(2)

STRUCTURE AND EXPRESSION OF THE RAT *N*-ACETYLGLUCOSAMINYLTRANSFERASE I GENE

TAKASHI FUKADA

1997

ABBREVIATIONS

cDNA	complementary DNA
CRE	cyclic adenosine 3',5'-monophosphate responsive element
dATP	deoxyadenosine 5'-triphosphate
DNA	deoxyribonucleic acid
EDTA	ethylenediaminetetraacetic acid
ER	endoplasmic reticulum
Glc	glucose
GlcNAc	*N*-acetylglucosamine
GnT	*N*-acetylglucosaminyltransferase
kb	kilobase
kDa	kilodalton
Man	mannose
MMLV	Moloney murine leukemia virus
mRNA	messenger RNA
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PIPES	piperazine-N,N'-bis(2-ethanesulfonic acid)
RACE	rapid amplification of cDNA ends
RNA	ribonucleic acid
rRNA	ribosomal RNA
RT-PCR	reverse transcription PCR
SDS	sodium dodecyl sulfate
UDP	uridine 5'-diphosphate
uORF	upstream ORF
UTR	untranslated region

CONTENTS

INTRODUCTION

Most proteins presented at the external surface of mammalian cells have carbohydrate moieties covalently linked through *N*- or *O*-glycosidic bond. Carbohydrate chains on proteins have much variety and heterogeneity because they contain rich variation of the component sugars, such as mannose (Man), galactose, *N*-acetylglucosamine (GlcNAc) and sialic acid. Further, sugars can link to each other with several sterically different linkages. However, in spite of this variety and heterogeneity, carbohydrate chains are not the mixture of random structures but are constrained within a complex but limited structural framework.

Recent studies showed that certain proteins often acquire specific carbohydrate structures, those serve as recognition markers or modulate the function of the proteins in many biological steps (1). For example, certain proteins on myeloid or endothelial cells have sialylated, fucosylated, and sometimes sulfated glycans those serve as recognition motifs for selectin family of cell adhesion molecules in case of inflammation and leukocyte homing (2). A portion of neural cell adhesion molecule (N-CAM) has polysialic acid and this unique carbohydrate structure modulates the homotypic adhesion of N-CAM (3). Monoglucosylated N-glycans can serve as a marker of misfolded proteins in endoplasmic reticulum (ER) and are reported to concern the retention of those proteins in ER (4). N-glycans are also reported to have some role in the polarized protein transprort in epithelial cells (5).

The biosynthesis of protein glycans is performed by a series of highly specific glycosyltransferases and glycosidases without templates, unlike protein and DNA. The synthesis of the conserved core structure of *N*-linked glycans is a well-characterized example (6). The all three subtypes of protein *N*-glycans, high-mannose, hybrid, and complex glycans, share the same core structures (Fig. 1) and are consequences of a serial procession of a common precursor $Glc_3Man_9GlcNAc_2$ (Fig. 2). The precursor glycan is assembled and

transferred to protein in ER and processed to $Man_5GlcNAc_2$ in ER and Golgi apparatus. Certain glycans leave the processing pathway at this step, to produce high-mannose glycans. The next step of the processing is catalyzed by *N*-acetylglucosaminyltransferase I (GnT-I; EC 2.4.1.101), which transfers a GlcNAc moiety to the $Man_5GlcNAc_2$. Hybrid glycans leave the processing pathway at this step. Mannosidase II removes two mannose residues from the remaining glycans and this allows GnT-II and other glycosyltransferases to process the glycans further to produce complex glycans.

Since the preceding transfer of a GlcNAc residue by GnT-I is absolutely required for the mannosidase II and thus for the following processing steps, mutant cells lacking GnT-I activity synthesize only high-mannose glycans (7, 8). This suggests that GnT-I activity would modulate the amount of complex and hybrid glycans relative to high-mannose glycans. Some sugar residue, such as sialic acid, fucose, and galactose for example, appear

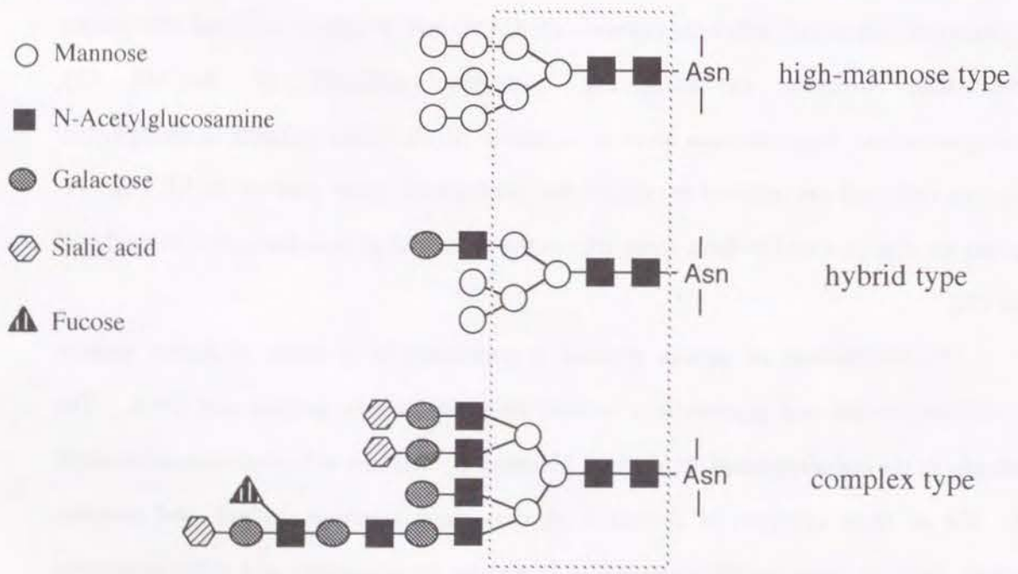


Fig. 1 Schematic diagram of three subtypes of N-glycans. A typical structure of each subtype is shown. The dotted box designates the shared core structure.

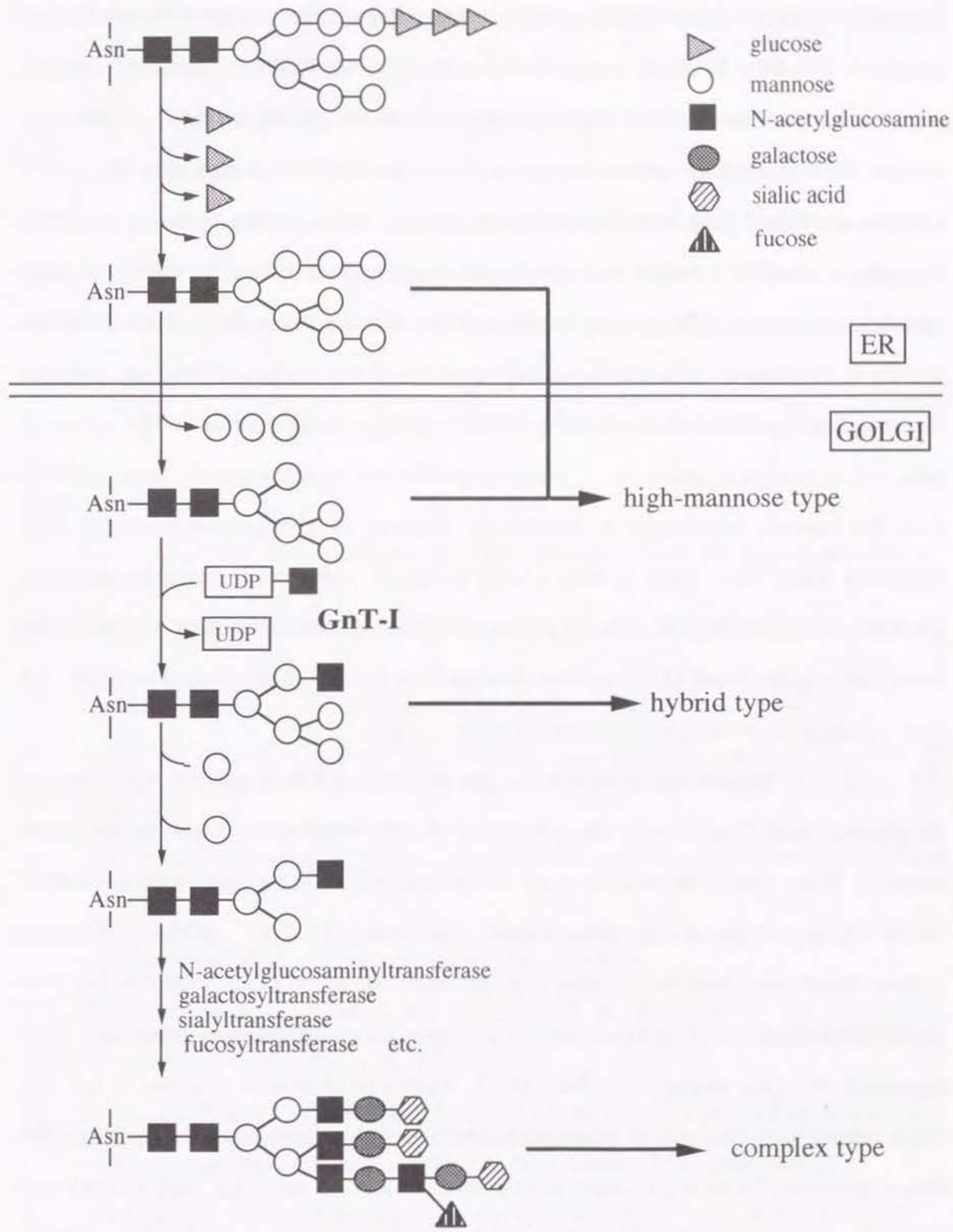


Fig. 2 Schematic diagram of the processing pathway of N-glycas. The precursor of 14 sugars attached to protein in ER and processed in ER and Golgi. See text.

only on complex and hybrid glycans and are included in biologically important carbohydrate structures like sialyl Lewis X and polysialic acid. On the contrary, terminally exposed oligomannose structures appear only on high-mannose and hybrid glycans. These facts suggest that GnT-I activity would regulate the glycan function through regulating the ratio of complex and hybrid glycans to high-mannose glycans. Although this regulation by GnT-I may not be specific, it would be a general mechanism which biases the effects of other specific modifications of *N*-glycans. Further, the fact that there are multiple lectins those are specific to mannose or galactose-containing structures in mammalian tissues may suggests the biological importance of the overall availability of these sugar moieties on the surface of cells and on biological molecules. Galactose-specific and mannose-specific lectins on liver cells, for example, are thought to concern the clearance of glycoprotein hormones from circulating blood (9). There is also a large family of widespread animal lectins called galectins, but their biological roles are still elusive (10). Another interesting example is the case of interleukin-1 and -2. They have lectin activity besides their cytokine activities and bind to high-mannose glycans on other molecules (11-15).

It is very important to determine the gene structure and the regulation mechanism of the glycosyltransferases because the activities of glycosyltransferases determine the glycan structure. Many glycosyltransferase were reported to express in a tissue-specific manner and/or to be regulated by physiological conditions (16-20). cDNAs of some glycosyltransferases have been cloned and regulated expression of their mRNA has been documented (21-27). Promoter structures of a few glycosyltransferase genes have been examined, too. (For example, see Ref. 28-32. More examples were reviewed in Ref. 33. Many other glycosyltransferase genes have been cloned, but most reports do not describe their expression. Those reports were also reviewed in Ref. 33 and 34.) GnT-I cDNA was cloned from human in 1990 (35) and later from rabbit (36) and mouse (37, 38). Tissue-specific expression of GnT-I mRNA has been reported for mouse (37-39). Partial structure of the GnT-I genomic gene reported for mouse (37) and human (40) showed that most part

of GnT-I mRNA including the entire coding region and 3'-untranslated region (UTR) was coded in the same single exon but the remaining 5'-distal region was not. Details of its gene regulation and the gene structure of the 5'-distal region of 5'-UTR are still unclear.

The author found that rat GnT-I mRNA had at least five forms of 5'-UTR and the difference in the 5'-UTR affected the translational efficiency of GnT-I mRNA. Tissue-specific usage of each 5'-UTR was found to significantly affect the GnT-I protein amount in rat liver and brain. The genomic structure of the rat GnT-I gene and its promoter structure are also discussed in this study.

References

1. Varki, A. (1993) "Biological roles of oligosaccharides: all of the theories are correct." *Glycobiology*, **3**, 97-130.
2. Varki, A. (1994) "Selectin ligands." *Proc. Natl. Acad. Sci. USA*, **91**, 7390-7397.
3. Acheson, A., Sunshine, J. L., and Rutishauser, U. (1991) "NCAM polysialic acid can regulate both cell-cell and cell-substrate interactions." *J. Cell Biol.*, **114**, 143-153.
4. Sousa, M., Parodi, J. (1995) "The molecular basis for the recognition of misfolded glycoproteins by the UDP-Glc : glycoprotein glucosyltransferase." *EMBO J.*, **14**, 4196-4203.
5. Scheiffele, P., Peranen, J., Simons, K. (1995) "*N*-glycans as apical sorting signals in epithelial cells." *Nature*, **378**, 96-98.
6. Kornfeld, R. and Kornfeld, S. (1985) "Assembly of asparagine-linked oligosaccharides." *Annu. Rev. Biochem.*, **54**, 631-664.
7. Robertson, M. A., Etchison, J. R., Robertson, J. S., Summers, D. F., and Stanley, P. (1978) "Specific changes in the oligosaccharide moieties of VSV grown in different lectin-resistant CHO cells." *Cell*, **13**, 515-526.
8. Li, E. and Kornfeld, S. (1978) "Structure of the altered oligosaccharide present in glycoproteins from a clone of chinese hamster ovary cells deficient in *N*-acetylglucosaminyltransferase activity." *J. Biol. Chem.*, **253**, 6426-6431.
9. Drickamer, K. (1991) "Clearing up glycoprotein hormones." *Cell*, **67**, 1029-1032.
10. Barondes, S. H., Cooper, D. N. W., Gitt, M. A., and Leffler, H. (1994) "Galectins — Structure and functions of a large family of animal lectins." *J. Biol. Chem.*, **269**, 20807-20810.

11. Brody, D. T. and Durum, S. K. (1989) "Membrane IL-1: IL-1α precursor binds to the plasma membrane via a lectin-like interactions." *J. Immunol.*, **143**, 1183-1187.

12. Sathyamoorthy, N., Decker, J. M., Sherblom, A. P., and Muchmore, A. (1991) "Evidence that specific high mannose structures directly regulate multiple cellular activities." *Mol. Cell. Biochem.*, **102**, 139-147.

13. Sherblom, A. P., Sathyamoorthy, N., Decker, J. M., and Muchmore, A. V. (1989) "IL-2, a lectin with specificity for high mannose glycopeptides." *J. Immunol.*, **143**, 939-944.

14. Muchmore, A., Decker, J., Shaw, A., and Wingfield, P. (1990) "Evidence that high mannose glycopeptides are able to functionally interact with recombinant tumor necrosis factor and recombinant interleukin 1." *Cancer Res.*, **50**, 6285-6290.

15. Muchmore, A. V., Sathyamoorthy, N., Decker, J., and Sherblom, A. P. (1990) "Evidence that specific high-mannose oligosaccharides can directly inhibit antigen-driven T-cell responses." *J. Leukocyte Biol.*, **48**, .

16. Narasimhan, S., Schachter, H., and Rajalakshmi, S. (1988) "Expression of *N*-acetylglucosaminyltransferase III in hepatic nodules during rat liver carcinogenesis promoted by orotic acid." *J. Biol. Chem.*, **263**, 1273-1281.

17. Biol, M.-C., Pintori, S., Mathian, B., and Louisot, P. (1991) "Dietary regulation of intestinal glycosyl-transferase activity: Relation between developmental changes and weaning in rats." *J. Nutr.*, **121**, 114-125.

18. Easton, E. W., Bolscher, J. G. M., and van den Eijnden, D. H. (1991) "Enzymatic amplification involving glycosyltransferase forms the basis for the increased size of asparagine-linked glycans at the surface of NIH 3T3 cells expressing the *N-ras* proto-oncogene." *J. Biol. Chem.*, **266**, 21674-21680.

19. Hahn, T. J. and Goochee, C. F. (1992) "Growth-associated glycosylation of transferrin secreted by HepG2 cells." *J. Biol. Chem.*, **267**, 23982-23987.

20. Lemaire, S., Derappe, C., Michalski, J.-C., Aubery, M., and Néel, D. (1994) "Expression of β1,6-branched *N*-linked oligosaccharides is associated with activation in human T4 and T8 cell populations." *J. Biol. Chem.*, **269**, 8069-8074.

21. Shaper, N. L., Shaper, J. H., Meuth, J. L., Fox, J. L., Chang, H., Kirsch, I. R., and Hollis, G. F. (1986) "Bovine galactosyltransferase: Identification of a clone by direct immunological screening of a cDNA library." *Proc. Natl. Acad. Sci. USA*, **83**, 1573-1577.

22. Wang, X., O'Hanlon, T. P., and Lau, J. T. Y. (1989) "Regulation of β-galactoside α2,6-sialyltransferase gene expression by dexamethasone." *J. Biol. Chem.*, **264**, 1854-1859.

23. Wen, D. X., Svensson, E. C., and Paulson, J. C. (1992) "Tissue-specific alternative splicing of the β-galactoside α2,6-sialyltransferase gene." *J. Biol. Chem.*, **267**, 2512-2518.

24. Shah, S., Lance, P., Smith, T. J., Berenson, C. S., Cohen, S. A., Horvath, P. J., Lau, J. T. Y., and Baumann, H. (1992) "*n*-butylate reduce the expression of β-galactoside α2,6-sialyltransferase in Hep G2 cells." *J. Biol. Chem.*, **267**, 10652-10658.

25. Gillespie, W., Kelm, S., and Paulson, J. C. (1992) "Cloning and expression of the Galβ1,3GalNAc α2,3-sialyltransferase." *J. Biol. Chem.*, **267**, 21004-21010.

26. Gillespie, W., Paulson, J. C., Kelm, S., Pang, M., and Baum, L. G. (1993) "Regulation of α2,3-sialyltransferase expression correlates with conversion of Peanut Agglutinin (PNA)$^+$ to PNA$^-$ phenotype in developing thymocytes." *J. Biol. Chem.*, **268**, 3801-3804.

27. Miyoshi, E., Nishikawa, A., Ihara, Y., Saito, H., Uozumi, N., Hayashi, N., Fusamoto, H., Kamada, T., and Taniguchi, N. (1995) "Transforming grouth factor β up-regulates expression of the *N*-acetylglucosaminyltransferase V gene in mouse melanoma cells." *J. Biol. Chem.*, **270**, 6216-6220.

28. Svensson, E. C., Conley, P. B., and Paulson, J. C. (1992) "Regulated expression of α2,6-sialyltransferase by the liver-enriched transcription factor HNF-1, DBP, and LAP." *J. Biol. Chem.*, **267**, 3466-3472.

29. Harduin-Lepers, A., Shaper, J. H., and Shaper, N. L. (1993) "Characterization of two *cis*-regulatory region in murine β1,4-galactosyltransferase gene." *J. Biol. Chem.*, **268**, 14348-14359.

30. Koyama, N., Miyoshi, E., Ihara, Y., Kang, R., Nishikawa, A., and Taniguchi, N. (1996) "Human *N*-acetylglucosaminyltransferase III gene is transcribed from multiple promoters." *Eur. J. Biochem.*, **238**, 853-861.

31. Kim, Y.-J., Park, J.-H., Kim, K.-S., Chang, J.-E., Ko, J. H., Kim, M.-H., Chung, D.-H., Chung, T.-W., Choe, I.-S., Lee, Y.-C., and Kim, C.-H. (1996) "Sequence analysis of the 5'-flanking region of the gene encoding human *N*-acetylglucosaminyltransferase III." *Gene*, **170**, 281-283.

32. Kang, R., Saito, H., Ihara, Y., Miyoshi, E., Koyama, N., Sheng, Y., and Taniguchi, N. (1996) "Transcriptional regulation of the *N*-acetylglucosaminyltransferase V gene in human bile duct carcinoma cells (HuCC-T1) is mediated by Ets-1." *J. Biol. Chem.*, **271**, 26706-26712.

33. Kleene, R. and Berger, G. (1993) "The molecular and cell biology of glycosyltransferases." *Biochim. Biophys. Acta*, **1154**, 283-325.

34. Field, M. C. and Wainwright, L. J. (1995) "Molecular cloning of eukaryotic glycoprotein and glycolipid glycosyltransferases: A survey." *Glycobiology*, **5**, 463-472.

35. Kumar, R., Yang, J., Larsen, R. D., and Stanley, P. (1990) "Cloning and expression of *N*-acetylglucosaminyltransferase I, the medial Golgi transferase that initiates complex *N*-linked cabohydrate formation." *Proc. Natl. Acad. Sci. USA*, **87**, 9948-9952.

36. Sarkar, M., Hull, E., Nishikawa, Y., Simpson, R. J., Moritz, R. L., Dunn, R., and Schachter, H. (1991) "Moleculer cloning and expression of cDNA encoding the

enzyme that controls conversion of high-mannose to hybrid and complex *N*-Glycans: UDP-*N*-acetylglucosamine : α-3-D-mannoside β-1,2-*N*-acetylglucosaminyltransferase I." *Proc. Natl. Acad. Sci. USA*, **88**, 234-238.

37. Pownall, S., Kozak, C. A., Schappert, K., Sarkar, M., Hull, E., Schachter, H., and Marth, J. D. (1992) "Molecular cloning and characterization of the mouse UDP-*N*-acetylglucosamine : α-3-D-mannoside β-1,2-*N*-acetylglucosaminyltransferase I gene." *Genomics*, **12**, 699-704.

38. Kumar, R., Yang, J., Eddy, R. L., Byers, M. G., Shows, T. B., and Stanley, P. (1992) "Cloning and expression of the murine gene and chromosomal location of the human gene encoding *N*-acetylglucosaminyltransferase I." *Glycobiology*, **2**, 383-393.

39. Yang, J., Bhaumik, M., Liu, Y., and Stanley, P. (1994) "Regulation of *N*-linked glycosylation. Neuronal cell-specific expression of a 5' extended transcript from the gene encoding *N*-acetylglucosaminyltransferase I." *Glycobiology*, **4**, 703-712.

40. Hull, E., Sarkar, M., Spruijt, M. P. N., Höppener, J. W. M., Dunn, R., and Schachter, H. (1991) "Organization and localization to chromosome 5 of the human UDP-*N*-acetylglucosamine : α-3-D-mannoside β-1,2-*N*-acetylglucosaminyltransferase I gene." *Biochem. Biophys. Res. Commun.*, **176**, 608-615.

Chapter 1. Cloning of a cDNA encoding GnT-I from rat liver and analysis of its expression in rat tissues

Introduction

N-glycosylation is a widely distributed protein modification in eucaryotes and is potentially important to many biological phenomena (1). GnT-I, a key enzyme of the complex *N*-glycan synthesis, initiates the processing pathway from the high mannose type to complex and hybrid type *N*-glycans by catalyzing the transfer of GlcNAc from UDP-GlcNAc to the oligomannosyl acceptor $Man_5GlcNAc_2Asn$ (2). Since its prior action is required for the subsequent processing steps catalyzed by enzymes such as GnT-II, III, and IV (3), mutant cells lacking GnT-I activity do not synthesize hybrid and complex *N*-glycans (4, 5).

In this study the author have cloned a GnT-I cDNA that contains the entire coding sequence from rat liver. The author found that rat GnT-I was highly homologous to the mouse, rabbit, and human enzymes in the nucleotide sequences and the deduced amino acid sequences. The author further estimated the mRNA expression in several rat tissues using this cDNA as the probe and found that the expression level of GnT-I mRNA varies greatly from tissue to tissue.

Materials and Methods

Amplification of a partial fragment of rat GnT-I cDNA by PCR.

Total RNA was isolated from rat liver by guanidium isothiocyanate method and poly(A)$^+$ RNA was purified by oligo(dT)-cellulose chromatography. First strand cDNA was synthesized from the poly(A)$^+$ RNA with an oligo(dT) primer and amplified by PCR

with primer H and B. Primer H (AGAAAGCTTGCTGTGCAGCCCGACC) corresponded to the rabbit GnT-I cDNA (6) nucleotides 557 to 581, containing a few base-changes to make a *Hin*dIII site. Primer B (AAGGGATCCAGTCATCCCAGAAGGC) was complementary to the rabbit nucleotides 930 to 911, also containing base-changes and additional nucleotides for a *Bam*HI site in its 5'-terminal region. The major product of this amplification was about 370 bp long, the size that was expected from the rabbit cDNA sequence. This fragment was cloned into pBluescript vector (Stratagene) and sequenced with 373A DNA sequencer (Applied Biosystems). It was proved to be 84% homologous to that of the rabbit GnT-I cDNA, suggesting that it was a part of rat GnT-I cDNA (data not shown).

Library construction and screening.

To isolate the full-length cDNA of rat GnT-I a rat liver cDNA library was constructed with λgt10 vector using cDNA Synthesis Plus and cDNA Cloning System λgt10 (Amersham).. Nine positive clones were obtained by the screening of 6 x 10^5 independent plaques using the cloned PCR fragment as a probe. The largest insert was 2.6 kb long and was considered to contain the entire coding region of GnT-I cDNA.

Northern hybridization.

Total RNA was isolated from rat tissues by the acid guanidium thiocyanate-phenol-chloroform method (7). This RNA (20 μg) was electrophoresed in formaldehyde-agarose gel and transferred to GeneScreen Plus (Du Pont). Hybridization was carried out at 60°C without formamide.

Results

Structure of rat GnT-I cDNA.

The rat GnT-I cDNA contained 157 bp of 5'-untranslated region, 1341 bp of coding region, and 1048 bp of 3'-untranslated region except for the poly(A) tail (Fig. 1). There was a typical polyadenylation signal (AATAAA) 20 bp upstream from the beginning of the tail. The entire nucleotide sequence was 92%, 75%, and 77% homologous to those of mouse (8), rabbit (6), and human (9), respectively.

The rat GnT-I cDNA encoded a protein of molecular weight 51.6 kDa containing 447 amino acids. The protein had no potential *N*-glycosylation site. A hydropathy plot analysis (not shown) suggested that there was a single transmembrane region near the *N*-terminus of this protein (Fig. 1). The deduced amino acid sequence of rat GnT-I was found to be 97%, 92%, and 91% homologous to those of mouse (8), rabbit (6), and human GnT-I (9), respectively (Fig. 2). This cDNA showed GnT-I activity when recombinantly expressed in E. coli after removal of transmembrane region (10).

Expression of GnT-I mRNA in rat tissues.

Two GnT-I transcripts, 3.0 kb and 3.3 kb, were detected by northern hybridization using the entire coding region of the cDNA as a probe in most rat tissues examined (Fig. 3). High expression of GnT-I mRNA was observed in spleen, lung, and liver. Brain also had moderately high expression, but colon, skin, and muscle were only weakly positive. Brain essentially expressed only the 3.3 kb transcript, while live almost exclusively expressed the 3.0 kb one. This mutually exclusive expression of two GnT-I transcript in brain and liver was also reported in mice (8).

```
CTCCAGGGAAGCAACAGAGAGAAGAAAAGAAAGGTACATGGCTACTTCCCACTCTCAGCCTAGAAGAATAAACCCTTAGATTGCGGGGGAACTGAGCCAG   100

GCAAGCCACAGGCAGCCTTGAGCGCCCCCTTGCCTGCCCTCCCCTGCGGGGGCCAGGATGCTGAAGAAGCAGTCTGCAGGGCTTGTGCTTTGGGGTGCTA   200
                                                  M  L  K  K  Q  S  A  G  L  V  L  W  G  A  I

TCATCTTTGTGGGCTGGAATGCCCTGCTGCTCCTCTTCTTCTGGACACGCCCAGCACCTGGCAGGCTGCCCTCAGACAGTGCCCTTGGTGATGACCCTGC   300
 I  F  V  G  W  N  A  L  L  L  L  F  F  W  T  R  P  A  P  G  R  L  P  S  D  S  A  L  G  D  D  P  A

CAGCCTCACCCGTGAGGTCATCCACCTGGCCGAGGACGCCGAGGCGGAGTTGGAGCGGCAGCGGGGACTACTGCAGCAGATCAAGGAGCATTATTCTCTG   400
 S  L  T  R  E  V  I  H  L  A  E  D  A  E  A  E  L  E  R  Q  R  G  L  L  Q  Q  I  K  E  H  Y  S  L

TGGAGGCAGAGGTGGAGAGTTCCCACCGTGGCCCCTCCAGCCTGGCCCCGTGTGCCTGGGACCCCCTCACCAGCTGTGATCCCCATTCTGGTCATTGCCT   500
W  R  Q  R  W  R  V  P  T  V  A  P  P  A  W  P  R  V  P  G  T  P  S  P  A  V  I  P  I  L  V  I  A  C

GTGACCGCAGCACTGTCCGGCGCTGCCTGGATAAGTTGTTGCACTATCGGCCCTCTGCTGAGCATTTCCCCATTATTGTCAGTCAAGACTGTGGGCATGA   600
 D  R  S  T  V  R  R  C  L  D  K  L  L  H  Y  R  P  S  A  E  H  F  P  I  I  V  S  Q  D  C  G  H  E

AGAGACAGCACAGGTCATTGCTTCTTATGGCACCGCTGTCACACACATCCGGCAGCCAGACCTGAGTAACATTGCCGTGCAGCCAGACCACCGTAAGTTC   700
 E  T  A  Q  V  I  A  S  Y  G  T  A  V  T  H  I  R  Q  P  D  L  S  N  I  A  V  Q  P  D  H  R  K  F

CAGGGTTACTACAAGATTGCCAGGCACTACCGCTGGGCGCTAGGCCAGATCTTCAACAAGTTCAAGTTCCCGGCTGCTGTGGTAGTGGAGGATGATCTGG   800
Q  G  Y  Y  K  I  A  R  H  Y  R  W  A  L  G  Q  I  F  N  K  F  K  F  P  A  A  V  V  V  E  D  D  L  E

AAGTGGCGCCGGACTTCTTCGAGTACTTCCAGGCCACCTACCCGCTGCTGAAAGCAGACCCCTCCCTTTGGTGTGTGTCGGCTTGGAATGATAATGGTAA   900
 V  A  P  D  F  F  E  Y  F  Q  A  T  Y  P  L  L  K  A  D  P  S  L  W  C  V  S  A  W  N  D  N  G  K

GGAGCAGATGGTGGACTCCAGCAAACCTGAGCTGCTCTATCGAACAGACTTTTTTCCTGGCCTTGGATGGCTGCTGTTGGCTGATCTCTGGGCAGAACTA  1000
 E  Q  M  V  D  S  S  K  P  E  L  L  Y  R  T  D  F  F  P  G  L  G  W  L  L  L  A  D  L  W  A  E  L

GAGCCCAAGTGGCCCAAGGCCTTCTGGGACGACTGGATGCGCAGACCTGAGCAGCGGAAGGGGCGGGCTTGTATTCGTCCAGAAATTTCAAGAACTATGA  1100
E  P  K  W  P  K  A  F  W  D  D  W  M  R  R  P  E  Q  R  K  G  R  A  C  I  R  P  E  I  S  R  T  M  T

CATTTGGTCGCAAGGGTGTGAGCCATGGGCAGTTCTTTGACCAGCATCTTAAATTCATCAAGCTGAACCAGCAGTTCGTGCCCTTCACCCAGTTGGACCT  1200
 F  G  R  K  G  V  S  H  G  Q  F  F  D  Q  H  L  K  F  I  K  L  N  Q  Q  F  V  P  F  T  Q  L  D  L

GTCGTACCTGCAGCGGGAGGCCTATGACCGGGACTTCCTTGCCCAGGTCTATGGTGCCCCCCAGCTGCAGGTGGAGAAAGTGAGGACCAATGATCGGAAG  1300
 S  Y  L  Q  R  E  A  Y  D  R  D  F  L  A  Q  V  Y  G  A  P  Q  L  Q  V  E  K  V  R  T  N  D  R  K

GAACTGGGGGAGGTGCGGGTACAGTACACTAGCAGAGACAGCTTTAAGGCCTTTGCTAAGGCCCTGGGTGTCATGGACGACCTCAAGTCTGGTGTCCCCA  1400
E  L  G  E  V  R  V  Q  Y  T  S  R  D  S  F  K  A  F  A  K  A  L  G  V  M  D  D  L  K  S  G  V  P  R

GAGCTGGCTACCGTGGCATTGTCACTTTCCAGTTCCGGGGTCGGCGTGTCCACCTGGCACCCCCAGAGACATGGAATGGCTATGATCCTAGCTGGAATTA  1500
  A  G  Y  R  G  I  V  T  F  Q  F  R  G  R  R  V  H  L  A  P  P  E  T  W  N  G  Y  D  P  S  W  N

GTGGCACCTGCCTTTCCCTCCTGGGTCTCCTTGCCGTATGATGAGCCGAGGCAGCCTGCAGGCCCTGGGCTATACCATTCTGCCAGTGTTTCCTCTTGGA  1600

TCTATAGATCTTCCCTTTTTGTCCTGGCCTTCCCCCAAGTGGCATTCTAGTGCACAAATCATAAGATGAGGGTTATACTCCTCTTGTCAAGGGAGTACTG  1700

TGTGGTATGTTCGGGGCATATTGAACAAGAAACCACTGTGTGGTGTGGGGAGGGTTGGACCTGTTGGACCAGGCTTGTAGTGTCCTGAGTTCTCTCCAAG  1800

GGCATCTGCGGAGAGCTTGGCCACTCCAGCTCTCCTGACCAGGCCTCTCCACCCTGATCTGGCTCCTGTTAGTACATGAGCCCACTTTATGTACCTTCCC  1900

CACTTTCTCCACCTTCCCCAGGGTGGGGCCGGATTATGGAAGAAGGAATAGATTTGTGGCCAACTGAGACTAACCAAAGGGATTCACTGTCAGAGTTAGA  2000

TTGCATGGCCGGGTTGCCGGGGTCACTGCCTCCTGGCTTCTCTTCCTGCCGACTTTGCCAGTGCTCTTCTTGTCCTGCAGTGGAGCAGTGTGTTGTGAGA  2100

TGGAAAGTAGGAGCTAAAGCAAGACCTCTCCCGTGGTGGAGCAGTTGTCAGGAACAATGACAGGCGGAAAGCCTACACTTGGGCACCCCTCTTCTTGCCC  2200

ATGGCCCCCCACTTCTGAAAGTTAGTTCCCTTTGCATTCTTTAGGGGGATTCAGTGGCAGCCTAGTGGGGGACAGTGGGCTCCACGCCTCTTTCCCCTGG  2300

GGTTGGTACAGCCTCCTCCCACAGGGCTTGTTCCTGGTGGCCATCTCAGCAAAGTCACCAGCAGCTGAGACCAGGACACCTGTGGCCCTTCCTAAGCATT  2400

CCCACATCTTCCCAGGCCGAGGGGCAGGTGGCAGGCTAGAGAGAAAGAAATTTGTGTGTTTTGTTTGTTGCCTGATCTTAGTTTCATGGAAGAAAATGGA  2500

ATCTACAGAATTATTTTCAAAAATAAAGTAAAGGCTGAATTGTCTGAAAAAAAAAAA                                              2557
```

Fig. 1 Nucleotide sequence of rat GnT-I cDNA. Deduced amino acid sequence is also shown. The polyadenylation signal (AATAAA) is underlined and the putative transmembrane region is doubly underlined.


```
rat     MLKKQSAGLV LWGAIIFVGW NALLLLFFWT RPAPGRLPSD SALGDDPASL     50
mouse   .....T.... .......... .......... .......... ..........     50
rabbit  .......... .....L..A. .......... ..V.S..... N..D......     50
human   .......... .....L..A. .......... ......P..V ...DG.....     50

rat     TREVIHLAED AEAELERQRG LLQQIKEHYS LWRQRWRVPT VAPPAWPRVP     100
mouse   .......... .......... .........A .......... ..........     100
rabbit  .....R..Q. ..V....... .....R..HA ..S...K... A....Q.H..     100
human   .....R..Q. ..V....... .....GD--A .SS..G.... A....Q....      98

rat     GTPSPAVIPI LVIACDRSTV RRCLDKLLHY RPSAEHFPII VSQDCGHEET     150
mouse   V....VQ... .......... .......... .....R.... ..........     150
rabbit  V..P...... .......... .......... .....L.... ..........     150
human   V..A...... .......... .......... .....L.... ..........     148

rat     AQVIASYGTA VTHIRQPDLS NIAVQPDHRK FQGYYKIARH YRWALGQIFN     200
mouse   .......... .......... .......... .......... ..........     200
rabbit  ........S. .......... .......... .......... .........H     200
human   ..A.....S. .......... S...P..... .......... .......V.R     198

rat     KFKFPAAVVV EDDLEVAPDF FEYFQATYPL LKADPSLWCV SAWNDNGKEQ     250
mouse   .......... .......... .......... .RT....... ..........     250
rabbit  N.NY...... .......... .......... .......... ..........     250
human   Q.R....... .......... ....R..... .......... ..........     248

rat     MVDSSKPELL YRTDFFPGLG WLLLADLWAE LEPKWPKAFW DDWMRRPEQR     300
mouse   .......... .......... .......... .......... ..........     300
rabbit  .......... .......... .....E.... .......... ..........     300
human   ...A.R.... .......... .....E.... .......... ..........     298

rat     KGRACIRPEI SRTMTFGRKG VSHGQFFDQH LKFIKLNQQF VPFTQLDLSY     350
mouse   .......... .......... .......... .......... ..........     350
rabbit  .....V.... .......... .......... .......... ..........     350
human   Q......... .......... .......... .......... .H........     348

rat     LQREAYDRDF LAQVYGAPQL QVEKVRTNDR KELGEVRVQY TSRDSFKAFA     400
mouse   ..Q....... .......... .........Q .......... ..........     400
rabbit  ..Q....... ..R....... .......... .......... .G........     400
human   .......... ..R....... .......... .......... .G........     398

rat     KALGVMDDLK SGVPRAGYRG IVTFQFRGRR VHLAPPETWN GYDPSWN       447
mouse   .......... .......... .......... ......Q..T .......       447
rabbit  .......... .......... ....L..... ......Q..D ......T       447
human   .......... .......... .......... ......L..E .......       445
```

Fig. 2 Comparison of amino acid sequences of GnT-I. A dot indicates that the corresponding amino acid is identical to that of rat GnT-I. The two hyphens designate the deletion of the corresponding amino acids in human GnT-I.

Discussion

In this study the author cloned a full-length GnT-I cDNA from rat liver and examined the expression of GnT-I mRNA in rat tissues. The cDNA was highly homologous to rabbit, mouse, and human GnT-I cDNA, implying that rat GnT-I also has the type II membrane protein topology, namely a short cytoplasmic tail, a membrane anchor domain, and a large lumenal domain which contains the catalytic domain. This idea was also supported by a result that the putative lumenal domain of rat GnT-I was enough to express the GnT-I activity in *E. coli* (10).

GnT-I has been reported to be retained in medial Golgi by a signal (or signals) in the transmembrane domain (11, 12). Comparison of amino acids in this domain between rat GnT-I and GnT-III (13), which is also retained in medial Golgi, showed that a short peptide sequence, Ala-Gly-Leu-Val/Cys-Leu, was common. However, the significance of this sequence is unclear since these amino acids are unlikely to have conspicuous features

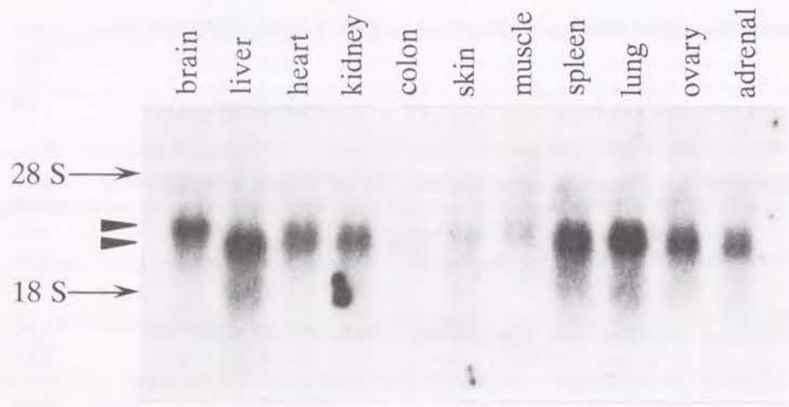


Fig. 3 Expression of GnT-I mRNA in various rat tissues. Total RNA (20 μg) was hybridized with the coding region of the GnT-I cDNA. The two arrowheads indicate the 3.3 kb and 3.0 kb GnT-I mRNA. The arrows show the positions of rRNAs.

and it is not known whether the retention in Golgi is mediated by receptor(s) for specific peptide sequence(s) or by other mechanisms, such as self-aggregation of the proteins.

High expression of GnT-I mRNA was observed in spleen and lung in rat. This expression pattern was roughly consistent with a reported result for mice (14). However, there is another report describing different expression pattern in mice (8), in which a high expression was observed in kidney and the expression in spleen was very low. This contrast may suggest the expression levels of some glycosyltransferases are dynamically changing even in normal tissues *in vivo*. Another observation which may support this idea is that the developmental change of rat intestinal fucosyltransferase activity during weaning can be affected by hormones (15) and dietary conditions (16). The transcription of the α2,6-sialyltransferase gene is also reported to be modulated by glucocorticoids in hepatoma cells and primary hepatocyte cultures (17). However, the biological significance of these phenomena remains to be elucidated.

To the contrast, moderately high expression of GnT-I mRNA in brain and liver was consistent among the author's result for rat and the two reports for mice (8, 14). It may suggest that GnT-I function should be important for the biological integrity of brain and liver.

It is intrigue that two differently sized GnT-I mRNA observed in rat tissues and that brain and liver had an unique pattern of expression of these transcripts. It is also the case for mouse (8, 14). It suggests the expression of GnT-I may have important consequences in these tissues and may be specifically regulated. These two transcripts may be produced by alternative usage of transcriptional start or stop sites, or by alternative splicing. Although obvious exon-intron structure has not been observed at least in the coding region in the mouse (8), splicing may occur in other regions. Analysis of the genomic structure of untranslated regions would be necessary.

References

1. Rademacher, T. W., Parekh, R. B., and Dwek, R. A. (1988) "Glycobiology." *Ann. Rev. Biochem.*, **57**, 785-833.

2. Kornfeld, R. and Kornfeld, S. (1985) "Assembly of asparagine-linked oligosaccharides." *Annu. Rev. Biochem.*, **54**, 631-664.

3. Schachter, H. (1986) "Biosynthetic controls that determine the branching and microheterogeneity of protein-bound oligosaccharides." *Biochem. Cell Biol.*, **64**, 163-181.

4. Robertson, M. A., Etchison, J. R., Robertson, J. S., Summers, D. F., and Stanley, P. (1978) "Specific changes in the oligosaccharide moieties of VSV grown in different lectin-resistant CHO cells." *Cell*, **13**, 515-526.

5. Li, E. and Kornfeld, S. (1978) "Structure of the altered oligosaccharide present in glycoproteins from a clone of chinese hamster ovary cells deficient in *N*-acetylglucosaminyltransferase activity." *J. Biol. Chem.*, **253**, 6426-6431.

6. Sarkar, M., Hull, E., Nishikawa, Y., Simpson, R. J., Moritz, R. L., Dunn, R., and Schachter, H. (1991) "Moleculer cloning and expression of cDNA encoding the enzyme that controls conversion of high-mannose to hybrid and complex *N*-Glycans: UDP-*N*-acetylglucosamine : α-3-D-mannoside β-1,2-*N*-acetylglucosaminyl-transferase I." *Proc. Natl. Acad. Sci. USA*, **88**, 234-238.

7. Chomczynski, P. and Sacchi, N. (1987) "Single-step method of RNA isolation by acid guanidium thiocyanate-phenol-chloroform extraction." *Anal. Biochem.*, **162**, 156-159.

8. Pownall, S., Kozak, C. A., Schappert, K., Sarkar, M., Hull, E., Schachter, H., and Marth, J. D. (1992) "Molecular cloning and characterization of the mouse UDP-*N*-acetylglucosamine : α-3-D-mannoside β-1,2-*N*-acetylglucosaminyltransferase I gene." *Genomics*, **12**, 699-704.

9. Kumar, R., Yang, J., Larsen, R. D., and Stanley, P. (1990) "Cloning and expression of *N*-acetylglucosaminyltransferase I, the medial Golgi transferase that initiates complex *N*-linked cabohydrate formation." *Proc. Natl. Acad. Sci. USA*, **87**, 9948-9952.

10. Nishiu, J., Kioka, N., Fukada, T., Sakai, H., and Komano, T. (1995) "Characterization of rat *N*-acetylglucosaminyltransferase I expressed in *Escherichia coli*." *Biosci. Biotech. Biochem.*, **59**, 1750-1752.

11. Tang, B. L., Wong, S. H., Low, S. H., and Hong, W. (1992) "The transmembrane domain of *N*-glucosaminyltransferase I contains a golgi retention signal." *J. Biol. Chem.*, **267**, 10122-10126.

12. Burke, J., Pettitt, J. M., Schachter, H., Sarkar, M., and Gleeson, P. A. (1992) "The transmembrane and flanking sequences of β1,2-*N*-acetylglucosaminyltransferase I specify *medial*-Golgi localization." *J. Biol. Chem.*, **267**, 24433-24440.

13. Nishikawa, A., Ihara, Y., Hatakeyama, M., Kangawa, K., and Taniguchi, N. (1992) "Purification, cDNA cloning, and expression of UDP-*N*-acetylglucosanime : β-D-mannoside β-1,4-*N*-acetylglucosaminyltransferase III from rat kidney." *J. Biol. Chem.*, **267**, 18199-18204.

14. Kumar, R., Yang, J., Eddy, R. L., Byers, M. G., Shows, T. B., and Stanley, P. (1992) "Cloning and expression of the murine gene and chromosomal location of the human gene encoding *N*-acetylglucosaminyltransferase I." *Glycobiology*, **2**, 383-393.

15. Chu, S. H. W. and Walker, W. A. (1986) "Developmental changes in the activity of sialyl- and fucosyl-transferases in rat small intestine." *Biochim. Biophys. Acta*, **883**, 496-500.

16. Biol, M.-C., Pintori, S., Mathian, B., and Louisot, P. (1991) "Dietary regulation of intestinal glycosyl-transferase activity: Relation between developmental changes and weaning in rats." *J. Nutr.*, **121**, 114-125.

17. Wang, X., O'Hanlon, T. P., and Lau, J. T. Y. (1989) "Regulation of β-galactoside α2,6-sialyltransferase gene expression by dexamethasone." *J. Biol. Chem.*, **264**, 1854-1859.

Chapter 2. Diversity in the 5'-untranslated region of rat GnT-I mRNA and its implications for the regulation of the gene expression.

Introduction

Among the protein *N*-glycosylation enzymes, GnT-I, which transfers a GlcNAc moiety from UDP-GlcNAc to the oligomannosyl acceptor $Man_5GlcNAc_2$, initiates the conversion of high-mannose *N*-glycans to hybrid and complex type structures (1). This conversion is important to many biological processes since GnT-I-deficient transgenic mouse dies *in utero* with pleiotropic defects (2-4). Furthermore, regulated expression of GnT-I can control the ratio of complex and hybrid *N*-glycans to high-mannose ones, and thus can regulate many biological interactions. This idea would be supported by the fact that there are many kinds of lectins those are specific to mannose- or galatose-containing glycans in mammalian tissues, although the roles of many animal lectins are still elusive (5, 6).

The author had cloned a GnT-I cDNA from rat liver and found that each rat tissue contained various amount of GnT-I mRNA (Chapter 1) and found that there were two transcripts of GnT-I different in size. In brain the larger transcript was expressed essentially as a sole GnT-I mRNA, while the shorter one was almost the only GnT-I mRNA expressed in liver. This tissue-specific expression of two different transcripts has been also observed in mouse brain and liver (7, 8), suggesting that these transcripts may play important roles in the regulation of the production of GnT-I enzyme. The mechanism of this regulation, however, has not been understood.

The author now report that the size difference of those transcripts resulted from the variation of the length of the 5'-untranslated region (UTR), probably due to alternative usage of transcriptional start sites. The author also found that 5'-UTR of GnT-I mRNA

regulates the translational efficiency of the transcripts, suggesting the GnT-I protein expression is regulated at multiple steps.

Materials and Methods

Library construction and screening.

Rat brain RNA was isolated by the acid guanidium thiocyanate-phenol-chroloform method (9) and poly(A)$^+$ RNA was purified with mRNA Purification Kit (Pharmacia). A cDNA library was constructed as described in chapter 1. The author screened 3 x 10^6 of plaques with a probe covering the entire coding region of rat liver GnT-I cDNA to obtain 21 positive clones. Nucleotide sequences were determined for 18 of them.

5'-rapid amplification of cDNA ends (5'-RACE).

5'-distal region of GnT-I cDNA was amplified from rat brain mRNA with 5'-AmpliFINDER RACE Kit (Clontech) following the manufacturer's instructions. Primer P1 (TCATCACCAAGGGCACTGTCTGAGG), which was complementary to the nucleotides 294 through 270 of the rat liver GnT-I cDNA (Chapter 1), was used for the cDNA synthesis. The primer for PCR, P2 (GTCGAATTCCAGCCCACAAAGATGA-TAGC), was complementary to the nucleotides 220 through 197 with additional bases on its 5'-end for an *Eco* RI site. The products were cloned into the *Eco RI* site of pBluescript (Stratagene). The nucleotide sequences were determined by a majority among multiple clones of the same class (see results) to correct mismatches caused by PCR mutation.

Northern hybridization.

Northern blotting and hybridization were performed as described in chapter 1.

Reverse-transcription PCR (RT-PCR).

cDNA was synthesized from 2 μg of total RNA by MMLV reverse transcriptase (Gibco BRL) with random hexamers and 1/50 of this cDNA was used for PCR. Primer MGF (AAGTGTACTCTCGCCATCCA) and MGR (CAGAGCTCCATAGAGCTTGA) were used for β_2-microglobulin mRNA (10) as an internal control. Primer 1AF (GTTGCGCGTTGATCTCC) and R (CAGACTGCTTCTTCAGCATCC) were for GnT-I mRNA including exon1A, primer 1BF (CATTGCTTGCTAGCAGGAAG) and R for GnT-I mRNA including exon1B, and primer 2F (TCATCACACTGTGCACATGG) and R for GnT-I mRNA including exon2. PCR was carried out in 10 μl reaction in the presence of 5 μCi of [α-^{32}P] dATP. After 20 cycles of amplification for microglobulin and exon 1A, 24 cycles for exon 1B and 2, the products were electrophoresed and analyzed by BAS 2000 image analyzer (FUJI PHOTO FILM co., Ltd.). The linearity of the amplification was confirmed by separate experiments.

Transfection and luciferase assay.

The luciferase expression plasmids were constructed to use a thymidine kinase promoter to transcribe the promoterless firefly luciferase gene. The thymidine kinase promoter was excised from pBLCAT2(11) by *Pvu* II/*Hin* dIII digestion and subcloned into the *Kpn* I (blunted)/*Xho* I site of pBluescript. This fragment was excised again by *Hin* dIII digestion and inserted into the *Hin* dIII site of the promoterless luciferase vector, pSV0ALΔ5' (12). The resulting plasmid was termed ptkLΔ5' and used as the positive control of luciferase expression.

To make the assay plasmids containing GnT-I 5'-UTR, the nucleotide sequence coding the first two amino acids of the luciferase was mutagenized from ATGGAA to ACTGCA, to make a *Pst* I site. The GnT-I 5'-UTRs were excised from the cloned 5'-RACE products by *Eco* RV (on the vector) and *Pst* I (on the 6th and 7th codon of GnT-I) and introduced into the mutagenized plasmid. The resulting plasmids, termed ptk1ALLuc

and ptk1ASLuc, express a modified luciferase which contains the first seven amino acids of GnT-I instead of the first two amino acids of luciferase. The resulted N-terminal sequence is MLKKQSADAK.

COS7 cells were transfected by DEAE-dextran method modified by Al-Moslih and Dubes(13). pact-β-gal, a β-galactosidase expression plasmid driven by β-actin promoter, was co-transfected as the internal control of the transfection efficiency. Luciferase activity was assayed with PiccaGene Luminescence Kit (Toyo Ink Co.).

Antibody production.

Recombinant rat GnT-I protein which lacks the transmembrane region and tagged with histidine hexamer (14) was produced in *E. coli*, purified, and used to immunize a rabbit. The crude serum was affinity-purified with the same recombinant GnT-I.

Protein preparation and western blotting.

Rat brain and liver were homogenized by Teflon homogenizer in PBS containing 1 mM EDTA, 100 μg/ml (*p*-aminophenyl)methanesulfonyl fluoride hydrochloride, 10 μg/ml leupeptin, and 2 μg/ml aprotinin. The tissue debris was removed by centrifugation at 700 g for 10 minutes. The membrane fraction was collected by ultracentrifugation of this supernatant at 100,000g for 2 hours. The pellet was suspended in PBS containing 1% Triton X-100, 0.1% SDS, and the proteinase inhibitors mentioned above, and stored at -80°C. These protein samples (50 μg) were separated by 10% SDS-PAGE and transferred to nitrocellulose filter. Immunodetection was carried out with ECL detection reagents (Amersham) following the manufacturer's instructions.

Results

Structure of rat brain GnT-I cDNA.

To know the functi on of the two differently sized transcripts, it is necessary to determine those structures. The author first tried to characterize the 3.3 kb transcript of GnT-I by cloning its cDNA from rat brain, where the GnT-I transcript of this form was preferentially expressed. Eighteen clones of GnT-I cDNA were partially sequenced from the both ends (data not shown). This preliminary analysis showed no sequence variety in the coding region among the brain cDNAs and the previously reported GnT-I cDNA obtained from rat liver. Judging from the length of the brain cDNA fragments and the alignment of their terminal sequences with the liver cDNA, it was unlikely that the coding region structure of the brain cDNA clones had any major variation, caused by alternative splicing, for example. Some clones had truncated 3' UTRs, which were shorter than that of the liver cDNA by 0.9 kb. As some GnT-I cDNA clones of this type were obtained also from rat liver and adrenal (data not shown), it is unlikely that this truncation of the 3' UTR is a brain-specific phenomenon. However, one of the brain clones, later found to belong to the class 1B-3 (see below), had an unique sequence of 31 bases in its 5'-UTR, which was totally different from the corresponding region of the liver GnT-I cDNA. Based on these results the author hypothesized that the variation of 5'-UTR caused the two differently sized transcripts.

To test this hypothesis, the 5'-UTRs of GnT-I mRNA were isolated from rat brain by 5'-RACE. The cloned RACE products could be divided into five classes (Fig. 1 *a*). The first class, class 1AS-3, consisted of clones which were similar to the liver cDNA. The 5'-UTR of the longest clone in this class was 175 bp and included the whole 5'-UTR of the liver cDNA. The second class, class 1AL-3 clones included the whole 5'-UTR of the class 1AS clones and had an additional sequence upstream to it. The 5'-UTR of the longest clone of this class was 623 bp. The third class, class 1AS-2-3, was similar to the class

1AS but had an additional sequence of 107 bp inserted between the bases 123 bp and 124 bp upstream to the start codon of GnT-I. The forth class, class 1B-3, consisted of the clones which had the same sequence as class 1AS and 1AL between the start codon and the base 123 bp upstream to it. The remaining 5'-distal sequences of these clones,

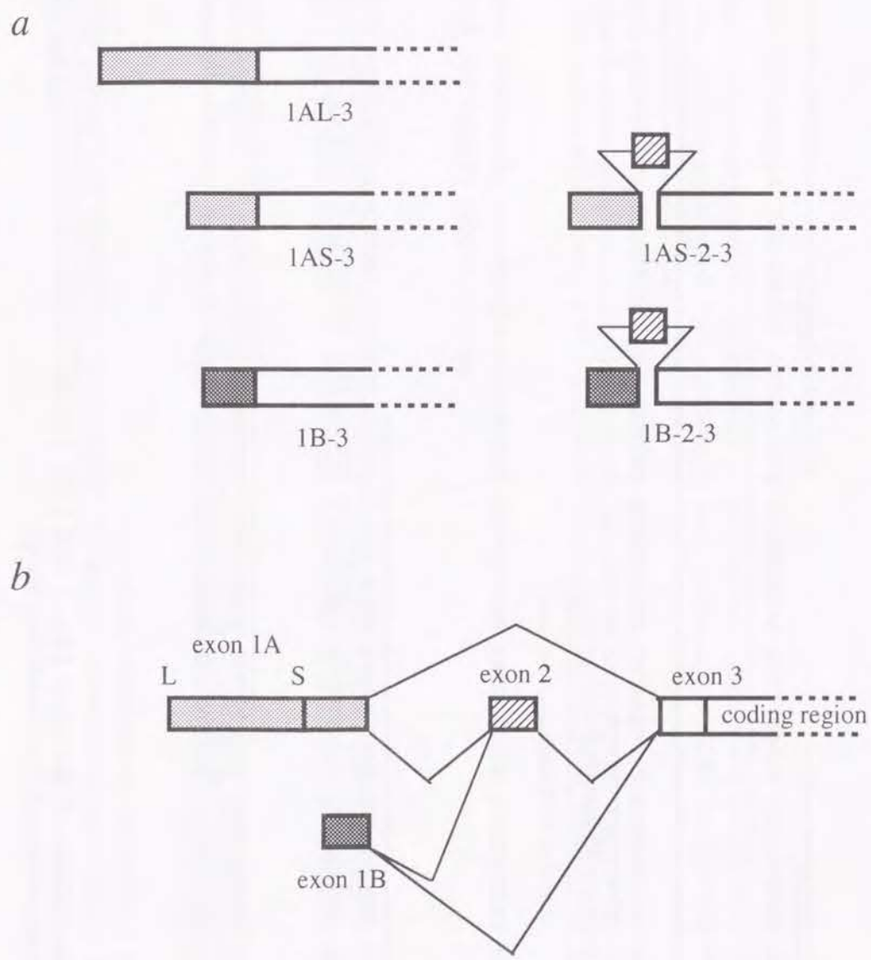


Fig. 1 The 5'-UTR structures of GnT-I mRNA detected by 5'-RACE. *a*, Isolated 5'-RACE clones (see text). *b*, A presumed genome structure of rat GnT-I gene based on the comparison of the five mRNA classes shown in *a*. It includes three transcriptional start sites and an alternatively spliced exon.

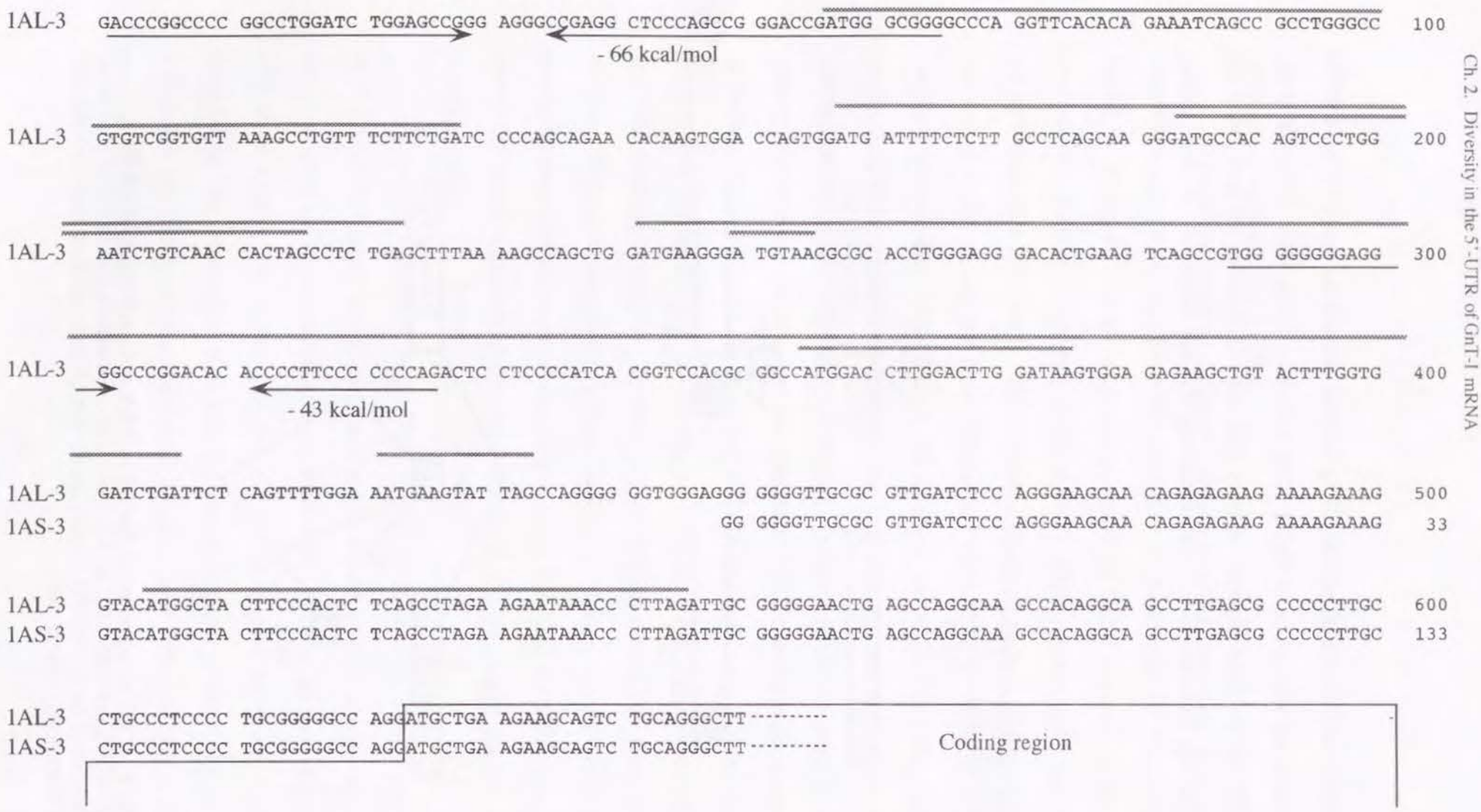


Fig. 2 Nucleotide sequences of the class 1AL-3 and 1AS-3 5'-RACE clones. The upper lines show uORFs. The paired arrows indicate potential secondary structures with calculated stabilizing energy.

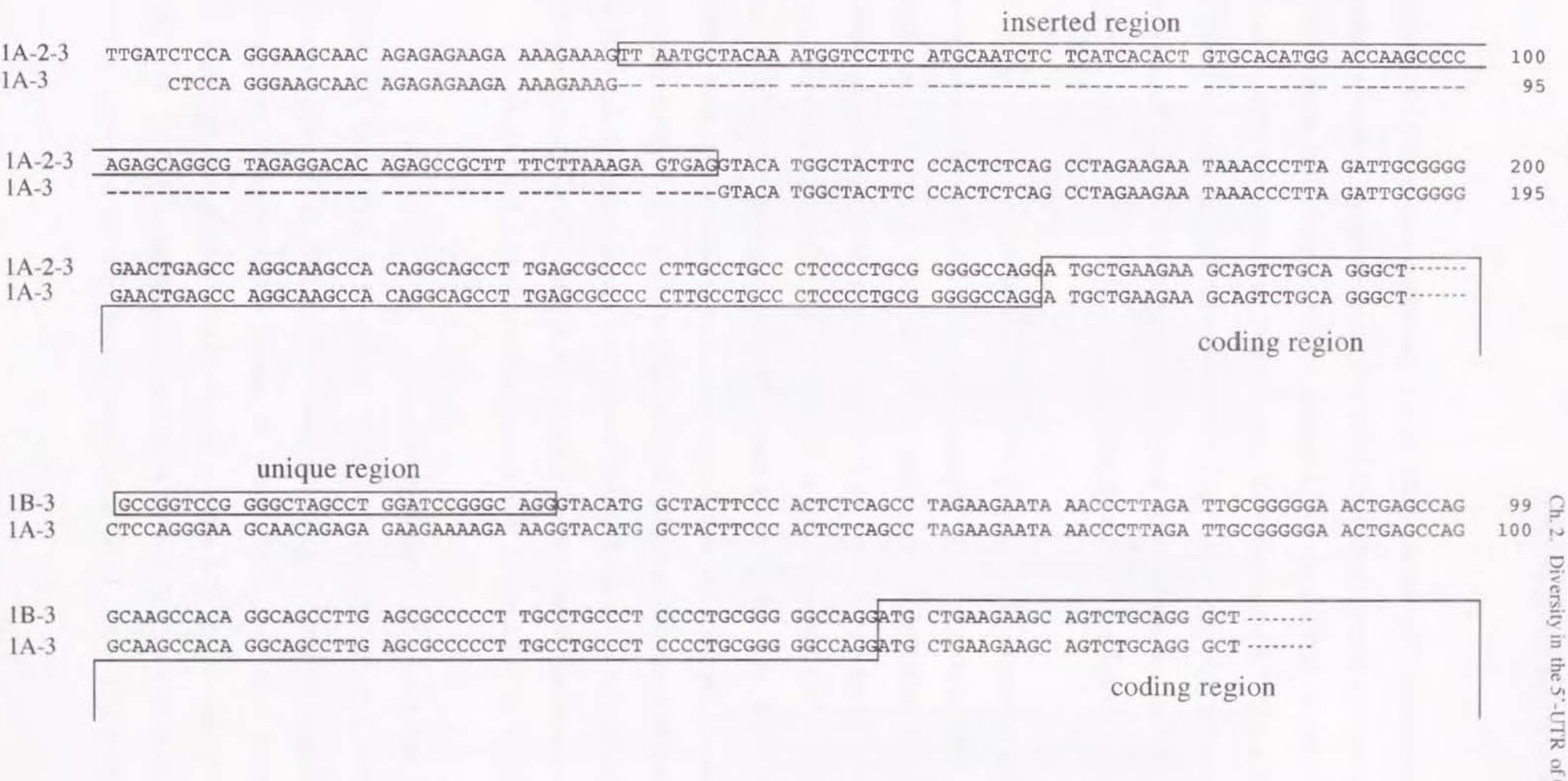


Fig. 3 Nucleotide sequence comparison of class 1A-3, 1A-2-3, and 1B 5'-RACE clones. Boxed are the inserted region of class 1A-2-3 clones and the unique region of class 1B-3 clones.

however, were unique. The last class, class 1B-2-3, was similar to class 1B-3 but the same 107 bp sequence was inserted as class 1AS-2-3 at the same position. A putative genome structure of the rat GnT-I gene was suggested by the comparison of these 5'-UTR structures (Fig. 1*b*). The actual nucleotide sequences of these 5'-UTRs are shown in Fig. 2 and 3. The author have also cloned genomic sequences covering the all exons shown in Fig. 1*b*. This result consisted with the suggested genome structure, except that the exon 1A was actually divided into two exons (Chapter 3).

Tissue-specific expression of the various 5'-UTRs of GnT-I mRNA.

The author also confirmed the expression of various 5'-UTRs of GnT-I mRNA in rat tissues by RT-PCR (Fig. 4). The results showed that class 1A-2-3 mRNA was much less abundant than class 1A-3. The only exception was found in testis, where the amount of class 1A-2-3 mRNA was comparable to that of class 1A-3. Class 1B-2-3 was much less abundant than class 1B-3 either, even in testis. These data suggested that the transcripts containing exon 2 are much less than those without exon 2, whatever first exon is used. Although the author could not determine the exact amount of class 1B-3 mRNA relative to class 1A, the analysis of the amplification kinetics (not shown) suggested that class 1B mRNA was comparably minor as class 1A-2-3 at least in brain. These data suggested that the major 5'-UTR of GnT-I mRNA in rat tissues is either class 1AS-3 or 1AL-3.

Class 1AL-3 mRNA is the 3.3 kb transcript in brain.

According to the difference in length between class 1AS-3 and 1AL-3 and estimated amounts of various 5'-UTRs, the longer GnT-I mRNA preferentially expressed in rat supposed to be class 1AL-3, while the shorter mRNA expressed in liver was supposed to be class 1AS-3. The author performed northern hybridization analysis with a probe specific to the class 1AL 5'-UTR to examine this assumption (Fig. 5). This 1AL-specific probe hybridized only to the longer transcript, and not to the shorter one. 1AL 5'-

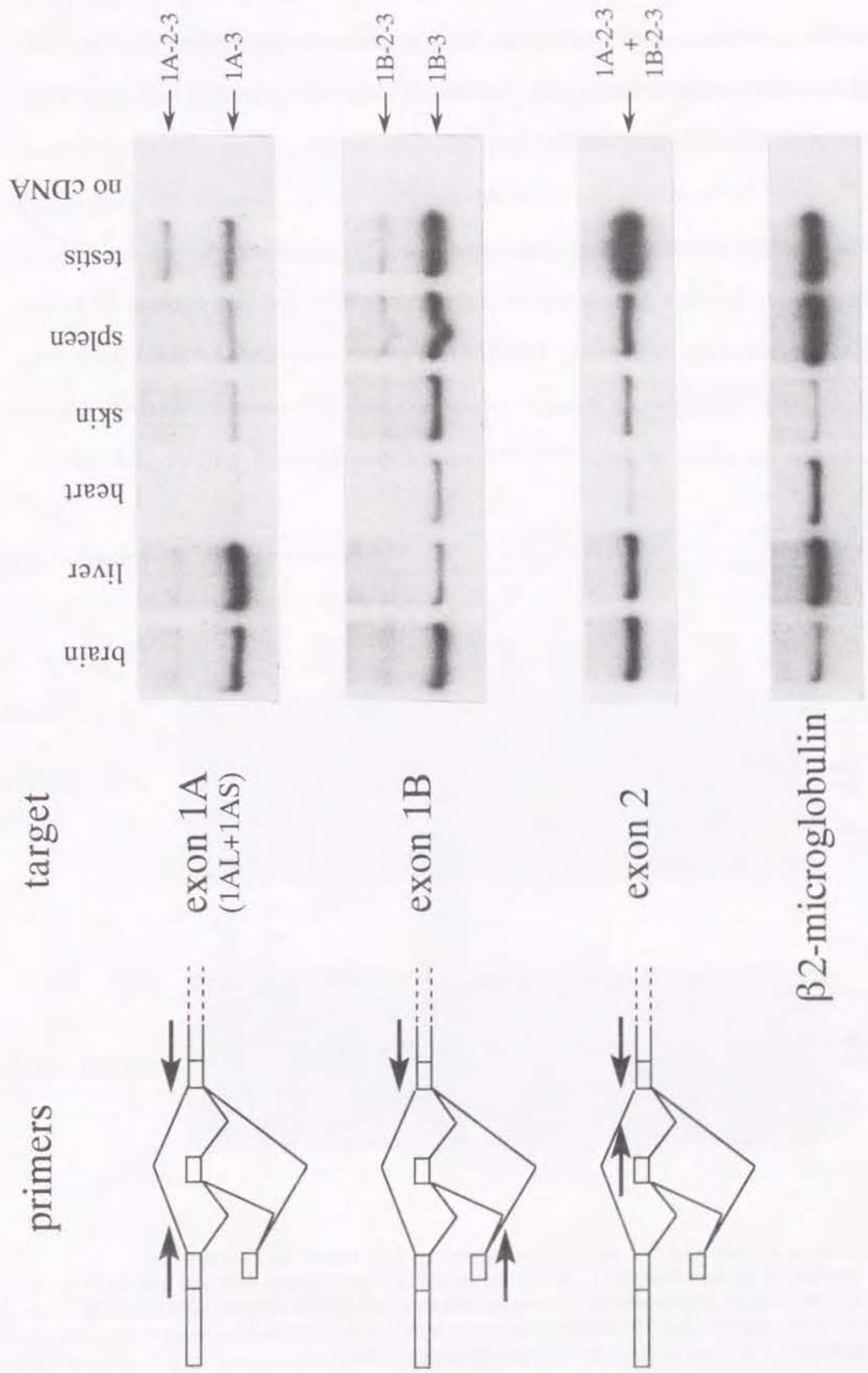


Fig. 4 Expression of GnT-I mRNA subtypes in various tissues. cDNA synthesis and amplification was performed as discribed in Materials and Methods. Note that either of the RNA with or without the exon 2 was detected by the primer sets used for the amplification of the exon 1A and 1B.

UTR was expressed strongly in brain, moderately in heart, kidney, spleen, lung and ovary, and scarcely in liver, colon, skin and muscle. Its expression was proportional to that of the entire GnT-I mRNA except in brain, liver, and heart. Heart had relatively high proportion of 1AL mRNA to the total GnT-I mRNA (Fig. 5*b*).

Class 1AL and class 1AS mRNAs were translated with different efficiency.

The 5'-UTR of class 1AL-3 mRNA have eight ORFs and two regions those can form very stable stem-loop structures. The 5'-UTR of the class 1AS-3 mRNA have only one ORF. As these structures are thought to interfere with the translational efficiency, the author examined the effect of these 5'-UTRs on the translation of mRNA of a reporter

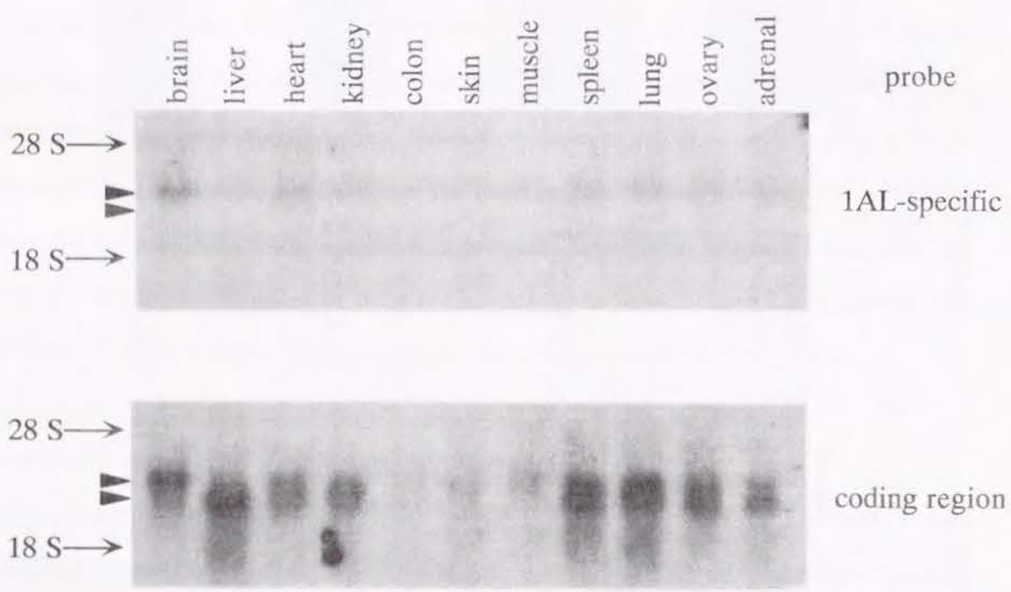
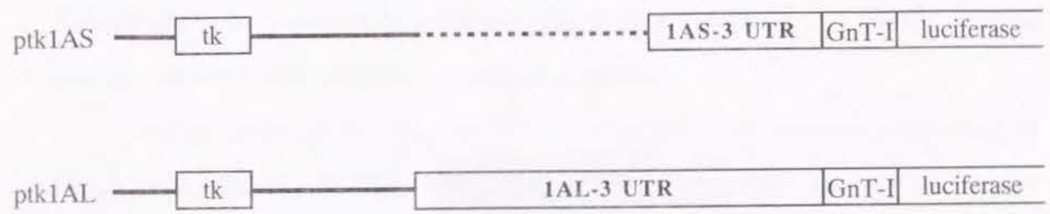


Fig. 5 Detection of the GnT-I mRNA containing the exon 1AL from rat tissues by northern hybridization. While the conding region probe detected two mRNA species, 3.3 kb and 3.0 kb (arrowheads), the exon 1AL-specific probe hybridized only to the 3.3 kb mRNA. Note that the 3.3 kb mRNA was most enriched in brain and virtually undetectable in liver. The arrows indicate the positions of rRNAs.

gene. The 5'-UTR and the first six codons of GnT-I were fused to the 5' end of the modified firefly luciferase gene which lacks the original first two codons. These fusion genes were transiently expressed in COS cells by the thymidine kinase promoter of herpes simplex virus (Fig. 6*a*). Compared with insertion of class 1AS 5'-UTR, insertion of class 1AL-3 5'-UTR reduced the luciferase expression by ~7-fold (Fig. 6*b*). This result suggested that GnT-I expression is regulated translationally as well as transcriptionally.

a



b

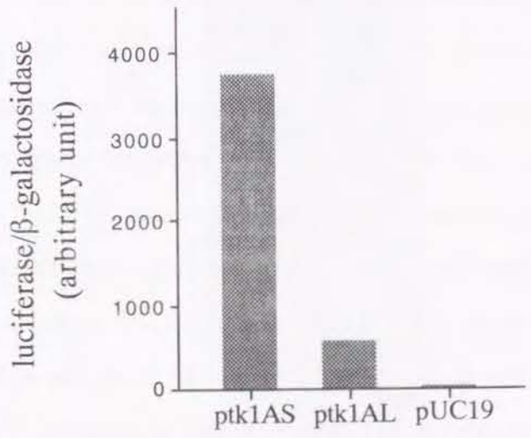


Fig. 6 Effect of GnT-I 5'-UTRs on translational efficiency of a reporter gene mRNA. *a*, Structure of the reporter vectors. The first two codons of the luciferase gene were replaced by the first six codons of GnT-I to fuse the GnT-I 5'-UTRs in ptk1AS and ptk1AL. tk; thymidine kinase promoter. *b*, Luciferase activity. COS cells were transfected with the vectors shown in *a* and transiently expressed luciferase activity was measured. Co-transfected β-galactosidase activity was used as an internal control.

Since the majority of GnT-I mRNA in rat brain was class 1AL while class 1AS was dominant in rat liver, rat brain would express much less GnT-I protein than liver if the translation of class 1AL mRNA was low *in vivo* as well as in the cultured cells. The immunodetection of GnT-I protein on a western blot confirmed this prediction. As shown in Fig. 7, the 48 kDa signal was detected only from liver on the author's conditions. This difference of the amount of GnT-I protein between brain and liver was much greater than the difference of GnT-I mRNA amount shown in Fig. 5. This result supported the idea that GnT-I expression was translationally regulated in rat brain.

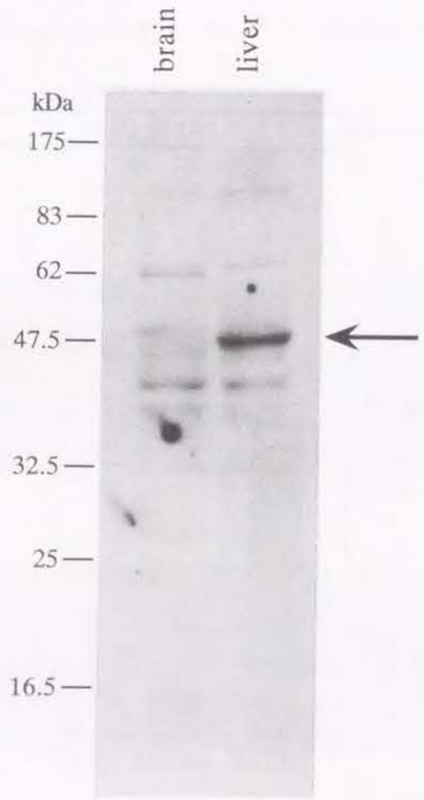


Fig. 6 Expression of GnT-I protein in rat brain and liver. Membrane proteins were precipitated by centrifugation from rat brain and liver homogenate at 100,000 g and 50 μg of protein was loaded to each lane. Western blotting and immunodetection was performed as described in Materials and Methods. The 48 kDa signal (arrow) detected by anti-GnT-I antiserum was only seen in liver lane.

Discussion

As a gate enzyme of complex *N*-glycan synthesis, regulated expression of GnT-I could change the proportion of complex and hybrid glycans to high-mannose glycans, thus could shift the whole repertoire of *N*-glycans which a cell can synthesis. The author's previous results showed that not only the amount of GnT-I mRNA are different in each rat tissue but also there are two size classes of mRNA which are differentially expressed in rat tissues (Chapter 1). In present study the author suggested that these size difference of rat GnT-I mRNAs were resulted from the variation of their 5'-UTR. GnT-I protein synthesis was also regulated by translational efficiency of mRNA.

Unlike *in vitro* translation, long 5'-UTR of mRNA does not necessarily reduce its translational efficiency in vivo. The effect of 5'-UTR of mRNA can vary in different tissues, so that can participate in tissue-specific regulation of gene expression. It can also be a part of the machinery with which cells respond to physiological conditions. Multiple mechanisms have been known for these translational regulation by 5'-UTR (15, 16). 5'-UTR can interfere with mRNA stability (17) and there are some trans-acting factors which were reported to bind certain sequences in 5'-UTR and control the translation of mRNA (18, 19). Upstream open reading frames (uORFs) in 5'-UTR are also shown to be important to translational regulation. The response of the *GCN4* gene expression to amino acid starvation, a well-known example of this kind of gene regulation in yeast, is governed by the uORFs exist in the 5'-UTR of its mRNA (reviewed in Ref. 20). uORFs are known to determine the tissue-specific expression of retinoic acid receptor β2, too (21). Long 5'-UTR with multiple uORFs is often found in mRNAs of strictly regulated genes like growth-factors and considered to be important to their regulation (15, 16). In this context, the 5'-UTR of class 1AL mRNA of GnT-I, which has eight uORFs, may participate in the regulation of the GnT-I gene in a complex manner. Although the author showed the class 1AL mRNA was not efficiently translated in rat brain as well as in cultured cells, its

translation may be up-regulated on certain physiological conditions or in other tissues. The author have found two GnT-I mRNA species different in length in human tissues as well (data not shown), which suggests the existence of conserved roles of the longer GnT-I mRNA.

The transcriptional regulation of the GnT-I gene expression is interesting as well. It should determine the total amount of GnT-I mRNA and the relative amounts of mRNA subclasses with different 5'-UTRs in a tissue-specific manner. Even class 1B mRNA, a minor subclass, could determine the protein amount in tissues like brain where the translation of the major mRNA is low. The examination of the promoter region for these GnT-I mRNA would be informative.

Mouse GnT-I has been reported to have two size classes of mRNA similar to rat GnT-I (15, 16). Yang et al have determined the 5'-UTR sequences of these mRNAs (*Mgat-1l* and *Mgat-1s*) (22). These mouse 5'-UTRs were divided by two introns and thus structurally resembled to the class 1A-2-3 mRNA of rat GnT-I. The region between the introns of the mouse 5'-UTR, however, was totally different from the putative exon2 in rat, despite that the other regions were ~90% homologous to the corresponding rat sequences. Moreover, the 22 bases on the 5'-terminal of *Mgat-1s*, which is not included to *Mgat-1l*, did not exist in the rat 1AS-3 mRNA. It suggests that the structure of the rat GnT-I gene may not be identical to that of mouse. Another possibility is that *Mgat-1s* was a splicing variant which has not been found, or does not exist, in rat. Analysis of the genom structure of the GnT-I gene both in rat and mouse would be indispensable.

References

1. Kornfeld, R. and Kornfeld, S. (1985) "Assembly of asparagine-linked oligosaccharides." *Annu. Rev. Biochem.*, **54**, 631-664.

2. Ioffe, E. and Stanley, P. (1994) "Mice lacking *N*-acetylglucosaminyltransferase I activity die at mid-gestation, revealing an essential role for complex or hybrid N-linked carbohydrates." *Proc. Natl. Acad. Sci. USA*, **91**, 728-732.

3. Metzler, M., Gertz, A., Sarkar, M., Schachter, H., Schrader, J. W., and Marth, J. D. (1994) "Complex asparagine-linked oligosaccharides are required for morphogenic events during post-implantation development." *EMBO J.*, **13**, 2056-2065.

4. Campbell, R. M., Metzler, M., Granovsky, M., Dennis, J. W., and Marth, J. D. (1995) "Complex asparagine-linked oligosaccharides in *Mgat1*-null embryos." *Glycobiology*, **5**, 535-543.

5. Drickamer, K. (1991) "Clearing up glycoprotein hormones." *Cell*, **67**, 1029-1032.

6. Barondes, S. H., Cooper, D. N. W., Gitt, M. A., and Leffler, H. (1994) "Galectins — Structure and functions of a large family of animal lectins." *J. Biol. Chem.*, **269**, 20807-20810.

7. Pownall, S., Kozak, C. A., Schappert, K., Sarkar. M., Hull, E., Schachter, H., and Marth, J. D. (1992) "Molecular cloning and characterization of the mouse UDP-*N*-acetylglucosamine : α-3-D-mannoside β-1,2-*N*-acetylglucosaminyltransferase I gene." *Genomics*, **12**, 699-704.

8. Kumar, R., Yang, J., Eddy, R. L., Byers, M. G., Shows, T. B., and Stanley, P. (1992) "Cloning and expression of the murine gene and chromosomal location of the human gene encoding *N*-acetylglucosaminyltransferase I." *Glycobiology*, **2**, 383-393.

9. Chomczynski, P. and Sacchi, N. (1987) "Single-step method of RNA isolation by acid guanidium thiocyanate-phenol-chloroform extraction." *Anal. Biochem.*, **162**, 156-159.

10. Mauxion, F. and Kress, M. (1987) "Nucleotide sequence of rat β_2-microglobulin cDNA." *Nucl. Acid. Res.*, **15**, 7638.

11. Luckow, B. and Schütz, G. (1987) "CAT constructs with multiple unique restriction sites for the functional analysis of eukaryotic promoters and regulatry elements." *Nucl. Acid Res.*, **15**, 5490.

12. De Wet, J. R., Wood, K. V., Deluca, M., Helinski, D. R., and Subramani, S. (1987) "Firefly luciferase gene: Structure and expression in mammalian cells." *Mol. Cell. Biol.*, **7**, 725-737.

13. Al-Moslih, M. I. and Dubes, G. R. (1973) "The kinetics of DEAE-dextran-induced cell sensitization to transfection." *J. Gen. Virol.*, **18**, 189-193.

14. Nishiu, J., Kioka, N., Fukada, T., Sakai, H., and Komano, T. (1995) "Characterization of rat *N*-acetylglucosaminyltransferase I expressed in *Escherichia coli*." *Biosci. Biotech. Biochem.*, **59**, 1750-1752.

15. Kozak, M. (1991) "An analysis of vertebrate mRNA sequences: Intimation of translational control." *J. Cell Biol.*, **115**, 887-904.

16. Kozak, M. (1992) "Regulation of translation in eukaryotic systems." *Annu. Rev. Cell Biol.*, **8**, 197-225.

17. Oliveira, C. C. and McCarthy, E. G. (1995) "The relationship between eukaryotic translation and mRNA stability — A short upstream open reading frame strongly inhibits translation initiation and greatly accelerates mRNA degradation in the yeast *Saccharomyces cerevisiae*." *J. Biol. Chem.*, **270**, 8936-8943.

18. McCarthy, J. E. G. and Kollmus, H. (1995) "Cytoplasmic mRNA-protein interactions in eukaryotic gene expression." *Trends Biochem. Sci.*, **20**, 191-197.

19. Sun, X.-L. and Antony, A. C. (1996) "Evidence that a specific interaction between an 18-base *cis*-element in the 5'-untranslated region of human folate receptor-α mRNA and a 46-kDa cytosolic *trans*-factor is critical for translation." *J. Biol. Chem.*, **271**, 25539-25547.

20. Geballe, A. P. and Morris, D. R. (1994) "Initiation codons within 5'-leaders of mRNAs as regulators of translation." *Trends Biochem. Sci.*, **19**, 159-164.

21. Zimmer, A., Zimmer, A. M., and Reynolds, K. (1994) "Tissue specific expression of the retinoic acid receptor-β2: Regulation by short open reading frames in 5'-noncoding region." *J. Cell Biol.*, **127**, 1111-1119.

22. Yang, J., Bhaumik, M., Liu, Y., and Stanley, P. (1994) "Regulation of *N*-linked glycosylation. Neuronal cell-specific expression of a 5' extended transcript from the gene encoding *N*-acetylglucosaminyltransferase I." *Glycobiology*, **4**, 703-712.

Chapter 3. Analysis of the genom structure of the rat GnT-I gene

Introduction

As a gate enzyme of complex *N*-glycan synthesis, regulated expression of GnT-I could change the ratio of complex and hybrid glycans to high-mannose glycans, thus could shift the whole repertoire of *N*-glycans which a cell can synthesis. Little is known, however, about the regulation of the GnT-I gene expression. There is no reported data concerning the comparison of the enzymatic activity of GnT-I among animal tissues as long as the author know. Tissue -specific expression of GnT-I mRNA has been documented for mouse (1, 2) and rat (Chapter 1 and 2). A part of genomic sequence have been reported for human (3) and mouse (1, 4) GnT-I, but neither of them describes the promoter region. In Chapter 2 the author found that there are at least five distinct 5'-UTR structures in rat GnT-I mRNA and they affect the translational efficiency of the mRNA. It suggested that alternative promoter usage may control GnT-I expression in a way coupled with translational regulation

In this study the author cloned the GnT-I gene from rat genome and examined the sequences upstream to the gene. This analysis revealed some sequences resembling transcriptional regulation motifs, suggesting tissue-specific and inductive nature of certain parts of GnT-I promoters.

Materials and Methods

Screening of genomic libraries.

Rat genomic libraries (Sprague-Dawley strain) were purchased from Clontech and Stratagene. λgntg22 and λ7 was isolated by hybridization with the coding region of GnT-I

cDNA and a part of the exon 2 excised from a 5'-RACE clone (Chapter 2). λ1A was by hybridization with a part of exon 1A.

S1 mapping.

The probe for class 1AL mRNA was made by connecting a part of a 5'-RACE clone and a genomic fragment at a position within the exon 1A-1 to bypass the intron between the exon 1A-1 and 1A-2. It began from the nucleotide 464 base upstream to the initiation codon and was 281 nucleotides long. The 3'-terminal sequence of 37 nucleotides was an unrelevant to GnT-I, which allows to distinguish the fully protected fragment from undigested probe. Single-strand DNA probes were uniformly labeled as described by Sanbrook et. al. (5). Poly(A)$^+$ RNA (5 μg) was heated at 85°C for 10 minutes with probe (1-5 x 10^4 cpm) in 30 μl of buffer (80% formamide, 0.4 M NaCl, 40 mM PIPES (pH 6.4), 1mM EDTA), hybridized overnight at 30°C, and quickly cooled on ice. Ice cold S1 nuclease buffer (300 μl) containing 2000 units of S1 nuclease (Gibco BRL) was added and incubated at 37°C for 1 hour. The reaction was stopped by EDTA and the digested samples were phenol-chloroform extracted, ethanol precipitated, and resolved on 4-5% sequencing gel.

Results

Genomic structure of the rat GnT-I gene.

The author screened rat genomic library with probes specific to each of the four putative exons (see Chapter 2) of the GnT-I gene, those had been presumed by 5'-RACE analysis. Three overlapping clones covering all of these exons were obtained (Fig. 1). Base sequences were determined in all exons and in the regions upstream to the exon 1A-1, 1A-2, and 1B. The presumed genomic structure of rat GnT-I was confirmed to be accurate with the only exception that the putative exon 1A was found to be divided into two exons, exon 1A-1 and 1A-2. Exon 1A-2 was 429bp long and included all of the 5'-UTR of the class

1AS 5'-RACE clones described in Chapter 2, suggesting this exon contains transcriptional start site(s) for the class 1AS mRNA (designated by an arrow labeled 'S' in Fig. 1). The author tried to determine the transcriptional start site for the class 1AL and 1AS mRNA by S1 mapping. Four major start sites were detected for the transcription of class 1AL mRNA. The author's extensive effort to determine the transcriptional start sites for the class 1AS mRNA was hampered by high background noise and non-specific signals (data not shown), perhaps due to the locally biased base composition of the 5'-UTR of rat GnT-I mRNA. The author did not try to determine the transcriptional start sites for class 1B mRNA since its expression had been found to be low (see Chapter 2). Therefore, the author assumed the 5' ends of the 5'-RACE clone of class 1AS and 1B as the transcriptional start sites for further analysis.

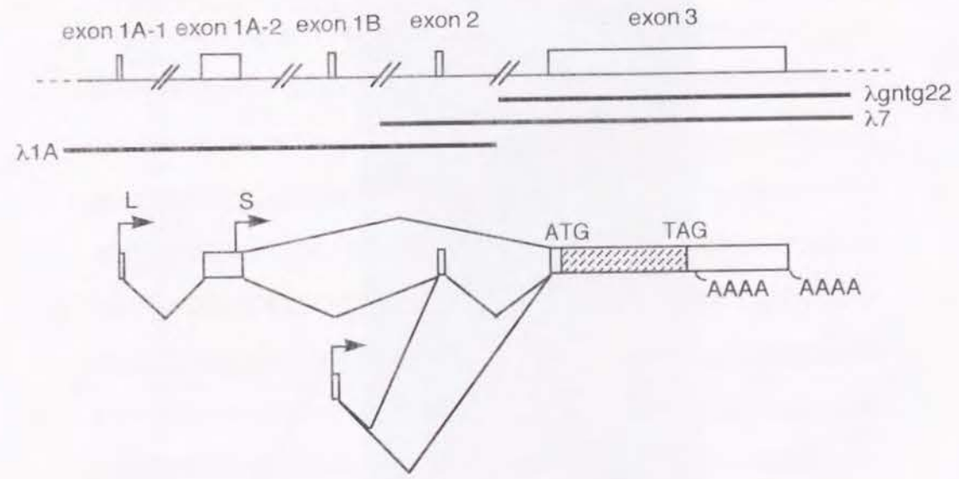


Fig. 1 Genom structure of rat GnT-I gene. The upper part of the figure shows the organization of the exons on rat genome and lambda clones covering these exons. The lower part is a schematic diagram of the splicing structure of GnT-I gene. The arrows indicate the transcriptional start sites (see text). The hatched box stands for the coding region of GnT-I, and the two polyadenylation sites are also indicated.

Sequence analysis of the upstream regions.

Fig. 3 shows the nucleotide sequence upstream to the transcriptional start sites for the class 1AL mRNA, and Fig. 4 shows the sequece upstream to the 5' ends of the 5'-RACE

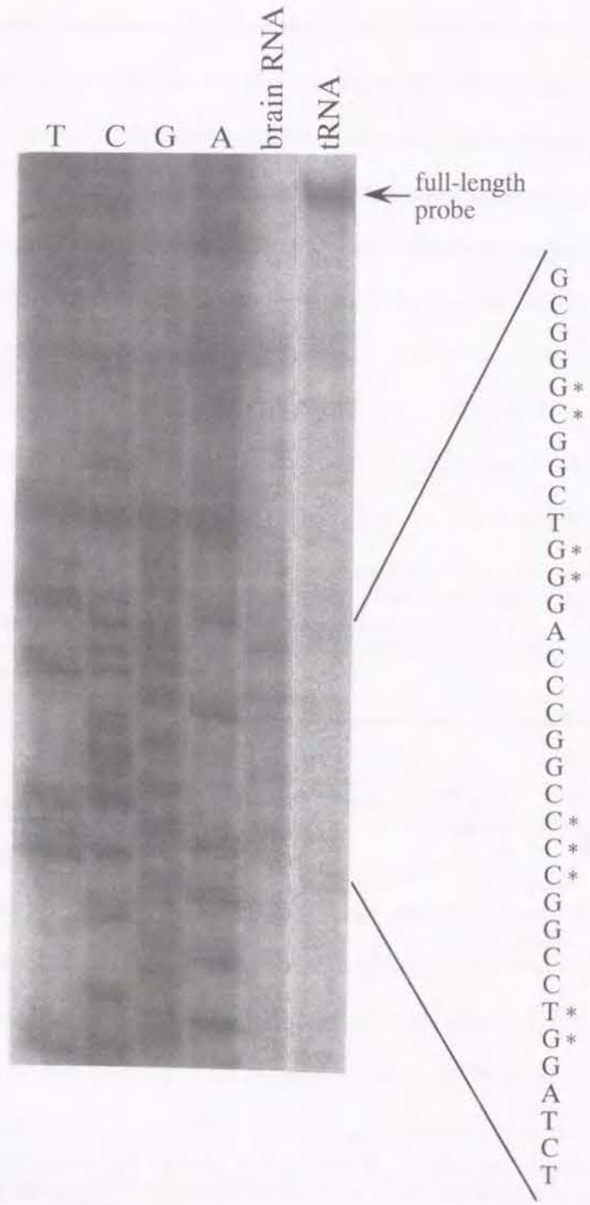


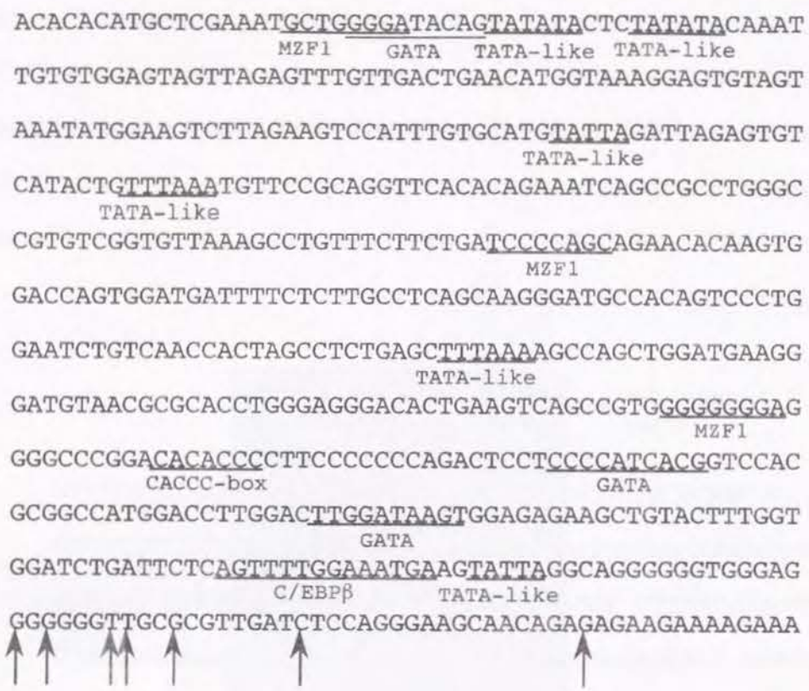
Fig. 2 S1 mapping of the 5'-end of rat brain GnT-I mRNA. S1 mapping was performed as described in Materials and Methods. Four major start sites were detected for the transcription of class 1AL mRNA from 5 μg of brain poly(A)$^+$ RNA (asterisks).

clones those belong to class 1AS and 1B. Potential regulatory motifs within those regions are also indicated. There was no typical TATAA sequence, but multiple TATA-like sequences existed upstream to the 5'-end of class 1AS mRNA. There were potential binding motifs of various transcription factors including an ubiquitous factor SP1, inducible factors like Rel family, MZF1, and CRE-BP, proto-oncogene products like c-myb and Ets family, and tissue-specific factors of GATA and Ikaros families.

```
GAGCTCAGCACAGTTCATGCTCTGTAAACAGTAACCGGGCGCTCTCCTTC

CTTTTAGAGATGTCCTACCTCATCCAGTGAGACAGTGGGCAATTGTGATG

ATTAAAACTTTTTCTGGCACTGTTACTCAGGAAGCTCAGCTCAGTAACTG
                          Ets                   c-myb
GTTTGTCATAACAGGCTGTCGATTACTGGCTGTTAGCATAAAGAGTGCCT
                                    TATA-like
AGAAGGCCTAAAAAGGGTCCGTGACCTCAGAAGCTACTTCAACTAGACAG
                         CRE                    c-myb
TTGCTAGGCGACCTTCCTGCATTGAGGTAAGAGGCAGGGCAGCCCGCACA

GGTTCCGCCTCCACTTCCAGCCCTAGTTCCTGCCGCCTCCGGCCGCAGGA

GCATCAAAGTGAGCGACTGGTCGGAGAACCGGCGGCGCAAGCGCGCTACT

TCCCGCGCCGCACTGAGTGACGCACACTCCCCCTCTAGGCCTCAGCCTCC

GCCCCGCCCCCCGGAGTACCGCTCCGGCCCTGAGCCTCCGCTACATATTC
  SP1
ACCGGGCCCTTCAGCCACGCCGGCCCTTCACCTTGAGAACCCTGTAGAAC

TGGTGCTGCCCTGGCGGCGTTGGTGATGGGCGGGCAAGAACCTGGCGAAT
                        GATA
GGGGCGGGGAAGTGGGGAACCAGCGGTTACCCAGGACCTACTTCCGGAGG
SP1  MZF1    MZF1                              Ets
ACAGGCGGAAGTGGCTGCGGTGCCTGGTGGGAGGGGAAAAGGAGCGAGAG
                          MZF1              **  MZF1
GGGGAGGGCCGGGCCGCAGGGGGTCGAGGGGGTGGGGCTTGCGGGCGGCT
**          ***      **
GGGACCCGGCCCCGGCCTGGATCTGGAGCCGGGAGGGCCGAGGCTCCCAG
```

Fig. 3 Nucleotide sequence upstream to GnT-I exon 1A-1. Asterisks indicate transcriptional start sites determined by S1 mapping. Potential binding sequences for transcription factors are underlined. TATA-like; sequence resembling TATAA-box.

a



b

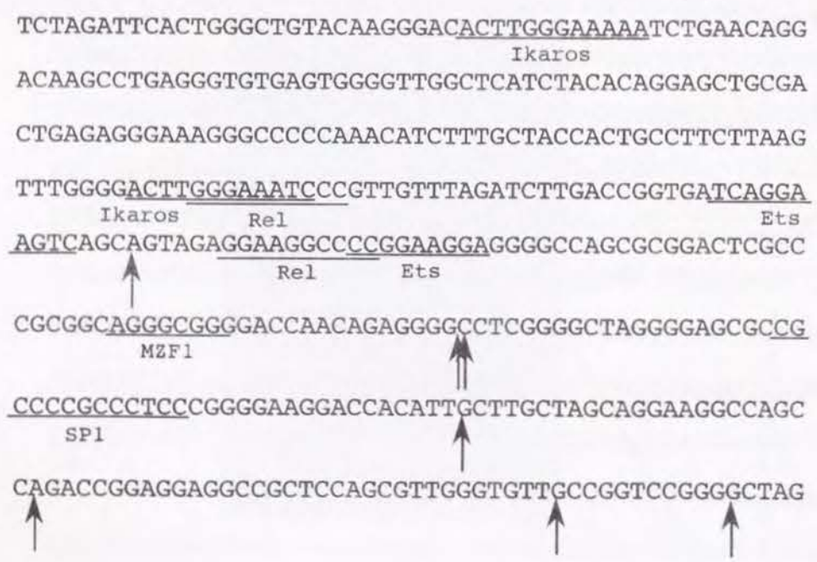


Fig. 4 Genomic seequences upstream to the 5'-end of class 1AS and 1B mRNA. Potential binding sequences for transcriptional factors are underlined. TATA-like; sequence resembling to TATAA-box. The arrows indicate the 5'-ends of the 5'-RACE clones those belong to class 1AS in *a*, and class 1B in *b*.

Discussion

In this study the author isolated the GnT-I gene from rat genome and determined its promoter and exon-intron structures. The rat GnT-I gene was transcribed from at least three regions and had five exons including alternatively spliced one. The regions upstream to the transcriptionl start sites contained no typical TATAA and CCAAT sequences but had many potential binding sites for various transcription factors.

Promoter structures have been reported for a few glycosyltransferases such as mouse β1,4-galactosyltransferase (6), rat α2,6-sialyltransferase (7), and human *N*-acetyl-glucosaminyltransferase III (8). It is worth to say that a model which includes tissue-specific usage of alternative promoter coupled with translational control by the 5'-UTR of mRNA was suggested for the mammary gland-specific induction of the β1,4-galactosyltransferase gene (6). This gene produces an ubiquitously expressed 4.1 kb mRNA, and 3.9 kb mRNA which is specifically up-regulated in lactating mammary glands. The translational efficiency of the 4.1 kb mRNA are thought to be low because its 5'-end can form a very stable secondary structure, so that induction of 3.9 kb mRNA in mammary glands would efficiently up-regulate the enzymatic activity. The promoter for 4.1 kb mRNA was reported to be controlled, at least to some extent, by SP1. SP1-dependent transcription is described for many other house-keeping promoters. The induction of the 3.9 kb mRNA was reported to be achieved by mammary gland-specific factors. An analogous model could be applicable to GnT-I assuming the promoter for the class 1AL mRNA as house-keeping one and that for the class 1AS mRNA as tissue-specific and/or inductive one. It is also analogous between β1,4-galactosyltransferase and GnT-I that the longer mRNA (class 1AL) was translated only inefficiently (Chapter 2) and its putative promoter had multiple SP1 binding sites.

The application of this model to GnT-I, however, may be contradictory to the observation that the expression of the class 1AL mRNA was strikingly tissue-specific in rat

brain and liver. It was strongly expressed in brain but only weakly in liver. This contrasts to the expression of the 4.1 kb mRNA of β1,4-galactosyltransferase which was reported to be very low in brain. Since SP1 level was reported to be low in mouse brain (9), transcription of the 3.3 kb GnT-I mRNA may be activated in brain by other factors. The existence of some potential sequence for tissue-specific and/or inductive regulation in the promoter for the class 1AL mRNA may also support the idea that this promoter is not solely house-keeping. It would be rather objectively expressed in some tissues including brain, although it could be house-keeping in other tissues.

The function of the potential regulation motifs in the GnT-I promoters is currently unclear. It is interesting that the upstream region of exon 1B has a binding sequence for Rel family factors including NF-κB (10). This family is essential to immune response and inflammation (11), and reported to participate in acute-phase response in liver (12) and astrocyte activation by cytokines in brain (13). Although GnT-I mRNA with exon 1B was minor in most rat tissues, it may be induced in these processes. This induction could be important to determining the enzymatic activity of GnT-I in brain, especially in astrocytes, since the major GnT-I mRNA in rat brain was translated only inefficiently and most non-neuronal cells in mouse brain were reported to express GnT-I mRNA only weakly (2). It is also intrigue that a subset of neurons has been reported to constitutively express NF-κB (14).

C/EBPβ is a liver-enriched activating factor and known to be induced in inflammation (15), and essential to biological defense (16). Its binding sequence was found also in the liver-specific promoter of α2,6-sialyltransferase and thought to support the high expression of this enzyme in liver(7). It may be the case for the relatively high expression of the class 1AS mRNA of GnT-I in rat liver. Potential CREs (17) exist in the upstream regions of *N*-acetylglucosaminyltransferase III (18) and β1,4-galactosyltransferase (19), but their significance has not been investigated. Ets family consists with transcription factors those bind purine-rich sequences (20) and concern cell proliferation, differentiation (especially in

lymphocytes), development, and carcinogenesis (21). Ets-1 has been reported to regulate the expression of the human *N*-acetylglucosaminyltransferase V gene (22). It is intrigue that potential binding sites for Ets family and c-myb co-exist in GnT-I promoter since Ets-2 and c-myb are known to cooperate in activation of mim-1 promoter (23). GATA family is a group of cell-specific and/or tissue-specific transcription factors found, for example, in certain blood cells, endothelial cells, neurons, kidney, placenta, and heart (24-26). Each member of the family has distinct expression pattern and are reported to have slightly different sequence preference in their binding (27, 28). CACCC-box is the binding sequence for Kruppel-like factor (KLF) family (29) and essential to the expression of the β-globin gene. This motif also exists in many other genes including the platelet-derived growth factor B chain (30) and P-selectin (31). It is interesting that the CACCC-box in the GnT-I promoter for class 1AS mRNA was adjacent to a potential GATA binding site since erythrocyte KLF and GATA-1 are reported to act synergistically (32). KLF family members are expressed in various tissues other than erythrocytes with distinct tissue-specificities (33, 34). MZF1 is a transcriptional activator cloned from human leukocyte and inducible by retinoic acid (35, 36). Homologous mRNA could be detected by northern hybridization in a murine myeloid cell line. Factors of Ikaros family are known to control the differentiation of lymphocytes (37).

Taking together, GnT-I promoters have potential binding motifs of various transcription factors known to work in hematopoietic cells. Since the hematopoietic cells isolated from GnT-I-deficient transgenic mice were shown to differentiate normally (38), GnT-I should concern the functions of differentiated hematopoietic cells rather than hematopoiesis itself. In this context it is intrigue that interleukin-1 and 2 are reported to have lectin activity and specifically bind to high mannose type *N*-glycans (39-43). Since interleukins can induce NF-κB and other inflammation-relevant transcription factors, it would be possible that GnT-I is a component of a feed-back system of cytokine functions.

References

1. Kumar, R., Yang, J., Eddy, R. L., Byers, M. G., Shows, T. B., and Stanley, P. (1992) "Cloning and expression of the murine gene and chromosomal location of the human gene encoding *N*-acetylglucosaminyltransferase I." *Glycobiology*, **2**, 383-393.

2. Yang, J., Bhaumik, M., Liu, Y., and Stanley, P. (1994) "Regulation of *N*-linked glycosylation. Neuronal cell-specific expression of a 5' extended transcript from the gene encoding *N*-acetylglucosaminyltransferase I." *Glycobiology*, **4**, 703-712.

3. Hull, E., Sarkar, M., Spruijt, M. P. N., Höppener, J. W. M., Dunn, R., and Schachter, H. (1991) "Organization and localization to chromosome 5 of the human UDP-*N*-acetylglucosamine : α-3-D-mannoside β-1,2-*N*-acetyl-glucosaminyltransferase I gene." *Biochem. Biophys. Res. Commun.*, **176**, 608-615.

4. Pownall, S., Kozak, C. A., Schappert, K., Sarkar, M., Hull, E., Schachter, H., and Marth, J. D. (1992) "Molecular cloning and characterization of the mouse UDP-*N*-acetylglucosamine : α-3-D-mannoside β-1,2-*N*-acetylglucosaminyl-transferase I gene." *Genomics*, **12**, 699-704.

5. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) "Molecular cloning — A laboratory manual", second edition, New York, Cold Spring Harbor Laboratory Press.

6. Harduin-Lepers, A., Shaper, J. H., and Shaper, N. L. (1993) "Characterization of two *cis*-regulatory region in murine β1,4-galactosyltransferase gene." *J. Biol. Chem.*, **268**, 14348-14359.

7. Svensson, E. C., Conley, P. B., and Paulson, J. C. (1992) "Regulated expression of α2,6-sialyltransferase by the liver-enriched transcription factor HNF-1, DBP, and LAP." *J. Biol. Chem.*, **267**, 3466-3472.

8. Koyama, N., Miyoshi, E., Ihara, Y., Kang, R., Nishikawa, A., and Taniguchi, N. (1996) "Human *N*-acetylglucosaminyltransferase III gene is transcribed from multiple promoters." *Eur. J. Biochem.*, **238**, 853-861.

9. Saffer, J. D., Jackson, S. P., and Annarella, M. B. (1991) "Developmental expression of Sp1 in the mouse." *Mol. Cell. Biol.*, **11**, 2189-2199.
10. Zabel, U., Schreck, R., and Baeuerle, P. A. (1991) "DNA binding of purified trascription factor NF-κB." *J. Biol. Chem.*, **266**, 252-260.

11. Baeuerle, P. A. and Henkel, T. (1994) "Function and activation of NF-κB in the immune system." *Annu. Rev. Immunol.*, **12**, 141-179.

12. Ron, D., Brasier, A. R., Wright, K. A., Tate, J. E., and Habener, J. F. (1990) "An inducible 50-kilodalton NF-κB-like protein and constitutive protein both bind the acute-phase response element of the angiotensinogen gene." *Mol. Cell. Biol.*, **10**, 1023-1032.

13. Friedman, W. J., Thakur, S., Seidman, L., and Rabson, A. B. (1996) "Regulation of nerve growth factor mRNA by interleukin-1 in rat hippocampal astrocytes is mediated by NF-κB." *J. Biol. Chem.*, **271**, 31115-31120.

14. Kaltschmidt, C., Kaltschmidt, B., Neumann, H., Wekerle, H., and Baeuerle, P. A. (1994) "Constitutive NF-κB activity in neurons." *Mol. Cell. Biol.*, **14**, 3981-3992.

15. Baumann, H., Morella, K. K., Campos, S. P., Cao, Z., and Jahreis, G. P. (1992) "Role of CAAT-enhancer binding protein isoforms in the cytokine regulation of acute-phase plasma protein genes." *J. Biol. Chem.*, **267**, 19744-19751.

16. Tanaka, T., Akira, S., Yoshida, K., Umemoto, M., Yoneda, Y., Shirafuji, N., Fujiwara, H., Suematsu, S., Yoshida, N., and Kishimoto, T. (1995) "Targeted Disruption of the NF-IL6 Gene Discloses Its Essential Role in Bacteria Killing and Tumor Cytotoxicity by Macrophages." *Cell*, **80**, 353-361.

17. Benbrook, D. M. and Jones, N. C. (1994) "Different binding specificities and transactivation of variant CRE's by CREB complexes." *Nucl. Acid Res.*, **22**, 1463-1469.

18. Kim, Y.-J., Park, J.-H., Kim, K.-S., Chang, J.-E., Ko, J. H., Kim, M.-H., Chung, D.-H., Chung, T.-W., Choe, I.-S., Lee, Y.-C., and Kim, C.-H. (1996) "Sequence analysis of the 5'-flanking region of the gene encoding human *N*-acetylglucosaminyltransferase III." *Gene*, **170**, 281-283.

19. Harduin Lepers, A., Shaper, N. L., Mahoney, J. A., and Shaper, J. H. (1992) "Murine β1,4-galactosyltransferase: Round spermatid transcripts are characterized by an extended 5'-untranslated region." *Glycobiology*, **2**, 361-368.

20. Woods, D. B., Ghysdael, J., and Owen, M. J. (1992) "Identification of nucleotide preferences in DNA sequences recognized specifically by c-ETS protein." *Nucl. Acid Res.*, **20**, 699-704.

21. Bhat, N. K., Fischinger, P. J., Seth, A., Watson, D. K., and Papas, T. (1996) "Pleiotropic functions of ETS-1 (Review)." *Int. J. Oncol.*, **8**, 841-846.

22. Kang, R., Saito, H., Ihara, Y., Miyoshi, E., Koyama, N., Sheng, Y., and Taniguchi, N. (1996) "Transcriptional regulation of the *N*-acetylglucosaminyltransferase V gene in human bile duct carcinoma cells (HuCC-T1) is mediated by Ets-1." *J. Biol. Chem.*, **271**, 26706-26712.

23. Dudek, H., Tantravahi, R. V., Rao, V. N., Reddy, E. S. P., and Reddy, E. P. (1992) "Myb and Ets proteins cooperate in transcriptional activation of the *mim-1* promoter." *Proc. Natl. Acad. Sci. USA*, **89**, 1291-1295.

24. Yamamoto, M., Ko, L. J., Leonard, M. W., Beug, H., Orkin, S. H., and Engel, J. D. (1990) "Activity and tissue-specific expression of the transcription factor NF-E1 multigene family." *Genes Dev.*, **4**, 1650-1662.

25. Dorfman, D. M., Wilson, D. B., Bruns, G. A. P., and Orkin, S. H. (1992) "Human transcription factor GATA-2: Evidence for regulation of preproendothelin-1 gene expression in endothelial cells." *J. Biol. Chem.*, **267**, 1279-1285.

26. Arceci, R. J., King, A. J., Simon, M. C., Orkin, S. H., and Wilson, D. B. (1993) "Mouse GATA-4: A retinoic acid-inducible GATA-binding transcription factor

expressed in endodermally derived tissues and heart." *Mol. Cell. Biol.*, **13**, 2235-2246.

27. Merika, M. and Orkin, S. H. (1993) "DNA-binding specificity of GATA family transcription factors." *Mol. Cell. Biol.*, **13**, 3999-4010.

28. Ko, L. J. and Engel, J. D. (1993) "DNA-binding specificities of GATA transcription factor family." *Mol. Cell. Biol.*, **13**, 4011-4022.

29. Miller, I. J. and Bieker, J. J. (1993) "A novel, erythroid cell-specific murine transcription factor that binds to the CACCC element and is related to the Kruppel family of nuclear proteins." *Mol. Cell. Biol.*, **13**, 2776-2786.

30. Liang, Y., Robinson, D. F., Dennig, J., Suske, G., and Fahl, W. E. (1996) "Transcriptional regulation of the SIS/PDGF-B gene in human osteosarcoma cells by the Sp family of transcription factors." *J. Biol. Chem.*, **271**, 11792-11797.

31. Pan, J. and McEver, R. P. (1993) "Characterization of the promoter for the human P-selectin gene." *J. Biol. Chem.*, **268**, 22600-22608.

32. Mkerika, M. and Orkin, S. H. (1995) "Functional synergy and physical interactions of the erythroid transcription factor GATA-1 with the Kruppel family proteins Sp1 and EKLF." *Mol. Cell. Biol.*, **15**, 2437-2447.

33. Miki, T., Kawamata, N., Hirosawa, S., and Aoki, N. (1994) "Gene involved in the 3q27 translocation associated with B-cell lymphoma, *BCL5*, encodes a Kruppel-like zinc-finger protein." *Blood*, **83**, 26-32.

34. Anderson, K. P., Kern, C. B., Crable, S. C., and Lingrel, J. B. (1995) "Isolation of a gene encoding a functional zinc finger protein homologous to erythroid Kruppel-like factor: Identification of a new multigene family." *Mol. Cell. Biol.*, **15**, 5957-5965.

35. Hromas, R., Collins, S. J., Hickstein, D., Raskind, W., Deaven, L. L., O'Hara, P., Hagen, F. S., and Kaushansky, K. (1991) "A retinoic acid-responsive human zinc finger gene, *MZF-1*, preferentially expressed in myeloid cells." *J. Biol. Chem.*, **266**, 14183-14187.

36. Morris, J. F., Hromas, R., and Rauscher III, F. J. (1994) "Characterization of the DNA-binding properties of myeloid zinc finger protein MZF1: Two independent DNA-binding domain recognize two DNA consensus sequences with a common G-rich core." *Mol. Cell. Biol.*, **14**, 1786-1795.

37. Molnár, A. and Georgopoulos, K. (1994) "The Ikaros gene encodes a family of fuctionally diverse zinc finger DNA-binding proteins." *Mol. Cell. Biol.*, **14**, 8292-8303.

38. Metzler, M., Gertz, A., Sarkar, M., Schachter, H., Schrader, J. W., and Marth, J. D. (1994) "Complex asparagine-linked oligosaccharides are required for morphogenic events during post-implantation development." *EMBO J.*, **13**, 2056-2065.

39. Brody, D. T. and Durum, S. K. (1989) "Membrane IL-1: IL-1α precursor binds to the plasma membrane via a lectin-like interactions." *J. Immunol.*, **143**, 1183-1187.

40. Sathyamoorthy, N., Decker, J. M., Sherblom, A. P., and Muchmore, A. (1991) "Evidence that specific high mannose structures directly regulate multiple cellular activities." *Mol. Cell. Biochem.*, **102**, 139-147.

41. Sherblom, A. P., Sathyamoorthy, N., Decker, J. M., and Muchmore, A. V. (1989) "IL-2, a lectin with specificity for high mannose glycopeptides." *J. Immunol.*, **143**, 939-944.

42. Muchmore, A., Decker, J., Shaw, A., and Wingfield, P. (1990) "Evidence that high mannose glycopeptides are able to functionally interact with recombinant tumor necrosis factor and recombinant interleukin 1." *Cancer Res.*, **50**, 6285-6290.

43. Muchmore, A. V., Sathyamoorthy, N., Decker, J., and Sherblom, A. P. (1990) "Evidence that specific high-mannose oligosaccharides can directly inhibit antigen-driven T-cell responses." *J. Leukocyte Biol.*, **48**, .

CONCLUSIONS

As a gate enzyme in the biosynthetic pathway of *N*-linked glycans, *N*-acetylglucosaminyltransferase I (GnT-I; EC 2.4.1.101) initiates the synthesis of hybrid and complex *N*-glycans and thus potentially can determine the ratio of hybrid and complex *N*-glycans to high-mannose *N*-glycans. In **Chapter 1**, the author cloned a GnT-I cDNA from rat liver. This cDNA had a primary structure highly homologous to that of mouse, rabbit, and human GnT-I cDNA and expressed GnT-I activity in *E. coli*. Expression of GnT-I mRNA in rat tissues was high in lung, spleen, and liver, moderately high in brain, and low in skin and muscle. There were two differently sized transcripts, 3.3 kb and 3.0 kb, in rat tissues. Brain preferentially expressed the 3.3 kb mRNA and liver essentially had only 3.0 kb mRNA, in spite that other tissues expressed the both mRNA. These results suggest the transcription of rat GnT-I is regulated in a tissue-specific manner.

In **Chapter 2**, the author investigated the implications of the existence of two different GnT-I transcripts for the regulation of GnT-I activity. cDNA cloning and 5'-RACE in brain showed the two GnT-I transcripts were different only in their 5'-UTRs, and that there were at least five classes of 5'-UTRs, named as class 1AL-3, 1AS-3, 1B-3, 1AS-2-3, 1B-2-3. Northern hybridization and RT-PCR showed class 1AL-3 and 1AS-3 were the 3.3 kb and 3.0 kb mRNA described in Chapter 1, respectively, and that these two classes were the major GnT-I 5'-UTRs in most tissues. The author found that class 1AL-3 5'-UTR reduced the translational efficiency of a reporter gene mRNA by ~7-fold in comparison with class 1AS-3. Rat brain, in which majority of the GnT-I mRNA is class 1AL-3, expressed much less GnT-I protein than liver in spite that moderately high expression of GnT-I mRNA was detected in brain. These results suggest GnT-I expression is regulated translationally as well as transcriptionally.

To investigate how the expression of the multiple classes of GnT-I mRNA is regulated, the author cloned the GnT-I gene from rat genome and examined its exon-intron structure and promoter sequences in **Chapter 3**. The rat GnT-I gene was consist of five exons, and the combination of three transcriptional start sites and one alternatively spliced exon produced the five classes of 5'-UTR. The exact transcriptional start sites were determined for the class 1AL mRNA by S1 mapping. The nucleotide sequences were determined in the putative promoter regions for class 1AL, 1AS, and 1B. There were many potential motifs for the binding of various transcription factors, including an ubiquitous factor SP1, inducible factors like Rel family, MZF1, and CRE-binding protein, proto-oncogene products like c-myb and Ets family, and tissue-specific factors of GATA and Ikaros families. Some combinations of the putative regulatory motifs in GnT-I promoters, GATA sequence and CACCC-box for example, is often observed in other eukaryotic promoters. These results suggest tissue-specific and/or inducive nature of some parts of GnT-I promoters. The transcriptional control of the GnT-I gene would dynamically regulate the GnT-I activity in a manner coupled with translational controls.

LIST OF PABLICATIONS

A. Yoshisue, H., Fukada, T., Yoshida, K., Sen, K., Kurosawa, S., Sakai, H., and Komano, T. (1993) "Transcriptional regulation of *Bacillus thuringiensis* ssp. *israelensis* mosquito larvicidal crystal protein gene *cryIVA*." *J. Bacteriol.*, **175**, 2750-2753.

B. Fukada, T., Iida, K., Kioka, N., Sakai, H., Komano, T. (1994) "Cloning of a cDNA encoding *N*-acetylglucosaminyltransferase I from rat liver and analysis of its expression in rat tissues." *Biosci. Biotech. Biochem.*, **58**, 200-201.

C. Nishiu, J., Kioka, N., Fukada, T., Sakai, H., Komano, T. (1995) "Characterization of rat *N*-acetylglucosaminyltransferase I expressed in *Escherichia coli*." *Biosci. Biotech. Biochem.*, **59**, 1750-1752.

D. Fukada, T., Nishiu, J., Kioka, N., Sakai, H., Komano, T. (1997) "Diversity of the 5'-untranslated region of rat *N*-acetylgucosaminyltransferase I mRNA and its implications for the regulation of the gene expression." in preparation.

E. Fukada, T., Ono, M., Sakata, S., Uchida, T., Kioka, N., Sakai, H., Komano, T. (1997) "Analysis of the genome structure of the rat *N*-acetylglucosaminyltransferase I gene." in preparation.

Chapter 1 is described in reference B.

Chapter 2 is described in reference D.

Chapter 3 is described in reference E.

ACKNOWLEDGMENTS

The author wishes to express his sincerest gratitude to Dr. Tohru Komano, professor of kyoto University, for his valuable suggestions and encouragement throughout this work.

The author wishes to express his grateful acknowledgment to Dr. Noriyuki Kioka, Kyoto University, for his valuable suggestions, discussions, technical advice throughout this work, and critical reading of this manuscript.

The author expresses his sincere appreciation to Dr. Hiroshi Sakai, associate professor of Kyoto University, and Dr. Kazumitsu Ueda for their useful discussions and technical advice.

The author appreciate to Dr. Yoshihiko Hosoi, Kinki University, and Dr. Masayasu Yamada, Kyoto University, for their technical advice and assistance in antibody production and for supplying rat tissues.

The author thanks to Dr. Makoto Takeuchi, Kirin Brewery Co., Ltd., for his helpful suggestions. The author also thanks to Dr. Sumihiro Hase, Osaka University, for supplying experimental materials.

Finally, the author is grateful to Mr. Kenji Iida, Jun Nishiu, Toshihiro Uchida, Masanobu Ono, Satoshi Okamoto, Shohei Sakata, Masahiko Akamatsu, and other members of Laboratory of Biochemistry, Department of Agricultural Chemistry, Kyoto University, for their encouragement and helpful suggestions.

T. Fukada

Takashi Fukada