Genomic Cloning and Expression of Three Murine UDP-galactose: β-N-Acetylgalcosamine β1,3-Galactosyltransferase Genes

(Received for publication, September 30, 1997, and in revised form, October 21, 1997)

Thierry Hennet‡‡, André Dinter‡, Peter Kuhnert‡, Taj S. Mattu, Pauline M. Ruddy, and Eric G. Berger‡

From the ‡Institute of Physiology, University of Zurich, 8057 Zurich, Switzerland, the ¶Institute of Veterinary Bacteriology, University of Bern, 3012 Berne, Switzerland, and the §Glycobiology Institute, Department of Biochemistry, Oxford University, Oxford OX1 3QU, United Kingdom

Based on the detection of expressed sequence tags that are similar to known galactosyltransferase sequences, we have isolated three novel UDP-galactose:β-N-acetylgalcosamine β1,3-galactosyltransferase (β3GalT) genes from a mouse genomic library. The three genes, named β3GalT-I, -II, and -III, encode type II transmembrane proteins of 326, 422, and 331 amino acids, respectively. The three proteins constitute a distinct subfamily as they do not share any sequence identity with other eucaryotic galactosyltransferases. Also, the entire protein-coding region of the three β3GalT genes was contained in a single exon, which contrasts with the genomic organization of the β1,4- and α1,3-galactosyltransferase genes. The three β3GalT genes were mainly expressed in brain tissue. The expression of the full-length murine genes as recombinant baculoviruses in insect cells revealed that the β3GalT enzymes share the same acceptor specificity for β-linked GlcNAc, although they differ in their $K_m$ for this acceptor and the donor UDP-Gal. The identification of β3GalT genes emphasizes the structural diversity present in the galactosyltransferase gene family.

The structural diversity of glycoconjugates results from the combined action of glycosidases and glycosyltransferases that are differentially expressed throughout tissues. Until recently, each carbohydrate linkage was believed to be catalyzed by one single glycosyltransferase. Deducing from this assumption, it was estimated that ~250 genes were required for shaping the glycoconjugate repertoire of vertebrates (1). This count is constantly corrected to a higher mark as many glycosyltransferase activities are found to be encoded by multiple related genes. This apparent redundancy likely reflects a selective preference of glycosyltransferase isozymes for related although distinct acceptor structures, as observed for α1,3-fucosyltransferase (2) and GalNAc α2,6-sialyltransferase (3–5) isozymes.

Various strategies have been applied to identify genes related to known glycosyltransferase genes. The use of low stringency hybridization was only successful to detect very closely related genes, as was the case for α1,3-fucosyltransferases (6, 7). Often, the comparison of related glycosyltransferase genes unveils conserved regions that correspond to residues essential for catalytic activity. Such comparisons performed on sialyltransferase genes led to the detection of a stretch of conserved sequences called the sialyl motif (8). By designing degenerate primers according to the sialyl motif, it has been possible to isolate additional sialyltransferase genes by polymerase chain reaction (PCR) (9). The recent availability of “single-pass” cDNA sequences, or expressed sequenced tags (ESTs), represents a tremendous amount of information that can be analyzed to retrieve sequences sharing limited but significant sequence similarity. This approach is particularly well suited to detect similarity at the level of the protein sequence, which would be difficult to isolate by DNA-based retrieval techniques such as cross-hybridization and PCR.

We have applied the EST screening procedure to identify genes encoding β1,3-galactosyltransferases, which refers to an activity yet uncharacterized at the genetic level. To this end, we have probed the EST division of the GenBank™/EMBL Data Bank with the sequence of a β-galactosyltransferase protein that was believed to encode a β1,3-galactosyltransferase (β3GalT) (10). This procedure led to the identification of novel open reading frames (ORFs) that were similar to the query β-galactosyltransferase, but distinct from the previously described α1,3-galactosyltransferase (11, 12) and β1,4-galactosyltransferase (13) genes. We have used probes designed after these ORFs to isolate the corresponding genes from a mouse genomic library. We have cloned three genes coding for proteins of 326, 422, and 331 amino acids. The three proteins are type II transmembrane proteins containing a single transmembrane domain of 17–19 amino acids. The sequence identity between the three proteins was the highest in the luminal domain, where it ranged from 35 to 51%. Several conserved motifs were detected in the catalytic domain. None of these motifs were present in α1,3- and β1,4-galactosyltransferases, although one motif was identified in some bacterial galactosyltransferases. The expression of the three murine genes as recombinant baculoviruses in SF9 insect cells confirmed them as encoding UDP-galactose:β-N-acetylgalcosamine β1,3-galactosyltransferase enzymes. The genomic organization of the murine β3GalT-I, -II, and -III genes differed from that of β1,4-galactosyltransferase (14) and α1,3-galactosyltransferase (15) genes. The entire protein-coding region of the murine β3GalT genes was contained in a single exon. Northern blot analysis

* This work was supported by Swiss National Science Foundation Grant 3100-046836.96 (to E. G. B. and T. H.) and partially supported by European Commission Grant BIO4-CT95-0138 (to E. G. B.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) AF029790–AF029792.

† To whom correspondence should be addressed: Inst. of Physiology, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland. Tel.: +41-1-635-5080; Fax: +41-1-635-6814.

‡§ From the Institute of Physiology, University of Zurich, 8057 Zurich, Switzerland, the Institute of Veterinary Bacteriology, University of Bern, 3012 Berne, Switzerland, and the Glycobiology Institute, Department of Biochemistry, Oxford University, Oxford OX1 3QU, United Kingdom

¶ From the Department of Biochemistry, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland, and the Bacteriology, University of Berne, 3012 Berne, Switzerland, and the Institute of Veterinary Bacteriology, University of Bern, 3012 Berne, Switzerland.

This paper is available on line at http://www.jbc.org
ESTs that were similar to a human β-galactosyltransferase gene (βGalT) were retrieved using a tblastn search of the αEST division of the GenBank™/EMBL Data Bank. The EST sequences were aligned, and those that overlapped were assembled to form larger ORFs. ESTs representative of each ORF, as well as the human βGalT gene, were used as DNA probes to screen a mouse genomic DNA library. Seven mouse genomic fragments were isolated and grouped according to their specificity for the DNA probes. ORF2 and ORF3 were found to be different parts of the same gene, as genomic fragments 5 and 14 hybridized to both DNA probes H70574 and R15977.

### EXPERIMENTAL PROCEDURES

**Materials**—The mouse 129/SvJ genomic library, Duralose-UV nitrocellulose membranes, and the Bluescript SKII plasmid were purchased from Stratagene. The Vent polymerase, restriction enzymes, and DNA-modifying enzymes were obtained from New England Biolabs Inc. The AmpliTaq FS Dye Terminator Cycle Sequencing kit, the CentriSep columns, and the ABI 310 Genetic Analyzer were from Applied Biosystems. The baculovirus expression system Bac-to-Bac and the ABI 310 Genetic Analyzer were from Perkin-Elmer. The mouse 129/SvJ genomic library was plated onto Luria-Bertani agar, and plaques were lifted onto Duralose-UV filters. Filters were hybridized to [32P]CTP-labeled DNA probes designed after the EST sequences that were similar to a human galactosyltransferase gene (17), and the Galβ1,4GlcNAcβ-pNP standard was purchased from Sigma. The HPLC system consisted of a Jour Research X-ACT degasser, two Waters 510 pumps, a Waters 717 autoinjector, a Jasco C0965 Column Heater, a Waters 2487 multiwavelength detector (214, 262, and 350 nm). The system was controlled using Waters Millennium (Version 2.1.5.2). UDP-[14C]Gal and Hybrid-N nylon membranes were from Amersham Corp. Sep-Pak C18 cartridges were obtained from Waters, and GFA glass-microfiber filters were from Whatman. [32P]CTP was purchased from Hartmann Analytics (Braunschweig, Germany).

**Genomic Cloning of βGalT Genes**—The mouse genomic library was plated onto Luria-Bertani agar, and plaques were lifted onto Duralose-UV filters. Filters were hybridized to [32P]CTP-labeled DNA probes designed after the EST sequences that were similar to a human galactosyltransferase gene (10). The genomic inserts from the phage clones were subcloned as M13mp18 inserts into the pBluescript vector. Sequenced products were further subcloned into pBluescript and sequenced.

**PCR cloning**—PCR cloning was done using the AmpliTaq FS Dye Terminator Cycle Sequencing kit with the vectors T3 and T7 matching primers flanking the cloned inserts in pBluescript. Sequenced products were purified with CentriSep columns and run on an ABI 310 Genetic Analyzer. Primer walking was used afterward for sequence determination throughout the complete coding region of every βGalT gene. The sequences of the coding and complementary strands were determined and edited using the Sequencher program (Gene Codes, Ann Arbor, MI).

**Cloning of Recombinant Baculoviruses and Expression in Sf 9 Cells**—Recombinant baculoviruses were generated using the transposon-mediated insertion system developed by Luckow et al. (18) as purchased from Life Technologies, Inc. The mouse βGalT-I, -II, and -III genes were amplified by PCR from the genomic clones using 5′-primers containing an EcoRI site (βGalT-I) or a BamHI site (βGalT-II and -III) just upstream of the ATG codon. The 3′-primers for the three βGalT genes contained a XhoI site following the stop codon. The conditions for PCR were 20 cycles at 95 °C for 45 s, 55 °C for 30 s, and 72 °C for 60 s using the Vent polymerase and 50 ng of plasmid DNA as template. The resulting PCR fragments were subcloned into the baculovirus donor vector FastBac1 opened at the EcoRI and XhoI sites (βGalT-I) or at the BamHI and XhoI sites (βGalT-II and -III). Following the instructions of the manufacturer, the FastBac1-βGalT plasmids were transformed into DH10Bac cells for transposition into the bacmid. The integration of the βGalT genes in the bacmids was confirmed by PCR using the 5′-primers specific for each βGalT-I gene and the m13r primer (5′-AACAGCTATGACCATGATTACG-3′), which binds to a bacmid region outside of the transposition element. Sf9 cells (106) were transfected with 1 μg of recombinant bacmid DNA using Cellfectin (Life Technologies, Inc.) and incubated for 3 days at 27 °C. The culture medium was saved as the primary baculovirus stock. The primary stocks were amplified three times to reach titers of ~108 plaque-forming units/ml. Sf9

### TABLE I

| ESTs similar to human βGalT | Assembled ORFs | DNA probes | Mouse genomic fragments |
|-----------------------------|---------------|-----------|------------------------|
| R13867, D81474, H14861      | ORF1: R13867  | D81474 (311 bp) | 6 (15.5 kbp)*          |
| R61672, R13064, R15977, T93289, H70585, H70574, Z43147, Z45582 | ORF2: T31289 H70585 H70574 Z45582 | R13064 (784 bp) | 5 (12.4 kbp) 14 (17.0 kbp) |
| R61672                      | ORF3: R15977  | R15977 (496 bp) | βGalT (981 bp) 2 (16.9 kbp) 15 (14.8 kbp) |

* kbp, kilobase pairs.

showed that the three genes were mainly expressed in brain tissue and also at lower levels in other tissues.

**FIG. 1.** Schematic alignment of EST sequences with the murine βGalT genes. The EST fragments similar to the βGalT genes were aligned with the corresponding βGalT genes. The protein-coding regions of the three murine βGalT genes are given as black rectangles. National Center for Biotechnology Information sequence identification numbers are provided to the left of the lines representing the EST fragments. The vertical dotted lines indicate 500-bp increments.
cells were infected at a multiplicity of 10 and further incubated at 27 °C for 72 h.

**Galactosyltransferase Activity Assays—**Baculovirus-infected Sf 9 cells were washed with phosphate-buffered saline and lyzed in 2% Triton X-100 for 15 min on ice. Nuclei were removed from the lysates by centrifugation at 500 × g. 10 ml of the lysates were assayed for 30 min at 37 °C in 50-ml reactions of 50 mM cacodylate buffer, pH 6.6, 10 mM MnCl2, 0.5 mM UDP-Gal, 1% Triton X-100, 10% Me2SO, and varying acceptors at different concentrations (see Table II). 105 cpm of UDP-[14C]Gal (164 pmol) were added to standard assays, whereas 2.5 × 105 cpm of UDP-[14C]Gal (410 pmol) were added when kinetic parameters were determined. The reactions were stopped by dilution with 500 ml of H2O and applied onto a Sep-Pak C18 cartridge. The cartridge was washed with 15 ml of H2O and eluted with 5 ml of methanol. The samples containing asialo-ovine submaxillary mucin as acceptor were stopped by adding 1 ml of cold 15% trichloroacetic acid and 5% phosphotungstic acid solution. The precipitates were collected on GFA glass-microfiber filters, which were washed with 5 ml of ice-cold ethanol and dried at 80 °C for 30 min. The amount of [14C]Gal in the methanol eluates and on the GFA filters was measured in a liquid scintillation counter (Rackbeta, Pharmacia Biotech Inc.).

**Analysis of Linkage Specificity of β3GalT Enzymes—**Sep-Pak C18 eluates of β3GalT-I, -II, and -III assays using unlabeled UDP-Gal with GlcNAc-β-pNP as acceptor were dried by centrifugal evaporation and redissolved in 100 ml of H2O. A 20-ml aliquot of this material was applied onto a Sep-Pak C18 cartridge. The cartridge was washed with 15 ml of H2O and eluted with 5 ml of methanol. The samples containing asialo-ovine submaxillary mucin as acceptor were stopped by adding 1 ml of cold 15% trichloroacetic acid and 5% phosphotungstic acid solution. The precipitates were collected on GFA glass-microfiber filters, which were washed with 5 ml of ice-cold ethanol and dried at 80 °C for 30 min. The amount of [14C]Gal in the methanol eluates and on the GFA filters was measured in a liquid scintillation counter (Rackbeta, Pharmacia Biotech Inc.).

**Exoglycosidase Digestion Conditions—**Streptococcus pneumoniae β-galactosidase (Oxford GlycoSciences) digestions were performed at 37 °C for 16–24 h in 100 mM sodium acetate, pH 6.0, and 27 milliunits/ml. An aliquot of the aqueous sample was mixed with acetonitrile at a ratio of 20:80 and applied to the NP-HPLC column.

**RNA Preparation and Northern Blotting—**Total RNA from 8-week-old ICR mouse tissues was isolated with guanidinium isothiocyanate, followed by centrifugation on cesium chloride cushions (22). 5 mg of total RNA from each tissue were separated on formaldehyde-agarose gels. RNA was either stained in ethidium bromide (0.5 μg/ml in 0.1 M ammonium acetate) or transferred to Hybond-N membranes by capillary elution. Filters were hybridized overnight at 42 °C with [32P]CTP-labeled probes corresponding to the full-length mouse β3GalT-I, -II, and -III genes. The filters were washed with 0.1 × SSC and 0.1% SDS up to 60 °C and exposed for 3 days at −20 °C using intensifying screens.

**RESULTS**

**Identification and Isolation of Mouse β3GalT Genes—**The translation initiation codon and stop codon are **boldface** and **underlined**. Amino acids are in single-letter notation. The amino acids corresponding to the putative transmembrane domain are **boldface** and **underlined**.

![FIG. 2. Nucleotide and predicted amino acid sequences of the murine β3GalT-I gene.](image)
using the tblastn algorithm (Version 1.4.9) (23). ESTs bearing the National Center for Biotechnology Information sequence identification numbers R13867, D81474, H14861, R61672, R13064, R15977, T31289, H70585, H70574, Z43147, and Z45582 were identified as being significantly similar to the query sequence (Table I). It is noteworthy that we have not detected any EST that is identical to the probed \( \beta_{x} \)GalT, indicating that this gene is not represented yet in the EST database. The retrieved EST sequences were aligned with the query sequence, and partially overlapping ESTs were assembled to build larger ORFs (Table I). DNA fragments corresponding to the \( \beta_{x} \)GalT gene and to EST probes D81474, H70574, and R15977 were amplified by PCR from a pool of human T-cell cDNAs. The resulting fragments were used as DNA probes in screening a mouse genomic DNA library. We could isolate seven clones, which represented three distinct genomic regions. The three fragments 6, 10, and 13 encompassed the same genomic region and hybridized to EST probe D81474. Both genomic fragments 5 and 14 hybridized to EST probes H70574 and R15977, indicating that the assembled ORF2 and ORF3 were parts of the same gene. DNA fragments 2 and 15 included the murine \( \beta_{x} \)GalT gene (Table I). The genomic regions hybridizing to the probes were sequenced, and the deduced genes were tentatively named \( \beta_{x} \)GalT-I, -II, and -III. The \( \beta_{x} \)GalT genes were aligned with the originally identified ESTs (Table I) and additional related ESTs that were retrieved from the EST database with the blastn algorithm (Fig. 1). This comparison confirmed ORF2 and ORF3 as being parts of the \( \beta_{x} \)GalT-III gene. Also, it appeared that the sequence of the human \( \beta_{x} \)GalT-III gene could be entirely determined by joining contiguous EST fragments. No EST corresponding to the \( \beta_{x} \)GalT-I gene could be identified in the EST database using the blastn, blastx, and tblastx programs.

### Primary Structure and Genomic Organization of Murine \( \beta_{x} \)GalT Genes

![Fig. 3](image)

**Fig. 3. Nucleotide and deduced amino acid sequences of the murine \( \beta_{x} \)GalT-II gene.** The first-in-frame ATG codon at 64 bp is given as the translation initiation codon and is **boldface** and **underlined**. The second in-frame ATG codon at 103 bp is **boldface**. The stop codon at 1330 bp is **boldface** and **underlined**. Single-letter notation is used for the amino acids. The amino acids corresponding to the putative transmembrane domain are **boldface** and **underlined**.
-III genes contain two in-frame ATG codons spaced by 12 and 11 amino acids, respectively. It is unclear whether the first, second, or both ATG codons are used as the start codon. It appears, however, that in both genes, the proximal ATG codons have stronger contexts for translation initiation (24). A Kyte-Doolittle hydropathy analysis detected single membrane-spanning domains of 19, 19, and 17 amino acids for the β3GalT-I, -II, and -III proteins, respectively (Figs. 2–4). The transmembrane domains are located close to the N terminus of each protein, which is typical of the type II transmembrane glycosyltransferase protein family. The protein-coding regions of β3GalT-I, -II, and -III are contained in a single exon. However, more proximal untranslated exons might exist, as is the case for other glycosyltransferase genes (25, 26). This assumption is supported by the fact that no TATA box sequence was detected in front of the β3GalT genes. Also, we noticed that the cDNAs of the three β3GalT genes contain 500 bp of untranslated sequence that differs from the genomic sequence 5' of the ATG codons (data not shown).

Homology between β3GalT Proteins—The three mouse β3GalT protein sequences were compared using the ClustalW alignment program (27) (Fig. 5). The greatest identity was observed in the catalytic domains of the proteins. In this region, β3GalT-I was 51% identical to β3GalT-II and 36% identical to β3GalT-III, whereas β3GalT-II was 35% identical to β3GalT-III. The β3GalT enzymes shared no sequence similarity in their N-terminal part corresponding to the cytoplasmic, transmembrane, and stem regions. We noticed that the larger size of β3GalT-II reflects a longer stem region, whereas the size of the catalytic domains is conserved among the three β3GalT enzymes. The positions of six cysteine residues in the catalytic domains of the three β3GalT proteins were conserved (Fig. 5). This supports the formation of up to three disulfide bridges that would confer a similar structure to the three β3GalT enzymes. Such related conservations of cysteine residues have been previously observed in the sialyltransferase protein family (28). Several conserved motifs emerged from the alignment of the β3GalT protein sequences (Fig. 5). Notably, we did not detect these same motifs in β1,4- and α1,3-galactosyltransferase sequences, indicating that there are no consensual galactosyl motifs in vertebrate galactosyltransferases. However, a variation of the β3GalT motif EDVYVG was recognized in the bacterial galactosyltransferases Lex1 (29), LgtB (30), and LgtE (31) (Fig. 6). This conservation was also detected by others and described as a putative galactosyltransferase motif (32).

Tissue Expression of β3GalT Genes—The expression pattern of the three β3GalT genes was investigated in adult mouse tissues. We have found the β3GalT genes to be predominantly expressed in brain tissue, whereas lower transcript levels were also detected in other tissues (Fig. 7). Beside brain tissue, the 7-kilobase β3GalT-I mRNA was detected at low levels in all tissues examined, indicating a possible constitutive expression. Two additional smaller transcripts were also observed in colon tissue. The β3GalT-II transcript was 3 kilobases long and was present in brain and heart tissue and could barely be detected in ovaries, colon, and lymph nodes. The expression of β3GalT-III yielded a 2.7-kilobase transcript that was restricted to brain, testes, ovaries, and uterus. In addition, the three β3GalT genes were expressed in lymphoid cells as detected by reverse transcriptase-PCR (data not shown).
Expression of β3GalT Genes in Sf9 Cells—Sf9 insect cells were infected with recombinant baculoviruses expressing the three β3GalT genes under the control of the polyhedrin promoter. The lysates of Sf9 cells harvested at 72 h post-infection were assayed for galactosyltransferase activity using various acceptor saccharides (Table II). The highest activity was detected with β-conjugated GlcNAc acceptors like GlcNAc-β-pNP and GlcNAc-β-benzyl (Bzl). By contrast, the α-conjugated Glc-NAc acceptors yielded a 10–20-fold lower activity compared with the β-conjugates. No significant galactosyltransferase activity was measurable toward GalNAc acceptors. This finding was corroborated by assays performed in the presence of the high molecular weight acceptor asialo-ovine submaxillary mucin, which carries multiple O-linked GalNAc monosaccharides (16). Also, the three β3GalT enzymes failed to transfer GlcNAc or GalNAc to the acceptors GlcNAc-β-bzl, Gal-β-bzl, and Gal-NAc-β-bzl, showing that the proteins are solely galactosyltransferases (data not shown). The activity of the three β3GalT proteins was strictly dependent on the divalent cation Mn^2+ (data not shown). Also, the proteins were more active in the presence of 1% Triton X-100 (data not shown). The enzymatic properties of the three β3GalT proteins were determined for the donor UDP-Gal and the acceptor GlcNAc-β-pNP (Table III).

When compared with β3GalT-I, we found that β3GalT-II and β3GalT-III exhibited 3- and 2-fold higher Km values for Glc-NAc-β-pNP, respectively. The opposite was observed with the Km for the donor UDP-Gal, for which β3GalT-I showed 4- and 2-fold greater Km values than β3GalT-II and β3GalT-III, respectively.

Analysis of Linkage Specificity of β3GalT—The products of the galactosyltransferase activity assays of the lysates from baculovirus-infected Sf9 cells expressing the three β3GalT genes under the control of the polyhedrin promoter. The lysates of Sf9 cells harvested at 72 h post-infection were assayed for galactosyltransferase activity using various acceptor saccharides (Table II). The highest activity was detected with β-conjugated GlcNAc acceptors like GlcNAc-β-pNP and GlcNAc-β-benzyl (Bzl). By contrast, the α-conjugated Glc-NAc acceptors yielded a 10–20-fold lower activity compared with the β-conjugates. No significant galactosyltransferase activity was measurable toward GalNAc acceptors. This finding was corroborated by assays performed in the presence of the high molecular weight acceptor asialo-ovine submaxillary mucin, which carries multiple O-linked GalNAc monosaccharides (16). Also, the three β3GalT enzymes failed to transfer GlcNAc or GalNAc to the acceptors GlcNAc-β-bzl, Gal-β-bzl, and Gal-NAc-β-bzl, showing that the proteins are solely galactosyltransferases (data not shown). The activity of the three β3GalT proteins was strictly dependent on the divalent cation Mn^2+ (data not shown). Also, the proteins were more active in the presence of 1% Triton X-100 (data not shown). The enzymatic properties of the three β3GalT proteins were determined for the donor UDP-Gal and the acceptor GlcNAc-β-pNP (Table III).

When compared with β3GalT-I, we found that β3GalT-II and β3GalT-III exhibited 3- and 2-fold higher Km values for Glc-NAc-β-pNP, respectively. The opposite was observed with the Km for the donor UDP-Gal, for which β3GalT-I showed 4- and 2-fold greater Km values than β3GalT-II and β3GalT-III, respectively.

Analysis of Linkage Specificity of β3GalT—The products of
Cloning and Expression of Three Murine GlcNAc βGalT Genes

Fig. 7. Differential expression of the three βGalT genes in murine tissues. Northern blot analysis was performed using 5 µg of total RNA from each tissue sample, as is visible on the ethidium bromide staining (bottom panel). The blots were probed with full-length βGalT gene probes as detailed under “Experimental Procedures.”

Table II

| Acceptor substrate specificity of the three mouse βGalT enzymes | 55SMock | βGalTI | βGalTII | βGalTIII |
|---------------------------------------------------------------|--------|--------|--------|--------|
| Acceptor                                                     | pmol/min/mg protein | pmol/min/mg protein | pmol/min/mg protein | pmol/min/mg protein |
| GlcNAc-β-pNP (10 mM)                                         | 14     | 10,616 | 407    | 313    |
| GlcNAc-β-Bal (10 mM)                                         | 8      | 6834   | 173    | 115    |
| GlcNAc-α-Bal (10 mM)                                         | 7      | 306    | 15     | 16     |
| GlcNAcβ1,3GlcNAc-α-pNP (2.5 mM)                              | 9      | 123    | 8      | 6      |
| GalNAc-β-pNP (10 mM)                                         | 10     | 89     | 21     | 11     |
| GalNAc-α-pNP (10 mM)                                         | 266    | 275    | 268    | 261    |
| Asialo-OSM (3 mg/ml)                                         | 35     | 38     | 34     | 35     |

a Lysate of SF9 cells infected with a wild-type baculovirus.
b Asialo-OSM, asialo-ovine α submaxillary mucin.

Table III

| Kinetic parameters of the three mouse βGalT enzymes         | βGalTI | βGalTII | βGalTIII |
|-------------------------------------------------------------|--------|--------|--------|
| Substrate                                                   | Kn     | vmax   | Kn     | vmax   | Kn     | vmax   |
| GlcNAc-β-pNP (10 mM)                                        | 1.18   | 29.1   | 32.3   | 0.12   | 23.3   | 0.19   |
| UDP-Gal                                                    | 2.3    | 2.62   | 0.8    | 0.08   | 1.5    | 0.21   |

Fig. 8. HPLC analysis of cell lysates from control and βGalT-expressing systems. A, NP-HPLC profiles of the lysate from βGalT-I-expressing cells (panel i) and the sugar standards GlcNAc-β-pNP (peak a), Galβ1,3/GlcNac-β-pNP (peak b), UDP (peak c), and UDP-Gal (peak d) (panel ii). In panel i, peak p indicates that the peak contains protein. B, reverse phase HPLC profiles of the Galβ1,4GlcNAc-β-pNP standard (panel i); the disaccharide peak (peak g) from A (panel i) for βGalT-I (panel ii), βGalT-II (panel iii), and βGalT-III (panel iv); and the Galβ1,3GlcNAc-β-pNP standard (panel v).

DISCUSSION

Based on the retrieval of ESTs sharing similarity with a β-galactosyltransferase gene sequence, we have identified three novel murine βGalT genes. The identification of multiple genes encoding a βGalT activity is consistent with recent similar findings made for other glycosyltransferase families (33). The presence of several isozymes can be interpreted as a way to maintain a greater potential to glycosylate different related acceptor structures, thereby increasing the versatility of glycosyltransferases in shaping the glycoconjugate repertoire.

The three βGalT proteins are heterologous in size, with βGalT-I and βGalT-III representing to date the smallest mammalian glycosyltransferases after the GalNAc α2,6-sialyltransferase-III (5). As observed for many glycosyltransferase genes, two in-frame ATG codons are present in the βGalT-II and βGalT-III genes. The two ATG codons are located 5’ of the region encoding the transmembrane domain, and an alternate initiation of translation would unlikely affect the localization and activity of the two proteins. The three βGalT proteins represent a distinct family from the β1,4- and α1,3-galactosyltransferase proteins, as no sequence similarity can be detected between βGalT enzymes and the other galactosyltransferases. This difference is emphasized by comparing the genomic organization of βGalT genes with that of β1,4- and α1,3-galactosyltransferase genes, which contain multiple exons coding for protein sequence (14, 15). These distinctions suggest that the family of βGalT genes separated early in evolution from the other galactosyltransferases and possibly expanded by gene duplication. It remains open as to whether the βGalT gene family comprises additional related members. The search for ESTs similar to the three murine βGalT sequences supports the notion of a gene family consisting of 7–10 members.

The alignment of the three βGalT protein sequences highlighted various conserved regions that were all located in the postulated catalytic domain. These βGalT putative motifs were absent in other vertebrate galactosyltransferase protein...
sequences. However, a short β3GalT motif was identified in bacterial galactosyltransferases, suggesting that these amino acids are essential for galactosyltransferase activity. Some of the β3GalT motifs were also present in the Drosophila proteins Fringe and Brainiac as outlined by Yuan et al. (32). Fringe (34) and Brainiac (35) have been described as secreted signaling molecules, although their mechanism of action has not been elucidated yet. The similarity between Fringe, Brainiac, and β3GalT enzymes suggests that two Drosophila proteins may represent additional members of the β3GalT family.

The acceptor specificity of the three β3GalT enzymes was restricted to β-conjugated GlcNAc. However, the high Km values measured for the acceptor GlcNAc-β-pNP indicate that the physiologic acceptors are likely more specific oligomeric structures. The β3GalT enzymes enable the formation of type 1 carbohydrate chains, which are widely distributed on glycolipids and glycoproteins throughout tissues. Being the acceptor of the Lewis enzyme, the type 1 disaccharide Galβ1,3GlcNAc is related to the occurrence of Leα and Leβ antigens (36, 37).

The acceptor specificity of the three β3GalT enzymes was elucidated yet. The similarity between Fringe, Brainiac, and β3GalT proteins may represent additional members of the β3GalT family.

The difference in sequence and genomic organization between the β3GalT genes highlights the structural diversity encountered in the galactosyltransferase gene family. The difference in sequence and genomic organization between β3GalT and other vertebrate galactosyltransferase genes suggests a distinct evolution. The structural information gained from the cloning of the three β3GalT genes provides a way to identify additional members of this branch of the galactosyltransferase family.

Acknowledgments—We thank Dr. Martine Malissard for advice and criticism throughout this work and Marianne Farah for technical assistance.

REFERENCES
1. Varki, A., and Marth, J. D. (1995) Semin. Dec. Biol. 6, 127–138
2. Legault, D. J., Kelly, R. J., Natsuka, Y., and Lowe, J. B. (1995) J. Biol. Chem. 270, 29987–29996
3. Karuosa, N., Hamamoto, T., Lee, Y.-C., Nakaoka, T., Kojima, N., and Tsuji, S. (1994) J. Biol. Chem. 269, 1402–1409
4. Karuosa, N., Kojima, N., Inoue, M., Hamamoto, T., and Tsuji, S. (1994) J. Biol. Chem. 269, 19048–19053
5. Sjöberg, E. R., Kitagawa, H., Glushka, J., van Halbeek, H., and Paulson, J. C. (1996) J. Biol. Chem. 271, 7450–7459
6. Weston, B. W., Smith, P. L., Kelly, R. J., and Lowe, J. B. (1992) J. Biol. Chem. 267, 4152–4160
7. Weston, B. W., Nair, R. P., Larsen, R. D., and Lowe, J. B. (1992) J. Biol. Chem. 267, 11504–11507
8. Kitagawa, H., and Paulson, J. C. (1994) J. Biol. Chem. 269, 1384–1401
9. Krohne, K. K. (July 5, 1994) Japanese Patent JP6181759
10. Yuasa, H., Shaper, J. H., van den Eijnden, D. H., Van Turen, A. J., and Shaper, N. L. (1990) J. Biol. Chem. 264, 14296–14297
11. Yamamoto, F., Clausen, H., White, T., Marken, J., and Hakomori, S. (1990) Nature 345, 229–233
12. Shaper, N. L., Shaper, J. H., Meath, J. L., Fox, J. L., Chang, H., Kirsch, I. R., and Hollis, G. F. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 1573–1577
13. Russo, R. N., Shaper, N. L., and Shaper, J. H. (1990) J. Biol. Chem. 265, 3324–3331
14. Joziasse, D. H., Shaper, N. L., Kim, D., van den Eijnden, D. H., and Shaper, J. H. (1992) J. Biol. Chem. 267, 5534–5541
15. Hill, H. D., Jr., Reynolds, J. A., and Hill, R. L. (1977) J. Biol. Chem. 252, 3791–3798
16. Malissard, M., Borsig, L., Di Marco, S., Guerri, M. G., Krägl, U., Wondrey, C., and Berger, E. G. (1996) Eur. J. Biochem. 239, 340–348
17. Luckow, V. A., Lee, S. C., Barry, G. F., and Olins, P. O. (1993) J. Virol. 67, 4566–4579
18. Guile, G. R., Rudd, P. M., Wing, D. R., Prime, S. B., and Dwek, R. A. (1996) J. Biol. Chem. 271, 7450–7459
19. Altschul, S. F., Boguski, M. S., Gish, W., and Wootton, J. C. (1994) Nature 345, 229–233
20. Jacob, G. S., and Scudder, P. (1994) J. Biol. Chem. 270, 1402–1409
21. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Guide to Genetic Manipulation, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
22. Kozak, M. (1986) Cell 44, 283–292
23. Hub, R., Sankar, S., Spruijt, M. P., Hoppener, J. W., Dunn, R., and Schachter, H. (1991) Biochem. Biophys. Res. Commun. 16, 608–615
24. Koyama, H., Nishikawa, A., and Taniguchi, N. (1996) Cell Biochem. 238, 853–861
25. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) Nucleic Acids Res. 22, 4673–4680
26. Drickamer, K. (1993) Glycobiology 1, 2–3
27. Cope, L. D., Yogev, R., Mertsola, J., Latimer, J. L., Hansen, M. S., McCracken, G. H., Jr., and Hansen, E. J. (1991) Mol. Microbiol. 5, 1113–1124
28. Jennings, M. P., Hood, D. W., Peak, I. R., Virji, M., and Moxon, E. R. (1991) Mol. Microbiol. 1, 783–790
29. Dwek, R. A. (1991) Cell 88, 9–11
30. Harduin-Lepers, A., Recchi, M.-A., and Delannoy, P. (1995) Glycobiology 5, 741–758
31. Irvine, K. D., and Wieschaus, E. (1994) Mol. Microbiol. 11, 199–234
32. Yuan, Y. P., Schultz, J., Mlodzik, M., and Bork, P. (1997) Nucleic Acids Res. 25, 673–680
33. Frick, S. (1986) Cell 44, 283–292