

## A $\beta$ -1,3-*N*-acetylglucosaminyltransferase with poly-*N*-acetylactosamine synthase activity is structurally related to $\beta$ -1,3-galactosyltransferases

DAPENG ZHOU\*, ANDRÉ DINTER\*, RICARDO GUTIÉRREZ GALLEGOT, JOHANNIS P. KAMERLING†, JOHANNES F. G. VLIAGENTHART†, ERIC G. BERGER\*, AND THIERRY HENNET\*‡

\*Institute of Physiology, University of Zürich, Winterthurerstrasse 190, CH 8057 Zürich, Switzerland; and †Bijvoet Center, Department of Bio-Organic Chemistry, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands

Edited by Stuart E. Kornfeld, Washington University School of Medicine, St. Louis, MO, and approved November 19, 1998 (received for review August 7, 1998)

**ABSTRACT** Human and mouse cDNAs encoding a new  $\beta$ -1,3-*N*-acetylglucosaminyltransferase ( $\beta$ 3GnT) have been isolated from fetal and newborn brain libraries. The human and mouse cDNAs included ORFs coding for predicted type II transmembrane polypeptides of 329 and 325 aa, respectively. The human and mouse  $\beta$ 3GnT homologues shared 90% similarity. The  $\beta$ 3GnT gene was widely expressed in human and mouse tissues, although differences in the transcript levels were visible, thus indicating possible tissue-specific regulation mechanisms. The  $\beta$ 3GnT enzyme showed a marked preference for Gal( $\beta$ 1–4)Glc(NAc)-based acceptors, whereas no activity was detected on type 1 Gal( $\beta$ 1–3)GlcNAc and O-glycan core 1 Gal( $\beta$ 1–3)GalNAc acceptors. The new  $\beta$ 3GnT enzyme was capable of both initiating and elongating poly-*N*-acetylactosamine chains, which demonstrated its identity with the poly-*N*-acetylactosamine synthase enzyme (E.C. 2.4.1.149), showed no similarity with the *i* antigen  $\beta$ 3GnT enzyme described recently, and, strikingly, included several amino acid motifs in its protein that have been recently identified in  $\beta$ -1,3-galactosyltransferase enzymes. The comparison between the new UDP-GlcNAc: $\beta$ Gal  $\beta$ 3GnT and the three UDP-Gal: $\beta$ GlcNAc  $\beta$ -1,3-galactosyltransferases-I, -II, and -III reveals glycosyltransferases that share conserved sequence motifs though exhibiting inverted donor and acceptor specificities. This suggests that the conserved amino acid motifs likely represent residues required for the catalysis of the glycosidic ( $\beta$ 1–3) linkage.

Glycosyltransferases are grouped in families defined by the type of carbohydrate donor transferred and the glycosidic linkage catalyzed (1). The relevance of this classification system is being confirmed at the genetic level as it is often found that genes encoding glycosyltransferases of the same family share similar primary and genomic structures (2–4). However, exceptions to this rule have been documented, as exemplified by  $\beta$ -1,6-*N*-acetylglucosaminyltransferase enzymes, which share little resemblance to each other (reviewed in ref. 5). Structurally related enzymes that transfer different donor molecules represent another kind of exception. For example, the blood group A and B enzymes differ by only four amino acids but catalyze either the  $\alpha$ -1,3 transfer of Gal or GalNAc to galactosylated acceptors, respectively (6). A similar case has been reported with the identification of a  $\beta$ -1,4-*N*-acetylglucosaminyltransferase from *Lymnaea stagnalis* that is homologous to mammalian  $\beta$ -1,4-galactosyltransferases (7). However, although the donor molecules differ, these homologous glycosyltransferases still share common acceptor sub-

strates. We now provide another exception to the list by presenting a glycosyltransferase that is structurally related to the family of  $\beta$ -1,3-galactosyltransferases ( $\beta$ 3GalT) but uses distinct donor and acceptor substrates.

Similarity searches in the expressed sequence tag (EST) division of the GenBank database have revealed the existence of several ORFs encoding proteins that share conserved motifs with  $\beta$ 3GalT enzymes (8–10). Among those ORFs, we have isolated a full-length cDNA that includes all of the conserved motifs identified in the four  $\beta$ 3GalT enzymes characterized to date. Surprisingly, this cDNA was found to encode a UDP-GlcNAc:Gal( $\beta$ 1–4)Glc(NAc)  $\beta$ 3GnT enzyme that was able to catalyze both the initiation and elongation of *N*-acetylactosamine chains, thereby enabling the formation of poly-*N*-acetylactosaminoglycans and showed no structural similarity to the *i*-antigen  $\beta$ 3GnT (iGnT) enzyme reported by others (11).

### MATERIALS AND METHODS

**Isolation of Human and Mouse  $\beta$ 3GnT cDNAs.** EST fragments similar to the mouse  $\beta$ 3GalT-I, -II, and -III enzymes were retrieved from the EST division of GenBank by using the TBLASTX algorithm (12). The human and mouse  $\beta$ 3GnT cDNAs were isolated from bacteriophage  $\lambda$  libraries of human fetal brain cDNA (CLONTECH) and mouse newborn brain cDNA (Stratagene) by using a 367-bp fragment derived from the AAL50140 EST as a probe. The probing fragment was generated by using PCR with 50 ng of human T cell cDNA as template with the primers 5'-GCGACTACTACCTGCCTACG-3' and 5'-CTCCCTTCTCTGGCAAGCACT-3' for 30 cycles at 95°C for 45 s, 58°C for 30 s, and 72°C for 45 s. Once purified and subcloned into the pBluescript vector (Stratagene), the cDNAs of interest were sequenced by primer walking by using the AmpliTaq FS Dye Terminator sequencing kit (Perkin-Elmer/Applied Biosystems).

**Cloning and Expression of Recombinant Baculoviruses.** A *EcoRI*-*XhoI* 1.9-kbp fragment isolated from a partial human  $\beta$ 3GnT cDNA lacking the transmembrane domain was subcloned into the FastBac-HTc vector (Life Technologies) linearized with *EcoRI*-*XhoI*. The mouse  $\beta$ 3GnT gene was sub-

This paper was submitted directly (Track II) to the *Proceedings* office. Abbreviations:  $\beta$ 3GalT,  $\beta$ -1,3-galactosyltransferase;  $\beta$ 3GnT,  $\beta$ -1,3-*N*-acetylglucosaminyltransferase; iGnT, *i*-antigen  $\beta$ -1,3-*N*-acetylglucosaminyltransferase; EST, expressed sequence tag; Mco, O(CH<sub>2</sub>)<sub>8</sub>CO<sub>2</sub>Me; Bn, benzyl; Lac-*N*-tetraose, Gal( $\beta$ 1–3)GlcNAc( $\beta$ 1–3)Gal( $\beta$ 1–4)Glc; Lac-*N*-neo-tetraose, Gal( $\beta$ 1–4)GlcNAc( $\beta$ 1–3)Gal( $\beta$ 1–4)Glc; TOCSY, total correlation spectroscopy; ROESY, rotating frame nuclear Overhauser effect spectroscopy.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. AF092050 and AF092051).

‡To whom reprint requests should be addressed. e-mail: thennet@physiol.unizh.ch.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

PNAS is available online at www.pnas.org.

cloned as a 1.0-kbp *StuI-StuI* fragment, representing the protein-coding region devoid of transmembrane domain, into the FastBac-HTc vector. Recombinant baculoviruses were generated by transposon-mediated recombination (13) following the recommendations of the manufacturer. Sf9 insect cells were infected at a multiplicity of 10 and further incubated at 27°C before assaying for β3GnT activity.

**Glycosyltransferase Activity Assays.** All donor and acceptor substrates were from Sigma, except for Gal(β1-4)GlcNAc(β1-OpNP) (pNP = *p*-nitrophenyl), which was purchased from Toronto Research Chemicals (Downsview, ON, Canada) and the O(CH<sub>2</sub>)<sub>8</sub>CO<sub>2</sub>Me-derivatized (Mco) acceptors, which were kindly provided by Markus Streiff (Novartis, Basel, Switzerland). Sf9 cells (5 × 10<sup>6</sup>) infected with wild-type or with recombinant β3GnT baculoviruses were lysed 72 hr postinfection in 1 ml of 2% Triton X-100 in phosphate-buffered saline on ice for 15 min. The cytoplasmic and solubilized membrane fractions were recovered by centrifugation at 300 × *g* for 5 min. β3GnT activity was assayed by using 10 μl of cell lysate in 50 μl of 50 mM cacodylate buffer (pH 7.0), 20 mM MnCl<sub>2</sub>, 5% Me<sub>2</sub>SO, 0.75 mM ATP, 0.5 mM UDP-GlcNAc including 5 × 10<sup>4</sup> cpm of UDP-[<sup>14</sup>C]GlcNAc (Amersham) and various acceptors (see Table 1). Galactosyltransferase and *N*-acetylgalactosaminyltransferase activity were assayed as above using UDP-Gal and UDP-GalNAc as donor molecules. Reaction products were quantified as described (9, 14).

**β3GnT Gene Expression in Human and Mouse Tissues.** The β3GnT mRNAs were detected by using Northern blot analysis with commercially available multiple-tissue poly(A)<sup>+</sup> RNA blots (CLONTECH). A *SacI-PstI* 668-bp fragment from the

mouse β3GnT cDNA and a 367-bp region of the human β3GnT cDNA between nucleotides 616 and 983 were labeled with [α-<sup>32</sup>P]CTP (Hartmann Analytics, Braunschweig, Germany) by random priming and hybridized to the poly(A)<sup>+</sup> RNA blots. Blots were washed in 0.1× SSC, 0.1% SDS heated to 55°C and exposed for 4 days between intensifying screens at -70°C.

**Flow Cytometry.** A 1.7-kbp fragment including the human β3GnT cDNA was subcloned into the pcDNA3.1 plasmid (Invitrogen) opened with *Bam*HI and *Xho*I. HeLa cells were transfected with 5 μg of pcDNA3.1-β3GnT plasmid by using LipofectAmine (Life Technologies). Three days later, cells were stained with *Lycopersicon esculentum* (tomato)-fluorescein isothiocyanate lectin (1 μg/ml; Vector Laboratories). Alternatively, transfected cells were stained with human anti-*i* antiserum as described (11). Fluorescence was analyzed on a FACScan flow cytometer and CELLQUEST software (Becton Dickinson).

**HPLC.** The carbohydrate-containing fraction of a 1-ml incubation using 10 mM Gal(β1-4)Glc[β1-OBn(benzyl)] as acceptor and 100 μl of β3GnT-expressing Sf9 cell lysate as enzyme source was prepurified through a Sep-Pak C<sub>18</sub> cartridge and purified by using HPLC with a Lichrosorb-NH<sub>2</sub> column (4.6 × 250 mm, Chrompack, Bergen op Zoom, The Netherlands) at a flow rate of 2 ml/min by using a Kratos SF 400 HPLC system (ABI Analytical, Kratos Division). The column was eluted isocratically with solvent A (aqueous 80% acetonitrile), followed by a linear gradient from 100% solvent A/0% solvent B (aqueous 20% acetonitrile) to 65% solvent A/35% solvent B in 18 min. The effluent was monitored at 257 nm by using a 757 absorbance detector (Applied Biosystems).

**Methylation Analysis.** Permethylylation was performed essentially as described by Ciucanu *et al.* (15). After hydrolysis with 2 M trifluoroacetic acid (2 hr, 120°C), the mixture of partially methylated monosaccharides was reduced with 0.5 M sodium borodeuteride (2 hr) and neutralized. Boric acid was removed, and the mixture was acetylated with acetic anhydride (3 hr, 120°C). The partially methylated alditol acetates were analyzed by GLC/electron-impact MS (16, 17) on a Fisons MD800/8060 system (electron energy = 70 eV) equipped with a CPSil5 fused silica capillary column (25 m × 0.32 mm, Chrompack) by using a temperature program of 130–240°C at 4°C/min.

**Matrix-Assisted Laser Desorption Ionization (MALDI) Mass Spectrometry.** Positive-ion mode MALDI-TOF (time-of-flight) MS analysis of permethylated oligosaccharide was performed on a Voyager-DE (PerSeptive Biosystems, Framingham, MA) instrument operating at an accelerating voltage of 24 kV (grid voltage 92.5%, ion guide wire voltage 0.01%) and equipped with a VSL-337ND-N<sub>2</sub> laser. The sample was dissolved in methanol (1 μg/μl) and subsequently mixed in the sample well with 2,5-dihydroxybenzoic acid (10 mg/ml in H<sub>2</sub>O) at a ratio of 1:2. Linear mass scans were recorded over 1,000 Da by using a pulse delay time of 90 ns. Recorded data were processed by using GRAMS/386 software (v. 3.04, Galactic Industries, Salem NH).

**NMR Spectroscopy.** Before analysis, the oligosaccharide sample was repeatedly exchanged in <sup>2</sup>H<sub>2</sub>O (99.9 atom % <sup>2</sup>H, Isotec) with intermediate lyophilization and finally dissolved in 450 μl of <sup>2</sup>H<sub>2</sub>O (99.96 atom % <sup>2</sup>H, Isotec). Resolution-enhanced <sup>1</sup>H one-dimensional and two-dimensional NMR spectra were recorded on a Bruker DRX-600 instrument (NSR Center, Nijmegen University) at probe temperatures of 300 K. Chemical shifts (δ) were expressed in parts per million relative to internal acetate (δ 1.908; acetone, δ 2.225). HO<sup>2</sup>H signal suppression was achieved by applying a water-eliminated Fourier transform pulse sequence (18) in one-dimensional <sup>1</sup>H experiments and by presaturation for 1 s in two-dimensional experiments. Two-dimensional total correlation spectroscopy (TOCSY) spectra were recorded by using MLEV-17 mixing

Table 1. Acceptor substrate specificity of the β3GnT enzyme

Acceptors	pmol/min/mg prot added	
	Sf9 mock*	β3GnT†
UDP-GlcNAc‡		
Gal(β1-4)GlcNAc(β1-OpNP)	6	1,478
Gal(β1-4)Glc(β1-OBn)	4	1,034
Gal(β1-4)GlcNAc(octyl)	5	284
Gal(β1-3)GlcNAc(octyl)	5	8
Gal(β1-3)GalNAc(octyl)	5	4
Gal(α1-OpNP)	4	89
Gal(β1-OpNP)	4	26
GalNAc(α1-OBn)	4	5
GalNAc(β1-OBn)	5	6
GlcNAc(α1-OBn)	4	4
GlcNAc(β1-OBn)	11	7
GlcNAc(β1-3)GalNAc(α1-OBn)	8	5
Gal(β1-4)Glc	13	502
Gal(β1-4)GlcNAc	14	714
Lac- <i>N</i> -neo-tetraose§	14	748
Lac- <i>N</i> -tetraose§	13	45
UDP-Gal‡		
GalNAc(α1-OBn)	170	190
GalNAc(β1-OBn)	15	18
GlcNAc(α1-OBn)	11	10
GlcNAc(β1-OBn)	23	22
UDP-GalNAc‡		
Gal(α1-OpNP)	3	5
Gal(β1-OpNP)	3	5
GlcNAc(α1-OBn)	3	3
GlcNAc(β1-OBn)	27	31

All acceptors were assayed at 5 mM except the (Mco)-derivatized acceptors, which were assayed at 2 mM.

\*Lysate of Sf9 cells infected with a wild-type baculovirus.

†Lysate of Sf9 cells infected with a recombinant baculovirus expressing the human β3GnT enzyme.

‡Carbohydrate donor.

§Lac-*N*-neo-tetraose, Gal(β1-4)GlcNAc(β1-3)Gal(β1-4)Glc; Lac-*N*-tetraose, Gal(β1-3)GlcNAc(β1-3)Gal(β1-4)Glc

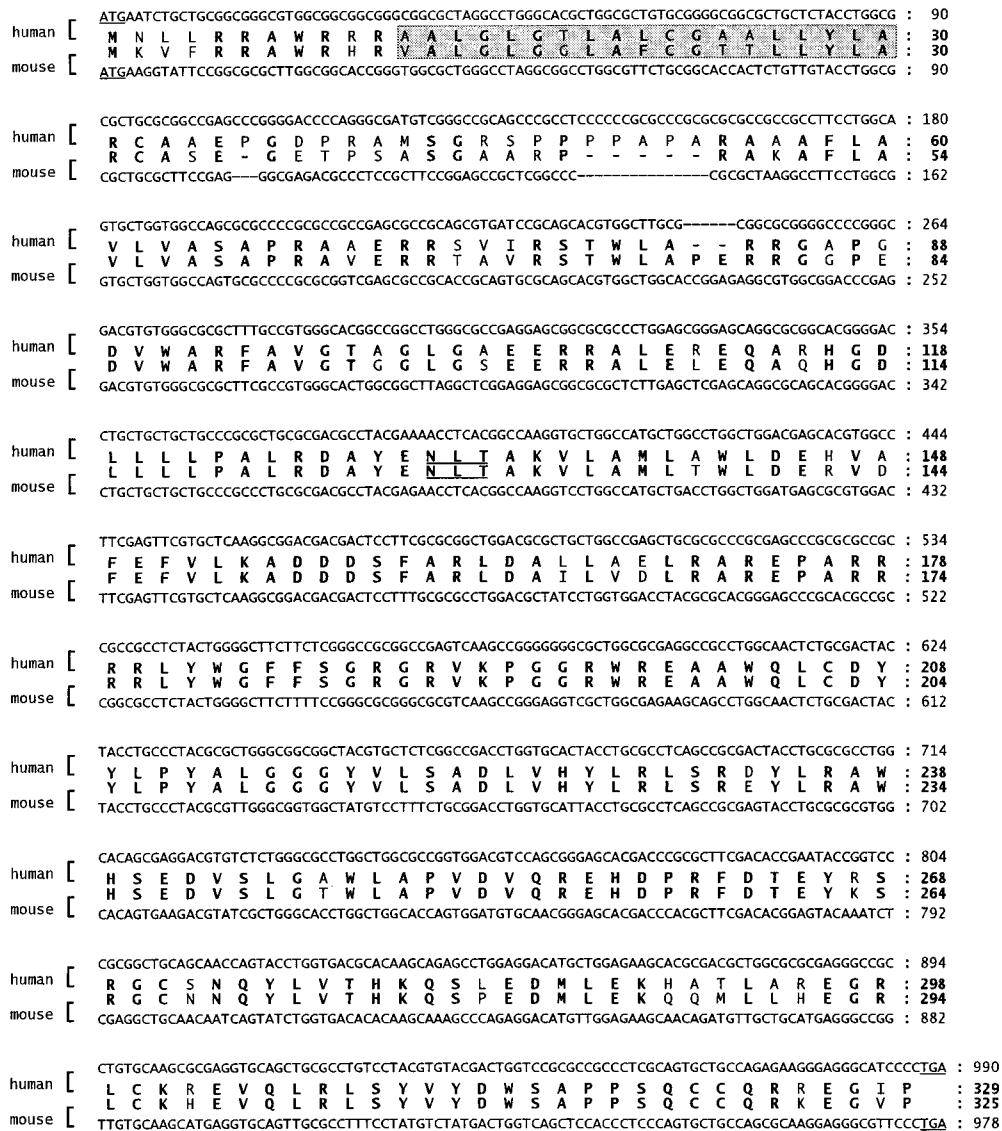


Fig. 1. Primary structure and deduced amino acid sequence of the human and mouse  $\beta$ 3GnT cDNA. Amino acids identical in both species are in boldface. The predicted transmembrane region is shaded and the single potential *N*-glycosylation site (N-X-[S/T]) is underlined.

sequences with effective spin-lock times between 20 and 100 ms. Two-dimensional rotating frame nuclear Overhauser effect spectroscopy (ROESY) spectra were recorded with a mixing time of 250 ms. The spin-lock field strength corresponded to a 90° pulse of about 120  $\mu$ s. <sup>1</sup>H one-dimensional and two-dimensional spectra were processed on Silicon Graphics IRIS workstations (Indigo 2 and O2) by using Bruker UXNMR software (Bijvoet Center, Department of NMR Spectroscopy).

## RESULTS

While searching for sequences related to the  $\beta$ 3GalT-I, -II, and -III enzymes (9) in the EST division of GenBank, we found several entries corresponding to distinct ORFs that shared sequence similarity with the  $\beta$ 3GalT enzymes. The 399-bp human EST fragment AA150140, which included the  $\beta$ 3GalT-motif EDVx[V/L]G, was therefore considered as a potential member of the  $\beta$ 3GalT gene family. The AA150140 EST was part of the Hs.48730 cluster of the UniGene collection (19), which represents sequences obtained from uterus, colon, tonsil, and brain tissues. By using the AA150140 fragment as a probe, we have isolated full-length cDNAs from human and

mouse brain cDNA libraries. The human and mouse cDNAs included each an ORF of 990 bp and 978 bp, which encoded proteins of 329 and 325 amino acids, respectively (Fig. 1). The encoded proteins were referred to as the human and mouse  $\beta$ 3GnT. Noteworthy, the human and mouse ORFs had an elevated G/C content of 75% and 67%, respectively. Two especially G/C rich clusters, reaching 82% and 74%, were both found in the human and mouse cDNAs between nucleotide positions 1–350 and 500–600. The human and mouse  $\beta$ 3GnT proteins shared 83% identity and 90% similarity, whereas the highest divergence was observed in the N-terminal 1/3 of the amino acid sequences. A Kyte–Doolittle hydrophathy analysis (20) revealed a potential transmembrane domain of 19 amino acids located at the N terminus of both human and mouse isoforms, thereby supporting a type II transmembrane topology typical of most glycosyltransferase enzymes (21).

A comparison between the  $\beta$ 3GnT studied here and the previously characterized  $\beta$ 3GalT proteins revealed that, whereas the overall identity was only between 15 and 19%, various motifs were conserved (Fig. 2). In addition to the EDVx[V/L]G stretch noticed in the AA150140 EST, all of the motifs identified previously in the four  $\beta$ 3GalT proteins were found at similar locations along the  $\beta$ 3GnT amino acid sequence. By contrast, the

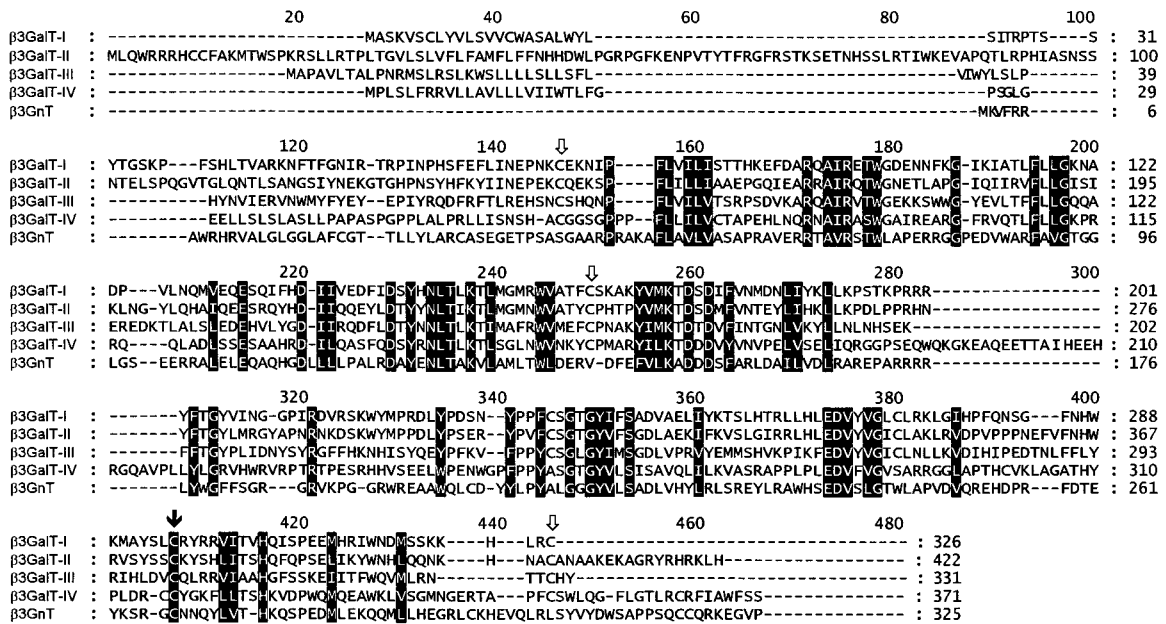


Fig. 2. CLUSTALW alignment of mouse  $\beta$ GalT-I, -II, -III, and -IV and mouse  $\beta$ GnT proteins. Conserved residues are shaded. The white arrows mark the positions of the cysteine residues conserved among  $\beta$ GalT proteins. The black arrow shows the position of the cysteines conserved in the five proteins.

cysteine residues of the  $\beta$ 3GnT protein could not be aligned with those of  $\beta$ 3GalT enzymes (Fig. 2).

The  $\beta$ 3GnT gene was expressed in all human tissues examined as determined by using Northern blot analysis (Fig. 3). Three transcripts were detected at 1.6, 2.4, and 3.3 kb, with the 3.3-kb mRNA producing the strongest hybridization signal. The intensity of the signal varied among the tissues examined, indicating differences in the level of  $\beta$ 3GnT gene expression in human tissues. In the mouse, a major 2.2-kb transcript was detected in most tissues and a minor 3.7-kb mRNA was detected in lung and kidney tissues (Fig. 3). No hybridization signal was detected in spleen and moderate-to-weak signals were found in liver and skeletal muscle, respectively.

The mouse and human  $\beta$ 3GnT proteins were expressed in Sf9 insect cells as recombinant baculoviruses. The  $\beta$ 3GnT enzyme was able to transfer GlcNAc to Gal( $\alpha$ 1-OpNP), Gal( $\beta$ 1-OpNP) but much more efficiently to Gal( $\beta$ 1-4)Glc(NAc)-based acceptors (Table 1). No activity was detected toward type 1 Gal( $\beta$ 1-3)GlcNAc-, O-glycan *core-1* Gal( $\beta$ 1-3)GalNAc-, and GlcNAc- and GalNAc-based acceptors (Table 1). The strict requirement for type 2 Gal( $\beta$ 1-4)Glc(NAc) acceptors indicated that this  $\beta$ 3GnT protein was distinct from the *core-3*  $\beta$ 3GnT enzyme (EC 2.1.4.147) and the

*elongation core-1*  $\beta$ 3GnT enzyme (EC 2.4.1.146). Although structurally related to the  $\beta$ 3GnT protein, the previously characterized  $\beta$ 3GalT-I to -III enzymes did not exhibit any *N*-acetylglucosaminyltransferase activity (9). The  $\beta$ 3GnT enzyme showed an absolute requirement for the divalent cation  $Mn^{2+}$  and a pH optimum at 7.0 (data not shown). The  $K_m$  values for the UDP-GlcNAc donor and for the Gal( $\beta$ 1-4)Glc( $\beta$ 1-OBn) acceptor were 0.2 mM and 6.6 mM, respectively. These values were in the range of those obtained with the  $\beta$ 3GnT poly-*N*-acetylglucosamine extension enzyme purified from calf serum by Kawashima *et al.* (22). This similitude suggested that this  $\beta$ 3GnT enzyme may indeed represent a poly-*N*-acetylglucosamine synthase. In fact,  $\beta$ 3GnT catalyzed both the initiation and the elongation of poly-*N*-acetylglucosamine chains as it transferred GlcNAc to the tetrasaccharide acceptor Gal( $\beta$ 1-4)GlcNAc( $\beta$ 1-3)Gal( $\beta$ 1-4)Glc (lac-*N*-neo-tetraose) as efficiently as to a single *N*-acetylglucosamine unit (Table 1). By contrast, the  $\beta$ 3GnT enzyme could not utilize the tetrasaccharide Gal( $\beta$ 1-3)GlcNAc( $\beta$ 1-3)Gal( $\beta$ 1-4)Glc (lac-*N*-tetraose) acceptor, confirming the adverse effect of the ( $\beta$ 1-3) linkage between the terminal Gal and the penultimate GlcNAc onto the elongation activity. Also, the  $\beta$ 3GnT enzyme did not exhibit significant Gal transferase and GalNAc transferase activity in comparison to the transferase activity measured with UDP-GlcNAc as donor (Table 1).

The relation between the  $\beta$ 3GnT enzyme and synthesis of polyglucosamine chains *in vivo* was investigated by analyzing the levels of i-antigen and tomato-lectin ligand on the surface of HeLa cells transiently transfected with an  $\beta$ 3GnT-expression vector. The i-antigen represents repeated ( $\beta$ 1-3)-linked *N*-acetylglucosamine units (23) and tomato lectin binds to *N*-acetylglucosamine oligomers (24). As shown by flow cytometry results, the overexpression of the  $\beta$ 3GnT cDNA in HeLa resulted in a slight increased reactivity for tomato lectin and a more pronounced elevation of i-antigen on the cell surface (Fig. 4), thus indicating the identity of this  $\beta$ 3GnT protein with the poly-*N*-acetylglucosamine synthase enzyme (EC 2.4.1.149).

To prove the nature of the glycosidic linkage in the product formed by this novel glycosyltransferase, a large-scale incubation was performed with Gal( $\beta$ 1-4)Glc( $\beta$ 1-OBn) as acceptor. The isolated product from the HPLC trisaccharide region was analyzed along different routes. Its molecular mass as deter-

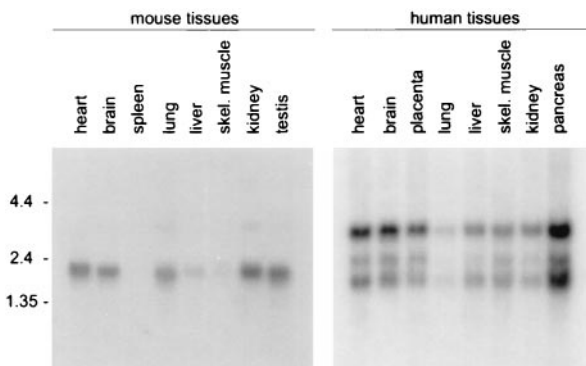


Fig. 3. Expression pattern of the  $\beta$ 3GnT gene in adult human and mouse tissues as determined by using Northern blot analysis. Each lane represents about 2  $\mu$ g of poly(A)<sup>+</sup> RNA. At the left, the size of the RNA markers is indicated in kilobases.

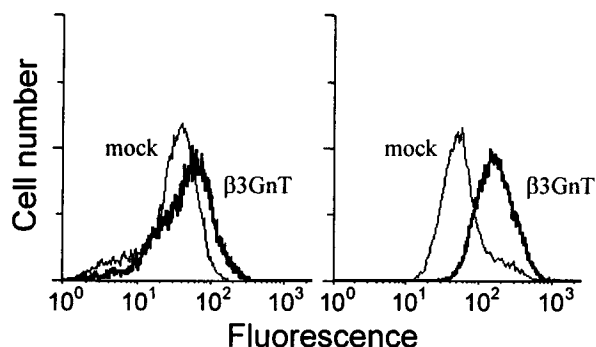


FIG. 4. Flow cytometry analysis of HeLa cells. HeLa cells transfected with an empty pcDNA3.1 vector (mock) or with the  $\beta$ 3GnT expression vector ( $\beta$ 3GnT) were stained with tomato lectin (Left) or anti-i antiserum (Right).

mined (after permethylation) by positive-ion mode MALDI TOF mass spectrometry indicated a trisaccharide with the bruttoformula HexNAcHex<sub>2</sub>Bn ( $m/z$  798,  $[M+Na]^+$ ). Methylation analysis revealed the presence of 4-substituted Glc, 3-substituted Gal, and terminal GlcNAc, demonstrating a (1-3)-linkage between GlcNAc and Gal. To confirm this conclusion, and to establish the anomeric configuration of the GlcNAc(1-3)Gal linkage, one-dimensional and two-dimensional  $^1H$  NMR experiments were performed. The most downfield signals in the one-dimensional  $^1H$  NMR spectrum (Fig. 5a) of the trisaccharide belonged to the benzyl group ( $\delta$  7.431, 4.930 ( $J_{gem}$  11.6 Hz), 4.755 ( $J_{gem}$  11.6 Hz)). By means of two-dimensional TOCSY (20 and 100 ms) and ROESY experiments, almost all monosaccharide resonances in the one-dimensional spectrum could be assigned (Table 2), and the three anomeric doublets present at  $\delta$  4.673 ( $^3J_{1,2}$  8.6 Hz), 4.546 ( $^3J_{1,2}$  8.0 Hz), and 4.423 ( $^3J_{1,2}$  8.0 Hz) were attributed to  $\beta$ -GlcNAc,  $\beta$ -Glc, and  $\beta$ -Gal, respectively. The ROESY spectrum (200 ms, Fig. 5b) permitted the determination of the two

Table 2.  $^1H$  NMR data [ $\delta$ , ppm;  $J$ , Hz (in parentheses)] of GlcNAc( $\beta$ 1-3)Gal( $\beta$ 1-4)Glc( $\beta$ 1-OBn), at 300 K, prepared with the human  $\beta$ 3GnT enzyme

Proton*	Residue			
	Benzyl	$\beta$ -Glc	$\beta$ -Gal	$\beta$ -GlcNAc
C <sub>6</sub> H <sub>5</sub>	7.431	—	—	—
CH <sub>2</sub>	4.930 (11.6)	—	—	—
	4.755 (11.6)	—	—	—
H-1	—	4.546 (8.0)	4.423 (8.0)	4.673 (8.6)
H-2	—	3.338 (8.3)	3.573 (8.9)	3.74
H-3	—	3.604	3.71	3.557
H-4	—	3.639	4.135 (2.8)	3.461
H-5	—	3.573	3.73	3.446
H-6a	—	3.974	nd	3.885
H-6b	—	3.79	nd	3.75
NAc	—	—	—	2.028

Based on data here, the assignment of H-6a and H-6b may have to be interchanged. nd, not determined.

\*In ppm relative to the signal of internal acetate ( $\delta$  1.908; acetone,  $\delta$  2.225).

glycosidic linkages. The anomeric track of the Gal residue revealed, besides two intraresidual crosspeaks [Gal H-1, H-3,  $\delta$  3.71; Gal H-1, H-2 (TOCSY transfer),  $\delta$  3.573], an interresidual crosspeak between Gal H-1 and Glc H-4 ( $\delta$  3.639). The anomeric track of the Glc residue revealed three intraresidual crosspeaks at  $\delta$  3.604 (Glc H-1, H-3), 3.573 (Glc H-1, H-5), and 3.338 [Glc H-1, H-2 (TOCSY transfer)]. Finally, the anomeric track of the GlcNAc residue showed besides three intraresidual crosspeaks [GlcNAc H-1, H-3,  $\delta$  3.557; GlcNAc H-1, H-5,  $\delta$  3.446; GlcNAc H-1, H-2 (TOCSY transfer),  $\delta$  3.74], an interresidual crosspeak between GlcNAc H-1 and Gal H-3 ( $\delta$  3.71). In summary, the combined results indicated the formed product to be GlcNAc( $\beta$ 1-3)Gal( $\beta$ 1-4)Glc( $\beta$ 1-OBn).

## DISCUSSION

We report the isolation of human and mouse  $\beta$ 3GnT cDNAs encoding a poly-*N*-acetylglucosaminyltransferase activity. The  $\beta$ 3GnT protein described here resembles  $\beta$ 3GalT enzymes, as both kinds of glycosyltransferases share several conserved motifs. These motifs are spread across the polypeptide chain and are not clustered in distinct regions as observed for example in sialyltransferases (25). This suggests that the conserved motifs do not directly represent a linear binding site for donor or acceptor molecules but rather shape a larger catalytic domain that directs the formation of a ( $\beta$ 1-3) glycosidic linkage. Preliminary data obtained from a mutagenesis study confirm the essential role of these conserved motifs in catalyzing the  $\beta$ -1,3-glycosyltransferase activity (A.D., unpublished results). The conservation between the  $\beta$ 3GnT and  $\beta$ 3GalT enzymes suggests a common evolutionary origin; structural requirements for the catalysis of a ( $\beta$ 1-3) glycosidic linkage probably maintained the motifs in the evolving polypeptides.

The establishment of a pattern of ( $\beta$ 1-3)-specific motifs should facilitate the recognition of further glycosyltransferases catalyzing a ( $\beta$ 1-3) linkage such as the O-glycan *core-1*  $\beta$ 3GalT and the *core-3*  $\beta$ 3GnT enzymes. In this respect, it is remarkable that the iGnT described recently (11) does not present any homology with the  $\beta$ 3GnT enzyme described here. A functional comparison between both enzymes is difficult to establish based on the data published (11). In fact, different expression systems were applied, which resulted in a great difference at the level of the  $\beta$ 3GnT activity measured. It appears that the  $\beta$ 3GnT reported here and the poly-*N*-acetylglucosaminyltransferase purified by Kawashima *et al.* (22) likely represent the same enzyme. Closely related enzymatic properties and identical acceptor substrate specificities strengthen this hypothesis in spite of the difference in

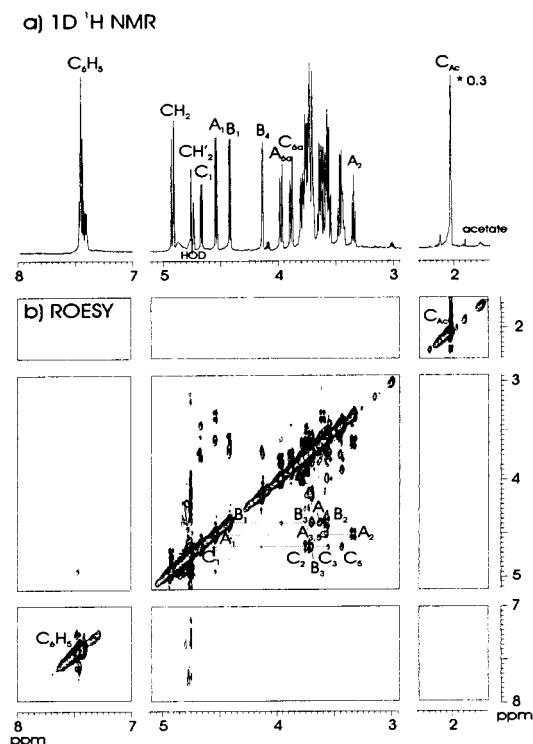


FIG. 5. 600-MHz one-dimensional (a) and two-dimensional (b) ROESY (200 ms)  $^1H$  NMR spectra of the isolated trisaccharide GlcNAc( $\beta$ 1-3)Gal( $\beta$ 1-4)Glc( $\beta$ 1-OBn) (C-B-A), recorded at 300 K.

size reported for both enzymes. The poly-*N*-acetylglucosamine extension enzyme purified from calf serum has an apparent  $M_r$  of 70,000, whereas the relative mass of  $\beta$ 3GnT averages 37,000. The cause of this discrepancy is unclear. Extensive posttranslational modifications such as glycosylation of the calf serum enzyme are unlikely to account for such an increase in mass. Alternatively, the secreted serum enzyme may be a disulfide-bridged homodimer. This possibility cannot be ruled out because the mass of the calf serum enzyme was determined under nonreducing conditions (22).

The existence of multiple enzymes capable of synthesizing poly-*N*-acetylglucosamine chains may indicate a selectivity at the level of the kind of glycoconjugates processed. Poly-*N*-acetylglucosamine chains occur on glycolipids and glycoproteins as well as on proteoglycans in the form of keratan sulfate (26). It is worth noting that keratan sulfate is the only glycosaminoglycan that is linked to its protein core via O- or N-linkage as found in glycoproteins. It shall be of interest to investigate whether  $\beta$ 3GnT and the iGnT enzymes can both synthesize poly-*N*-acetylglucosamine chains on proteoglycans with the same efficiency as on glycoproteins.

The extent of poly-*N*-acetylglucosamine chains on tumor cells has directly been related to their metastatic potential (27). It has been previously implied that the branching enzymes  $\beta$ -1,6-*N*-acetylglucosaminyltransferase-V (28) and core-2  $\beta$ -1,6-*N*-acetylglucosaminyltransferase (29) control the generation of poly-*N*-acetylglucosamine chains by providing the ( $\beta$ 1-6) branch favored for initiation of poly-*N*-acetylglucosamine formation (30, 31). The present Northern blot analysis revealed a variability at the level of  $\beta$ 3GnT transcripts detected in human and mouse tissues. This observation suggests that the expression of the  $\beta$ 3GnT gene may be regulated to some extent by tissue-specific promoters. This raises the possibility that the formation of poly-*N*-acetylglucosaminoglycans may in part be modulated by the availability of the  $\beta$ 3GnT enzyme itself. Along this line, it will be interesting to investigate the relation between the expression of the  $\beta$ 3GnT gene in tumor cells and the emergence of a metastatic phenotype.

We thank Minoru Fukuda for generously providing us with the anti i-antigen antiserum. This work was funded by a Roche Research Foundation grant to T.H., by the Swiss National Science Foundation Grant 3100-046836.96 to E.G.B. and T.H., and by the European Union Grant BIO4-CT95-0138 to J.P.K., J.F.G.V., and E.G.B.

1. Kleene, R. & Berger, E. G. (1993) *Biochim. Biophys. Acta* **1154**, 283–325.
2. Harduin-Lepers, A., Recchi, M.-A. & Delannoy, P. (1995) *Glycobiology* **5**, 741–758.
3. Reguigne-Arnould, I., Coullin, P., Mollicone, R., Faure, S., Fletcher, A., Kelly, R. J., Lowe, J. B. & Oriol, R. (1995) *Cytogenet. Cell Genet.* **71**, 158–162.
4. Hagen, F. K. & Nehrke, K. (1998) *J. Biol. Chem.* **273**, 8268–8277.
5. Van den Eijnden, D. H. & Joziase, D. H. (1993) *Curr. Opin. Struct. Biol.* **3**, 711–721.
6. Yamamoto, F., Clausen, H., White, T., Marken, J. & Hakomori, S. (1990) *Nature (London)* **345**, 229–233.
7. Bakker, H., Agterberg, M., Van Tetering, A., Koeleman, C. A., Van den Eijnden, D. H. & Van Die, I. (1994) *J. Biol. Chem.* **269**, 30326–30333.
8. Miyazaki, H., Fukumoto, S., Okada, M., Hasegawa, T. & Furukawa, K. (1997) *J. Biol. Chem.* **272**, 24794–24799.
9. Hennes, T., Dinter, A., Kuhnert, P., Mattu, T. S., Rudd, P. M. & Berger, E. G. (1998) *J. Biol. Chem.* **273**, 58–65.
10. Kolbinger, F., Streiff, M. B. & Katopodis, A. G. (1998) *J. Biol. Chem.* **273**, 433–440.
11. Sasaki, K., Kurata-Miura, K., Ujita, M., Angata, K., Nakagawa, S., Sekine, S., Nishi, T. & Fukuda, M. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 14294–14299.
12. Altschul, S. F., Boguski, M. S., Gish, W. & Wootton, J. C. (1994) *Nat. Genet.* **6**, 119–129.
13. Luckow, V. A., Lee, S. C., Barry, G. F. & Olins, P. O. (1993) *J. Virol.* **67**, 4566–4579.
14. Malissard, M., Borsig, L., Di Marco, S., Grütter, M. G., Kragl, U., Wandrey, C. & Berger, E. G. (1996) *Eur. J. Biochem.* **239**, 340–348.
15. Ciucanu, I. & Kerek, F. (1984) *Carbohydr. Res.* **131**, 209–217.
16. Kamerling, J. P. & Vliegthart, J. F. G. (1989) in *Clinical Biochemistry—Principles, Methods, Applications*, ed. Lawson, A. M. (de Gruyter, Berlin), Vol. 1, pp. 176–263.
17. Jansson, P.-E., Kenne, L., Liedgren, H., Lindberg, B. & Lönngren, J. (1976) *Chem. Commun. (Stockholm Univ.)* **8**, 1–74.
18. Hård, K., Van Zadelhoff, G., Moonen, P., Kamerling, J. P. & Vliegthart, J. F. G. (1992) *Eur. J. Biochem.* **209**, 895–915.
19. Schuler, G. D. (1997) *J. Mol. Med.* **75**, 694–698.
20. Kyte, J. & Doolittle, R. F. (1982) *J. Mol. Biol.* **157**, 105–132.
21. Paulson, J. C. & Colley, K. J. (1989) *J. Biol. Chem.* **264**, 17615–17618.
22. Kawashima, H., Yamamoto, K., Osawa, T. & Irimura, T. (1993) *J. Biol. Chem.* **268**, 27118–27126.
23. Feizi, T., Childs, R. A., Watanabe, K. & Hakomori, S.-I. (1979) *J. Exp. Med.* **149**, 975–980.
24. Merkle, R. K. & Cummings, R. D. (1987) *J. Biol. Chem.* **262**, 8179–8189.
25. Livingston, B. D. & Paulson, J. C. (1993) *J. Biol. Chem.* **268**, 11504–11507.
26. Bhavanandan, V. P. & Davidson, E. A. (1992) in *Proteoglycans: Structure, Synthesis, Function*, eds. Allen, H. J. & Kisailus, E. C. (Dekker, New York), pp. 167–202.
27. Dennis, J. W., Laferte, S., Waghorne, C., Breitman, M. L. & Kerbel, R. S. (1987) *Science* **236**, 582–585.
28. Shoreibah, M., Perng, G. S., Adler, B., Weinstein, J., Basu, R., Cupples, R., Wen, D., Browne, J. K., Buckhaults, P., Fregien, N. & Pierce, M. (1993) *J. Biol. Chem.* **268**, 15381–15385.
29. Bierhuizen, M. F. & Fukuda, M. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 9326–9330.
30. Heffernan, M., Lotan, R., Amos, B., Palcic, M., Takano, R. & Dennis, J. W. (1993) *J. Biol. Chem.* **268**, 1242–1251.
31. Higgins, E. A., Siminovitch, K. A., Zhuang, D. L., Brockhausen, I. & Dennis, J. W. (1991) *J. Biol. Chem.* **266**, 6280–6290.