Control of O-Glycan Branch Formation

MOLECULAR CLONING AND CHARACTERIZATION OF A NOVEL THYMUS-ASSOCIATED CORE 2 β1,6-N-ACETYLGLUCOSAMINYLTRANSFERASE

(Received for publication, December 21, 1999)

Tilo Schwiendek‡, Jiunn-Chern Yeh‡, Steven B. Levery‡, Birgit Keck‡, Gerard Merkx‡*, Ad Geurts van Kessel‡*, Minoru Fukuda‡, and Henrik Clausen‡

From the ‡School of Dentistry, University of Copenhagen, Nørre Allé 20, 2200 Copenhagen N, Denmark, the ‡Glycosobiology Program, Cancer Research Center, The Burnham Institute, La Jolla, California 92037, the University of Georgia, Complex Carbohydrate Research Center, Athens, Georgia 30602, and the **Department of Human Genetics, University Hospital Nijmegen, 6500 HB Nijmegen, The Netherlands

Core 2 O-glycan branching catalyzed by UDP-N-acetyl-α-D-glucosamine: acceptor β1,6-N-acetylglucosaminyltransferases (β6GlcNAc-Ts) is an important step in mucin-type biosynthesis. Core 2 complex-type O-glycans are involved in selectin-mediated adhesion events, and O-glycan branching appears to be highly regulated. Two homologous β6GlcNAc-Ts functioning in O-glycan branching have previously been characterized, and here we report a third homologous β6GlcNAc-T designated C2GnT3. C2GnT3 was identified by BLAST analysis of human genome sequence sequences. The catalytic activity of C2GnT3 was evaluated by in vitro analysis of a secreted form of the protein expressed in insect cells. The results revealed exclusive core 2 β6GlcNAc-T activity. The product formed with core 1-para-nitrophenyl was confirmed by 1H NMR to be core 2-para-nitrophenyl. In vivo analysis of the function of C2GnT3 by coexpression of leukosialin (CD43) and a full coding construct of C2GnT3 in Chinese hamster ovary cells confirmed the core 2 activity and failed to reveal I activity. The C2GnT3 gene was located to 5q12, and the coding region was contained in a single exon. Northern analysis revealed selectively high levels of a 5.5-kilobase C2GnT3 transcript in thymus with only low levels in other organs. The unique expression pattern of C2GnT3 suggests that this enzyme serves a specific function different from other members of the β6GlcNAc-T gene family.

Control of mucin-type O-glycosylation involves an initiation step followed by a processing step. The initiation step is complex and is carried out by a large family of homologous UDP-GalNAc:polypeptide GalNAc-transferases (1). The polypeptide GalNAc-transferase isoforms have distinct enzymatic properties and are differentially expressed, thus presumably allowing for a high level of control in determining sites of O-glycan attachments in proteins. The processing step involves elongation, branching, and terminal modification of the O-glycans (2).

It is apparent that essential steps in O-glycan elongation and branching are catalyzed by multiple glycosyltransferase isoforms from families of homologous glycosyltransferases. The main biosynthetic pathway utilized to synthesize complex-type O-glycans is through core 2 branching. Two UDP-GlcNAc: Galβ1–3GalNAcβ6GlcNAc-Ts, C2GnT1 and C2GnT2, have been identified and cloned thus far (3–5). Both function in vitro and in vivo as core 2 synthases, but C2GnT2 can also make the core 4 structure (4, 5). The existence of multiple core 2 and related β6GlcNAc-T isoforms was predicted by analyses of enzyme activities in cells and organs (2, 6, 7). Elongation of core 2 O-glycans with poly-N-acetyllactosamine repeats (Galβ1–4GlcNAcβ1–3β3Glc) may be carried out by alternative action of multiple β4-galactosyltransferases and β3GlcNAc-Ts (8–11). In the case of β4-galactosyltransferases one isoform, β4Gal-T4, appears to have superior kinetic properties for catalysis of the core 2 O-glycan poly-N-acetyllactosamine elongation (9). Two β3GlcNAc-Ts elongate core 2 O-glycans, although comparative analyses have not been performed (10, 12). Although the catalytic functions of these glycosyltransferase isoforms are not completely understood, the available data suggest the hypothesis that individual enzyme isoforms display unique kinetic properties and specific catalytic activities with regard to glycan structure and types of glycoconjugate. This is further supported by the finding that glycosyltransferase isoforms have different tissue expression patterns.

Core 2 O-glycan branching is a key step in mucin-type O-glycosylation (2). Core 2 O-glycans such as sialyl-Lea and associated structures serve as ligands for selectin- and galectin-mediated cell-cell adhesion events that play important roles in T-cell development (13), lymphocyte trafficking (14, 15), the inflammatory process (16, 17), and cancer metastasis (18). Some evidence indicates that a major control point for synthesis of these ligands is the core 2 branching event catalyzed by β6GlcNAc-T activities (15, 19). Marked changes in core 2 branching and β6GlcNAc-T activities are associated with T-cell maturation and malignant transformation (18, 20, 21). Multi-

1 C2GnT1 (C2GnT, C2GnT-L), C2GnT2 (C2/G4nT, C2GnT-M), and IgNRT represent human β6GlcNAc-Ts with GenBank access numbers M79347, AF038650/AF102452, and Z19550, respectively.
2 The abbreviations used are: β6GlcNAc-T, UDP-N-acetyl-α-D-glucosamine:βGalαGlcNAc β1,6-N-acetylgalactosaminyltransferase; β3GlcNAc-T, UDP-N-acetyl-α-D-glucosamine:Galβ1,3-N-acetylgalactosaminyltransferase; EST, expressed sequence tag; HMBC, heteronuclear multiple bond correlation; HSQC, heteronuclear single quantum correlation; MES, 4-morpholinoethanesulfonic acid; PCR, polymerase chain reaction; TOCSY, total correlation spectroscopy; UTR, untranslated region; CHO, Chinese hamster ovary.
ple core 2 β6GlcNAc-T isoforms exist; hence, it is important to identify and characterize these to understand the molecular genetic basis for their differential regulation and activity. The first core 2 β6GlcNAc-T identified by transection cloning, C2GnT1, is widely expressed, functions only in core 2 synthesis (3, 5), and resembles the leukocyte β6GlcNAc-T activity (also designated C2GnT-L) identified by Williams and Schachter (22). The second core 2 β6GlcNAc-T, C2GnT2, was cloned by an EST cloning strategy and has a broader acceptor specificity (4, 5). C2GnT2 resembles the mucous β6GlcNAc-T activity (designated C2GnT-M) (23, 24) and is mainly expressed in mucous-secreting organs (4, 5). C2GnT1 was predicted to control synthesis of core 2 selectin ligands in leukocytes and lymphoid tissues; however, mice deficient in C2GnT1 exhibited partial reduction in selectin ligand production, and there were no significant changes in lymphocyte homing properties (17). One possible explanation for these results would be the existence of additional core 2 β6GlcNAc-Ts. C2GnT2 does not appear to be a candidate for this gene as its expression pattern is restricted and mainly associated with mucous epithelia (4, 5, 17).

In the present report we describe the cloning and characterization of a third homologous core 2 β6GlcNAc-T designated C2GnT3. C2GnT3 exhibits exclusive core 2 acceptor specificity, and it shows a unique tissue distribution with high expression found only in thymus. C2GnT3 is predicted to play an important role in T-cell development and lymphocyte homing.

**EXPERIMENTAL PROCEDURES**

Cloning and Sequencing of C2GnT3—Expression of C2GnT3 in Insect Cells—A construct designed for expression of the full coding sequence of C2GnT3 in CHO cells was prepared by PCR using P1 DNA and the primer pair TSHC115 (5′-GGCGAATTCATTGGGTGGTTTTCTCAA-3′) and the antisense primer TSHC121 (5′-AGCGAATTCTTACTATCATGATGTGGTAGTG-3′). The sequence of the 5′-untranslated region was determined by analysis of human genome survey sequences. The open reading frame of 1359 base pairs was represented in the clone and sequenced in full. Confirmatory sequencing was performed on a cDNA clone. The coding regions of human C2GnT1 and C2GnT2 were previously found to be organized in one exon (4, 25). The putative coding sequence identified in the genome survey sequence data base was incomplete but likely to be localized in one single exon as well. The additional 3′-sequence of the open reading frame was therefore obtained by sequencing a genomic P1 clone (844/B1) obtained from Research Genetics Inc.) contained a novel partial open reading frame with significant sequence similarity to C2GnTs. The coding region of C2GnT3-sol was cotransfected with Baculo-Gold™ DNA (PharMingen) and mainly associated with mucous epithelia (4, 5, 17).

**Cloning and Sequencing of C2GnT3**—Insect cells, as described previously (26). Controls included.

**Expression of C2GnT3 in CHO Cells**—A novel member of the β6GlcNAc-T gene family, designated C2GnT3, was identified by analysis of human genome survey sequences. The open reading frame of 1359 base pairs encodes a protein of 453 amino acids with four potential N-linked glycosylation sites. A type II domain structure with an N-terminal cytoplasmic domain of 11 residues, a transmembrane segment of 21 residues, and a 30.00 ppm for 1H and 13C, respectively). Evaluation of the chromosomal slides was performed as described above and washed as described previously (30). Evaluation of the chromosomal slides was performed as described (4).

**RESULTS**

Identification and Cloning of C2GnT3—A novel member of the β6GlcNAc-T gene family, designated C2GnT3, was identified by analysis of human genome survey sequences. The open reading frame of 1359 base pairs encodes a protein of 453 amino acids with four potential N-linked glycosylation sites. A type II domain structure with an N-terminal cytoplasmic domain of 11 residues, a transmembrane segment of 21 residues, and a 30.00 ppm for 1H and 13C, respectively). Evaluation of the chromosomal slides was performed as described above and washed as described previously (4).

**Human O-Glycan β6GlcNAc-transferrases Share Significant...**

**Expression of C2GnT3 in CHO Cells**—A construct designed for expression of the full coding sequence of C2GnT3 in CHO cells was prepared by PCR using P1 DNA and the primer pair TSHC115 and TSHC121 with BamHI and EcoRI restriction sites, respectively (Fig. 1). The PCR product was cloned directionally between the BamHI and EcoRI site of pcDNA3 (Invitrogen), and the expression cassette was fully sequenced. CHO cells were transiently transfected with plasmids pcDNA3-C2GnT3-full, pcDNA3.1-C2GnT-M, pcDNA1-C2GnT-L, and pcDNA1-IgG1 and selected for neomycin resistance. Evaluation of in vivo core 2 synethase activity was performed by cotransfection with pcGFP-leskiosin (28). Core 2 O-glycan-modified leskiosin (CD43) was detected by immunofluorescence staining with monoclonal antibody T305, and large I antigen was detected with anti-I serum (Ma) using procedures described previously (3, 5, 29).

**Enzymatic Assays and Product Characterization**—Assayed were performed with semi-purified C2GnT3 in 50-μl reaction mixtures containing 100 mm MES (pH 6.5), 0.1% Nonidet P-40, 150 μM UDP-[14C]GlcNAc (2.000 cpm/nmol) (Amersham Pharmacia Biotech), and the indicated concentrations of acceptor substrates (Sigma and Toronto Research Laboratories Ltd., see Table I for structures). Reaction products were quantified by chromatography on Dowex AG1-X8. Transfer of N-acetylgalactosamine to glycoprotein acceptors was evaluated by acid precipitation and filtration through glass fiber filters as described (8).

Complete glycosylation of core 1-para-nitrophenyl was performed in a reaction mixture consisting of 0.6 milliliters of C2GnT3 (specific activity determined with core 1-benzyl), 2 mg of core 1-para-nitrophenyl, 100 mm MES (pH 7.0), 5 mm EDTA, 4.6 μmol of UDP-GlcNAc, and 100 milliliters of alkaline phosphatase in a final volume of 200 μl. The glycosylation of core 1-para-nitrophenyl was monitored by thin layer chromatography and run for 1 h until complete. The reaction product was purified, deuterium exchanged, and dissolved in D2O as described (4). One-dimensional 1H NMR, two-dimensional 1H-H correlation spectra, 1H-detected, 13C-decoupled, phase-sensitive gradient 1H-13C HSQC, and HMB experiments were performed as described previously (Ref. 4 and references cited therein). One-dimensional reference 13C NMR spectra were acquired using direct detection on a Varian Unity Inova 500 MHz spectrometer. A 1-μg sample of core 1-para-nitrophenyl (Sigma) was used for 30.00 ppm for 1H and 13C, respectively). Expression Analysis—A human RNA Master Blot was obtained from CLONTECH. The cDNA fragment of soluble C2GnT3 (base pairs 115–1359) was used as a probe for hybridization. The probe was random primer-labeled using [α-32P]dATP and the Strip-EZ DNA labeling kit (Ambion). The membrane was probed overnight and washed according to the protocol of the manufacturer. A Northern blot of multiple human tissues (MTN II from CLONTECH) was probed as described above and washed as described previously (4).

**Chromosomal Localization of C2GnT3 by Fluorescence in Situ Hybridization**—BAC DNA from the genomic clone CIT-HSP-2288B17 was used for fluorescence in situ hybridization on normal human lymphocyte metaphase chromosomes using procedures described previously (30). Evaluation of the chromosomal slides was performed as described (4).

**RESULTS**

Identification and Cloning of C2GnT3—A novel member of the β6GlcNAc-T gene family, designated C2GnT3, was identified by analysis of human genome survey sequences. The open reading frame of 1359 base pairs encodes a protein of 453 amino acids with four potential N-linked glycosylation sites. A type II domain structure with an N-terminal cytoplasmic domain of 11 residues, a transmembrane segment of 21 residues, and a 30.00 ppm for 1H and 13C, respectively). Evaluation of the chromosomal slides was performed as described (4).
Sequence Similarities—Fig. 2 shows a multiple amino acid sequence alignment (ClustalW) of C2GnT3, C2GnT2, C2GnT1, and IGnT. C2GnT3 shows a higher overall amino acid sequence identity to human C2GnT1 and C2GnT2 (42%) than to human IGnT (39%). High sequence similarity with several well conserved motifs can be found in three regions (A, B, and C) in the putative catalytic domains of the four human proteins as defined previously (29). The spacing of nine cysteine residues is conserved in all four \( \beta 6 \)-GlcNAc-Ts (Fig. 2). There is one conserved potential \( \text{N} \)-linked glycosylation site located in the stem region of C2GnT3, C2GnT2, and C2GnT1. This \( \text{N} \)-linked glycosylation site was shown to be essential for the catalytic function of C2GnT1 (32).

Chromosomal Localization and Genomic Organization—Fluorescence in situ hybridization of the genomic clone CIT-HSP-2288B17 to human metaphase chromosomes located the C2GnT3 gene to 5q12 (Fig. 3). The coding region of C2GnT3 is contained in a single exon, similar to the genomic structure of the human C2GnT1 and C2GnT2 genes (4, 25).

Kinetic Properties of Recombinant C2GnT3—Transfection of Sf9 cells with pAcGP67-C2GnT3-sol resulted in marked increase in \( \beta 6 \)-GlcNAc-T activity. Secreted C2GnT3 exhibited significant catalytic activity with disaccharide derivatives of \( \text{O} \)-linked core 1 (Gal\( \beta 1\)–3GalNAc\( \alpha 1\)-R) as substrates (Table I). In contrast, no activity was detected with the core 3 substrate (GlcNAc\( \beta 1\)–3GalNAc\( \alpha 1\)-R) indicating a lack of core 4 synthase activity. Similarly, no activity was detected with \( \alpha \)-D-GalNAc-1-para-nitrophenyl, suggesting that C2GnT3 does not synthesize core 6 (GlcNAc\( \beta 1\)–6GalNAc\( \alpha 1\)-R). Several substrates were tested to detect I activity, and low activity was found with high concentrations of GlcNAc\( \beta 1\)–3Gal-methyl but not with lacto-N-neo-tetraose or para-lacto-N-hexaose, which indicated that C2GnT3 exhibits low in vitro distal IGnT activity. C2GnT3 activity was also evaluated with glycoprotein acceptors. There was high activity with asialoglycophorin A and asialofetuin and lower activity with bovine asialo-submaxillary mucin (Table II). No activity was detected with human \( \alpha 1\)-acid glycoprotein, fetuin, IgG, and transferrin. C2GnT3 exhibited strict donor substrate specificity for UDP-GlcNAc and did not utilize UDP-galactose, UDP-\( \text{N} \)-acetylgalactosamine, UDP-glucose, or UDP-xylose with the acceptor substrates tested here (data not shown) (4). The in vitro activity of C2GnT3 was enhanced by several detergents including Nonidet P-40. In our hands recombinant secreted forms of C2GnT1 and C2GnT2 expressed in Sf9 cells are inactivated by detergents (4).

C2GnT3 Controls Core 2 Branching in Vivo—C2GnT1 and C2GnT2 generate core 2 \( \beta 1,6\)-N-acetylglucosaminyltransferase 3
indicating C2GnT3 does not exhibit IGnT activity in vivo. The low activity found in vitro with GlcNAc β-1–3Gal β-1-Me (Table I) appears to be in conflict with this result, but in vitro evaluation of I branching activity is complicated by lack of suitable complex acceptor substrates. Furthermore, analysis of in vitro and in vivo I branching activity with C2GnT2 clearly revealed discrepancies. Thus, Yeh et al. (5) found very low in vitro activity of C2GnT2 with a tetrasaccharide substrate, and Schwientek et al. (4) found no I branching activity with simpler substrates. In contrast, in vivo analysis of I branching activity of C2GnT2 demonstrated a good I branching function similar to IGnT (Fig. 4 (5)).

**Product Characterization by 1H and 13C NMR Spectroscopy**—The product generated by C2GnT3 using Galβ-1–3GalNAc-1-para-nitrophenyl was characterized by NMR spectroscopy.
phenyl and Gal
b(35), all 1H assignments agreed with those published previously for Gal\(\beta\)1–3(GlcNAc\(\beta\)1–6)GalNAc1–para-nitrophenyl, showing all nonexchangeable monosaccharide ring methine and exocyclic methylene resonances. Residue designations for Gal\(\beta\)1–3 (\(\beta\)), Gal\(\beta\)1–6 (\(\beta\)), and GalNAc1–1 (\(\alpha\)) are followed by proton designations (1–6). Resonances from an as yet undetermined alternate product are marked by asterisks.

**Fig. 5.** Sections of a one-dimensional 1H NMR spectrum of the C2GnT product, Gal\(\beta\)1–3(GlcNAc\(\beta\)1–6)GalNAc1–para-nitrophenyl, showing all nonexchangeable monosaccharide ring methine and exocyclic methylene resonances. Residue designations for Gal\(\beta\)1–3 (\(\beta\)), Gal\(\beta\)1–6 (\(\beta\)), and GalNAc1–1 (\(\alpha\)) are followed by proton designations (1–6). Resonances from an as yet undetermined alternate product are marked by asterisks.

**Table II**

| Substrate | C2GnT3* |
|-----------|---------|
| Asialo-fetuin | 2 mg/ml 10 mg/ml |
| Asialo-glycoprotein A | 97.1 ND* |
| Asialo-bovine submaxillary mucin | 1.7 7.2 |
| \(\alpha\)-Acid glycoprotein | <1.0 <1.0 |
| Fetuin | <1.0 <1.0 |
| Immunoglobulin G | <1.0 <1.0 |
| Transferin | <1.0 <1.0 |
| Hen egg albumin | <1.0 <1.0 |

a Enzyme sources were as described in Table I.
b ND, not determined.

copy. Comparison of a one-dimensional 1H NMR spectrum of the product (Fig. 5) with that of the substrate (data not shown) clearly showed an additional H-1 resonance (4.457 ppm) from a sugar residue linked in the \(\beta\)-configuration \((J_{1,2} = 7–9 \text{ Hz})\). Some resonances from a minor product (not unreacted substrate) were also observed (Fig. 5, asterisks). Complete assignments for all 1H and 13C resonances of both major product and substrate were obtained from sequential TOCSY and HSQC experiments (Table III). With the exception of some H-6R and H-6S pairs (assigned by comparison of their 3J,5,6 coupling constants with those of the corresponding benzylglycosides (35)), all 1H assignments agreed with those published previously for both Gal\(\beta\)1–3(GlcNAc\(\beta\)1–6)GalNAc1–para-nitrophenyl and Gal\(\beta\)1–3GalNAc1–para-nitrophenyl (18). Linkage assignments in the product, previously made on the basis of interglycosidic nuclear Overhauser enhancements (18), were confirmed unambiguously by observation of appropriate interglycosidic H1–C1–O1–Cx and C1–O1–Cx–Hx correlations in an HMBC spectrum (data not shown). As observed previously for a biosynthetic core 4-para-nitrophenyl product (4), the newly formed GlcNAc\(\beta\)1–6GalNAc linkage in the putative core 2-para-nitrophenyl product was clearly demonstrated by strong cross-peaks correlating the \(\beta\)-GlcNAc H-1 at 4.457 ppm with \(\alpha\)-GalNAc C-6, and the corresponding \(\beta\)-GlcNAc C-1 at 100.95 ppm with both \(\alpha\)-GalNAc H-6 resonances. The structure of the minor product is under investigation.

**Expression of C2GnT3 in Human Tissues and Cancer Cell Lines**—mRNA analyses with human mRNA from 50 adult normal organs revealed a highly selective expression pattern with high levels of a C2GnT3 transcript (approximately 5.5 kilobases) exclusively in thymus (Figs. 6 and 7). Low transcript levels were also found in pancreas, peripheral blood leukocytes, placenta, small intestine, and stomach, whereas expression in kidney, liver, spleen, lung, and lymph node was barely detectable. Expression of C2GnT3 in lymphoid cancer cells was determined in a panel of human cancer cell lines. The lymphoblastic leukemia cell line MOLT-4 expressed the C2GnT3 transcript, whereas it was not detected in HL-60, HeLaS3, K-562, Raji, SW480, A549, and C361 (data not shown). The size of the C2GnT3 transcript was similar to the largest of three transcripts of C2GnT1 (5.4 kilobases) (3). Multiple transcripts of C2GnT1 and -T2 are proposed to result from differential usage of polyadenylation signals, but this has not been confirmed (3).

**DISCUSSION**

At least three human \(\beta\)GalNAc-Ts are involved in O-glycan branching. The C2GnT3 presented in this report is similar to C2GnT1 in that it functions solely in core 2 O-glycan branching (3). In contrast, C2GnT2 has more diverse functionality in core 2 and core 4 O-glycan branching as well as in synthesis of I poly-N-acetyllactosamine structures (4, 5). The catalytic activities of all three enzymes have been studied in vitro with saccharide and glycoprotein substrates and in vivo in CHO cells using immunoprecipitation for product identification.

**Table III**

| Substrate | C2GnT3 |
|-----------|--------|
| Asialo-fetuin | 15.5 51.3 |
| Asialo-glycoprotein A | 97.1 ND* |
| Asialo-bovine submaxillary mucin | 1.7 7.2 |
| \(\alpha\)-Acid glycoprotein | <1.0 <1.0 |
| Fetuin | <1.0 <1.0 |
| Immunoglobulin G | <1.0 <1.0 |
| Transferin | <1.0 <1.0 |
| Hen egg albumin | <1.0 <1.0 |

a Enzyme sources were as described in Table I.
b ND, not determined.

c copy. Comparison of a one-dimensional 1H NMR spectrum of the product (Fig. 5) with that of the substrate (data not shown) clearly showed an additional H-1 resonance (4.457 ppm) from a sugar residue linked in the \(\beta\)-configuration \((J_{1,2} = 7–9 \text{ Hz})\). Some resonances from a minor product (not unreacted substrate) were also observed (Fig. 5, asterisks). Complete assignments for all 1H and 13C resonances of both major product and substrate were obtained from sequential TOCSY and HSQC experiments (Table III). With the exception of some H-6R and H-6S pairs (assigned by comparison of their 3J,5,6 coupling constants with those of the corresponding benzylglycosides (35)), all 1H assignments agreed with those published previously for both Gal\(\beta\)1–3(GlcNAc\(\beta\)1–6)GalNAc1–para-nitrophenyl and Gal\(\beta\)1–3GalNAc1–para-nitrophenyl (18). Linkage assignments in the product, previously made on the basis of interglycosidic nuclear Overhauser enhancements (18), were confirmed unambiguously by observation of appropriate interglycosidic H1–C1–O1–Cx and C1–O1–Cx–Hx correlations in an HMBC spectrum (data not shown). As observed previously for a biosynthetic core 4-para-nitrophenyl product (4), the newly formed GlcNAc\(\beta\)1–6GalNAc linkage in the putative core 2-para-nitrophenyl product was clearly demonstrated by strong cross-peaks correlating the \(\beta\)-GlcNAc H-1 at 4.457 ppm with \(\alpha\)-GalNAc C-6, and the corresponding \(\beta\)-GlcNAc C-1 at 100.95 ppm with both \(\alpha\)-GalNAc H-6 resonances. The structure of the minor product is under investigation.

**Expression of C2GnT3 in Human Tissues and Cancer Cell Lines**—mRNA analyses with human mRNA from 50 adult normal organs revealed a highly selective expression pattern with high levels of a C2GnT3 transcript (approximately 5.5 kilobases) exclusively in thymus (Figs. 6 and 7). Low transcript levels were also found in pancreas, peripheral blood leukocytes, placenta, small intestine, and stomach, whereas expression in kidney, liver, spleen, lung, and lymph node was barely detectable. Expression of C2GnT3 in lymphoid cancer cells was determined in a panel of human cancer cell lines. The lymphoblastic leukemia cell line MOLT-4 expressed the C2GnT3 transcript, whereas it was not detected in HL-60, HeLaS3, K-562, Raji, SW480, A549, and C361 (data not shown). The size of the C2GnT3 transcript was similar to the largest of three transcripts of C2GnT1 (5.4 kilobases) (3). Multiple transcripts of C2GnT1 and -T2 are proposed to result from differential usage of polyadenylation signals, but this has not been confirmed (3).

**DISCUSSION**

At least three human \(\beta\)GalNAc-Ts are involved in O-glycan branching. The C2GnT3 presented in this report is similar to C2GnT1 in that it functions solely in core 2 O-glycan branching (3). In contrast, C2GnT2 has more diverse functionality in core 2 and core 4 O-glycan branching as well as in synthesis of I poly-N-acetyllactosamine structures (4, 5). The catalytic activities of all three enzymes have been studied in vitro with saccharide and glycoprotein substrates and in vivo in CHO cells using immunoprecipitation for product identification.

**Table III**

| Substrate | C2GnT3 |
|-----------|--------|
| Asialo-fetuin | 15.5 51.3 |
| Asialo-glycoprotein A | 97.1 ND* |
| Asialo-bovine submaxillary mucin | 1.7 7.2 |
| \(\alpha\)-Acid glycoprotein | <1.0 <1.0 |
| Fetuin | <1.0 <1.0 |
| Immunoglobulin G | <1.0 <1.0 |
| Transferin | <1.0 <1.0 |
| Hen egg albumin | <1.0 <1.0 |

a Enzyme sources were as described in Table I.
b ND, not determined.
the EST data bases, but the specific reason for this is not clarified. The C2GnT3 gene identified here was shown to produce an mRNA transcript of approximately 5.5 kilobases, mainly in the thymus (Fig. 6). The available 3′-UTR sequence information did not contain a consensus polyadenylation signal, and BLAST searches against the EST data base did not provide ESTs derived from this sequence. It is conceivable that further 3′ sequence contains the polyadenylation signal.

The in vitro kinetic properties of C2GnT3 resembled those of C2GnT1 (4, 5, 28) and infer its exclusive function in core 2 O-glycan synthesis. C2GnT3 efficiently utilized glycoproteins carrying unsubstituted core 1 structures including asialo-glycophorin A and asialo-fetuin. Asialo-bovine submaxillary mucin, with only 6% fucosylated or unsubstituted core 1, was a poor substrate (36, 37). A similar pattern of activity toward glycoproteins was recently obtained with recombinant mouse C2GnT1 (38). Both C2GnT1 and C2GnT3 exhibited core 2 synthase activity in vivo with CD43, suggesting that both enzymes function with the same glycoprotein acceptor (33, 34). Glycosphingolipid substrates have not been tested with any of the human β6GlcNAc-Ts, mainly because of a lack of availability of these complex glycolipids. However, a kidney-associated form of murine C2GnT1 was shown to regulate expression of hybrid-type lacto-globoseries glycolipids (Gal[b1-3][GlcNAc[b1-6]GlcNAc[b1-3]Gal[b1-4]Gal[b1-4]Glc[b1-Cer]) (39). To our knowledge β6GlcNAc branching has only been found in (neo-)lacto-series glycolipids in man.

Several previous studies described developmentally regulated core 2 GlcNAc-T activities during various cellular events (18, 20, 21, 34, 40), but these have only been correlated with specific β6GlcNAc-T gene expression in a few instances (13, 41). Nakamura et al. (42) found that enzyme activity and transcript levels of C2GnT1 are coordinately down-regulated during 12-O-tetradecanoylphorbol-13-acetate treatment of the human leukemia cell line KM3. Changes in C2GnT1 expression in the developing mouse embryo were shown by in situ hybridization (43). A recent study found a 100-fold difference in enzyme activity but unchanged levels of murine C2GnT1 mRNA upon induced retrodifferentiation of the embryonal cell lines PSA-5E and PYS-2 (44). Thus, the specific contribution of individual β6GlcNAc-T isoforms to core 2 and I branching activities in cells and tissues are still unclear.

Northern analysis indicated highly restricted expression of C2GnT3 mainly to the thymus (Fig. 6). C2GnT1 is variably expressed in most organs tested, and the highest levels are found in thyroid, spleen, and mucosal tissues (5). Expression of C2GnT1 in mucosal tissues resembles the expression pattern of C2GnT2 (4, 5). C2GnT3 is also weakly expressed in small intestine and stomach. C2GnT1 was found by Northern analysis to be weakly expressed in thymus (5), and in situ hybridization localized C2GnT1 expression to thymocytes (13). Thus, among the known core 2 O-glycan synthase genes, C2GnT3 appears to be the dominant core 2 synthase in thymus. One possibility is that C2GnT3 has important functions in synthesis of core 2 O-glycans on cortical thymocytes (13). Core 2 structures on thymocyte surface proteins are ligands for galectin-1, which participates in interactions between thymocytes and thymic epithelium (13). Both C2GnT1 and C2GnT3 are weakly expressed in peripheral blood leukocytes and lymph nodes; however, further studies are required to define specific cell types. C2GnT1 is highly expressed in spleen indicating strong expression in B-cells, which is in agreement with the finding that C2GnT1 expression correlates with O-glycan

---

3 E. P. Bennett, T. Schwientek, and H. Clausen, unpublished observations.
branch formation and sialyl-Le\(^a\) expression in the human pre-B lymphocytic leukemia cell line KM3 (42). In contrast, strong expression of C2GnT3 in thymus indicates association with T-cells. In situ hybridization and/or immunohistology with appropriate probes are required for further clarification.

The \(\beta\)6GlcNAc-T gene family consists of at least four glycosyltransferases with acceptor substrate specificity for galactose and/or N-acetylgalactosamine and one pseudogene (45). The genomic organizations of the three C2GnT genes are similar with the coding regions contained in a single exon. Intronless in the 5'-UTR of C2GnT1 and C2GnT2 have been found (25). 4

In contrast, the coding region of C2GnT3 is contained in three exons (25). The genes are located at different chromosomal loci as follows: C2GnT1 is at 6p24; C2GnT2 is at 9q21; C2GnT3 is at 15q22; and C2GnT3 is at 5q12 (4, 5, 29). The relatively high sequence similarity among the coding regions of the core 2 synthases and their similar genomic organizations indicate that these arose through gene duplication followed by sequence divergence. The more complex genomic organization of C2GnT combined with lower sequence similarity to C2GnTs suggests that C2GnT represents a distinct member of the \(\beta\)6GlcNAc-T family that diverged early in evolution. Surprisingly, a total of six homologues of the C2GnT gene family have been identified in the Caenorhabditis elegans genome, but functional characterizations of these genes have not been reported (46). The characteristics of the \(\beta\)6GlcNAc-T gene family resemble those of other homologous glycosyltransferase gene families in terms of genes and primary structures of the proteins. Moreover, there are differences in kinetic properties and expression patterns of individual isoforms of each family. Thus, it appears likely that gene duplication and divergence was influenced by functional requirements for synthesis of different glycoconjugates. Another possible influence was the need to obtain differential regulation of glycoconjugate synthesis in different cell types during various biological and developmental processes.

The biological roles of core 2 O-glycans are among the most extensively studied in the carbohydrate field. Understanding the regulatory mechanisms of core 2 O-glycosylation is of high interest. Identification and characterization of the core 2 synthase genes is a first step. The finding that C2GnT3 is a thymus-associated core 2 synthase is intriguing and holds promise for future studies of its role in the immune system.

Acknowledgment—We thank Dr. Michael A. Hollingsworth for helpful suggestions and critical reading of the manuscript.

REFERENCES

1. Clausen, H., and Bennett, E. P. (1996) Glycobiology 6, 635–646
2. Brockhausen, I., and Schachter, H. (1997) in Glycosciences (Gabius, H.-J., and Gabius, S., eds) pp. 79–113, Chapman and Hall, Weinheim, Germany
3. Bierhuizen, M. F., and Fukuda, M. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 9226–9230
4. Schwientek, T., Nomoto, M., Lavery, S. B., Merkx, G., van Kessel, A. G., Bennett, E. P., Hollingsworth, M. A., and Clausen, H. (1999) J. Biol. Chem. 274, 4504–4512
5. Yeh, J.-C., Ong, E., and Fukuda, M. (1999) J. Biol. Chem. 274, 3215–3221
6. Brockhausen, I., Kuhns, W., Schachter, H., Matta, K. L., Sutherland, D. R., and Baker, M. A. (1991) Cancer Res. 51, 1257–1263
7. Schwientek, T., Almeida, R., Levery, S. B., Bennett, E., and Clausen, H. (1998) J. Biol. Chem. 273, 29331–29340
8. Ujita, M., McAuliffe, J., Schwientek, T., Almeida, R., Hindsgaul, O., Clausen, H., and Fukuda, M. (1998) J. Biol. Chem. 273, 34843–34849
9. Sasaki, K., Kurata-Miura, K., Ujita, M., Angata, K., Nakagawa, S., Sekine, S., Nishii, T., and Fukuda, M. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 14294–14299
10. Zhou, D., Henion, T., Altmann, F., Berger, E. G., and Hennet, T. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 406–411
11. Zhou, D., Henion, T., Altmann, F., Berger, E. G., and Hennet, T. (1999) Glycobiology 9, 1121 (abstr.)
12. Zhou, D., Henion, T., Altmann, F., Berger, E. G., and Hennet, T. (1999) J. Biol. Chem. 273, 29331–29340
13. Baum, L. G., Pang, M., Perillo, N. L., Wu, T., Delegeane, A., Uittenbogaart, C. H., Fukuda, M., and Seilhamer, J. J. (1995) J. Biol. Chem. 270, 14294–14299
14. Zhou, D., Dinter, A., Gutierrez, G. B., Kamerling, J. P., Vliegenthart, J. F., Berger, E. G., and Hennet, T. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 14294–14299
15. Zhou, D., Henion, T., Altmann, F., Berger, E. G., and Hennet, T. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 406–411
16. Zhou, D., Henion, T., Altmann, F., Berger, E. G., and Hennet, T. (1999) J. Biol. Chem. 273, 29331–29340
17. Elies, L. G., Tsuboi, S., Petryniak, B., Lowe, J. B., Fukuda, M., and Marth, J. D. (1998) Immunity 9, 881–890
18. Yousefi, S., Higgins, E., Dasling, Z., Pellex-Kruger, A., Hindsgaul, O., and Dennis, J. W. (1991) J. Biol. Chem. 266, 1772–1782
19. Fukuda, M. (1996) Cancer Res. 56, 2237–2244
20. Piller, F., Piller, V., Fox, R. I., and Fukuda, M. (1998) J. Biol. Chem. 263, 15146–15150
21. Yang, J. M., Byrd, J. C., Siddiki, B. B., Chung, Y. S., Okuno, M., Sowa, M.,

---

\(^4\) T. Schwientek and H. Clausen, unpublished observations.
Kim, Y. S., Matta, K. L., and Brockhausen, I. (1994) Glycobiology 4, 873–884.
22. Williams, D., and Schachter, H. (1980) J. Biol. Chem. 255, 11247–11252.
23. Ropp, P. A., Little, M. R., and Cheng, P. W. (1991) J. Biol. Chem. 266, 23863–23871.
24. Kuhns, W., Rutz, V., Paulsen, H., Matta, K. L., Baker, M. A., Barner, M., Granovsky, M., and Brockhausen, I. (1993) Glycoconj. J. 10, 381–384.
25. Bierhuizen, M. F., Maemura, K., Kudo, S., and Fukuda, M. (1995) Glycobiology 5, 417–425.
26. Bennett, E. P., Hassan, H., Mandel, U., Mirgorodskaya, E., Roepstorff, P., Burchell, J., Taylor-Papadimitriou, J., Hollingsworth, M. A., Merkx, G., van Kessel, A. G., Eiber, H., Steffensen, R., and Clausen, H. (1998) J. Biol. Chem. 273, 30472–30481.
27. Wandel, 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.
28. Bierhuizen, M. F., Maemura, K., and Fukuda, M. (1994) J. Biol. Chem. 269, 4473–4479.
29. Bierhuizen, M. F., Mattei, M. G., and Fukuda, M. (1993) Genes Dev. 7, 468–478.
30. Almeida, R., Amado, M., David, L., Levery, S. B., Holmes, R. H., Merkx, G., van Kessel, A. G., Rygaard, E., Hassan, H., Bennett, E., and Clausen, H. (1997) J. Biol. Chem. 272, 31979–31991.
31. Kozak, M. (1999) Annu. Rev. Cell Biol. 8, 197–225.
32. Toki, D., Sarkar, M., Yip, B., Reck, F., Joziassie, D., Fukuda, M., Schachter, H., and Brockhausen, I. (1997) Biochem. J. 325, 63–69.
33. Fox, R. I., Hueniken, M., Fong, S., Behar, S., Royston, I., Singhal, S. K., and Thompson, L. (1983) J. Immunol. 131, 762–767.
34. Saitoh, O., Filler, P., Fox, R. I., and Fukuda, M. (1991) Blood 77, 1491–1499.
35. Polle-Kruger, A., Meyer, B., Stuoke-Prill, R., Sinnwell, V., Matta, K. L., and Brockhausen, I. (1993) Glycoconj. J. 10, 365–380.
36. Spiro, R. G., and Bhoyroo, V. D. (1974) J. Biol. Chem. 249, 5704–5717.
37. Martensson, S., Levery, S. B., Fang, T. T., and Bendahl, B. (1998) Eur. J. Biochem. 258, 603–622.
38. Zeng, S., Dinter, A., Eisenkrautzer, D., Biselli, M., Wondrey, C., and Berger, E. G. (1997) Biochem. Biophys. Res. Commun. 237, 653–658.
39. Sekine, M., Hashimoto, Y., Suzuki, M., Inagaki, F., Takio, K., and Suzuki, A. (1994) J. Biol. Chem. 269, 31143–31148.
40. Higgin, E. A., Simonovitch, K. A., Zhunag, D. L., Brockhausen, I., and Dennis, J. W. (1991) J. Biol. Chem. 266, 6280–6290.
41. Shimodaira, K., Nakayama, J., Nakamura, H., Hasebe, O., Katsuyama, T., and Fukuda, M. (1997) Cancer Res. 57, 5201–5206.
42. Nakamura, M., Kudo, T., Narimatsu, H., Furukawa, Y., Kikuchi, J., Asakura, S., Yang, W., Iwase, S., Hatake, K., and Miura, Y. (1998) J. Biol. Chem. 273, 26779–26789.
43. Granovsky, M., Fode, C., Warren, C. E., Campbell, R. M., Marth, J. D., Pierce, M., Fregien, N., and Dennis, J. W. (1995) Glycobiology 5, 797–806.
44. VanderElst, I. E., and Datti, A. (1998) Glycobiology 8, 731–740.
45. Bierhuizen, M. F., Maemura, K., and Fukuda, M. (1995) Glycoconj. J. 12, 857–864.
46. Warren, C. E., Krizus, A., Partridge, E., Lau, J. Y. T., Roy, P. J., Zhang, L., Culotti, J. G., and Dennis, J. W. (1999) Glycobiology 9, 1111 (abstr.).
Control of O-Glycan Branch Formation: MOLECULAR CLONING AND CHARACTERIZATION OF A NOVEL THYMUS-ASSOCIATED CORE 2 β 1,6-N-ACETYLGUCOSAMINYLTRANSFERASE

Tilo Schwientek, Jiunn-Chern Yeh, Steven B. Levery, Birgit Keck, Gerard Merkx, Ad Geurts van Kessel, Minoru Fukuda and Henrik Clausen

J. Biol. Chem. 2000, 275:11106-11113.
doi: 10.1074/jbc.275.15.11106

Access the most updated version of this article at http://www.jbc.org/content/275/15/11106

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 45 references, 29 of which can be accessed free at http://www.jbc.org/content/275/15/11106.full.html#ref-list-1