAT AAT GTC TCC CAA GAG GCT ATC GCC ATG AAG CGG ATG CTA GAG ATG  972
Ser Gly His Asn Val Ser Gln Glu Ala Ile Ala Met Lys Arg Met Leu Glu Met  324

GGA GCT GTC AAG AAC CTC ACA TAC TTC TGA                                1002
Gly Ala Val Lys Asn Leu Thr Tyr Phe ---                                 333
```

Fig. 2. Nucleotide and deduced amino acid sequences of mST 3Gal IV. The nucleotide and deduced amino acid sequences are numbered from the presumed start codon and initiation methionine, respectively. The *boxed* amino acid residues correspond to a putative hydrophobic transmembrane domain. Sialyl motif L and S are *underlined*. The asterisks indicate potential N-glycosylation sites (Asn-X-Ser/Thr). The N-terminal amino acid residue used to construct the fusion protein is indicated by the *arrow*. The nucleotide sequence of mST3Gal IV has been submitted to the GeneBank™/EMBL Data Bank with accession number X95809.

To determine the sizes of mST3Gal III and IV, and their tissue distribution, total RNAs were isolated from several mouse tissues and analyzed by Northern hybridization using the 1.0 kb and 0.9 kb cDNAs amplified by PCR for the construction of expression plasmids as the probes, respectively. As shown in Fig. 4, there appears to be a single mRNA band of approximately 2.0 kb for the mST3Gal III evident in heart, liver and adult brain, while expression of mST3Gal IV mRNA was detected in all tissues tested except for testis as a single band with the size of about 2.1 kb (Fig. 5). However, the level was the highest in liver, while very low levels were found in thymus and spleen.

Enzymatic characterization of mST3Gal III and IV produced as soluble forms

In order to verify that the cloned genes indeed repre-

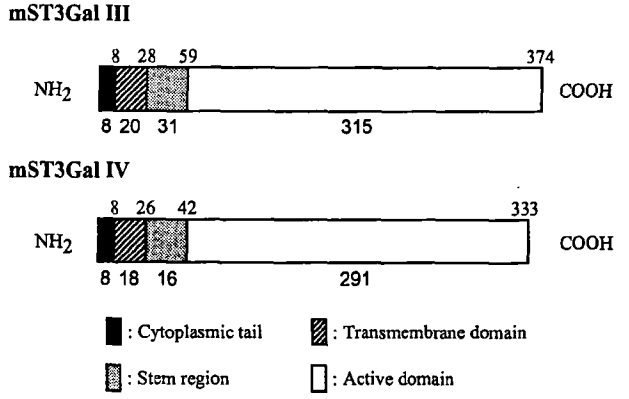

Fig. 3. Putative domain structures of mST3Gal III and IV. The numbers above and below the bars indicate the end position and number of amino acid residues of each domain, respectively.

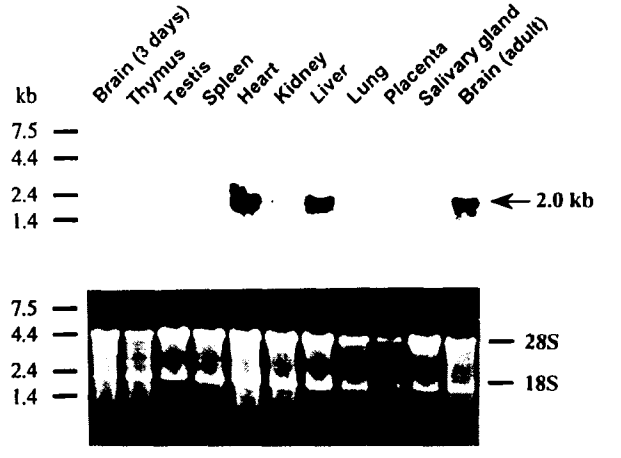

Fig. 4. Northern blot analysis of the mST3Gal III gene. Total RNAs (5 μg) were prepared from various mouse tissues and hybridized with a radiolabeled cDNA probe (1.0 kb) of mST 3Gal III (*upper panel*). *Lower panel* shows the formaldehyde-agarose gel picture stained with ethidium bromide.

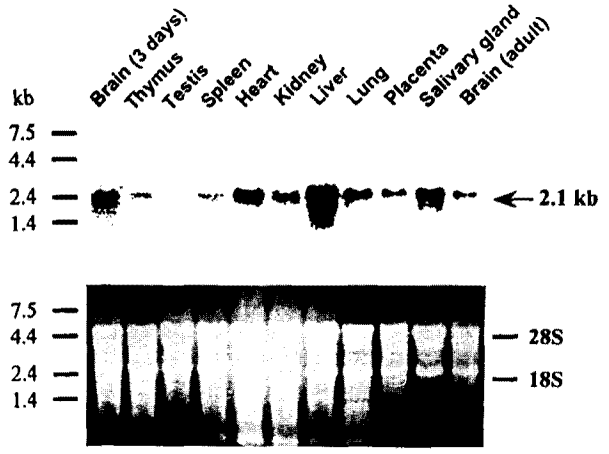

Fig. 5. Northern blot analysis of the mST3Gal IV gene. Total RNAs (5 μg) were prepared from various mouse tissues and hybridized with a radiolabeled cDNA probe (0.9 kb) of mST 3Gal IV (*upper panel*). *Lower panel* shows the formaldehyde-agarose gel picture stained with ethidium bromide.

sent mouse ST3Gal III and IV, we used the expression vector, pcDSA. By using this vector plasmid placement in-frame cDNA fragments with cloning sites just downstream from the COOH terminus of the IgG-binding domain of protein A enables the cDNA products to be secreted as active proteins fused with that domain, and therefore allows a fast and easy isolation of these secreted products from the culture medium by affinity chromatography on IgG-Sepharose (Lee *et al.*, 1996). When the expression plasmids, pcDSA-mST3Gal III and IV, were expressed in COS-7 cells, the enzymes were secreted from cells, adsorbed to IgG-Sepharose as fusion proteins and exhibited sialyltransferase activities. No radioactive incorporation was observed in the medium from the cells transfected with pcDSA vector alone (data not shown). The enzymatic properties of the rat liver ST3Gal III were previously found to utilize β-galactoside acceptor containing either the Galβ1,3GlcNAc or Galβ1,4GlcNAc terminal sequences as acceptor substrates (Wen *et al.*, 1992), while those of the human ST 3GalIV utilized β-galactoside acceptor containing either the Galβ1,3GalNAc or Galβ1,4GlcNAc terminal sequences as acceptor substrates (Kitagawa and Paulson, 1994). As shown in Table 1, the soluble recombinant mST3Gal III was capable of utilizing β-galactoside acceptor containing either the Galβ1,3GlcNAc or Galβ1, 4GlcNAc terminal sequences and exhibited the highest activity toward lacto-N-tetraose (Galβ1,3GlcNAcβ1, 3Galβ1,4Glc). However, although mST3Gal IV utilized acceptor substrates containing either the Galβ1, 3GlcNAc or Galβ1,4GlcNAc terminal sequences like mST3Gal III, the enzyme exhibited the highest activity toward lacto-N-neotetraose (Galβ1,4GlcNAcβ1,3Galβ1,

Table 1. Comparison of oligosaccharide acceptor specificity and of mST3Gal III and IV

	Relative rate (%)	
	mST3Gal III	mST3Gal IV
Galβ1,3GalNAc	11	14
Galβ1,3GlcNAc	58	43
Galβ1,4GlcNAc	34	63
Galβ1,3GlcNAcβ1,3Galβ1,4Glc (Lacto-*N*-tetraose)	100	32
Galβ1,4GlcNAcβ1,3Galβ1,4Glc (Lacto-*N*-neotetraose)	24	100

Relative rates for each enzyme are calculated as a percentage of the incorporation obtained with the best substrate.

4Glc), and the previously reported Galβ1,3GalNAc was not utilized as a good acceptor substrate. These results show that both enzymes preferentially utilize the tetrasaccharides over the disaccharides.

Discussion

Based on the amino acid sequences of the sialylmotif L of sialyltransferases cloned previously (Wen *et al.*, 1992; Kitagawa and Paulson, 1994), we have isolated and characterized full-length cDNAs encoding mouse Galβ1,3(4)GlcNAc α2,3-sialyltransferase (mST3Gal III) and Galβ1,4(3)GlcNAc α2,3-sialyltransferase (mST3Gal IV) from mouse brain cDNA library. The following results indicate that the two cloned cDNAs code for mST3Gal III and mST3Gal IV, respectively: (I) comparison of the deduced amino acid sequences of mST 3Gal III and IV with those of rat ST3Gal III (Wen *et al.*, 1992) and human ST3Gal IV (Kitagawa and Paulson, 1994) revealed 98% and 89%, respectively. However, there is no significant similarity between these enzymes and the other sialyltransferases cloned so far; (ii) compared with rat ST3Gal III and human ST3Gal IV, these enzymes are predicted to have a similar domain structure, and moreover, the number of amino acid residues of each domain is nearly consistent with that in the case of rat ST3Gal III and human ST3Gal IV, respectively; (iii) these enzymes expressed in COS cells revealed the sialyltransferase activity toward β-galactoside acceptor containing either the Galβ1,3GlcNAc or Galβ1,4GlcNAc terminal sequences as acceptor substrates. However, mST 3Gal IV showed a difference in acceptor substrate specificity when compared to human ST3Gal IV. Human ST 3Gal IV exhibited a much higher activity toward Galβ1, 3GalNAc than Galβ1,4GlcNAc, and no activity was detected toward Galβ1,3GlcNAc (Kitagawa and Paulson, 1994). However, mST3Gal IV cloned in this study exhibited strong activity toward Galβ1,4GlcNAc, followed

by Galβ1,3GlcNAc, and very low activity was detected toward Galβ1,3GalNAc (Table 1). This difference is at present difficult to explain, but one possibility may include a difference in the construction of expression plasmid. It is remarkable that mST3Gal III and IV utilize tetrasaccharide more preferentially than disaccharides as acceptor substrates.

Galβ1,3(4)GlcNAc α2,3- and Galβ1,4(3)GlcNAc α2,3-sialyltransferases are responsible for the formation of the NeuAcα2.3 Galβ1,3GlcNAc and NeuAcα2,3 Galβ1,4GlcNAc structures found to terminate complex type N-linked oligosaccharides of glycoproteins (Weinstein *et al.*, 1982). These enzymes involved in the biosynthesis of the sialyl Lewis x (sLe^x) (NeuAcα2,3 Galβ1,4(Fucα1,3) GlcNAc) determinant are of particular interest because the determinant has been reported to be a ligand for the three known selectins (E-, P-, and L-selectins) (Lowe *et al.*, 1990; Phillips *et al.*, 1990; Polley *et al.*, 1991; Foxall *et al.*, 1992; Lasky, 1992), which are cell adhesion molecules involved in the recruitment of leukocytes into lymphoid tissues and sites of inflammation, and cancer metastatis.

Sialyltransferases so far cloned have been shown to exhibit remarkable tissue-specific expression, which is correlated with the existence of cell type-specific carbohydrate structures (Paulson and Colley, 1989). Northern blotting indicated that the expression of mST 3Gal III mRNA with a size of about 2.0 kb is tissue-specific, being prominent in heart, liver and adult brain, while that of the other tissues was scarcely detected (Fig. 4). The pattern of mRNA expression for this enzyme is considerably different from those for rat ST 3Gal III, of which mRNA expression of about 2.5 kb was strong in kidney and brain, and low in heart and lung (Wen *et al.*, 1992). On the other hand, mRNA expression of mST3Gal IV with a size of about 2.1 kb was detected in all tissues tested except for testis, and the level was the highest in liver (Fig. 5). The pattern of mRNA expression for this enzyme is also considerably different from those for human ST3Gal IV, of which mRNA expression of approximately 2.0 kb was particularly abundant in placenta, followed by heart and kidney (Kitagawa and Paulson, 1994). These differences in expression pattern might be correlated with the existence of tissue- and cell type-specific carbohydrate structures as suggested in previous reports (Paulson *et al.*, 1989; Kitagawa and Paulson, 1994).

Acknowledgements

This work was supported by grants of G7 project (HS220M) and a special project of KRIBB (Matrix program; KM1081) from the Korean Ministry of Science and Technology.

References

Brandely, B. K., Swiedler, S. J. and Robbins, P. W. (1990) *Cell* **63**, 861.

Broquet, P., Baubichon-Cortay, H., George, P. and Louisot, P. (1991) *Int. J. Biochem.* **23**, 17735.

Foxall, C., Watson, S. R., Dowbenko, D., Fennie, C., Lasky, L. A., Kiso, M., Hasegawa, A., Asa, D. and Brandley, B. K. (1992) *J. Cell Biol.* **117**, 895.

Kimber, S. J. (1989) *Biochem. Soc. Trans.* **17**, 23.

Kitagawa, H. and Paulson, J. C. (1994) *J. Biol. Chem.* **269**, 17872.

Kojima, N., Lee, Y.-C., Hamamoto, T., Kurosawa, N. and Tsuji, S. (1994) *Biochemistry* **33**, 5772.

Kono, M., Yoshida, Y., Kojima, N. and Tsuji, S. (1996) *J. Biol. Chem.* **271**, 29366.

Kornfeld, R. and Kornfeld, S. (1985) *Annu. Rev. Biochem.* **54**, 631.

Lasky, L. A. (1992) *Science* **258**, 964.

Lee, Y.-C., Kim, C.-H. and Tsuji, S. (1996) *Mol. and Cells* **6**, 552.

Lee, Y.-C., Kojima, N., Wada, E., Kurosawa, N., Nakaoka, T., Hamamoto, T. and Tsuji, S. (1994) *J. Biol. Chem.* **269**, 10028.

Lee, Y.-C., Kurosawa, N., Hamamoto, T., Nakaoka, T. and Tsuji, S. (1993) *Eur. J. Biochem.* **216**, 377.

Lowe, J. B., Stoolman, L. M., Nair, R. P., Larsen, R. D., Berhend, T. L. and Marks, R. M. (1990) *Cell* **63**, 475.

Paulson, J. C. (1989) *Trends Biochem. Sci.* **14**, 272.

Paulson, J. C. and Colley, K. J. (1989) *J. Biol. Chem.* **264**, 17615.

Paulson, J. C., Weinstein, J. and Schauer, A. (1989) *J. Biol. Chem.* **264**, 10934.

Phillips, M. L., Nudelman, E., Gaeta, F. C. A., Perez, M., Singhal, A. K., Hakomori, S. and Paulson, J. C. (1990) *Science* **250**, 1130.

Polley, M. J., Phillips, M. L., Wayner, E., Nudelman, E., M., Singhal, A. K., Hakomori, S. and Paulson, J. C. (1991) *Proc. Natl. Acad. Sci. USA.* **88**, 6224.

Roos, E. (1984) *Biochim. Biophys. Acta* **738**, 263.

Sanger, F., Nicklen, S. and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463.

Tsuji, S. (1996) *J. Biochem.* **120**, 1.

Varki, A. (1992) *Curr. Opin. Cell Biol.* **4**, 257.

Weinstein, J., de Souza-e-Silva, U. and Paulson, J. C. (1982) *J. Biol. Chem.* **257**, 13845.

Wen, D. X., Livingston, B. D., Medzihradszky, K. F., Kelm, S., Burlingame, A. L. and Paulson, J. C. (1992) *J. Biol. Chem.* **267**, 21011.