*Glycoconjugate Journal* (1997) **14:** 407–411

# Murine hepatic β-galactoside α2,6-sialyltransferase gene expression involves usage of a novel upstream exon region

Yi Ping Hu*, Martin Dalziel and Joseph T.Y. Lau‡

*Department of Molecular and Cellular Biology, Roswell Park Cancer Institute, Elm and Carlton Streets, Buffalo, NY 14263, USA*

**ST6Gal I (β-galactoside α2,6-sialyltransferase, SiaT-1, ST6N, EC 2.4.99.1) mediates the attachment of the α2,6-sialyl linkage common on N-linked glycans. Previous work suggests substantial inter-species conservation in *SIAT1*, the gene encoding ST6Gal I. In human and in rat, hepatic-specific *SIAT1* transcription is initiated at Exon I. Here we report a surprising departure in the structural organization of the murine ST6Gal I gene. By a combination of primer extension analysis, 5′-RACE analysis, and analysis of genomic sequences, we show that the murine hepatic ST6Gal I mRNA contains a novel region 5′ of Exon I. This novel sequence is encoded on a discrete upstream exon, Exon H. In contrast to human and rat hepatic ST6Gal I, the murine mRNA is transcriptionally initiated at the start of Exon H. Differential mRNA blot analysis indicates that transcripts containing Exon H sequences are preferentially expressed in liver.**

***Keywords*: β-galactoside α2,6-sialyltransferase, ST6Gal I**

## Introduction

ST6Gal I (β-galactoside α2,6-sialyltransferase, SiaT-1, ST6N, EC 2.4.99.1) mediates the attachment of the α2,6-sialyl linkage common on N-linked glycans. Tissue differences in ST6 Gal I mRNAs result, at least to a large degree, from usage of a number of physically distinct transcriptional promoters contained within a single gene, *SIAT1* [1, 2]. In human and in rat, hepatic *SIAT1* transcription is initiated at the first nucleotide of Exon I. The 5′ flanking region to Exon I is competent as a promoter in the transcription of a heterologous gene [3–5]. Although Exon I encodes only the 5′-untranslated region, substantial sequence conservation exists between human and rat [2]. A cursory examination of the murine hepatic ST6Gal I mRNA sequence (MUSGAS) deposited in Genbank [6] revealed a divergence at the 5′-terminus from known human and rat Exon I sequence. Here we report that murine hepatic mRNA, in contrast to that previously documented in human and rat, is transcribed from a site distal to Exon I. Thus, the extreme 5′ region of murine hepatic ST6Gal I mRNA includes sequence contribution from an additional exon, Exon H, that precedes sequence contribution from Exon I.

## Materials and methods

### Construction and screening of the mouse genomic DNA library

High molecular weight ( > 100 kb) DNA was isolated from the liver of 129SvTer mouse. Partial MboI DNA digests were inserted into the *Bam*HI site of lambda DASH II vector and packaged using the Gigapack II packaging system from Stratagene (La Jolla). Clones containing sequences for Exons H, I, and II were identified and isolated using Probes 3, 2, and 1 (see below), respectively. Appropriate genomic regions were subcloned into plasmid vectors and analysed.

### 5′-RACE analysis

The 5′-Amplifinder RACE kit from Clontech was used. Fifty μg of total RNA from liver of 129SvTer mouse was annealed to the primer mST1-P1 (5′-GATGATGGCAAA CAGGAGAA-3′, also see Figure 2) and reverse transcribed. The resultant cDNA was ligated to the anchor primer as per instructions and subjected to 35 cycles of PCR amplification using the anchor primer and mST1-P4 (5′-TCTGGCTAA TCCTTCTGGGGC-3′, also see Figure 2) at 94 °C, 45 s; 60 °C, 45 s; and 72 °C, 60 s. An extension incubation of 72 °C

* Current address: Second Military Medical University, Department of Biology, Shanghai 200433, China.
‡ Corresponding author. Email: lau@sc3101.med.buffalo.edu; Fax: (716) 845-8389; Tel: (716)845-8914.

0282–0080 

for 10 min follows before termination of PCR reaction. The PCR product was cloned into the plasmid vector pCRII (Invitrogen, San Diego, CA).

### Northern blot analysis and generation of hybridization probes

A mouse multiple tissue Northern blot obtained from Clontech Laboratories that contains 2 μg of poly ($A^+$) mRNA per lane from each of eight different mouse tissues was used. Probes generated against different regions of *SIAT1* genomic and cDNA regions were used and are described below. After analysis with each probe, the blot was stripped by an incubation in 0.5% SDS at 95 °C for 10 min. Complete removal of the hybridized probe was assessed by overnight exposure of the stripped blot to a phosphor-storage screen (Molecular Dynamics). Hybridization probe fragments 1, 2 and 3 (see below) were generated by PCR amplification. Genomic DNA and cDNA were used. Hybridization was performed by the method of Church and Gilbert [7].

Probe 1, specific for Exon II, is a 425 bp fragment generated using the primer pair mST1-P5 and mST1-P6 (5′-CTCTCCTGCCCGTACCATTC-3′, 5′-TACGCTGAA CCTGACTCCTG-3′, respectively). Probe 2 for Exon I is a 127 bp fragment generated using the primer pair mST1-P3 and mST1-P4 5′-TGGTTTTTGATCATCCTGTG-3′, 5′-TCTGGCTAATCCTTCTGGGC-3′, respectively). Probe 3 for Exon H was a 144 bp fragment generated from a 5′-RACE clone containing Exon H using the primer pair mST1-P8 5′-TGAAGTGCCCAGAAGTC-3′) and the SP6 vector primer. These three fragments were internally labeled by five additional cycles of PCR in which the non-radioactive dATP was replaced with 50 μCi of $\alpha[^{32}P]$-dATP (3000 Cimmol$^{-1}$, Amersham). Probe 4 is a 317 bp *Hinf*I fragment isolated from mouse genomic DNA that corresponds to the region $^{-}$23 to approximately $^{-}$336 upstream from the mapped transcription start site.

### Primer extension analysis

Twenty μg of total SvTer mouse liver RNA was annealed to the complementary primer, mST1-P8 (see Figure 2) that has been end-labeled with $\gamma[^{32}P]$-ATP (3000 Ci mmol$^{-1}$, Amersham). Annealing was performed at 28 °C overnight in 30 μl of hybridization buffer (40 mM PIPES, pH 6.4, 1 mM EDTA, 0.4 M NaCl, and 80% formamide). After precipitation with ethanol, the samples were resuspended in 20 μl of reverse transcriptase buffer (50 mM Tris-HCl, pH7.6, 60 mM KCl, 10 mM $MgCl_2$, 1 mM each of dNTP, and 1 mM DTT). Reverse transcription was carried out with 200 U of Superscript II RNase H$^-$ reverse transcriptase (Gibco-BRL) in the presence of 50 U of RNAsin (Gibco-BRL) for 2 h at 37 °C. The reactions were terminated by addition of 1.0 μl RNase (10 mg ml$^{-1}$) and 1.0 μl of 0.5 M EDTA. After further incubation at 37 °C for 30 min, the products were phenol/chloroform extracted, precipitated with ethanol, and analysed on sequencing gel.

## Results and discussion

### 5′-RACE analysis of murine hepatic ST6Gal I mRNA

In order to elucidate the 5′ terminal sequence of the mouse liver ST6Gal I mRNA isoform, first strand cDNA was reverse transcribed from mouse liver mRNA using the specific primer mST1-P1 that is complementary to a coding region within Exon II (see diagram in Figure 2). A synthetic anchor primer was ligated 3′ to the first strand cDNA. ST6Gal I specific sequences were amplified by PCR using the primer pair, mST1-P4 and the anchor primer. As shown in Figure 1, the 5′-RACE product migrates as a predominately single band between 290–300 bp. This RACE product should represent the ST6Gal I mRNA sequence 5′ of the mST1-P4 complementary site. Twenty-five clones of the RACE product were obtained. Five clones were randomly selected and subjected to sequence analysis. These five clones contain the identical sequence and differ among each other only in the length of the 5′ extension. The positions of the 5′-most nucleotide in these five clones are denoted by double underline marks in Figure 2. The 5′ terminus of the sequence reported by Hamamoto *et al.* [6] is marked by an arrow (Figure 2).

### Genomic analysis of the murine ST6Gal I gene

A lambda mouse genomic library was constructed. Four clones were isolated on the basis of positive hybridization probes 1, 2, and 3 (see Materials and methods). One, λMGST1, contains the genomic region encompassing Exons I and II. Comparative sequence analysis of the exon regions in λMGST1 indicate complete conservation of exon boundaries between the murine, rat, and human genes. Two

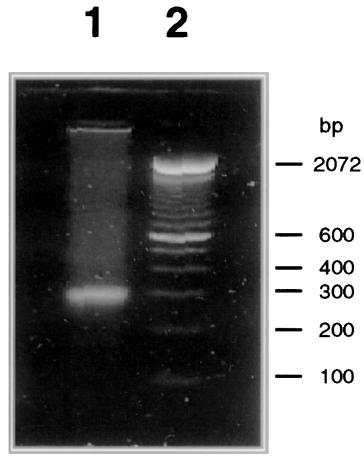


**Figure 1.** 5′-RACE analysis of murine hepatic ST6Gal I mRNA. First strand cDNA was reverse transcribed from mouse liver mRNA and subjected to RACE analysis as described in Materials and methods. Shown is the ethidium bromide stained agarose gel analysis of the RACE product migrating as a band of approximately 300 bps (lane 1). Molecular weight standards are shown in lane 2.

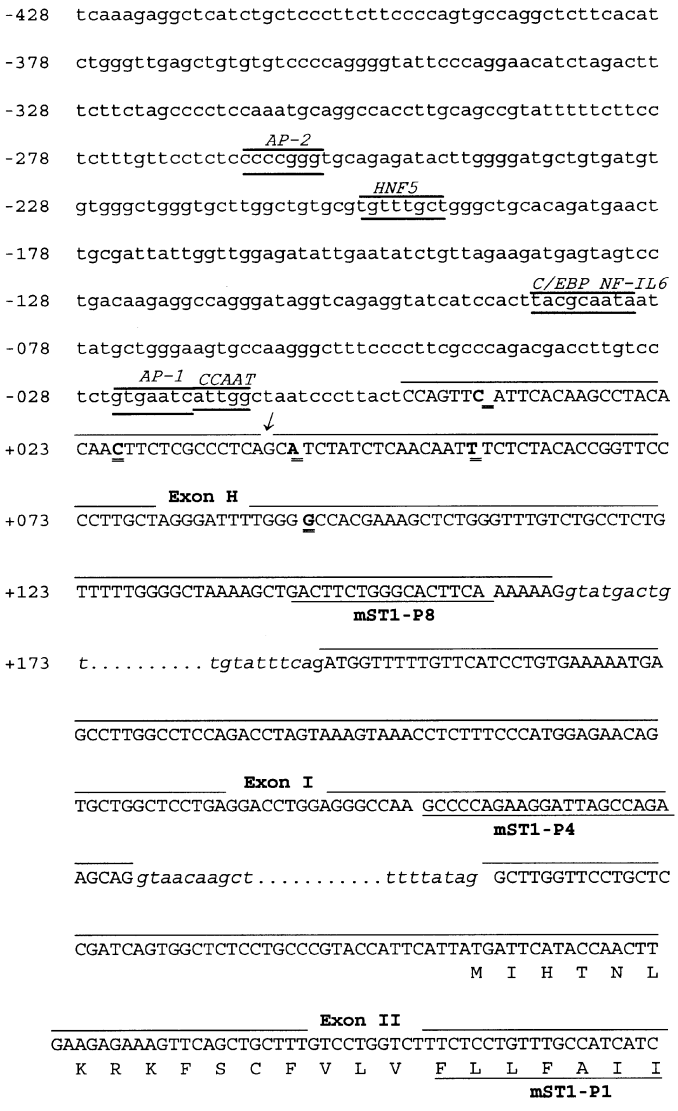


**Figure 2.** Sequence diagram of the exon regions encoding the 5′ terminus of murine hepatic ST6Gal I mRNA. Genomic sequence of Exons H, I, and II are shown along with partial genomic sequences that flank these exon regions. Sequence regions that comprise the transcribed exons are in capital letters; intron and flanking sequences are not capitalized. Oligonucleotide primers mST1-P8, P4, and P1 are denoted by the underlined regions. 5′-terminus of sequence MUSGAS reported by Hamamoto *et al.* [6] is marked by an arrow. Transcriptional start site as determined by primer extension analysis is denoted as Nucleotide Position No. 1. The 5′-termini of the five 5′-RACE clones are marked by double underlines. Consensus regions to transcription cis-elements in the 5′-flanking region are also shown.

other clones, λMGST1ExHa and λMGST1ExHb, do not contain Exons I and II, but they contain an exon region encoding the 5′-most sequence information predicted by the 5′-RACE products. Together, λMGST1ExHa and λMGST1ExHb encode 10.5 kb of genomic information in the 5′-flanking region of Exon H and approximately 3.4 kb on the 3′-flanking region of Exon H. Southern blot analysis using Exon H as probe resulted in a single band that indicates Exon H is a unique region in the murine genome (data not shown). A fourth clone, λMGST2, hybridized positively to Exon II specific sequences but not to Exon I or Exon H sequences. λMGST2 apparently carries a fortuitously cross hybridizing segment(s) to our probe since: i) after non exon II cutting restriction digests, λMGST2 displayed multiple positive bands (save for Not I) when hybridized with the exon II PCR probe, whereas λMGST1 gave only one positive signal when the same enzymes were used; ii) the absence of a *Pst*I restriction site within the Exon II positive fragments, as predicted by cDNA sequence data and also by λMGST1; iii) the complete difference in restriction pattern between inserts of λMGST1 and λMGST2; and iv) the inability of a number of synthetic oligonucleotide primers to murine Exon II to prime λMGST2 DNA in sequencing and PCR reactions (data not shown).

Figure 3 is a schematic of the murine ST6Gal I Exons H, I, and II region as identified on λMGST1 and λMGST1ExHa. A 3.3 kb intron separates Exon I and Exon II. At least 10–13 kb of sequence, and likely more, separate Exon H and Exon I; there is 3.4 kb of sequence on the 3′ flank of Exon H on λMGST1ExHa, and there is approximately 10.0 kb of sequence on the 5′ flank of Exon I on λMGST1. However, the precise distance separating Exon H and Exon I is not known since they have been isolated on separate clones. Probes against Exon H hybridized to λMGST1ExHa and λMGST1ExHb but not λMGST1 or λMGST2 (data not shown).

## Transcriptional initiation point of murine hepatic ST6Gal I mRNA isoform

Primer extension analysis was performed to pinpoint the precise transcription initiation point for liver ST6Gal I mRNA synthesis. As shown in Figure 4, extension of mouse liver mRNA using the primer mST1-P8 resulted in a single band that indicates transcription initiates 154 nts 5′ of the binding site for mST1-P8 (denoted as nucleotide position 1 in Figure 2). The result is fully consistent with the 5′-RACE data obtained in Figure 1. S1 nuclease analysis was also performed using a complementary probe spanning Exon H and 5′-flanking region. The result of the S1 protection assay is consistent with transcription initiation site as predicted by the primer extension analysis (data not shown).

The genomic sequence of the region flanking the 5′-end of the transcription start site is depicted in Figure 2. Notable is the absence of an obvious TATA box nor is it GC rich. As with a number of other genes sharing these features, accurate transcription may depend on transcription from an Inr or similar start site element (reviewed in [8]). Earlier studies in rat reported the glucocorticoid responsiveness in hepatic expression of ST6Gal I [4, 9, 10]. Glucocorticoid responsiveness in the murine ST6Gal I hepatic promoter may be exerted via the C/EBP consensus region at position ⁻89 and

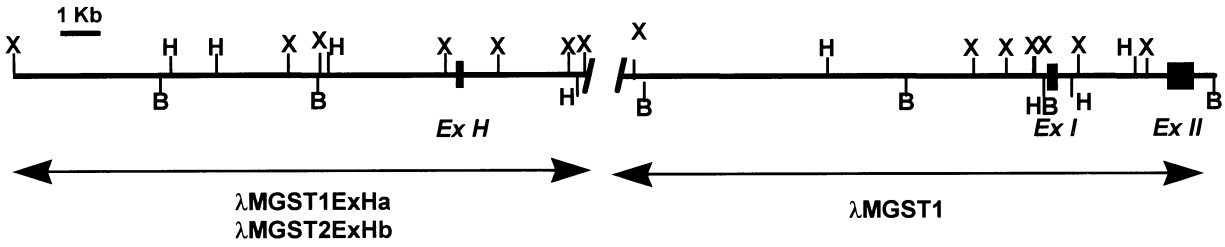


**Figure 3.** Restriction map and schematic of the murine ST6Gal I gene region encoding Exons H, I, and II. Shown is a partial restriction map of the genomic regions carried on lambda clones λMGST1EXHa, λMGST1EXHb, and λMGST1EX1. X, *Xba*I; H, *Hind*III, B, *Bam*HI. The size standard for 1 kb is shown in the upper left corner.

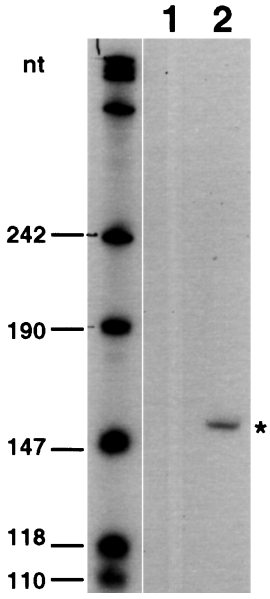


**Figure 4.** Primer extension analysis of murine hepatic ST6Gal I mRNA. Primer extension analysis was performed as described in Materials and methods. Shown are the autoradiogram results of the analysis as separated in a denaturing gel. Lane 1 shows the result of primer extension using 20 μg of tRNA as a negative control template. Lane 2 shows the result using 20 μg of total mouse liver RNA.

**Table I** Differential expression[a] of SIAT1 exonic regions among murine tissues

| *Tissue* | *Exon H* | *Exon I* | *Exon II* | *ExH/ExI* | *ExH/ExII* |
|---|---|---|---|---|---|
| Heart | 2.0 | 12.7 | 10.6 | 0.16 | 0.19 |
| Brain | 4.7 | 13.2 | 13.1 | 0.36 | 0.36 |
| Spleen | 1.9 | 5.4 | 6.3 | 0.36 | 0.30 |
| Lung | 3.6 | 19.7 | 20.2 | 0.18 | 0.18 |
| Liver | 100.0 | 100.0 | 100.0 | 1.00 | 1.00 |
| Muscle | 3.8 | 6.6 | 6.7 | 0.58 | 0.57 |
| Kidney | 1.3 | 9.8 | 11.0 | 0.14 | 0.12 |
| Testis | 1.4 | 9.3 | 10.9 | 0.15 | 0.13 |

[a]Expressed in normalized units with signal from liver set to 100 for each exonic region

the HNF-5 consensus region at position $^{-}$204 [11]. Other noteworthy potential cis-acting transcription elements include a CCAAT box at position $^{-}$18, AP-1 site at position $^{-}$25, NF-IL6 site at position $^{-}$83, and AP-2 site at position $^{-}$264.

## Differential Northern blot analysis of murine hepatic ST6Gal I mRNA

In human and rat, the expression of the hepatic form of ST6Gal I mRNA is restricted to liver and intestinal epithelial cells at certain stages of development [12]. Here we demonstrate that murine liver expresses an mRNA form divergent from that observed in human and rat liver. In order to determine that the Exon H-containing mRNA is the hepatic equivalent of the human and rat forms, RNA from a number of murine tissues was differentially hybridized with probes specific for Exon H, Exon I, and Exon II. As shown in Table I, Exon H signal is clearly highest in liver; negligible levels are seen in the other tissues examined. When the hybridization signals are expressed as a ratio between Exon H and either Exon I or Exon II, it is clear that Exon H-containing mRNA is most abundantly expressed in liver. Exon H-containing ST6Gal I mRNA also seems to be present in low but significant levels in muscle, spleen and brain, although the overall levels of ST6Gal I mRNA in these tissues are substantially lower than that observed in liver. Among the other tissues examined, lung and heart express significant levels of ST6Gal I mRNA. However, as evident by the ratio of ExH to ExI signals (Table 1), Exon H-containing mRNA constitutes a relatively minor component in the pool of ST6Gal I mRNAs in these tissues.

Thus, the data indicate that the predominant liver mRNA for ST6Gal I in mouse, in contrast to human and rat, contains an additional exon, Exon H, at the 5′ terminus. Two lines of observation strongly suggest that the Exon H-containing form is the murine equivalent of the

human/rat hepatic-specific ST6Gal I mRNA form. First, 5′-RACE analysis indicates that this is the predominant, if not the only, ST6Gal I mRNA expressed in murine liver. Second, differential blot analysis indicates that the Exon H-containing mRNA form is preferentially enriched in liver, with a substantially lower level of expression in muscle, spleen, and brain.

## Acknowledgements

This work is supported by Grant GM38193 from the National Institutes of Health to J.T.Y.L. and by institute Core Grant CA16056-21 to Roswell Park Cancer Institute.

## References

1 Wang XC, O'Hanlon TP, Young RF, Lau JTY (1990) *Glycobiology* **1**: 25–31.
2 Wang XC, Vertino-Bell A, Eddy RL, Byers MG, Jani-Sait SN, Shows TB, Lau JTY (1993) *J Biol Chem* **268**: 4355–61.
3 Lo NW, Lau JTY (1996) *Glycobiology* **6**: 271–79.
4 Wang XC, Smith TJ, Lau JTY (1990) *J Biol Chem* **265**: 17849–53.
5 Svensson EC, Conley PB, Paulson JC (1992) *J Biol Chem* **267**: 3466–72.
6 Hamamoto T, Kawasaki M, Kurosawa N, Nakaoka T, Lee YC, Tsuji S (1996) *Bioorg Med Chem* **1**: 141–45.
7 Church GC, Gilbert W (1984) *Proc Natl Acad Sci USA* **81**: 1991–95.
8 Smale ST, Baltimore D (1989) *Cell* **57**: 103–13.
9 van Dijk W, Boers W, Sala M, Lasthuis A, Mookerjea S (1986) *Biochem Cell Biol* **64**: 79–84.
10 Kaplan HA, Woloski BMRNJ, Hellman M, Jamieson JC (1983) *J Biol Chem* **258**: 11505–9.
11 Grange T, Roux J, Rigaud G, Pictet R (1990) *Nucleic Acids Res* **19**: 131–39.
12 Vertino-Bell A, Ren J, Black JD, Lau JTY (1994) *Devel Biol* **165**: 126–36.

Received 27 August 1996, revised 28 October 1996 and 8 November 1996