Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 9326–9330, October 1992
Biochemistry

Expression cloning of a cDNA encoding UDP-GlcNAc:Galβ1-3-GalNAc-R (GlcNAc to GalNAc) β1-6GlcNAc transferase by gene transfer into CHO cells expressing polyoma large tumor antigen

(glycosyltransferase/core 2 O-glycan/branching enzyme/oligosaccharide biosynthesis)

MARTI F. A. BIERHUIZEN AND MINORU FUKUDA

La Jolla Cancer Research Foundation, Cancer Research Center, 10901 North Torrey Pines Road, La Jolla, CA 92037

Communicated by Marilyn G. Farquhar, June 23, 1992

ABSTRACT **A cDNA encoding UDP-GlcNAc:Galβ1-3GalNAc-R (GlcNAc to GalNAc) β1-6GlcNAc transferase (EC 2.4.1.102), which forms critical branches in O-glycans, has been isolated by an expression cloning approach using Chinese hamster ovary (CHO) cells. Increased activity of this enzyme and the concomitant occurrence of the O-glycan core 2 structure [Galβ1-3(GlcNAcβ1-6)GalNAc] has been observed in a variety of biological processes, such as T-cell activation and immunodeficiency due to the Wiskott–Aldrich syndrome and AIDS. Since CHO cells do not express this enzyme, CHO cell lines were established to stably express polyoma large tumor (T) antigen, which enables transient expression cloning. Because the antibody used was found to detect most efficiently the oligosaccharide products attached to leukosialin, the CHO cells were also stably transfected with leukosialin cDNA. By using this particular CHO cell line, a cDNA that encodes a protein determining the formation of the core 2 structure was isolated from an HL-60 cDNA library. The cDNA sequence predicts a protein with type II membrane topology, as has been found for all other mammalian glycosyltransferases cloned to date. The expression of the presumed catalytic domain as a fusion protein with the IgG binding domain of protein A enabled us to demonstrate unequivocally that the cDNA encodes the core 2 β-1,6-*N*-acetylglucosaminyltransferase, the enzyme responsible for the formation of Galβ1-3(GlcNAcβ1-6)GalNAc structures. No activity with this enzyme was detected toward the acceptors for other β1-6GlcNAc transferases.**

Most O-glycosidic oligosaccharides in mammalian glycoproteins are linked via GalNAc to the hydroxyl groups of serine or threonine. Although less well studied than N-glycans, recent studies suggest that O-glycans also have important biological roles. In particular, the appearance of the O-linked oligosaccharides with the core 2 branch (see Fig. 1), Galβ1-3(GlcNAcβ1-6)GalNAc, has been demonstrated in many biological processes, such as human T-cell activation (1), leukemias (2, 3), and immunodeficiency due to AIDS (2) and Wiskott–Aldrich syndrome (4). In all instances, it was discovered that these changes are associated with an increase in the activity of UDP-GlcNAc:Galβ1-3GalNAc-R (GlcNAc to GalNAc) β1-6GlcNAc transferase (EC 2.4.1.102), or core 2 β1-6GlcNAc transferase (C2GnT) in short (1–4). In human leukocytes, these O-glycans are carried by a major sialoglycoprotein, leukosialin (or CD43) (5). Leukosialin containing the hexasaccharides exhibits a larger molecular mass (≈135 kDa) than that exhibited by leukosialin with the tetrasaccharides (≈105 kDa) (ref. 6; see Fig. 1). A monoclonal antibody, T305, originally raised against human T-lymphocytic leukemia cells, was found to specifically react with leukosialin of high molecular mass (2, 4). Since the T305 binding can be abolished by neuraminidase treatment (7), the T305 antibody was judged to bind the hexasaccharides (2, 4).

Our earlier studies indicated that poly(*N*-acetyllactosamine) repeats are almost exclusively extended from the branch formed by the C2GnT (8). Consistent with these results, Yousefi *et al.* (9) demonstrated in metastatic tumor cells that the core 2 enzyme regulates the levels of poly(*N*-acetyllactosamine) synthesis in O-linked oligosaccharides. Thus it appears that the formation of the core 2 structure is critical for the formation of poly(*N*-acetyllactosamine) in O-glycans.

In this paper, we report our systematic efforts to clone cDNA for human C2GnT by transient expression cloning (10, 11) in polyoma large tumor (T) antigen-expressing Chinese hamster ovary (CHO) cells (12). Since the monoclonal antibody T305 was used for selection, we first describe the establishment of a CHO cell line expressing both human leukosialin as well as polyoma large T antigen. Subsequently, we describe the molecular cloning of cDNA for C2GnT and show that the expressed protein specifically forms the core 2 branch.[†]

MATERIALS AND METHODS

Transient Transfection of Cells. Cells were transfected with plasmids by using a modification of the lipofection procedure (13). For each 100-mm tissue culture dish with 20–40% confluent cells, 50 μg of lipofectin reagent (Bethesda Research Laboratories) and 20 μg of CsCl gradient-purified plasmid DNA were diluted separately to 1.5 ml with Opti-MEM I (GIBCO), mixed, and then added to the cells. The cells were incubated for 6 hr, and then 10 ml of complete medium was added. After incubation for 16 hr at 37°C, the medium was replaced with 10 ml of fresh medium, and the cells were incubated for an additional 48 hr at 37°C.

Construction of Stably Transfected CHO Cell Lines Expressing the Polyoma Virus T Antigen and Human Leukosialin. *Expression vectors*. The plasmid vector harboring the polyoma virus early genes under the control of the simian virus 40 early promoter, designated here as pPSVE1-PyE, was constructed as described (14) with a slight modification. First, plasmid pPSVE1 was prepared by using pPSG4 DNA (ATCC no. 37337) and simian virus 40 viral DNA (Bethesda Research Laboratories). pPyLT-1 DNA (ATCC no. 41043) was then digested with *Eco*RI/*Hin*cII to generate the COOH-terminal coding sequences for the polyoma virus large T antigen. An *Eco*RI site at the 3′ end was introduced by blunt-end ligation of this *Eco*RI/*Hin*cII polyoma large T fragment with phos-

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.

Abbreviations: C2GnT, UDP-GlcNAc:Galβ1-3GalNAc-R (GlcNAc to GalNAc) β1-6GlcNAc transferase (core 2 β1-6GlcNAc transferase); GnT, β1-6GlcNAc transferase; T antigen, tumor antigen.

[†]The sequence reported in this paper has been deposited in the GenBank data base (accession no. M97347).

phorylated *Eco*RI linkers (Stratagene). Plasmid pPSVE1-PyE was then generated by ligating the COOH-terminal coding sequence for the T antigen into the unique *Eco*RI site of plasmid pPSVE1.

The plasmid vector pZIPNEO-leu was constructed by introducing the *Eco*RI insert of PEER-3 cDNA (15), which contains the complete coding sequence for human leukosialin, into the unique *Eco*RI site of plasmid pZIPNEO (16). pZIPNEO was kindly provided by Channing Der of this institute.

Isolation of stably transfected CHODG44 cells. Polyoma large T antigen- and human leukosialin-expressing cell lines were established by cotransfecting CHODG44 cells with pZIPNEO-leu and pPSVEI-PyE in a 1:4 molar ratio using a calcium phosphate technique (17) and by subsequent selection for G418 resistance. Polyoma virus large T antigen-mediated replication of plasmids in these cell lines was assessed by measurement of the methylation status of the recombinant DNA (12, 14). Finally, one clonal CHO cell line, named CHO-Py·leu, was chosen for transient expression cloning.

Isolation of a Human C2GnT cDNA Clone. A cDNA library, pcDNAI-HL-60, was constructed from poly(A)$^+$ RNA isolated from HL-60 promyelocytic cells and the mammalian expression vector pcDNAI (Invitrogen). Plasmid DNA was transfected into CHO-Py·leu cells using lipofection as described above. After a 64-hr expression period, the transfected cells were detached at 37°C in phosphate-buffered saline/5 mM EDTA, pH 7.4 for 30 min. The detached cells were pooled, centrifuged, and resuspended in cold phosphate-buffered saline/10 mM EDTA/5% fetal calf serum, pH 7.4, containing T305 monoclonal antibody as ascites fluid in a 1:200 dilution. After a 1-hr incubation on ice, the cells were washed and panned (10) on dishes coated with goat anti-mouse IgG (Sigma). T305 monoclonal antibody was kindly provided by R. I. Fox at the Scripps Research Foundation, La Jolla, CA.

Plasmid DNA molecules were recovered from adherent cells by the Hirt procedure (18). Hirt DNA was first digested with *Dpn* I to remove plasmids that were not replicated in transfected cells and then transformed into the host *Escherichia coli* MC1061/P3. Plasmid DNA was prepared again and used for an additional round of screening by the same procedure. *E. coli* transformants containing plasmids recovered from this second enrichment were plated to yield eight pools of ≈500 colonies each, and replica plates were made (19). Plasmid DNA was prepared from replica plates and transfected separately into the CHO-Py·leu cells, and the transfectants were screened by panning as described above. One of the plasmid pools was selected, and three subsequent rounds of sib selection with sequentially smaller, active pools identified a single plasmid that determined the expression of the T305 antigen at the cell surface.

Construction and Analysis of the Protein A–C2GnT$_c$ Fusion Vector. The cDNA fragment encoding the putative catalytic plus stem domains of C2GnT was prepared by PCR and fused with cDNA encoding a signal peptide sequence and the IgG binding domain of protein A (11). Thus the cDNA fragment encoding amino acid residues 38–428 of C2GnT was ligated into the unique *Eco*RI site of pPROTA (20), yielding plasmid pPROTA-C2GnT$_c$. Plasmids pPROTA and pPROTA-C2GnT$_c$ were separately transfected into COS-1 cells, and after a 64-hr expression period the cell supernatants were collected and processed exactly as described before (11).

DNA Sequence and RNA Blot Analysis. The nucleotide sequences were determined (21) by using oligonucleotides synthesized according to the flanking sequences and obtained sequences within the insert. The protein sequence was analyzed using the PCGENE program to search Genpro release 6.6. RNA blots (19) were hybridized with the ^{32}P-radiolabeled (22) *Eco*RI insert of pPROTA-C2GnT$_c$ at 42°C.

β1-6GlcNAc Transferase (GnT) Assays and Product Characterization. GnT assays were performed essentially as described (2, 9, 23). In all instances the reaction mixtures contained 50 mM Mes (pH 7.0), 0.5 μCi (1 Ci = 37 GBq) of UDP-[^{3}H]GlcNAc in 1 mM UDP-GlcNAc, 0.1 M GlcNAc, 10 mM Na_2EDTA, 1 mM acceptor, and 25 μl of cell lysate, cell supernatant, or IgG-Sepharose matrix in a total reaction volume of 50 μl. The reaction mixtures were incubated for 1 hr at 37°C, followed by C_{18} Sep-Pak (Waters) processing as described (24). C2GnT and C4GnT were assayed by using the acceptors *p*-nitrophenyl-Galβ1-3GalNAc and *p*-nitrophenyl-GlcNAcβ1-3GalNAc, respectively (Toronto Research Chemicals). UDP-GlcNAc:GlcNAcβ1-2Man GnTV was assayed by using GlcNAcβ1-2Manα1-6Glcβ-O-$(CH_2)_7CH_3$ as an acceptor. UDP-GlcNAc:GlcNAcβ1-3Galβ1-4GlcNAc (GlcNAc to Gal) GnT, the blood group I enzyme (25), was assayed by using GlcNAcβ1-3Galβ1-4GlcNAcβ1-6Manα1-6Manβ1-O-$(CH_2)_8COOCH_3$ and Galβ1-4GlcNAcβ1-3Galβ1-4GlcNAcβ1-3Galβ1-4GlcNAc-β1-O-$(CH_2)_7CH_3$ as acceptors. These synthetic acceptors were kindly provided by Olé Hindsgaul, University of Alberta, Edmonton, Canada.

The radiolabeled C2GnT reaction product was purified by C_{18} Sep-Pak column chromatography as described above and then analyzed by HPLC on a column (0.4 × 25 cm) of Lichrosorb NH_2 bonded silica. The column was developed isocratically with acetonitrile/water, 83:17 (vol/vol) under the conditions described (26).

RESULTS

Establishment of CHO Cell Lines That Stably Express Polyoma Large T Antigen and Leukosialin. Since our preliminary results indicated that COS-1 cells express a low, but detectable, level of C2GnT, it was necessary to find other recipient cells for expression cloning. In addition, the monoclonal antibody T305 was found to react primarily with the hexasaccharides (Fig. 1) attached to leukosialin. As shown previously (27), CHO cells express the tetrasaccharides but lack the hexasaccharides. Furthermore, no C2GnT activity could be detected in CHO cell lysates. Therefore, we decided to establish CHO cell lines that stably express human leukosialin and polyoma large T antigen, using a procedure similar to that described by Heffernan and Dennis (12). Expression of the polyoma large T antigen allows the replication of a plasmid vector harboring the polyoma replication

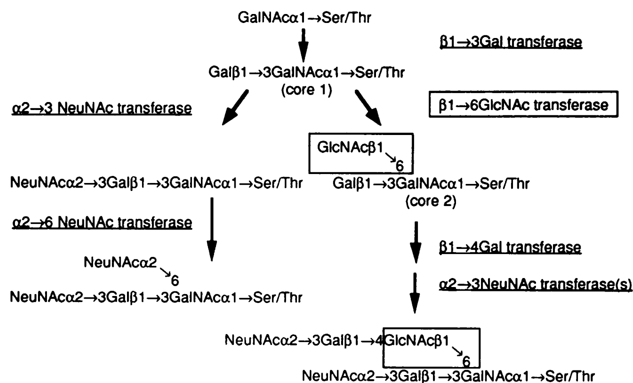

FIG. 1. Structures and biosynthesis of O-glycans. The core 1 structure can be converted to core 2 by the addition of a GlcNAcβ1-6 residue, eventually forming the hexasaccharide (the bottom right). The C2GnT and the linkage formed by the enzyme are indicated by boxes. If C2GnT is absent, core 1 is eventually converted to the disialoform by sequential addition of (α2-3)- and (α2-6)-linked sialic acid residues (the bottom left). NeuNAc, *N*-acetylneuraminic acid.

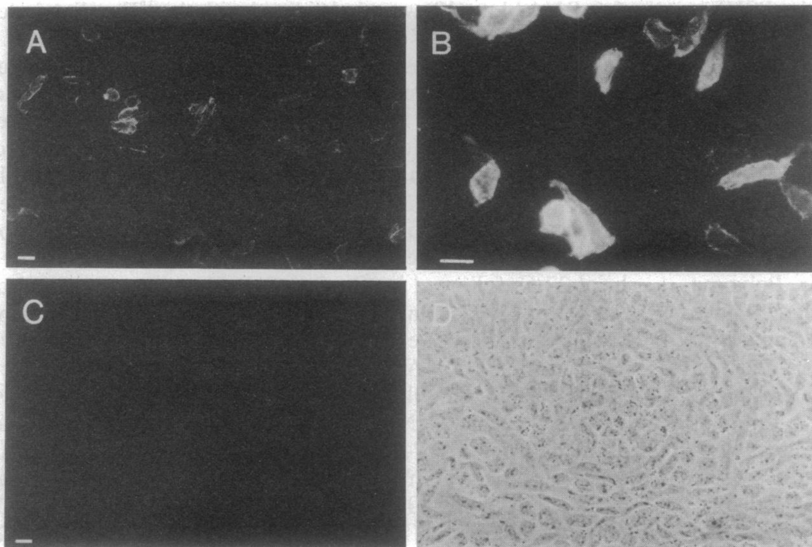

FIG. 2. Expression of T305 antigen by pcDNAI-C2GnT. Subconfluent CHO-Py·leu cells were mock-transfected with pcDNAI (*C* and *D*) or transfected with pcDNAI-C2GnT (*A* and *B*). Sixty-four hours after transfection, the cells were fixed and examined by incubation with mouse T305 monoclonal antibody followed by fluorescein isothiocyanate-conjugated sheep anti-mouse IgG (*A*–*C*) as described (28). (*D*) Phase-contrast micrograph of the same field shown in *C*. Two different areas are shown in *A* and *B*. (Bar = 20 μm.)

origin, whereas the expression of leukosialin would greatly increase the affinity of the monoclonal antibody binding to CHO cells expressing the hexasaccharides. After stable transfection of plasmid vectors encoding the polyoma early genes and leukosialin, cells were first selected for leukosialin expression with anti-leukosialin antibodies. Selected leukosialin-positive clones, obtained after limiting dilution, were then assayed for replication of a pcDNAI-based vector, the vector eventually used in the expression cloning. Four out of six clones tested were found to replicate plasmids, and one CHO cell line (CHO-Py·leu) was chosen for further studies, because it also expressed a significant amount of leukosialin.

Isolation of a cDNA Clone That Determines the Expression of the T305 Antigen in CHO Cells. Since HL-60 cells contain a significant amount of the C2GnT (2), a cDNA expression library from HL-60 cells was prepared and screened for cDNAs that expressed the T305 antigen. After several selection rounds, one plasmid named pcDNAI-C2GnT was isolated that directed the expression of the cell surface molecules recognized by T305. As shown in Fig. 2, CHO-Py·leu cells transfected with pcDNAI-C2GnT express the antigen recognized by T305, whereas the same cells transfected with pcDNAI itself showed no staining at all. These results indicate that pcDNAI-C2GnT directs the expression of a determinant on leukosialin recognized by monoclonal antibody T305, presumably NeuNAcα2-3Galβ1-3(NeuNAcα2-3Galβ1-4GlcNAcβ1-6)GalNAc (NeuNAc = *N*-acetylneuraminic acid).

The Amino Acid Sequence Deduced from cDNA Predicts a Protein with Type II Transmembrane Topology. The cDNA insert of 2105 base pairs in size contains a single open reading frame in the sense orientation with respect to the pcDNAI promoter sequences (Fig. 3). This reading frame predicts a protein of 428 amino acids in length, with a molecular mass of 49,790. Hydropathy analysis predicts that this protein is a type II transmembrane molecule as has been shown for all mammalian glycosyltransferases cloned to date (29). In this topology, the very short cytoplasmic NH_2-terminal segment of 9 amino acid residues is followed by a 23-amino acid transmembrane domain that is flanked by basic amino acid residues. The COOH-terminal sequence, presumably consisting of stem and catalytic domains, is large and most likely faces the lumen of the Golgi complex.

There are three potential N-glycosylation sites (see asterisks in Fig. 3), one of which may not be utilized because of

```
-219   GTGAAGTGCTCAGAATGGGGCAGGATGTCACCTGGAATCAGCACTAAGTGATTCAGACTTTCCTTACTTTTAAATGTGCTGCTCTTCATTTCAAGATGC   -121

CGTTGCAGCTCTGATAAATGCAAACTGACAACCTTCAAGGCCACGACGGAGGGAAAATCATTGGTGCTTGGAGCATAGAAGACTGCCCTTCACAAAGGAAATCCCTGATTATTGTTTGAA   -1

ATGCTGAGGACGTTGCTGCGAAGGAGACTTTTTTCTTATCCCACCAAATACTACTTTATGGTTCTTGTTTTATCCCTAATCACCTTCTCCGTTTTAAGGATTCATCAAAAGCCTGAATTT   120
 M  L  R  T  L  L  R  R  R  L  F  S  Y  P  T  K  Y  Y  F  M  V  L  V  L  S  L  I  T  F  S  V  L  R  I  H  Q  K  P  E  F    40
                            =====================================================================

GTAAGTGTCAGACACTTGGAGCTTGCTGGGGAGAATCCTAGTAGTGATATTAATTGCACCAAAGTTTTACAGGGTGATGTAAATGAAATCCAAAAGGTAAAGCTTGAGATCCTAACAGTG   240
 V  S  V  R  H  L  E  L  A  G  E  N  P  S  S  D  I  N  C  T  K  V  L  Q  G  D  V  N  E  I  Q  K  V  K  L  E  I  L  T  V    80
                                  *                 *

AAATTTAAAAAGCGCCCTCGGTGGACACCTGACGACTATATAAACATGACCAGTGACTGTTCTTCTTTCATCAAGAGACGCAAATATATTGTAGAACCCCTTAGTAAAGAAGAGGCGGAG   360
 K  F  K  K  R  P  R  W  T  P  D  D  Y  I  N  M  T  S  D  C  S  S  F  I  K  R  R  K  Y  I  V  E  P  L  S  K  E  E  A  E   120
                                           *

TTTCCAATAGCATATTCTATAGTGGTTCATCACAAGATTGAAATGCTTGACAGGCTGCTGAGGGCCATCTATATGCCTCAGAATTTCTATTGCGTTCATGTGGACACAAAATCCGAGGAT   480
 F  P  I  A  Y  S  I  V  V  H  H  K  I  E  M  L  D  R  L  L  R  A  I  Y  M  P  Q  N  F  Y  C  V  H  V  D  T  K  S  E  D   160

TCCTATTTAGCTGCAGTGATGGGCATCGCTTCCTGTTTTAGTAATGTCTTTGTGGCCAGCCGATTGGAGAGTGTGGTTTATGCATCGTGGAGCCGGGTTCAGGCTGACCTCAACTGCATG   600
 S  Y  L  A  A  V  M  G  I  A  S  C  F  S  N  V  F  V  A  S  R  L  E  S  V  V  Y  A  S  W  S  R  V  Q  A  D  L  N  C  M   200

AAGGATCTCTATGCAATGAGTGCAAACTGGAAGTACTTGATAAATCTTTGTGGTATGGATTTTCCCATTAAAACCAACCTAGAAATTGTCAGGAAGCTCAAGTTGTTAATGGGAGAAAAC   720
 K  D  L  Y  A  M  S  A  N  W  K  Y  L  I  N  L  C  G  M  D  F  P  I  K  T  N  L  E  I  V  R  K  L  K  L  L  M  G  E  N   240

AACCTGGAAACGGAGAGGATGCCATCCCATAAAGAAGAAAGGTGGAAGAAGCGGTATGAGGTCGTTAATGGAAAGCTGACAAACACAGGGACTGTCAAAATGCTTCCTCCACTCGAAACA   840
 N  L  E  T  E  R  M  P  S  H  K  E  E  R  W  K  K  R  Y  E  V  V  N  G  K  L  T  N  T  G  T  V  K  M  L  P  P  L  E  T   280

CCTCTCTTTTCTGGCAGTGCCTACTTCGTGGTCAGTAGGGAGTATGTGGGGTATGTACTACAGAATGAAAAAATCCAAAAGTTGATGGAGTGGGCACAAGACACATACAGCCCTGATGAG   960
 P  L  F  S  G  S  A  Y  F  V  V  S  R  E  Y  V  G  Y  V  L  Q  N  E  K  I  Q  K  L  M  E  W  A  Q  D  T  Y  S  P  D  E   320

TATCTCTGGGCCACCATCCAAAGGATTCCTGAAGTCCCGGGCTCACTCCCTGCCAGCCATAAGTATGATCTATCTGACATGCAAGCAGTTGCCAGGTTTGTCAAGTGGCAGTACTTTGAG   1080
 Y  L  W  A  T  I  Q  R  I  P  E  V  P  G  S  L  P  A  S  H  K  Y  D  L  S  D  M  Q  A  V  A  R  F  V  K  W  Q  Y  F  E   360

GGTGATGTTTCCAAGGGTGCTCCCTACCCGCCCTGCGATGGAGTCCATGTGCGCTCAGTGTGCATTTTCGGAGCTGGTGACTTGAACTGGATGCTGCGCAAACACCACTTGTTTGCCAAT   1200
 G  D  V  S  K  G  A  P  Y  P  P  C  D  G  V  H  V  R  S  V  C  I  F  G  A  G  D  L  N  W  M  L  R  K  H  H  L  F  A  N   400

AAGTTTGACGTGGATGTTGACCTCTTTGCCATCCAGTGTTTGGATGAGCATTTGAGACACAAAGCTTTGGAGACATTAAAACACTGACCATTACGGGCAATTTTATGAACAAGAAGAAGG   1320
 K  F  D  V  D  V  D  L  F  A  I  Q  C  L  D  E  H  L  R  H  K  A  L  E  T  L  K  H end                                  428

ATACACAAAACGTACCTTATCTGTTTCCCCTTCCTTGTCAGCGTCGGGAAGATGGTATGAAGTCCTCTTTGGGGCAGGGACTCTAGTAGATCTTCTTGTCAGAGAAGCTGCATGGTTTCT   1440

GCAGAGCACAGTTAGCTAGAAAGGTGATAGCATTAAATGTTCATCTAGAGTTAATAGTGGGAGGAGTAAAGGTAGCCTTGAGGCCAGAGCAGGTAGCAAGGCATTGTGGAAAGAGGGGAC   1560

CAGGGTGGCTGGGGAAGAGGCCGATGCATAAAGTCAGCCTGTTCCAAGTGCTCAGGGACTTAGCAAAATGAGAAGATGTGACCTGTGCCAAAACTATTTTGAGAATTTTAAATGTGACCA   1680

TTTTTCTGGTATG[AATAAA]CTTACAGCAACAAATAATCAAAGATACAATTAATCTGATATTATATTTGTTGAAATAGAAATTTGATTGTACTATAAATGATTTTTGTAAATAATTTATAT   1800

TCTGCTCTAATACTGTACTGTGTAGTGTGTCTCCGTATGTCATCTCAGGGAGCTTAAAATGGGCTTGATTTAACATTGAAAAAAAA   1886
```

FIG. 3. DNA and translated amino acid sequences of C2GnT. The open reading frame and full-length nucleotide sequences of clone C2GnT are shown. The signal/membrane-anchoring domain is doubly underlined. A consensus polyadenylylation signal is boxed. Potential N-glycosylation sites are marked with asterisks. The sequences are numbered relative to the translation initiation site.

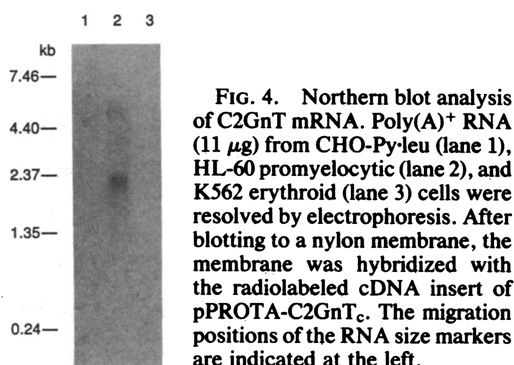

FIG. 4. Northern blot analysis of C2GnT mRNA. Poly(A)$^+$ RNA (11 μg) from CHO-Py·leu (lane 1), HL-60 promyelocytic (lane 2), and K562 erythroid (lane 3) cells were resolved by electrophoresis. After blotting to a nylon membrane, the membrane was hybridized with the radiolabeled cDNA insert of pPROTA-C2GnT$_c$. The migration positions of the RNA size markers are indicated at the left.

the presence of a proline residue next to asparagine. Although a consensus sequence for the polyadenylylation signal, AATAAA, is present at nucleotides 1694–1699, the cDNA actually ends at nucleotide 1878. It is possible that a more downstream sequence acts as the polyadenylylation signal (Fig. 3). No significant similarities were found between this sequence and other sequences in our protein data base. Sequence comparison with other glycosyltransferases, including GlcNAc transferase I (30, 31), did not reveal any similarity either.

Northern Blot Analysis. Fig. 4 demonstrates the Northern blot analysis of mRNAs isolated from the recipient CHO-Py·leu cells, HL-60 promyelocytic cells, and K562 erythroleukemic cells. The results show that the major band in the Northern blot analysis of HL-60 cells and the isolated cDNA insert are essentially identical in size, ≈2.1 kilobases (kb). In addition, two transcripts of ≈3.3 kb and 5.4 kb in size could be detected. They might be produced by differential usage of polyadenylylation signals. In contrast, no signal was obtained from poly(A)$^+$ RNA of K562 cells (Fig. 4, lane 3), which have been shown to lack the hexasaccharide and instead synthesize the tetrasaccharide (6). No signal was obtained from the recipient cells either (Fig. 4, lane 1).

Expression of Catalytically Active C2GnT. To confirm that our cDNA encodes C2GnT, CHO-Py·leu cells were transiently transfected with pcDNAI or pcDNAI-C2GnT. It was found that pcDNAI-C2GnT directs specifically the expression of the core 2 enzymatic activity (764 pmol per mg of protein per hr).

To unequivocally establish that this cDNA encodes C2GnT, sequences corresponding to the putative catalytic plus stem domains of this protein were fused in-frame with a

Table 1. Determination of enzymatic activities directed by pPROTA-C2GnT$_c$

Enzyme	Acceptors and linkages formed	Activity in conditioned medium, pmol/hr
C2GnT	*GlcNAcβ1*↘ 6 Galβ1→3GalNAc	718
C4GnT	*GlcNAcβ1*↘ 6 GlcNAcβ1→3GalNAc	<10
GnTV	*GlcNAcβ1*↘ 6 GlcNAcβ1→2Man	<10
I-GnT	*GlcNAcβ1*↘ 6 GlcNAcβ1→3Gal	<10

COS-1 cells were transfected with pPROTA-C2GnT$_c$ and the conditioned media were incubated with IgG-Sepharose. The proteins bound to the IgG-Sepharose were assayed for GnT activity by using appropriate acceptors. The linkages formed are indicated by italics. Similar results were obtained in three independent experiments. GnTV, UDP-GlcNAc:GlcNAcβ1-2Man GnT; I-GnT, UDP-GlcNAc:GlcNAcβ1-3Galβ1-4GlcNAc (GlcNAc to Gal) GnT.

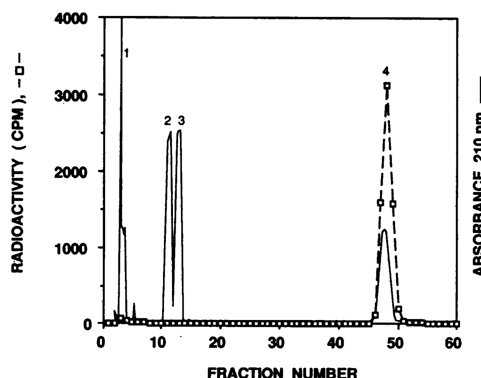

FIG. 5. HPLC characterization of the C2GnT product. The radiolabeled C2GnT product (≈10,000 cpm) was prepared (see *Materials and Methods*) using *p*-nitrophenyl-Galβ1-3GalNAc as an acceptor and the product of pPROTA-C2GnT$_c$, bound to IgG-Sepharose, as the source of the enzyme. The radiolabeled product was fractionated on a Lichrosorb NH_2 column. Numbers and the absorbance profile denote the migration positions of the *p*-nitrophenyl derivatives: 1, GalNAc; 2, Galβ1-3GalNAc; 3, GlcNAcβ1-3GalNAc; 4, Galβ1-3(GlcNAcβ1-6)GalNAc, which were obtained from Toronto Research Chemicals.

signal peptide and the IgG binding domain of *Staphylococcus aureus* protein A. A significant amount of C2GnT activity was detected in the medium of COS-1 cells transfected with pPROTA-C2GnT$_c$, and the activity generated by pPROTA-C2GnT$_c$ could be bound to IgG-Sepharose beads (Table 1). No activity was detected in the supernatant after incubation with IgG-Sepharose (data not shown).

Determination of the Specificity of C2GnT. So far, four different GlcNAcβ1-6 linkages have been reported—i.e., Galβ1-3(GlcNAcβ1-6)GalNAc, GlcNAcβ1-3(GlcNAcβ1-6)GalNAc, GlcNAcβ1-3(GlcNAcβ1-6)Gal, and GlcNAcβ1-2(GlcNAcβ1-6)Man. To determine whether these different structures are also synthesized by the enzyme cloned in the present study, the enzymatic activities were determined by using five different acceptors. The fused protein bound to IgG beads was found to display only an activity toward the acceptor for core 2 formation (Table 1). In addition, CHO cells were stably transfected with pcDNAI-C2GnT, and the same activities were tested in the cell lysates. Again, only an activity toward the acceptor for core 2 formation was substantially increased, and all the other GnT activities were the same as in control cells. To confirm that the coded protein synthesizes the expected product, the radiolabeled product was analyzed by HPLC. As shown in Fig. 5, the enzyme yielded the expected product Galβ1-3(GlcNAcβ1-6)GalNAc. These results indicate that the C2GnT present in myeloid cells is exclusively responsible for the formation of the GlcNAcβ1-6 branch on Galβ1-3 GalNAc.

DISCUSSION

In the present study, we report the isolation of a cDNA clone encoding C2GnT and the determination of its substrate specificity. For this cloning, we established a CHO cell line that stably expresses the polyoma large T antigen, which allows the replication of a plasmid vector harboring a polyoma replication origin in the transfected cells (12, 14). In addition, the CHO cell line was made to simultaneously express leukosialin in order to facilitate the detection of the oligosaccharide product by the T305 monoclonal antibody.

It has been generally accepted that each glycosyltransferase catalyzes only one enzymatic reaction to form a specific linkage, with one notable exception for the Lewis fucosyltransferase, which can synthesize both α1-3 and α1-4 linkages (11, 32). Such formation of a specific linkage is

usually associated with the formation of specific oligosaccharides in conjunction with other glycosyltransferases. Very often, lectins or antibodies are available for detecting those oligosaccharide structures. In addition to defects in glycosylation (33), CHO cell mutants have been selected for various defects in cellular metabolism, loss of cell surface molecules, and resistance to cytotoxic drugs (for example, see refs. 33 and 34). Thus the currently used approach should also be widely applicable to isolate cDNA clones coding for proteins having various cellular functions.

Although this C2GnT of myeloid cells has not been purified, a similar enzyme was purified from bovine tracheal epithelium (35). The apparent molecular mass of this purified enzyme is ≈69,000 Da. The amino acid sequence predicted from the cDNA sequence indicates that the polypeptide portion is ≈50 kDa for C2GnT. The deduced amino acid sequence also reveals that there are two to three N-glycosylation sites. Thus N-glycosylation and O-glycosylation with other possible posttranslational modifications could account for the larger apparent size of the enzyme, assuming that these two enzymes have a similar size in their polypeptide portion(s). However, it is also possible that the size of these two polypeptides differs significantly (see below).

Although more than half a dozen cDNA clones encoding glycosyltransferases have been isolated, so far there has been no report on the cloning of cDNA for an enzyme exclusively involved in O-glycan synthesis. Besides the cDNA isolated in the present study, cDNA has been obtained for only one other GlcNAc transferase (30, 31). Although these enzymes utilize the same nucleotide sugar, UDP-GlcNAc, no homology in primary amino acid sequence could be detected between these two enzymes. In O-glycans, GlcNAcβ1-6 linkages can be found in both core 2, Galβ1-3(GlcNAcβ1-6)GalNAc, and core 4, GlcNAcβ1-3(GlcNAcβ1-6)GalNAc, structures (36). In addition, GlcNAcβ1-6 linkages can be found in the side chains of poly(N-acetyllactosamine) forming the blood group I structure (25) and in the side chain attached to α-mannose of the N-glycan core structure, forming a tetraantennary saccharide (37). The enzymes responsible for all these linkages share the same unique property that Mn^{2+} is not required for their activity. It was thus suggested that these activities may be carried by the same enzyme (25). However, the present study clearly demonstrates that the C2GnT in myeloid cells is specific for the formation of O-glycan core 2. This result is consistent with the recent report that myeloid cell lysates contain the enzymatic activity for core 2 formation but not core 4 formation (3).

Recent studies suggest that there are at least two different GnTs in tracheal epithelium, which are differentially eluted from affinity adsorbent (35). One of them was, in fact, found to contain the activities to form core 2, core 4, and blood group I structures. These results suggest that there is at least one other GnT in epithelium, which can form core 2, core 4, and blood group I structures. Thus it is possible that there is a family of GnTs differing in acceptor specificity but yet forming the same linkage. A similar situation has been reported for the α1-3 fucosyltransferases (38). It will be of great interest to isolate cDNAs encoding C2GnTs from other tissues and to see if there are additional GnTs that are homologous to C2GnT.

We thank Dr. John B. Lowe for the kind gift of pPROTA and helpful discussions throughout the work, Dr. Naoko Arai for useful discussions, the members of our laboratory for useful advice and support, Dr. Harry Schachter for critical reading of the manuscript, Yaron Hakakha for technical assistance, and Henny Bierhuizen for secretarial assistance. This work was supported by Grants CA33000 and CA33895 awarded by the National Cancer Institute. M.F.A.B. was initially supported by a postdoctoral fellowship of the Netherlands Organization for Scientific Research (NWO).

1. Piller, F., Piller, V., Fox, R. I. & Fukuda, M. (1988) *J. Biol. Chem.* **263**, 15146–15150.
2. Saitoh, O., Piller, F., Fox, R. I. & Fukuda, M. (1991) *Blood* **77**, 1491–1499.
3. Brockhausen, I., Kuhns, W., Schachter, H., Matta, K. L., Sutherland, D. R. & Baker, M. A. (1991) *Cancer Res.* **51**, 1257–1263.
4. Piller, F., Le Deist, F., Weinberg, K. I., Parkman, R. & Fukuda, M. (1991) *J. Exp. Med.* **173**, 1501–1510.
5. Carlsson, S. R. & Fukuda, M. (1986) *J. Biol. Chem.* **261**, 12779–12786.
6. Carlsson, S. R., Sasaki, H. & Fukuda, M. (1986) *J. Biol. Chem.* **261**, 12787–12795.
7. Sportsman, J. R., Park, M. M., Cheresh, D. A., Fukuda, M., Elder, J. H. & Fox, R. I. (1985) *J. Immunol.* **135**, 158–164.
8. Fukuda, M., Carlsson, S. R., Klock, J. C. & Dell, A. (1986) *J. Biol. Chem.* **261**, 12796–12806.
9. Yousefi, S., Higgins, E., Daoling, Z., Pollex-Krüger, A., Hindsgaul, O. & Dennis, J. W. (1991) *J. Biol. Chem.* **266**, 1772–1782.
10. Seed, B. & Aruffo, A. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 3365–3369.
11. Kukowska-Latallo, J. F., Larsen, R. D., Nair, R. P. & Lowe, J. B. (1990) *Genes Dev.* **4**, 1288–1303.
12. Heffernan, M. & Dennis, J. W. (1991) *Nucleic Acids Res.* **19**, 85–92.
13. Felgner, P. L., Gadek, T. R., Holm, M., Roman, R., Chan, H. W., Wenz, M., Northrop, J. P., Ringold, G. M. & Danielsen, M. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 7413–7417.
14. Muller, W. J., Naujokas, M. A. & Hassell, J. A. (1984) *Mol. Cell. Biol.* **4**, 2406–2412.
15. Pallant, A., Eskenazi, A., Mattei, M.-G., Fournier, R. E. K., Carlsson, S. R., Fukuda, M. & Frelinger, J. G. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 1328–1332.
16. Cepko, C. L., Roberts, B. E. & Mulligan, R. C. (1984) *Cell* **37**, 1053–1062.
17. Graham, F. L. & Van der Eb, A. J. (1973) *Virology* **52**, 456–467.
18. Hirt, B. (1967) *J. Mol. Biol.* **26**, 365–369.
19. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), 2nd Ed.
20. Sanchez-Lopez, R., Nicholson, R., Gesnel, M.-C., Matrisian, L. M. & Breathnach, R. (1988) *J. Biol. Chem.* **263**, 11892–11899.
21. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
22. Feinberg, A. P. & Vogelstein, B. (1983) *Anal. Biochem.* **132**, 6–13.
23. Lee, N., Wang, W.-C. & Fukuda, M. (1990) *J. Biol. Chem.* **265**, 20476–20487.
24. Palcic, M. M., Ripka, J., Kaur, K. J., Shoreibah, M., Hindsgaul, O. & Pierce, M. (1990) *J. Biol. Chem.* **265**, 6759–6769.
25. Piller, F., Cartron, J.-P., Maranduba, A., Veyrieres, A., Leroy, Y. & Fournet, B. (1984) *J. Biol. Chem.* **259**, 13385–13390.
26. Schachter, H., Brockhausen, I. & Hull, R. (1989) *Methods Enzymol.* **179**, 351–397.
27. Sasaki, H., Bothner, B., Dell, A. & Fukuda, M. (1987) *J. Biol. Chem.* **262**, 12059–12076.
28. Williams, M. A. & Fukuda, M. (1990) *J. Cell Biol.* **111**, 955–966.
29. Paulson, J. C. & Colley, K. J. (1989) *J. Biol. Chem.* **264**, 17615–17618.
30. Kumar, R., Yang, J., Larsen, R. D. & Stanley, P. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 9948–9952.
31. Sarkar, M., Hull, E., Nishikawa, Y., Simpson, R. J., Moritz, R. L., Dunn, R. & Schachter, H. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 234–238.
32. Prieels, J.-P., Monnom, D., Dolmans, M., Beyer, T. A. & Hill, R. L. (1981) *J. Biol. Chem.* **256**, 10456–10463.
33. Stanley, P. (1984) *Annu. Rev. Genet.* **18**, 525–552.
34. Yayon, A., Klagsbrun, M., Esko, J. D., Leder, P. & Ornitz, D. M. (1991) *Cell* **64**, 841–848.
35. Ropp, P. A., Little, M. R. & Cheng, P.-W. (1991) *J. Biol. Chem.* **266**, 23863–23871.
36. Brockhausen, I., Matta, K. L., Orr, J. & Schachter, H. (1985) *Biochemistry* **24**, 1866–1874.
37. Cummings, R. D., Trowbridge, I. S. & Kornfeld, S. (1982) *J. Biol. Chem.* **257**, 13421–13427.
38. Weston, B. W., Nair, R. P., Larsen, R. D. & Lowe, J. B. (1992) *J. Biol. Chem.* **267**, 4152–4160.