Cloning of a Novel Member of the UDP-Galactose:β-N-Acetylglucosamine β1,4-Galactosyltransferase Family, β4Gal-T4, Involved in Glycosphingolipid Biosynthesis*

(Received for publication, April 27, 1998, and in revised form, August 20, 1998)

Tilo Schwientek‡, Raquel Almeida‡‡, Steven B. Levery§, Eric H. Holmes¶, Eric Bennett‡, and Henrik Clausen**

From the ‡School of Dentistry, University of Copenhagen, Nørre Allé 20, 2200 Copenhagen N, Denmark, the §Institute of Molecular Pathology and Immunology of University of Porto, Rua Dr. R. Frias s/n, 4200 Porto, Portugal, the ¶University of Georgia, Complex Carbohydrate Research Center, Athens, Georgia 30602, and the **Department of Cell Surface Biochemistry, Northwest Hospital, Seattle, Washington 98125

A novel putative member of the human UDP-galactose:β-N-acetylglucosamine β1,4-galactosyltransferase family, designated β4Gal-T4, was identified by BLAST analysis of expressed sequence tags. The sequence of β4Gal-T4 encoded a type II membrane protein with significant sequence similarity to other β1,4-galactosyltransferases. Expression of the full coding sequence and a secreted form of β4Gal-T4 in insect cells showed that the gene product had β1,4-galactosyltransferase activity. Analysis of the substrate specificity of the secreted form revealed that the enzyme catalyzed glycosylation of glycolipids with terminal β-GlcNAc; however, in contrast to β4Gal-T1, -T2, and -T3, this enzyme did not transfer galactose to asialo-agalactofetuin, asialo-agalacto-transferrin, or ovalbumin. The catalytic activity of β4Gal-T4 with monosaccharide acceptor substrates, N-acetylglucosamine as well as glucose, was markedly activated in the presence of α-lactalbumin. The genomic organization of the coding region of β4Gal-T4 was contained in six exons. All intron/exon boundaries were similarly positioned in β4Gal-T1, -T2, and -T3. β4Gal-T4 represents a new member of the β4-galactosyltransferase family. Its kinetic parameters suggest unique functions in the synthesis of neolactoseries glycosphingolipids.

A family of human UDP-galactose:β-N-acetylglucosamine β1,4-galactosyltransferases (β4Gal-Ts)1 was recently identified (1–3). Four genes within this family encode β4-galactosyltransferases, which form the Galβ1→4GlcNAc linkage (2, 4–6). The kinetic parameters and expression patterns of these enzymes differ and they are predicted to show some degree of overlap in biological function (2, 3, 6). Two β4-galactosyltransferases, β4Gal-T1 and -T2, can function as lactose synthases in the presence of α-lactalbumin (2, 3, 7), whereas β4Gal-T3 and β4Gal-T52 are largely insensitive to α-lactalbumin modulation (2, 6, 8). β4Gal-T1, -T2, and -T3 catalyze transfer of galactose to lactoseries glycosphingolipids, but β4Gal-T3 only efficiently catalyzes synthesis of the first N-acetylactosamine unit in lactoseries glycolipids (2). In contrast, β4Gal-T5 was reported to be inactive with a glycolipid substrate (Lc,Cer)2 as well as with the glycoprotein acceptor, asialo-agalacto-transferrin (6). A rat lactosylceramide synthase was recently purified and cloned by Nomura et al. (9), and it appears to represent the ortholog of the human member of the gene β4-galactosyltransferase family designated β4Gal-T6 (10). Thus, the formation of Galβ1→4GlcNAc linkages in different glycoconjugates and their varying oligosaccharide structures may be catalyzed by different β4-galactosyltransferases.

Analysis of ESTs suggested the existence of additional members of the human β4Gal-T gene family (1–3), and recently, Lo et al. (10) compared the full coding sequences of six homologous human genes and suggested a nomenclature based on sequence similarity: β4Gal-T1 (5, 11, 12), β4Gal-T2 (2), β4Gal-T3 (2), β4Gal-T5 (6), and β4Gal-T6 (9). So far, all genes except one, localized at chromosome 3q13.3 and designated β4Gal-T7, have been expressed and shown to represent β4-galactosyltransferases. β4Gal-T4 is the subject of this report.

The greater α-lactalbumin modulation, it is likely that the snail 4GalNAc-transferase activity is not responsive to α-lactalbumin modulation (14). In contrast, a snail β4GalNAc-transferase activity with acceptor substrate specificity similar to β4Gal-T1 exhibits sensitivity to α-lactalbumin modulation of the acceptor specificity (15). The donor substrate specificity of β4Gal-T1 is modulated by α-lactalbumin to include UDP-GalNAc, and the donor substrate specificity of the snail β4GalNAc-transferase activity is modulated to include UDP-Gal, albeit at much less efficiencies (15, 16). Given the similarities in donor substrate specificities and α-lactalbumin modulation, it is likely that the snail β1,4GalNAc-transferase will be homologous to the mammalian β4Gal-T gene family (15, 17). The GalNAcβ1→4GlcNAcβ1-R structure exists in man, but is mainly associated with N-linked glycolipids.

**To whom correspondence should be addressed: School of Dentistry, Nørre Allé 20, DK-2200 Copenhagen N, Denmark, Tel.: 45-35326835, E-mail: henrik.clausen@odont.ku.dk.

† The abbreviations used are: β4Gal-T, UDP-galactose:β-N-acetylglucosamine β1,4-galactosyltransferase; EST, expressed sequence tag; PCR, polymerase chain reaction.

1 The designations of glycosphingolipids are abbreviated according to the recommendations of the IUPAC-IUB Commission on Nomenclature (57).

2 The designations of glycosphingolipids are abbreviated according to the recommendations of the IUPAC-IUB Commission on Nomenclature (57).
glycans found on glycoprotein hormones (18). A putative glycoprotein $\beta 4$GalNAc-transferase with selective activity for $N$-glycans associated with a specific peptide sequence has been characterized (19); however, this enzyme may be unrelated to the $\beta 4$Gal-T gene family.

In the present study, we used human EST sequence information to identify and clone a novel member of the $\beta 4$Gal-T gene family, designated $\beta 4$Gal-T4. $\beta 4$Gal-T4 is an active UDP-Gal:GlcNAc $\beta 1,4$Gal-transferase with specificity for glycolipid substrates; however, it does not catalyze glycosylation of several glycoprotein acceptors, which are good substrates for other $\beta 4$Gal-transferases. $\beta 4$Gal-T4 exhibits $\alpha$-lactalbumin modulation that is similar to a previously characterized snail $\beta 1,4$Gal-Nac-transferase activity (15). The data demonstrate that members of the $\beta 4$Gal-T gene family have distinct functions in galactosylation of different glycoconjugates, and suggest that $\beta 4$Gal-T4 mainly plays a role in glycolipid biosynthesis.

**EXPERIMENTAL PROCEDURES**

*Identification of $\beta 4$Gal-T4—*The BLASTn and tBLASTn were used with the coding sequence of human $\beta 4$Gal-T3 to search the dbEST database at The National Center for Biotechnology Information (NCBI) as described previously (2). Overlapping segments of EST sequences were compiled and compared with known members of the human $\beta 4$Gal-T family. cDNA clones of ESTs with the longest inserts (Fig. 1) were obtained from Genome Systems Inc.

*Cloning and Sequencing of the Full Coding Sequence of $\beta 4$Gal-T4—*A large number of overlapping ESTs derived from a putative gene were identified and assembled by using Unigene (NCBI, transcript map A0043536). Five ESTs representing nearly the full coding sequence were selected (Fig. 1). Sequencing of EST clone 489768 revealed an open reading frame that encoded a sequence similar to $\beta 4$Gal-T3, except that the 5' sequence was shorter and the clone lacked a translational initiation codon. The genomic organizations of $\beta 4$Gal-T1, -T2, and -T3 genes were previously shown to be identical (2). Since the 5' sequence available from the $\beta 4$Gal-T4 EST composite was incomplete but likely to extend into the first coding exon, the 5' position of the open reading frame was obtained by sequencing a genomic P1 clone. Confirmatory sequencing was performed on a cDNA clone obtained by PCR on total cDNA from the human MKN45 gastric cancer cell line with the sense primer TSHC 25 (5' - GTCTACUAGGGATGTTTTC - 3') and the antisense primer TSHC 12 (5' - CCAGTGTGAGCCAAAGTGGTAC- 3'), for 30 cycles at 95°C, 15 s, 55°C, 20 s, 72°C, 2 min 30 s. The entire sequence was confirmed by sequencing genomic P1 clones. The composite sequence contained an open reading frame of 1032 base pairs encoding a putative protein with a type II domain structure (Fig. 2).

*Genomic Organization of $\beta 4$Gal-T4—*A human foreskin genomic P1 library (DuPont Merck Pharmaceutical Co. Human Foreskin Fibroblast P1 Library) was screened using the primer pairs TSHC14 (5' - GAGCGGCAAGCCGCATTTCG - 3') and TSHC15 (5' - GGATGAATGATGACC - 3'), which encode the full-length $\beta 4$Gal-T4. Two clones for $\beta 4$Gal-T4 (DPMC-HFF1-934 (G7) and DPMC-HFF1-1025 (A6)) were obtained from Genome Systems Inc. DNA from P1 phage were prepared as recommended by Genome Systems Inc. The entire coding sequence of the $\beta 4$Gal-T4 gene was sequenced in full in automated sequencing (ABI377, Perkin-Elmer) with dye terminator chemistry. Intron/exon boundaries were determined by comparison with the complete cDNA sequence, optimizing for the gt/ag rule (20). The 1477 base pairs of the $\beta 4$Gal-T4 was determined using 3' EST mapping data (NCBI).

*Expression of $\beta 4$Gal-T4 in Insect Cells—*An expression construct designed to encode amino acid residues 42–344 of $\beta 4$Gal-T4 was prepared by PCR using EST clone 489768, and the primer pair TSHC30 (5' - AGCGGATCCTAAGAAGAAGGAGTTACG- 3') and TSHC36 (5' - AGCGAATGCCGGTCATGAC - 3'), which are compatible with the gt/ag rule (20). The chromosomal localization of $\beta 4$Gal-T4 was determined using 3' EST mapping data (NCBI).

Expression constructs designed to encode amino acid residues 42–344 of $\beta 4$Gal-T4 was prepared by PCR using EST clone 489768, and the primer pair TSHC30 (5' - AGCGGATCCTAAGAAGAAGGAGTTACG- 3') and TSHC36 (5' - AGCGAATGCCGGTCATGAC - 3'), which are compatible with the gt/ag rule (20). The chromosomal localization of $\beta 4$Gal-T4 was determined using 3' EST mapping data (NCBI).

Expression of $\beta 4$Gal-T4 in Insect Cells—An expression construct designed to encode amino acid residues 42–344 of $\beta 4$Gal-T4 was prepared by PCR using EST clone 489768, and the primer pair TSHC30 (5' - AGCGGATCCTAAGAAGAAGGAGTTACG- 3') and TSHC36 (5' - AGCGAATGCCGGTCATGAC - 3'), which are compatible with the gt/ag rule (20). The chromosomal localization of $\beta 4$Gal-T4 was determined using 3' EST mapping data (NCBI).
A Novel Member of the β4-Galactosyltransferase Family, β4Gal-T4

β4Gal-T4

Tris (pH 7.5), 2 mM MnCl₂, 0.1% Triton CF-54, and 250 nmol), with 10 mM MnCl₂, 0.1% Triton X-100, 90 mM MnCl₂, 0.1% Triton CF-54, and 100 secreted enzyme in 100-

-32P]dCTP and an oligonucleotide labeling kit (Amersham Pharmacia Biotech). The probe was random primer-labeled using

The hydrophobic segment representing the putative transmembrane domain is

A potential polyadenylation signal is indicated in boldface underlined type.

Fig. 2. Nucleotide sequence and predicted amino acid sequence of human β4Gal-T4. The amino acid sequence is shown in single-letter code. The hydrophobic segment representing the putative transmembrane domain is underlined with a double line (Kyte & Doolittle, window of 8; Ref. 29). Three consensus motifs for N-glycosylation are indicated by asterisks. The location of the primers used for preparation of the expression constructs are indicated by single underlining.

RESULTS

Identification and Cloning of Human β4Gal-T4—The strategy outlined in Fig. 1 produced a novel gene with significant sequence similarity to β4Gal-T3 and other members of the β4Gal-T gene family. A multiple sequence alignment of six human β4Gal-transferases is shown in Fig. 3. The β4Gal-T4 gene has highest sequence similarity to β4Gal-T3. Sequence similarities among the six human genes are found predominantly in the central regions; there were no significant similarities in the NH₂-terminal regions. Several sequence motifs in the putative catalytic domains are conserved among all the transferases (1). Importantly, four cysteine residues are conserved in all β4Gal-transferases; a fifth cysteine residue in the β4Gal-T1 expressed in insect cells (Calbiochem) was used as controls. Reaction products were quantified by chromatography on Dowex-1. Assays with glycoprotein acceptors were performed with the standard reaction mixture modified to contain 100 mM 

The strat-

Glycolipid Substrate Specificity of β4Gal-T4—Analysis of acceptor specificity with glycolipid substrates was performed with semipurified enzyme in 100-mM reaction mixtures containing 25 mM Tris (pH 7.5), 4 mM MnCl₂, 0.1% Triton X-100, 90 μM UDP-

A human multiple tissue Northern blot, MTN I (CLONTECH), was probed overnight at 42 °C as described previously (21), and washed twice for 10 min each at room temperature with 2 × SSC, 0.1% SDS; and once for 10 min with 0.1 × SSC, 0.1% SDS at 55 °C.

RESULTS

Identification and Cloning of Human β4Gal-T4—The strategy outlined in Fig. 1 produced a novel gene with significant sequence similarity to β4Gal-T3 and other members of the β4Gal-T gene family. A multiple sequence alignment of six human β4Gal-transferases is shown in Fig. 3. The β4Gal-T4 gene has highest sequence similarity to β4Gal-T3. Sequence similarities among the six human genes are found predominantly in the central regions; there were no significant similarities in the NH₂-terminal regions. Several sequence motifs in the putative catalytic domains are conserved among all the transferases (1). Importantly, four cysteine residues are conserved in all β4Gal-transferases; a fifth cysteine residue in the C-terminal end of β4Gal-T1 is substituted by a tyrosine in the other transferases (Fig. 3) (3). N-Linked glycosylation sites are not generally conserved in glycosyltransferase species homologues or within different members of glycosyltransferase gene families; however, a single N-linked site in the C-terminal regions of β4Gal-T2, -T3, -T4, -T5, and -T6 is conserved (Fig. 4). Similarly, a single site in the central region of the putative catalytic domains of four β2Gal-transferases was conserved (26).

The predicted coding region of β4Gal-T4 has a single initiation codon in agreement with Kozak's rule (27), which precedes a sequence encoding a potential hydrophobic transmembrane segment (Figs. 2–4). The predicted coding sequence indicates that β4Gal-T4 is a type II transmembrane glycoprotein with an N-terminal cytoplasmic domain of 14 residues, a transmembrane segment of 20 residues, and a stem region and catalytic domain of 310 residues with three potential N-linked glycosylation sites (28). One N-linked site is located in the putative membrane segment of 20 residues, and a stem region and catalytic domain of 310 residues with three potential N-linked glycosylation sites (28). One N-linked site is located in the putative

The predicted coding region of β4Gal-T4 has a single initiation codon in agreement with Kozak's rule (27), which precedes a sequence encoding a potential hydrophobic transmembrane segment (Figs. 2–4). The predicted coding sequence indicates that β4Gal-T4 is a type II transmembrane glycoprotein with an N-terminal cytoplasmic domain of 14 residues, a transmembrane segment of 20 residues, and a stem region and catalytic domain of 310 residues with three potential N-linked glycosylation sites (28). One N-linked site is located in the putative

The predicted coding region of β4Gal-T4 has a single initiation codon in agreement with Kozak's rule (27), which precedes a sequence encoding a potential hydrophobic transmembrane segment (Figs. 2–4). The predicted coding sequence indicates that β4Gal-T4 is a type II transmembrane glycoprotein with an N-terminal cytoplasmic domain of 14 residues, a transmembrane segment of 20 residues, and a stem region and catalytic domain of 310 residues with three potential N-linked glycosylation sites (28). One N-linked site is located in the putative

The predicted coding region of β4Gal-T4 has a single initiation codon in agreement with Kozak's rule (27), which precedes a sequence encoding a potential hydrophobic transmembrane segment (Figs. 2–4). The predicted coding sequence indicates that β4Gal-T4 is a type II transmembrane glycoprotein with an N-terminal cytoplasmic domain of 14 residues, a transmembrane segment of 20 residues, and a stem region and catalytic domain of 310 residues with three potential N-linked glycosylation sites (28). One N-linked site is located in the putative
cytoplasmic sequence and therefore may not be utilized. A hydropathy plot (29) of β4Gal-T4 indicated that the putative stem region was highly hydrophilic similar to β4Gal-T1, -T2, and -T3 (Fig. 4). In contrast, β4Gal-T5 and -T6 have unusually long hydrophobic regions at the putative signal anchor sequences, which are not clearly defined. A comparison of four members of a β3Gal-transferase family showed that one member with exclusive substrate specificity for glycolipids, the GM1 synthase β3Gal-T4, differed from the other three members of the family by having a unique hydrophobic stem region (26). A hydrophobic stem region is not found in all glycosyltransferases acting on glycolipids (30). The 3′-untranslated region contains a polyadenylation signal at base pair 1796 (1761).

Genomic Organization and Chromosomal Localization—The coding region of β4Gal-T4 was found in six exons, similar to β4Gal-T1, -T2, and -T3 (Figs. 1 and 5) (2, 31, 32). Comparison of the intron/exon boundaries of four of the six human β4Gal-transferases cloned to date revealed that the five introns in the coding regions are placed identically. The central coding exons of all β4Gal-T genes are nearly identical in length. β4Gal-T4 does not appear to have intronic sequences in the 5′-untranslated region (150 bases sequenced), which is similar to β4Gal-T1. β4Gal-T2 and -T3 have an intron 53 and 34 bases, respectively, 5′ of the initiating ATG (2) (note that the initiating ATG of β4Gal-T2 was changed in an erratum; Ref. 2). The chick homologue of β4Gal-T2 also has an intron in the 5′-untranslated region (3). The 3′ ESTs of β4Gal-T4 were linked in transcript map A004F36 to chromosome 3q13.3 near the D3S1558 microsatellite marker at 136 centimorgans (NCBI). Thus, all β4Gal-transferases are localized to different loci with β4Gal-T1 on 9p13 (33), β4Gal-T2 on 1p32-p33, and β4Gal-T3 at 1q23 (2).

Expression of β4Gal-T4—Expression of a soluble construct of β4Gal-T4 in insect cells resulted in marked increase in galactosyltransferase activity with a number of βGlcNAc containing acceptor substrates, compared with uninfected cells or cells infected with a control construct (Table I). All identified substrates had βGlcNAc at the nonreducing end. Of the simple saccharide derivatives tested, only the disaccharide GlcNAcβ1–6GlcNAca1-benzyl was better than monosaccharide derivatives. β4Gal-T4 did have significant activity with disaccharide derivatives representing N-linked and O-linked core structures (GlcNAcβ1–6Manα1-Me, GlcNAcβ1–2Man, and GlcNAcβ1–3GalNAca1-pNP) and with the biantennary pentasaccharide. In contrast, no activity was found with three glycoproteins that served as substrates for bovine milk β4Gal-T and human β4Gal-T2 and -T3 (Table II). Interestingly, β4Gal-T2 showed the highest relative activity with glycoprotein substrates. Previously, it was found that this enzyme has a low apparent $K_m$.
for βGlcNAc-benzyl and UDP-Gal (2). Analysis with glycolipid substrates showed that β4Gal-T4 had good activity with Lc₅Cer and 4-fold lower activity with nLc₅Cer (Table III). Lower activity with the longer lactoseries glycolipids was previously found to be more pronounced for β4Gal-T3, which had 10-fold lower activity with nLc₅Cer (2). β4Gal-T4 had higher apparent $K_m$ for UDP-Gal (31 μM) than recombinant bovine β4Gal-T1 (20 μM) (Table IV). No significant differences in activity of the full coding construct was found with the simple saccharide derivatives (data not shown).

**Fig. 4.** Comparison of human β4Gal-transferases. A, Kyte and Doolittle (1) (window of 8) hydropathy plots. The position of conserved sequence motifs as shown in Fig. 3 are indicated with dotted lines. TM indicates the putative transmembrane regions. Arrows indicate the positions of the four conserved cysteine residues. **B,** schematic depiction of β4Gal-transferases aligned for the conserved cysteine residues. Potential N-glycosylation sites are indicated by trees.
A Novel Member of the β4-Galactosyltransferase Family, β4Gal-T4

Although the activities of both human β4Gal-T1 and -T2 with GlcNAc concentrations above apparent \( K_a \) are inhibited by α-lactalbumin (2, 7, 34), β4Gal-T4 showed a marked increase in N-acetyllactosamine synthase activity in the presence of α-lactalbumin (Fig. 6A). Two-fold activation was achieved at 0.25 mg/ml and almost 8-fold at 20 mg/ml, which is substantially higher than previously observed for the bovine milk β4Gal-T activity and β4Gal-T2, which required 400 μg/ml and 100 μg/ml, respectively, to achieve maximum lactose synthase activity (2). As shown in Fig. 7, β4Gal-T4 was not inhibited at high concentrations of either βGlcNAc-benzyl or free N-acetylgalactosamine, which is in contrast to other β4Gal-transferases (2, 40). β4Gal-T4 showed strict donor substrate specificity for UDP-Gal and did not utilize UDP-GalNAc or UDP-GlcNAc with the acceptor substrates tested (data not shown). The soluble and full coding constructs exhibited the same modulation of activity by α-lactalbumin (data not shown).

Structure characterization of the product formed with \( \text{NcL}_4\text{Cer} \) by 1H NMR showed that the β4Gal-T4 forms the Galβ1-4GlcNAc linkage. One-dimensional 1H NMR spectroscopy showed that the product was a single glycosphingolipid product with a spectrum virtually identical to that of \( \text{NcL}_4\text{Cer} \) acquired previously under identical conditions (25), and distinct from that of Lc4Cer (23). In the downfield region of the spectrum (Fig. 8), four distinct β-anomeric resonances (3\( \beta_1 \), 2\( \beta_1 \)) were observed at chemical shifts 4.169 ppm (3\( \beta_1 \)), 4.263 ppm (3\( \beta_2 \)), 4.312 ppm (3\( \beta_3 \)), and 4.663 ppm (J uneven 3\( \beta_1 \), 2\( \beta_1 \)), with a coupling constant of 7.7 Hz, whereas that of \( \text{NcL}_4\text{Cer} \) is found at 4.140 ppm (3\( \beta_1 \), 2\( \beta_1 \)), 4.263 ppm (3\( \beta_2 \), 2\( \beta_2 \)), and 4.663 ppm (J uneven 3\( \beta_1 \), 2\( \beta_1 \)). Under these conditions, the resonance for H-1 of the terminal Gal\( \beta_1 \)-3\( \beta_4 \)Gal is found at 4.140 ppm (3\( \beta_1 \), 2\( \beta_1 \)), whereas that of β-GlcNAc III-1 is found at 4.780 ppm (3\( \beta_1 \), 2\( \beta_1 \)).

Expression Pattern of β4Gal-T4—Since a large number of ESTs from human adult organs has been identified, the cDNA library sources from which these are derived may provide information about the expression pattern. Based on this information, β4Gal-T4 is expressed in brain, central nervous system, colon, heart, lung, muscle, ovary, placenta, testis, and uterus.

Northern analysis with mRNA from eight human adult organs showed expression in most adult organs with highest levels observed in heart, placenta, kidney, and pancreas (Fig. 9). The transcript size of β4Gal-T4 was approximately 2.5 kilobase, which is similar to the transcript sizes of 2.2 kilobase for β4Gal-T2 and -T3. Two transcripts of 3.9 and 4.1 kilobase from β4Gal-T1 have been fully characterized, and shown to be differentially regulated (41, 42).

**DISCUSSION**

The human β4Gal-transferase gene family includes at least six members, which are involved in the synthesis of the N-acetyllactosamine disaccharide in oligosaccharides and glycoconjugates (2, 4–6, 9, 10). This large number of enzymes covering a single glycosidic linkage suggests either a high degree of redundancy in functions, or it may suggest that the enzymes have different functions. The high degree of divergence in primary sequence of the enzymes, studies of the acceptor substrate specificities of recombinant β4Gal-Ts (2, 6), and the findings that mice deficient in β4Gal-T1 exhibit a severe phenotype (43, 44), clearly point to different functional roles for each enzyme. Hence, the regulation of β1–4-galactosylation is
This is in contrast to the bovine milk substrate specificities, it was expected that glycoproteins were used at 10 mg/ml. Glycergate acceptor for utilized both glycolipid substrates (2). No natural glycoconjugates and hormo-logs, whereas 4Gal-T5 and -T6, are especially related (10). Asialo-agalacto-transferrin were not determined in the present study. Ovalbumin contains a single N-glycan with considerable heterogeneity, and mainly one potential acceptor sequence, GlcNAcβ1–2Manα1–3Man and GlcNAcβ1–2Manα1–6Man acceptor sites (46). Fetuin contains three complex N-glycans of the biantennary form or of the 2,4-branched tris-terogenetics, and a pentasaccharide representing complex N-glycans served as a substrate (Table I). However, the activities with these structures were less than with monosaccharide derivatives, suggesting that indeed these oligosaccharides do not represent the glycoconjugate substrates. Fetuin also contains three O-glycans, of which some are of the complex type (49). Although β4Gal-T4 catalyzed glycosylation of the disaccharide structure GlcNAcβ1–3GalNAcα1–pNP (the O-linked core 3 structure), the enzyme apparently did not work with the O-linked acceptors of asialo-agalacto-fetuin (Tables I and II). If β4Gal-T4 functions with glycoprotein acceptors, it may be with more complex structures. Preliminary studies with O-GlcNAc glycopolypeptides indicate that most enzymes can catalyze transfer to this type of protein glycosylation (50), but β4Gal-T4 showed the poorest activity. Collectively, it appears likely that the main function of β4Gal-T4 is in the biosynthesis of neolacto-series glycosphingolipids.

The catalytic efficiency of β4Gal-T4 with simple sugars and sugar derivatives was poor compared with β4Gal-T1, -T2, and -T3 (Tables I and II). At the concentrations tested, only β4Gal-T4 did not show substrate inhibition (Fig. 7). Interestingly, the activity of β4Gal-T4 with GlcNAc was activated by α-lactalbumin to levels comparable to the activities with GlcNAc of other β4Gal-Ts in the absence of α-lactalbumin (Fig. 6). A similar, although weaker, effect was previously found for the milk β4Gal-T (T1) activity with concentrations of GlcNAc acceptor below K_m (7, 35–37), and this was also observed in the present study (Fig. 7). At high GlcNAc acceptor concentrations, both β4Gal-T1 and -T2 are strongly inhibited by α-lactalbumin, and this was suggested to reflect enhanced acceptor substrate accessibility (Fig. 7) (2, 35–37, 51). Correspondingly, interaction of α-lactalbumin with β4Gal-T1 modulates the acceptor specificity from GlcNAc to Glc, thus forming the basis for lactose synthesis in mammary glands (7, 52). β4Gal-T2 is also efficiently induced to catalyze synthesis of lactose in the presence of α-lactalbumin (2), a feature that is also found in the chick orthologs (3). In contrast, β4Gal-T3 and β4Gal-T5 are largely insensitive to α-lactalbumin modulation, although weak inhibition of N-acetyllactosamine synthesis was observed (2, 6, 8). α-Lactalbumin induced β4Gal-T4 to catalyze synthesis of lactose, but the catalytic efficiency was low (Fig. 6). In contrast to the effect of α-lactalbumin on β4Gal-T1 activity toward monosaccharide acceptors, α-lactalbumin acts as a competitive inhibitor of β4Gal-T1 activity with extended acceptor substrates (e.g. β-O-GlcNAc1-benzyl or N,N'-diacytethylchitobiose; see Refs. 34–37 and 51), and this effect was also observed in the present study (Fig. 7B). Surprisingly, α-lactalbumin did not show a

### Table II

| Substrate specificities with glycolipid acceptors | β4Gal-T4 | β4Gal-T2 | β4Gal-T3 |
|-----------------------------------------------|----------|----------|----------|
| Henry egg albumin                              | 0.0      | 2.0      | 1.0      |
| Asialo-agalacto-fetuin                         | 0.1      | 2.8      | 0.7      | 0.8      |
| Asialo-agalacto-transferrin                    | 0.1      | 2.3      | 0.5      | 0.8      |
| Asialo-fetuin                                  | 0.0      | 0.2      | 0.0      | 0.1      |

* β-n-GlcNAc-1-Bzl was used at 0.25 mM with β4Gal-T2, 0.625 mM with bovine milk β4Gal-T, 2 mM with β4Gal-T3, and 20 mM with β4Gal-T4; glycoproteins were used at 10 mg/ml.

### Table III

| Substrate specificities with glycolipid acceptors | β4Gal-T4 |
|-------------------------------------------------|----------|
| GlcCer (Glcβ1-Cer)                              | ND       |
| LacCer (Galβ1–Glcβ1-Cer)                        | ND       |
| Gbβ (Galβ1–4Galβ1–4Glcβ1-Cer)                   | ND       |
| Gbγ (GalNAcβ1–3Galα1–4Galβ1–4Glcβ1-Cer)         | ND       |
| Gb6 (GalNAcβ1–4Galβ1–4Glcβ1-Cer)                | ND       |
| O2α2 (GalNAcβ1–4NeuAc2–3Galβ1–4Glcβ1-Cer)       | ND       |
| O2α2 (Galβ1–3GalNAcβ1–4NeuAc2–3Galβ1–4Glcβ1-Cer) | ND     |
| LeC (Galβ1–3GalNAcβ1–4Glcβ1–4Glcβ1-Cer)          | 0.59     |
| nLeC (Galβ1–GlcNAcβ1–3Galβ1–4Glcβ1-Cer)         | ND       |
| nLeC (GalNAcβ1–3Galβ1–4GlcNAcβ1–3Galβ1–4Glcβ1-Cer) | 0.13  |

* Assayed using 100 μg of Triton CF 54/100 μl of reaction mixture (enzyme partially purified as described under "Experimental Procedures").

**TABLE II**

| Acceptor substrate | β4Gal-T4 |
|--------------------|----------|
| nmol/min/ml | nmol/min/ml | nmol/min/ml | nmol/min/μg |
| β-n-GlcNAc-1-Bzl    | 3.1      | 3.5      | 3.9      | 3.4      |
| Hen egg albumin    | 0.0      | 2.0      | 1.0      | 0.7      |
| Asialo-agalacto-fetuin | 0.1      | 2.8      | 0.7      | 0.8      |
| Asialo-agalacto-transferrin | 0.1      | 2.3      | 0.5      | 0.8      |
| Asialo-fetuin       | 0.0      | 0.2      | 0.0      | 0.1      |

* ND, not detectable.
similar inhibitory effect on the activity of β4Gal-T4 with β-D-GlcNAc-1-benzyl at any concentration tested (Fig. 7B). The obtained data therefore suggest that α-lactalbumin interacts differently with β4Gal-T1 and β4Gal-T4.

β4Gal-T1 is strongly expressed in lactating mammary glands (41), and it has been reported that β4Gal-T5 is weakly expressed as well (8). However, Lo et al. (10) found that, of the six members of the family identified so far, only β4Gal-T1 is expressed in murine lactating glands. β4Gal-T2 is unlikely to be expressed in mammary glands since mice deficient in β4Gal-T1 do not produce lactose in milk (43, 44). Interestingly, the interaction of β4Gal-T4 with α-lactalbumin may be weaker or different than β4Gal-T1, since affinity chromatography with α-lactalbumin-Sepharose did not bind β4Gal-T4 under the same conditions under which β4Gal-T1 binds (53) (data not shown).

The distinct response of β4Gal-T4 to α-lactalbumin resembles the response reported for a snail UDP-GalNAc:βGlcNAc β1-4-N-acetylgalactosaminyltransferase activity (15). The β4GalNAc-transferase activity with GlcNAc concentrations below \( K_m \) was activated nearly 3-fold in the presence of α-lactalbumin; ■, bovine milk β4Gal-transferase without α-lactalbumin; □, bovine milk β4Gal-transferase with 2.5 mg/ml α-lactalbumin; ●, β4Gal-T4 without α-lactalbumin; ○, β4Gal-T4 with 2.5 mg/ml α-lactalbumin.

FIG. 6. α-Lactalbumin modulation of β4Gal-T4 activity. A, activity with GlcNAc in the presence of increasing amounts of α-lactalbumin. 0.1 milliunit (measured with βGlcNAc-benzyl) of β4Gal-T4 and bovine milk β4Gal-T were used with 50 mM GlcNAc as acceptor substrate. Purified bovine milk enzyme or media from High Five™ cells expressing the secreted form of β4Gal-T4 were used as enzyme sources. B, activity with glucose in the presence of increasing amounts of α-lactalbumin. ■, bovine milk Gal-transferase; ○, β4Gal-T4.

FIG. 7. N-Acetyllactosamine synthase activation of β4Gal-T4 by α-lactalbumin: lack of inhibition by high acceptor substrate concentrations. A, α-GlcNAc; B, β-α-GlcNAc-1-benzyl. ■, bovine milk β4Gal-transferase without α-lactalbumin; □, bovine milk β4Gal-transferase with 2.5 mg/ml α-lactalbumin; ●, β4Gal-T4 without α-lactalbumin; ○, β4Gal-T4 with 2.5 mg/ml α-lactalbumin.

### Table IV

| Substrate                          | \( K_m \) (mM) | \( V_{\text{max}} \) (pmol/min) | \( K_m \) (mM) | \( V_{\text{max}} \) (pmol/min) |
|------------------------------------|----------------|-------------------------------|----------------|-------------------------------|
| UDP-Gal                            | 0.021 (± 0.002) | 149.0                         | 0.020 (± 0.002) | 101.5                         |
| β-D-GlcNAc-1-Bzl                   | 2.46 (± 0.15)   | 106.0                         | 0.10 (± 0.02)   | 135.0                         |
| β-D-GlcNAc-(1–6)-α-D-GlcNAc-1-Bzl  | 0.41 (± 0.05)   | 108.0                         | 0.04 (± 0.01)   | 111.0                         |

A Novel Member of the β4-Galactosyltransferase Family, β4Gal-T4
N-acetyllactosamine synthase in the presence of the same concentration of α-lactalbumin, the lactose synthase activity is lower than the N-acetyllactosamine synthase activity (15) (Fig. 6). The snail β4GalNAc-transferase activity also resembles bovine milk β4Gal-transferase by showing modulation by α-lactalbumin of broader donor substrate specificity to include UDP-Gal. The equivalent was not found for β4Gal-T1 and -T2 homologues. Related to this, only two putative members of the β4Gal-transferase gene family have been identified in Caenorhabditis elegans, ce1 (GenBank accession no. Z29085) and ce2 (GenBank accession no. X98132) (1, 55), and these exhibit most of the highly conserved motifs found in the chick and mammalian enzymes. The gene designated ce2 shows the highest sequence similarity to β4Gal-T5 and -T6, and the least to β4Gal-T1. ce2 contains all four cysteine residues conserved among β4Gal-T1, -T2, -T3, -T4, -T5, and -T6. The gene designated ce1 shows highest similarity to a more distant member of the human β4Gal-transferase gene family, which has not been fully characterized yet. ce1 does not contain the four conserved cysteine residues, and shows several differences in other conserved motifs among the β4Gal-transferase family. The evolutionary trait of the β4Gal-transferase gene family thus remains to be clarified.

Acknowledgments—We thank Drs. Michael A. Hollingsworth and Kiyoshi Furukawa for many helpful suggestions and critical reading of the manuscript.

REFERENCES
1. Van Die, I., Bakker, H., and Van den Eijnden, D. H. (1997) Glycobiology 7, v–ix
2. Almeida, R., Amado, M., David, L., Lavery, S. B., Holmes, E. H., Merkx, G., van Kessel, A. G., Rygaard, E., Hassan, H., Bennett, E. P., and Clausen, H. (1997) J. Biol. Chem. 272, 31979–31992
3. Shaper, N. L., Meurer, J. A., Jeziorsse, D. H., Chou, T. D., Smith, E. J., Schnaar, R. R., and Shaper, J. H. (1997) J. Biol. Chem. 272, 31389–31399
4. Shaper, N. L., Shaper, J. H., Meuth, J. L., Fox, J. L., Chang, H., Kirsch, I. R., and Hollis, G. F. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 1573–1577
5. Narimatsu, H., Sinha, S., Brew, K., Okayama, H., and Qasba, P. K. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 4720–4724
6. Sato, T., Furukawa, K., Bakker, H., Van den Eijnden, D. H., and Van Die, I. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 472–477
7. Brew, K., Vanaman, T. C., and Hill, R. L. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 491–497
8. Sato, T., Aoki, N., Matsuda, T., and Furukawa, K. (1998) Biochem. Biophys.
9. Nomura, T., Takizawa, M., Aoki, J., Arai, H., Inoue, K., Wakisaka, E., Yoshizuka, N., Imokawa, G., Doehme, N., Takio, K., Hattori, M., and Matsue, N. (1996) J. Biol. Chem. 271, 13570–13577
10. Lo, N., Shaper, J. H., Pevsner, J., and Shaper, N. L. (1998) Glycobiology 8, 517–528
11. Shaper, N. L., Hollis, G. F., Douglas, J. G., Kirsch, I. R., and Shaper, J. H. (1988) J. Biol. Chem. 263, 10420–10428
12. Masri, K. A., Appert, H. E., and Fukuda, M. N. (1988) Biochem. Biophys. Res. Commun. 157, 657–665
13. Bakker, H., Agterberg, M., Van Tetering, A., Koeleman, C. A., Van den Eijnden, D. H., and Van Die, I. (1994) J. Biol. Chem. 269, 30326–30333
14. Bakker, H., Schoenmakers, P. S., Koeleman, C. A., Joziasse, D. H., Van Die, I., and Van den Eijnden, D. H. (1997) Glycobiology 7, 539–548
15. Neeleman, A. P., and Van den Eijnden, D. H. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 10111–10116
16. Do, K. Y., Do, S. I., and Cummings, R. D. (1995) J. Biol. Chem. 270, 18447–18451
17. Van den Eijnden, D. H. (1997) Biochem. Soc. Trans. 25, 887–893
18. Smith, P. L., and Baenziger, J. U. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 329–333
19. Mengeling, B. J., Manzella, S. M., and Baenziger, J. U. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 502–506
20. Breathnach, R., and Chambon, P. (1981) Annu. Rev. Biochem. 50, 349–383
21. Bennett, E. P., Hassan, H., and Clausen, H. (1996) J. Biol. Chem. 271, 17066–17072
22. Wandall, H. H., Hassan, H., Mirgorodskaya, E., Kristensen, A. K., Roepstorff, P. Bennett, E. P., Nielsen, P. A., Hollingsworth, M. A., Burchell, J., Taylor-Papadimitriou, J., and Clausen, H. (1997) J. Biol. Chem. 272, 23503–23514
23. Holmes, E. H., and Levery, S. B. (1990) Arch. Biochem. Biophys. 274, 633–647
24. Levery, S. B., Nudelman, E. D., and Hakomori, S. (1992) Biochemistry 31, 1069–1080
25. Levery, S. B., Nudelman, E., Kannagi, R., Symington, F. W., Andersen, N. H., Clausen, H., Baldwin, M., and Hakomori, S. (1989) Carbohydr. Res. 176, 121–144
26. Amado, M., Almeida, R., Carneiro, F., Levery, S. B., Holmes, E. H., Nomoto, M., Butterworth, M. A., Hassan, H., Schwientek, T., Nielsen, P. A., Bennett, E. P., and Clausen, H. (1996) J. Biol. Chem. 271, 12770–12778
27. Kozak, M. (1992) Annu. Rev. Cell Biol. 8, 197–225
28. Paulson, J. C., and Colley, K. J. (1989) J. Biol. Chem. 264, 17615–17618
29. Kyte, J., and Doolittle, R. F. (1982) J. Mol. Biol. 157, 105–132
30. Nagata, Y., Yamashiro, S., Yodoi, J., Lloyd, K. O., Shiku, H., and Furukawa, K. (1992) J. Biol. Chem. 267, 12082–12089
31. Hollis, G. F., Douglas, J. G., Shaper, N. L., Shaper, J. H., Stafford-Hollis, J. M., Evans, R. J., and Kirsch, I. R. (1989) Biochem. Biophys. Res. Commun. 162, 1069–1075
32. Mengle-Gaw, L., McCoy-Haman, M. F., and Tiemeier, D. C. (1991) Biochem. Biophys. Res. Commun. 176, 1269–1276
33. Shaper, N. L., Shaper, J. H., Bertness, V., Chang, H., Kirsch, I. R., and Hollis, G. F. (1986) Somat. Cell. Mol. Genet. 12, 633–636
34. Schanbacher, P. L., and Ehner, K. E. (1970) J. Biol. Chem. 245, 5057–5061
35. Powell, J. T., and Brew, K. (1976) J. Biol. Chem. 251, 3653–3663
36. Bell, J. E., Beyer, T. A., and Hill, R. L. (1976) J. Biol. Chem. 251, 3003–3013
37. Khatra, B. S., Herries, D. G., and Brew, K. (1974) Eur. J. Biochem. 44, 537–560
38. Roden, L., and Smith, R. (1966) J. Biol. Chem. 241, 5949–5954
39. Roden, L. (1981) in Biochemistry of Glycoproteins and Proteoglycans (Lennarz, W. J., ed) pp. 267–371, Plenum Press, New York
40. Shur, B. D. (1982) J. Biol. Chem. 257, 6871–6878
41. Harduin-Lepers, A., Shaper, J. H., and Shaper, N. L. (1993) J. Biol. Chem. 268, 14348–14359
42. Rajput, B., Shaper, N. L., and Shaper, J. H. (1996) J. Biol. Chem. 271, 5131–5142
43. Lu, Q., Hasty, P., and Shur, B. D. (1997) Dev. Biol. 181, 257–267
44. Asano, M., Furukawa, K., Kido, M., Matsumoto, S., Umesaki, Y., Kochibe, N., and Iwakura, Y. (1997) EMBO J. 16, 1850–1857
45. Kobata, A. (1984) in Biology of Carbohydrates (Ginsburg, V., and Robbins, P. W., eds) pp. 87–161, Wiley-Interscience, New York
46. Beyer, T. A., Rearick, J. I., Paulson, J. C., Prieels, J. P., Sadler, J. E., and Hill, R. L. (1970) J. Biol. Chem. 245, 12531–12534
47. Yet, M. G., Chin, C. C., and Wold, F. (1988) J. Biol. Chem. 263, 111–117
48. Do, K. Y., Fregien, N., Pierce, M., and Cummings, R. D. (1994) J. Biol. Chem. 269, 23546–23544
49. Sprio, R. G., and Bhoyroo, V. D. (1974) J. Biol. Chem. 249, 5704–5717
50. Torres, C. R., and Hart, G. W. (1984) J. Biol. Chem. 259, 3308–3317
51. Grobler, J. A., Wang, M., Pike, A. C. W., and Brew, K. (1994) J. Biol. Chem. 269, 5106–5114
52. Brodbeck, U., Denton, W. L., Tanahashi, N., and Ehner, K. E. (1967) J. Biol. Chem. 242, 1391–1397
53. Bendaik, B., Ward, L. D., and Simpson, R. J. (1993) Eur. J. Biochem. 216, 405–417
54. Bakker, H., Van Tetering, A., Agterberg, M., Smit, A. B., van den Eijnden, D. H., and Van Die, I. (1997) J. Biol. Chem. 272, 18580–18585
55. Wilson, R., Ainscough, R., and Anderson, K. e. (1994) Nature 368, 32–38
56. Aoki, D., Appert, H. E., Johnson, D., Wong, S. S., and Fukuda, M. N. (1990) EMBO J. 9, 3171–3178
57. IUPAC-IUB Commission on Biochemical Nomenclature (1977) Lipids 12, 455–463