Molecular cloning of a human cDNA encoding β-1,4-galactosyltransferase with 37% identity to mammalian UDP-Gal:GlcNAc β-1,4-galactosyltransferase

(3′-rapid amplification of cDNA ends/PCR/glycosyltransferase/tissue distribution)

TAKESHI SATO*, KIYOSHI FURUKAWA*†‡, HANS BAKKER§, DIRK H. VAN DEN EIJNDEN†§, AND IRMA VAN DIE†§

*Department of Biosignal Research, Tokyo Metropolitan Institute of Gerontology, Itabashi-ku, Tokyo 173, Japan; and §Department of Medical Chemistry, Vrije Universiteit, van der Boechorststraat 7, 1081 BT Amsterdam, The Netherlands

ABSTRACT A cDNA encoding a β-1,4-galactosyltransferase named β-1,4-GalT II was cloned from a cDNA library of the human breast tumor cell line, MRK-nu-1. Initially, a 860-bp PCR fragment was obtained from MRK-nu-1 mRNA by 3′-rapid amplification of cDNA ends by using two nested degenerate oligonucleotide primers based on a highly conserved amino acid sequence found in the catalytic domain of mammalian β-1,4-galactosyltransferases and *Lymnaea stagnalis* β-1,4-*N*-acetylglucosaminyltransferase (β-1,4-Glc-NAcT), both of which utilize the same sugar acceptor. This subsequently was used as a probe to isolate a 4.7-kb cDNA that contained an ORF of 1,164 bp predicting a polypeptide of 388 aa. Its deduced amino acid sequence shows an identity of 37% with that of the previously characterized human β-1,4-galactosyltransferase (referred to as β-1,4-GalT I) and of 28% with that of *L. stagnalis* β-1,4-GlcNAcT. Study of the properties of the β-1,4-GalT II fused to protein A expressed as a soluble form in COS-7 cells revealed that it is a genuine β-1,4-GalT but has no lactose synthetase activity in the presence of α-lactalbumin. Northern blot analysis of 24 human tissues showed that they all express the β-1,4-GalT II transcript, although the levels varied. These results indicate that human cells contain another β-1,4-GalT.

Cell surface carbohydrates have been demonstrated to play important roles in many biological events. In particular, complex-type sugar chains on mammalian glycoconjugates containing Galβ1 → 4GlcNAc outer branches carrying polysialic acid and HNK-1 epitopes or LewisX and sialyl LewisX determinants seem to be involved in adhesion processes such as cell-to-cell and cell-to-matrix interactions (reviewed in refs. 1–4). The two enzymes controlling the biosynthesis of complex-type oligosaccharides, UDP-GlcNAc:Man β-1,2-*N*-acetylglucosaminyltransferase (GlcNAcT-I) and UDP-Gal:GlcNAc β-1,4-galactosyltransferase (β-1,4-GalT) [EC 2.4.1.38], are among the best-studied glycosyltransferases (5–9). Targeted inactivation of the mouse GlcNAcT-I gene has been shown to result in death of the mice at mid-embryonic stages, whereas ablation of the β-1,4-GalT gene resulted in growth retardation and early death of the mice, indicating that the absence of complex-type N-linked oligosaccharides carrying the carbohydrate antigens described above and/or terminal *N*-acetylglucosamine or galactose residues themselves have important functions in embryonic development and further growth of the individual animals (10–13). The β-1,4-GalT-deficient mice, however, were able to synthesize a reduced amount of complex-type oligosaccharides containing galactose and also show low residual β-1,4-GalT activity. It is possible that such mice survive the gestation because of a partial compensation of the defective glycosylation by an additional, so far unidentified, β-1,4-GalT. Most β-1,4-GalTs characterized so far show lactose synthetase activity in the presence of α-lactalbumin (14, 15). No lactose was found in the milk of the mutant mice (13), indicating that the predicted additional β-1,4-GalT may not have this property. Interestingly, the occurrence of β-1,4-GalTs that appear to be unresponsive to α-lactalbumin has been described in porcine trachea and in *Schistosoma mansoni* (16, 17).

Mammalian β-1,4-GalTs are structurally related to a *Lymnaea stagnalis* UDP-GlcNAc:GlcNAc β-1,4-*N*-acetylglucosaminyltransferase (β-1,4-GlcNAcT) (18), an enzyme that shows an acceptor–substrate requirement resembling that of β-1,4-GalT (19). By comparing the amino acid sequences deduced from mammalian β-1,4-GalT cDNAs (8, 20, 21), two putative chicken GalT cDNAs (22), *L. stagnalis* β-1,4-GlcNAcT cDNA (18), and two related *Caenorhabditis elegans* genes (23), several conserved amino acid sequences have been discovered in the putative catalytic domain (24). To explore whether additional β-1,4-GalTs occur in mammalian cells, a 3′-rapid amplification of cDNA ends (RACE) approach was used with degenerate PCR primers based on one of the conserved amino acid boxes (Y/F)(W/F)GWG(G/A)EDD(D/E). In the present study we describe the isolation of a human cDNA encoding a β-1,4-GalT, the properties and tissue distribution of which differ from those of the well known mammalian β-1,4-GalT.

MATERIALS AND METHODS

Materials, Cells, and Oligonucleotides. UDP-Gal (48.8 Ci/mmol; 1 Ci = 37 GBq), UDP-GalNAc (24 Ci/mmol), and UDP-GlcNAc (18.9 Ci/mmol) were purchased from DuPont/NEN and diluted with unlabeled UDP-sugars (Sigma) to give the desired specific activity. GlcNAcβ1 → 2Manα1 → 3Manβ1 → O(CH₂)₈COOCH₂-*m*-nitrophenol (*m*NP), Glc-NAcβ1 → 4(GlcNAcβ1 → 2)Manα1 → 3Manβ1 → O(CH₂)₈-COOCH₂-*m*NP, and GlcNAcβ1 → 6(GlcNAcβ1 → 2)Manα1 → 3Manβ1 → O(CH₂)₈COOCH₂-*m*NP, which were chemically synthesized and donated by T. Kitajima, Towa Kasei Co. (Japan), and *p*-nitrophenyl-*N*-acetyl-1-thio-β-D-glucosaminide (GlcNAcβ-S-*p*NP) was purchased from Sigma. Asialo (As), agalacto (Ag)-human transferrin was prepared as

described previously (25). The human breast tumor cell line, MRK-nu-1, was established by M. Sekiguchi at the Institute of Medical Science, University of Tokyo (Japan). The oligonucleotides, TS11: 5′-GTCCCGGAATTCCGGGCAAGGCAT-TCTGATCCGGGACAACGTGAGA-3′, TS9: 5′-CCGGAA-TTCCGGCTCCTCTCAGTACTCGTTCACCTGAGCCAG-CTC-3′, HB25: 5′-TAIT(G/T)(C/G/T)GGITGGGG-3′, and ID10: 5′-GGITGGGGIIIIGA(A/G)GA(C/T)GA(C/T)GA--3′, were synthesized by Nippon Flour Mills (Kanagawa, Japan).

Isolation of a β-1,4-GalT-Related PCR Fragment by 3′-RACE. Poly(A)$^+$ RNA was prepared from MRK-nu-1 cells with a Fast Track mRNA Isolation kit (Invitrogen) and used for first-strand cDNA synthesis. Subsequently, this cDNA was used as a template with two nested primers, HB25 and ID10, in an adapted 3′-RACE protocol (GIBCO/BRL). A first-round of PCR [95°C, 3 min (94°C, 1 min; 50°C, 1 min; and 72°C, 1 min) × 25, and 72°C, 15 min] was carried out by using a tailed oligo(dT) primer already present in the cDNA sample mixture (25 nM) and primer HB25 (400 nM). In the second round of PCR, an aliquot of this reaction was subjected to PCR [95°C, 3 min (94°C, 1 min; 63°C, 1 min; and 72°C, 1 min) × 30, and 72°C, 15 min] by using primer ID10 (400 nM) and universal amplification primer (80 nM). Subsequently, the PCR fragments were separated by agarose gel electrophoresis, isolated from the gel, and purified with a Prep-A-Gene DNA purification kit (Bio-Rad), cloned into pBluescript SK (Stratagene), and sequenced.

Isolation of a Human cDNA Clone. A cDNA library was constructed from the poly(A)$^+$ RNA by using the λ-ZAP Express Vector Cloning kit (Stratagene). Approximately 3 × 10^5 independent clones were screened by using the 860-bp PCR fragment labeled with [α-^{32}P]dCTP and by using a Ready to Go DNA labeling kit (Pharmacia) as a probe and hybridization conditions (60°C) as described previously (18). The nucleotide sequence of the cDNA was determined by the deoxynucleotide chain termination method (26) with an Auto Read Sequencing kit (Pharmacia).

Northern Blot Analysis of Human Tissues. Human multiple tissue Northern blots, obtained from CLONTECH, were hybridized with part (nucleotides 145-1164) of the [α-^{32}P]dCTP-labeled β-1,4-GalT II cDNA. Blots were subsequently reprobed with part (nucleotides 594-1200) of human β-1,4-GalT I cDNA, which was kindly supplied by M. N. Fukuda (The Burnham Institute, La Jolla, CA) and β-actin cDNA under the conditions as recommended by the manufacturer.

Construction of a Protein A/β-1,4-GalT II cDNA Fusion Plasmid. A plasmid containing a fusion between part of the protein A sequence (27) and the β-1,4-GalT II cDNA was constructed as follows. A cDNA fragment in which the putative N-terminal cytoplasmic and transmembrane sequences had been removed was obtained by PCR by using the isolated cDNA as a template and the oligonucleotide primers TS11 and TS9. The resultant PCR fragment was sequenced, digested with *Eco*RI, and cloned into the *Eco*RI site of pPROTA (28), which resulted in the hybrid construct.

Expression of the Fusion Plasmid in COS-7 Cells. Plasmids pPROTA and pPROTA/β-1,4-GalT II were transiently transfected into COS-7 cells by electroporation (29). COS-7 cells were cultured in DMEM containing 10% FCS for 24 h and then in DMEM containing transferrin (5 μg/ml), insulin (5 μg/ml), and sodium selenite (5 ng/ml). The culture media were harvested after 70 h and incubated with IgG-Sepharose beads (100 μl slurry for 4 ml media) at 4°C for 16 h. The enzyme-containing beads were subsequently collected by centrifugation, washed 10 times with PBS, and suspended in 100 μl of 0.1 M Mes buffer (pH 7.0) containing 4 mM 5′-AMP.

The full coding sequence (nucleotides 1–1164) of the β-1,4-GalT II cDNA was ligated into expression vector pVL1393 (PharMingen), and plasmid pVL1393/β-1,4-GalT II was co-transfected with BaculoGold DNA (PharMingen) into Sf-9 cells grown at 27°C in Grace's medium containing 10% FCS. The cells were sonicated in 0.1 M Mes buffer (pH 7.0) containing 4 mM 5′-AMP and 1.25% Triton X-100, and the cell homogenates were used for assay of lactose synthetase activity.

Glycosyltransferase Assay and Product Characterization. Glycosyltransferase assays were performed by incubating 0.1 M Mes buffer (pH 7.0) containing 4 mM 5′-AMP, 250 μM UDP-[^{3}H]sugar (6.64 Ci/mol), 1 mM GlcNAcβ-S-*p*NP or acceptor as indicated, 20 mM MnCl$_2$, and beads-associated enzyme in a total volume of 50 μl. After incubation at 37°C, the product was isolated by using Sep-Pak C$_{18}$ cartridges (Waters) and radioactivity incorporated into the product was determined (30). The ^{3}H-labeled products obtained from the acceptors were subjected to digestion with diplococcal β-galactosidase, which cleaves the Galβ1 → 4GlcNAc linkage but not the Galβ1 → 3GlcNAc linkage (31). The galactosylated GlcNAcβ1 → 2Manα1 → 3Manβ1 → O(CH$_2$)$_8$COOCH$_3$ was subjected to RCA-I-agarose column chromatography (32), and the bound fraction was characterized by methylation analysis (33). Lactose synthesis in the presence of α-lactalbumin and lacto-*N*-neotetraosylceramide synthesis by protein A/β-1,4-GalT II fusion protein were assayed as described previously (14, 15, 34–36). Bovine β-1,4-GalT I was prepared from bovine mammary gland as described previously (37) and incubated with the acceptors listed above.

RESULTS

Isolation of a Human cDNA Related to β-1,4-GalT. By 3′-RACE of human breast tumor cell line mRNA several PCR fragments were obtained as was revealed by agarose gel electrophoresis. One of the fragments (1,000 bp) most likely was derived from the known human β-1,4-GalT transcript, because restriction fragment analysis revealed a pattern expected for the 3′ part of this cDNA (20). A smaller 860-bp fragment was cloned and sequenced. This PCR fragment showed 39% identity at the amino acid level with human β-1,4-GalT. To isolate a full-length human cDNA clone, a cDNA library was constructed from MRK-nu-1 cells. In a plaque hybridization assay this cDNA library was screened with the α-^{32}P-labeled PCR fragment as a probe. This resulted in the isolation of a cDNA of approximately 4.7 kb. Determination of the nucleotide sequence revealed that the cDNA consists of a 112-bp G/C-rich 5′ noncoding region, followed by an ORF of 1,164 bp and a 3.4-kb long 3′ noncoding region (Fig. 1). The deduced amino acid sequence of the ORF predicts a polypeptide of 388 aa with eight potential N-glycosylation sites (Fig. 1). A hydropathy plot of the amino acid sequence (38) revealed one hydrophobic segment located at amino acid positions 28–48, corresponding to a potential transmembrane domain, suggesting a type-II membrane topology of the polypeptide as typically found in glycosyltransferases (39). Comparison of the deduced amino acid sequence with those of other glycosyltransferases revealed that it has 37% identity with human β-1,4-GalT (ref. 20, Fig. 2) and other mammalian β-1,4-GalTs, and 28% identity with *L. stagnalis* β-1,4-GlcNAcT (18). Four cysteine residues occur at highly conserved positions in both the protein and human β-1,4-GalT (indicated by solid circles in Fig. 2). In β-1,4-GalT, two of these four cysteine residues (amino acids 129 and 246, respectively) have been reported to form a disulfide bond that is considered to be a prerequisite for the enzyme to interact with α-lactalbumin (40). These results indicate that the cDNA encodes a protein named β-1,4-GalT II related to the mammalian β-1,4-GalT described previously.

Expression of a Protein A/β-1,4-GalT II Fusion Plasmid and Characterization of the Secreted Protein. To obtain evidence that the isolated cDNA encodes a glycosyltrans-

```
-112        GGGAGGCGGTGGCCGAGGCCCAGGCGGTGGCGGCGGCGGCCCAGGAGGCGGCGGACGGGGAGCTGCGGGAGCAGGCCCGGGCCTGGCTCTCTAGCGGCCGCCTGGCTGCAGC

1     ATGCGCGCCCGCCGGGGCGTGCTGCGGCTGCCGCGCCGCTCGCTGCTCGCCGCGCTCTTCTTCTTTTCTCTCTCGTCCTCGCTGCTGTACTTCGTCTATGTGGCGCCCGGCATAGTGAAC
1      M  R  A  R  R  G  L  L  R  L  P  R  R  S  L  L  A  A  L  F  F  F  S  L  S  S  S  L  L  Y  F  V  Y  V  A  P  G  I  V  N

121   ACCTACCTCTTCATGATGCAAGCCCAAGGCATTCTGATCCGGGACAACGTGAGGAACAATCGGTGCTCAGGTTTATGAGCAGGTGCTTCGGAGTGCTTATGCCAAGAGGAACAGCAGTGTA
41     T  Y  L  F  M  M  Q  A  Q  G  I  L  I  R  D  N  V  R  T  I  G  A  Q  V  Y  E  Q  V  L  R  S  A  Y  A  K  R  N  S  S  V
                                                                                                                  *

241   AATGACTCAGATTATCCTCTTGACTTGAACCACAGTGAAACCTTCCTGCAAACTACAACATTTCTTCCTGAAGACTTCACCTACTTTGCAAACCATACCTGCCCTGAAAGACTCCCTTCC
81     N  D  S  D  Y  P  L  D  L  N  H  S  E  T  F  L  Q  T  T  T  F  L  P  E  D  F  T  Y  A  N  H  T  C  P  E  R  L  P  S
       *                          *                                                        *

361   ATGAAGGGCCCAATAGACATAAACATGAGTGAAATTGGAATGGATTACATTCATGAACTCTTCTCCAAAGACCCAACCATCAAGCTCGGAGGTCACTGGAAGCCTTCTGATTGCATGCCT
121    M  K  G  P  I  D  I  N  M  S  E  I  G  M  D  Y  I  H  E  L  F  S  K  D  P  T  I  K  L  G  G  H  W  K  P  S  D  C  M  P
                        *

481   CGGTGGAAGGTGGCGATCCTTATCCCCTTCCGGAACCGCCACGAGCACCTCCCAGTCCTGTTCAGACACCTGCTTCCCATGCTCCAGCGCCAGCGCCTGCAGTTTGCATTTTATGTGGTT
161    R  W  K  V  A  I  L  I  P  F  R  N  R  H  E  H  L  P  V  L  F  R  H  L  L  P  M  L  Q  R  Q  R  L  Q  F  A  F  Y  V  V

600   GAACAAGTTGGTACCCAACCCTTTAATCGAGCCATGCTTTTCAACGTTGGCTTTCAAGAAGCAATGAAAGACTTGGATTGGGACTGTTTGATTTTTCATGATGTAGATCACATACCGGAA
201    E  Q  V  G  T  Q  P  F  N  R  A  M  L  F  N  V  G  F  Q  E  A  M  K  D  L  D  W  D  C  L  I  F  H  D  V  D  H  I  P  E

721   AGTGATCGCAACTATTATGGATGTGGACAGATGCCGAGGCATTTGCAACCAAATTGGATAAGTATATGTATCTGCTTCCTTATACCGAGTTCTTTGGCGGAGTGAGTGGCTTAACAGTG
241    S  D  R  N  Y  Y  G  C  G  Q  M  P  R  H  F  A  T  K  L  D  K  Y  M  Y  L  L  P  Y  T  E  F  F  G  G  V  S  G  L  T  V

841   GAACAATTTCGGAAAATCAATGGCTTTCCTAATGCTTTCTGGGGTTGGGGTGGAGAAGATGACGACCTCTGGAACAGAGTACAGAATGCAGGCTATTCTGTGAGCCGGCCAGAGGGTGAC
281    E  Q  F  R  K  I  N  G  F  P  N  A  F  W  G  W  G  G  E  D  D  D  L  W  N  R  V  Q  N  A  G  Y  S  V  S  R  P  E  G  D

961   ACAGGAAAGTACAAGTCCATTCCTCATCACCATCGAGGAGAAGTCCAGTTTCTTGGAAGGTATGCTCTGCTGAGGAAGTCAAAAGAACGGCAAGGGCTGGATGGCCTCAACAACCTGAAC
321    T  G  K  Y  K  S  I  P  H  H  H  R  G  E  V  Q  F  L  G  R  Y  A  L  L  R  K  S  K  E  R  Q  G  L  D  G  L  N  N  L  N

1081  TACTTTGCAAACATCACATACGACGCCCTTGTATAAAAAACATAACTGTCAACCTGACACCCGAGCTGGCTCAGGTGAACGAGTACTGAGAGGAGAGAATGTACGTTTGCTTTACCCACCGC
361    Y  F  A  N  I  T  Y  D  A  L  Y  K  N  I  T  V  N  L  T  P  E  L  A  Q  V  N  E  Y END
       *              *                      *              *

1201  CACCAAGAAAGCAGTCCGATGAGATTTTTTTTTGGAGGGGGGAGGGTCTACACAGCAAGAGAACAGAAATACTGTGTCTCATGAAGGATCACAGAGTTCAGGGGGAAAATGTGACAGCAC

1321  ACGCACAAACGCCTTCACTGGATCAGCCGCTGGAACTGAGGGAGTGAGCTTGGGGACTTCCTTCGTCAGCACTGGCTTTCTGTTTTCACAAGACAGACGTCTGTCCCGCTGCTCTCTCCC

1441  CATCTCCTACCCCACATCCTGTCTTAGCCGCAGTCTCCAGAACCCATGATGAACTGTGATCTGCCGTGGTCCTGCCGTGGTCCTGCCGTGGAGCCTGTCCCTACACATGACCTTGGAGCC

1561  TCTTGGCCTTCAGAGCAGAGGCAAACCCCACCACAGGGCAGCTGCGTTTTAGGAAGAGCAAATGAACTCCACACCATTCTTCTAGATCTCTGGTGTTCTCTTTGGTTTCATTTTTTTTAA

1681  AAATTACCTTCTTTGGGTGGGGATTGAGGGTGGAGGGGAGGGTGTTTGGGAAAGATAAATAGACATAAATATATAACAATCACTTCTTGAAGAAGTATAATTGTAAATAAGCCATGTAAA
```

FIG. 1. Nucleotide sequence of the 5′-terminal part of β-1,4-GalT II cDNA and its deduced amino acid sequence. The putative transmembrane domain is underlined. The potential N-glycosylation sites are indicated by asterisks. This nucleotide sequence has been submitted to the DDBJ/GenBank/EMBL Data Bank with accession no. AB004550.

ferase, a hybrid construct consisting of part of protein A (28) fused to the stem region of β-1,4-GalT II was transiently transfected to COS-7 cells. This was expected to result in the secretion of an enzymatically active soluble fusion protein (PROTA/β-1,4-GalT II), in the culture medium, that is similar to what has been observed previously with other glycosyltransferases (41–43). The soluble fusion protein was isolated from the medium by using IgG-Sepharose beads, which bind to the protein A part of the fusion protein. The activity of bead-associated fusion protein was assayed by using different sugar-donor and sugar-acceptor substrates. In standard enzyme assays, when using different UDP-sugar donors and GlcNAcβ-S-pNP as an acceptor, PROTA/β-1,4-GalT II showed a GalT

activity, whereas no enzyme activity could be detected when using UDP-GalNAc or UDP-GlcNAc as a sugar donor (Table 1). This GalT activity was stimulated in the presence of 20 mM MnCl₂, and optimal activity was obtained at pH 7.0 (data not shown). When an IgG-bead suspension that was incubated with pPROTA-transfected COS-7 cell medium was assayed, no detectable activity to the acceptor, regardless of the type of sugar donor used, was found with these beads (Table 1).

To characterize the product obtained with PROTA/β-1,4-GalT II, [³H]galactosylated GlcNAcβ-S-pNP was digested with diplococcal β-galactosidase, which specifically cleaves the Galβ1 → 4GlcNAc linkage but not the Galβ1 → 3GlcNAc linkage (31). After Sep-Pak C₁₈ column chromatography, all

```
β-1,4-GalT I    MRLREPLLSGAAMPGASLQRACRLLVAVCALHLGVTLVYYLAGRDLSRLPQLVGVSTPLQ    60
β-1,4-GalT II   MRARRGLLRLPRRSLLAALFFFSLSSSLLYFVYVAPGIVNTYLFMMQAQGILIRDNVRTI   60

β-1,4-GalT I    GGSNSAAAIGQSSGELRTGGARPPPPLGASSQPRPGGDSSPVVLSGPGPASNLTSVPVPH   120
β-1,4-GalT II   G----AQVYEQV---LRSAYAK-----RNSSVNDSDYPLDLNH-SETFLQT-TIFLPEDF   106

β-1,4-GalT I    TT-ALSLPACPEESPLLVGPMLIEFNMPV-DLELVAKQ-NPNVKMGGRYAPRDCVSPHKV   177
β-1,4-GalT II   TYFANHT--CPERLPSMKGPIDINMSEIGMDYIHELFSKDPTIKLGGHWKPSDCMPRWKV   164

β-1,4-GalT I    AIIIPFRNRQEHLKYWLY-YLHPVLQRQQLDYGIYGIYVINQAGDTIFNRAKLLNVGFQE   237
β-1,4-GalT II   AILLIPFRNRHEHLPV-LFRHLLPMLQRQRLQFA-F--YVVEQVGTQPFNRAMLFNVGFQE  221

β-1,4-GalT I    ALKDYDYTCFVFSDVDLIPMNDHNAYRCFSQPRHISVAMDKFGFSLPYVQYFGGVSALSK   297
β-1,4-GalT II   AMKDLDWDCLIFHDVDHIPESDRNYYGCGQMPRHFATKLDKYMYILPYTEFFGGVSGLTV  281

β-1,4-GalT I    QQFLTINGFPNNYWGWGGEDDDIFNRLVFRGMSISRPNAVVGRCRMI--RHSRDKKNEPN   355
β-1,4-GalT II   EQFRKINGFPNAFWGWGGEDDDLWNRVQNAGYSVSRPEGDTGKYKSIPHHH-RGEVQF--  338

β-1,4-GalT I    PQRFDRIAHTKETMLSDGLNSLTYQVLDVQRY-PLYTQITVDIGTHS               400
β-1,4-GalT II   LGRYALLRKSKERQGLDGLNNLNY-FANI-TYDALYKNITVNL-TEELAQVNEY        388
```

FIG. 2. Comparison of the deduced amino acid sequences of human β-1,4-GalT I and β-1,4-GalT II. Gaps introduced to optimize the alignment are indicated by bars. Identical amino acid residues in the sequence alignment are boxed. The putative transmembrane domains are underlined. Cysteine residues are indicated with solid circles.

Table 1. Glycosyltransferase activity in the PROTA/β-1,4-GalT II fusion protein produced by COS-7 cells

Transfected plasmid	Glycosyltransferase activity, pmol/min·mg protein		
	GalT	GalNAcT	GlcNAcT
pPROTA/β-1,4-GalT II	302.3 ± 16.9	0.5	0.2
pPROTA (mock)	0.1	0.1	0.1

The values represent the averages with standard errors obtained by three independent assays. GalT, GalNAcT, and GlcNAcT indicate galactosyltransferase, *N*-acetylgalactosaminyltransferase, and *N*-acetylglucosaminyltransferase, respectively.

the radioactivity was recovered in the pass-through fraction of the column (data not shown), indicating that the galactose was cleaved off efficiently. Galactosylated GlcNAcβ1 → 2Manα1 → 3Manβ1 → O(CH$_2$)$_8$COOCH$_2$-*m*NP was subjected to RCA-I-agarose column chromatography. RCA-I binds with high affinity to oligosaccharides terminating with β-1,4-linked galactose and binds poorly to those terminating with β-1,3-linked galactose (32, 44). The bound fraction subsequently was subjected to methylation analysis, which showed the presence of 2,3,4,6-tetra-O-methyl galactitol and 3,6-di-O-methyl 2-*N*-methylacetamido-2-deoxyglucitol at a similar ratio in addition to 3,4,6-tri-O-methyl and 2,4,6-tri-O-methyl mannitols (data not shown). These results indicate that the PROTA/β-1,4-GalT II fusion protein catalyzes transfer of galactose from UDP-Gal to the acceptor in β-1,4-linkage.

GlcNAcβ1 → 6(GlcNAcβ1 → 2)Manα1 → 3Manβ1 → O(CH$_2$)$_8$COOCH$_2$-*m*NP containing terminal β1 → 6- and/or β1 → 2-linked *N*-acetylglucosamine could be galactosylated as effectively as GlcNAcβ-S-*p*NP. However, the galactosylation of GlcNAcβ1 → 2Manα1 → 3Manβ1 → O(CH$_2$)$_8$COOCH$_2$-*m*NP and GlcNAcβ1 → 4(GlcNAcβ1 → 2)Manα1 → 3Manβ1 → O(CH$_2$)$_8$COOCH$_2$-*m*NP was less efficient (Table 2). In contrast, little enzyme activity was found by using Galβ-*p*NP, *N*-acetylglucosamine, AsAg-human transferrin, or lacto-*N*-triaosylceramide (data not shown). These results indicate that PROTA/β-1,4-GalT II is able to galactosylate some specific acceptors that have a terminal β-linked *N*-acetylglucosamine. Because β-1,4-GalT I is also involved in lactose synthesis (15), it was investigated whether α-lactalbumin was able to induce PROTA/β-1,4-GalT II to act on glucose. However, no lactose synthetase activity was observed in the presence of α-lactalbumin (0.5–5 mg/ml).

The reason why PROTA/β-1,4-GalT II has no lactose synthetase activity could be because of the presence of part of the protein A sequence. To investigate this, the recombinant β-1,4-GalT II with the full coding sequence was expressed by transfection of pVL1393/β-1,4-GalT II in Sf-9 cells that showed little β-1,4-GalT activity toward GlcNAcβ-S-*p*NP and

Table 2. Partial acceptor specificity of β-1,4-GalT II

Acceptor	Relative β-1,4-GalT II activity, %
1. GlcNAcβ-S-*p*NP	100*
2. GlcNAc	7
3. GlcNAcβ1 → 2Manα1 → 3Manβ1 → R[†]	26
4. GlcNAcβ1 ↘ 4 Manα1 → 3Manβ1 → R GlcNAcβ1 ↗ 2	34
5. GlcNAcβ1 ↘ 6 Manα1 → 3Manβ1 → R GlcNAcβ1 ↗ 2	109

The activity of the β-1,4-GalT II was determined at 1 mM of sugar acceptor.
*One hundred percent activity corresponds to 274.5 pmol/min·mg protein toward 1 mM GlcNAcβ-S-*p*NP.
[†]R: O(CH$_2$)$_8$COOCH$_2$-*m*NP.

GlcNAc. Assay of lactose synthetase activity when using the cell homogenates as an enzyme source revealed that the gene-transfected cells, which acquired a high level of β-1,4-GalT activity to GlcNAcβ-S-*p*NP but not GlcNAc, contain no activity toward glucose in the presence of increasing concentrations of α-lactalbumin (unpublished data). These results showed that the presence of part of the protein A sequence in PROTA/β-1,4-GalT II is not responsible for the lack of lactose synthetase activity because the recombinant protein with only the full coding sequence also did not have such activity. Thus, β-1,4-GalT II is a transferase that has no lactose synthetase activity.

Expression of β-1,4-GalT II mRNA in Human Tissues. The expression of β-1,4-GalT II mRNA in a variety of human tissues was examined by Northern blot analysis. Commercially available Northern blots containing poly(A)$^+$ RNA from human tissues were hybridized with β-1,4-GalT II cDNA as a probe. A single transcript of approximately 4.7 kb was detected in all 24 tissues examined, although the levels of the transcript varied among different tissues (Fig. 3 *Top*). When the same blots were hybridized with human β-1,4-GalT I cDNA as a probe, a single transcript of 4.2 kb was detected in most tissues except brain and lung (Fig. 3 *Bottom*). Comparison of the 4.7-kb and 4.2-kb transcript levels revealed that relatively high levels of the β-1,4-GalT II message are found in brain, lung, skeletal muscle, and bone marrow, in which no detectable or low amounts of the β-1,4-GalT I message are observed. In thyroid, more β-1,4-GalT I mRNA was observed than β-1,4-GalT II mRNA, whereas in spleen and bone marrow, more β-1,4-GalT II mRNA was observed than β-1,4-GalT I mRNA. Thus, the β-1,4-GalT II gene and β-1,4-GalT I gene are both widely expressed in human tissues but at different relative levels.

DISCUSSION

In the present study we have described the isolation of a cDNA clone encoding a novel β-1,4-GalT from a cDNA library of the human breast tumor cell line, MRK-nu-1. The deduced amino acid sequence of this cDNA shows 37% identity with that of the previously described human β-1,4-GalT cDNA. To discriminate these transferases, we named the present β-1,4-GalT as β-1,4-GalT II whereas we refer to the well described β-1,4-GalT as β-1,4-GalT I in the present paper. The isolated 4.7-kb cDNA most likely represents a full-length cDNA because Northern blot analysis revealed a single transcript of the similar size. Hydropathy plot analysis of the deduced amino acid sequence indicated a hydrophobic segment (amino acids 28–48) that may be a transmembrane domain. However, topology refinement based on the helical transmembrane structures (45) revealed alternatively that the region containing amino acids 17–33 could represent that domain. In either case, however, the putative transmembrane domain is located near the N terminus of the protein, indicating a type-II membrane topology that is typical for glycosyltransferases.

The cloning strategy employed in this study was based on the occurrence of several highly conserved regions in the putative catalytic domains of mammalian β-1,4-GalTs and *L. stagnalis* β-1,4-GlcNAcT (18, 24). In particular, a highly conserved region of 10 aa, the (Y/F)(W/F)GWG(G/A)EDD(D/E) box, has been very useful in this study. By 3'-RACE with degenerate primers based on this box, a 860-bp PCR fragment was obtained from human breast tumor cell line mRNA, which subsequently was used as a probe to isolate a human cDNA. Indeed, the cloned β-1,4-GalT II cDNA shows most of the conserved regions as described above, allowing its identification as a novel member of the β-1,4-GalT gene family (24). The 3'-RACE used thus has proven to be a powerful technique to isolate novel members of this gene family. Related cDNA fragments have been obtained from mRNA preparations
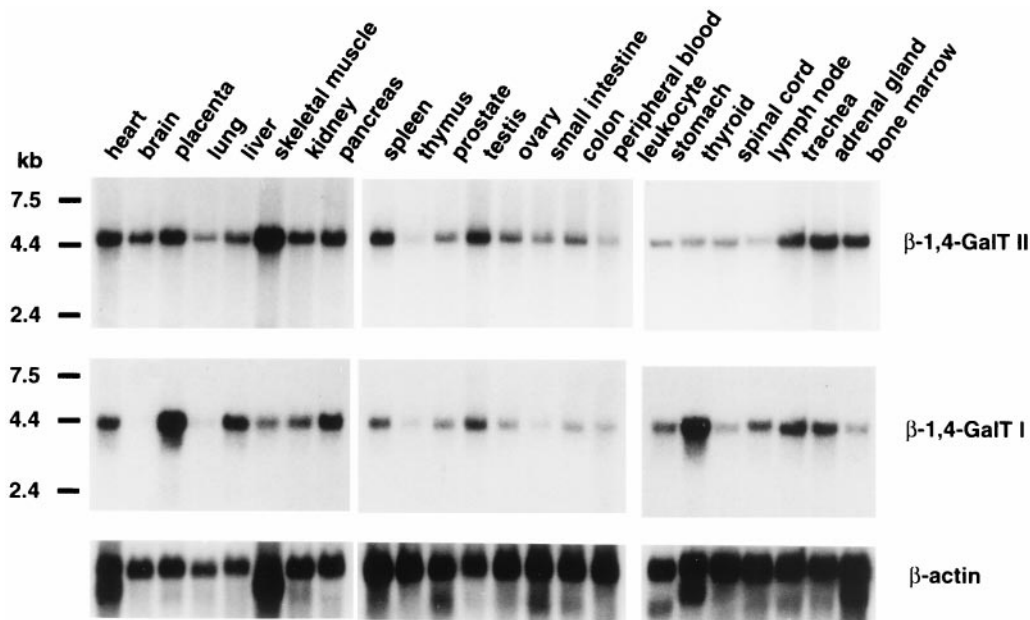
FIG. 3.   Northern blot analysis of β-1,4-GalT II and β-1,4-GalT I transcripts in human tissues. Commercially available mRNA blots containing 2 μg of poly(A)$^+$ RNA/lane were probed with ^{32}P-labeled β-1,4-GalT II cDNA, human β-1,4-GalT I cDNA, or β-actin cDNA. The sizes of RNA markers are indicated in kilobases (kb).

derived from bovine mammary gland and mouse brain (unpublished data) and also have been reported as expressed sequence tags (accession nos. AA013728 and AA015244), suggesting the existence of a similar gene in cow and mouse. These data suggest that the β-1,4-GalT I and β-1,4-GalT II genes form separate lineages in mammals and may have evolved from a common ancestor gene by duplication event that happened before radiation of mouse, cow, and man.

Initial acceptor specificity studies suggest that differences may be expected between the two enzymes in their affinity for oligosaccharide substrates because the sequence identity between them is only 37%. Quite interestingly, β-1,4-GalT II seems to galactosylate acceptors containing GlcNAcβ1 → 6(GlcNAcβ1 → 2)Man structure more efficiently than those containing GlcNAcβ1 → 2Man and GlcNAcβ1 → 4(GlcNAcβ1 → 2)Man structures. These results suggest that β-1,4-GalT II prefers terminal *N*-acetylglucosamine residue, which is β1 → 6 linked as found in tri- and tetraantennary complex-type oligosaccharides. This property is also found in the genetically related *L. stagnalis* β-1,4-GlcNAcT (19). A marked difference was observed in lactose synthesis because the novel β-1,4-GalT showed no ability to transfer galactose from UDP-Gal to glucose in the presence of α-lactalbumin. Under the same conditions bovine mammary gland β-1,4-GalT I synthesized lactose. In addition, β-1,4-GalT II may not be involved in lacto-series glycolipid synthesis because no appreciable transfer of galactose was observed to lacto-*N*-triaosylceramide. Further studies on the acceptor specificity as well as kinetics of this novel β-1,4-GalT, however, are necessary to get more insight in the specific biological function of this enzyme.

Northern blot analysis revealed that β-1,4-GalT II is expressed in most human tissues like β-1,4-GalT I, although the relative expression levels of the transcript vary. This raises the question as to the specific *in vivo* function of either enzyme. The β-1,4-GalT I has been proposed to have a housekeeping function in addition to a specific function in the mammary gland, where it is involved in lactose synthesis on interaction with α-lactalbumin (9, 15). In β-1,4-GalT I-deficient mice, there was residual galactosylation of N-linked oligosaccharides, which indeed might result from the action of β-1,4-GalT II. However, the milk from such mutant mice does not contain lactose (13), which is consistent with the unresponsiveness of

the latter enzyme to α-lactalbumin. Remarkably, in human brain a significant level of β-1,4-GalT II transcript was present whereas no β-1,4-GalT I message was found. Similarly, mouse brain has been reported to express low levels of β-1,4-GalT I transcript (46). Because mouse brain glycoproteins have been reported to contain N-linked oligosaccharides with Galβ1 → 4GlcNAc groups that may carry the HNK-1 epitope or polysialic acid (4), it might be expected that β-1,4-GalT II is the enzyme that is involved in the galactosylation of brain glycoproteins rather than β-1,4-GalT I. In other tissues both enzymes may participate in the cooperative galactosylation of different substrates in a way that is, so far, unknown. This might be important, for example, in the case of B cells producing IgG molecules, the galactosylation of which can vary over the course of rheumatoid arthritis and other inflammatory diseases (reviewed in refs. 47 and 48). In this context it is of interest to note that β-1,4-GalT II shows considerable expression in bone marrow, a tissue that is rich in B cells. Furthermore, in the lung of β-1,4-GalT I-deficient mice, in which no transcript nor activity of the transferase could be detected, high levels of β-1,4-galactosylated proteins were found (12). This suggests that these proteins can be β-1,4-galactosylated in an alternative pathway. In human lung, much higher β-1,4-GalT II transcript levels than β-1,4-GalT I transcript levels are observed. A similar expression pattern in mouse lung would account for the presence of such β-1,4-linked galactose on the glycoproteins in this tissue. The discovery of this novel β-1,4-GalT will lead to further understanding of the biological significance of protein galactosylation.

Note Added in Proof. During processing of this manuscript, related but not identical human β-1,4-GalTs have been described by Almeida, R., *et al.* (49). These β-1,4-GalTs have been named β-4Gal-T2 and β-4Gal-T3. Therefore, we propose to refer to β-1,4-GalT II described here as β-1,4-GalT IV in the future.

1. Hakomori, S. (1991) *Curr. Opin. Immunol.* **3,** 646–653.
2. Varki, A. (1993) *Glycobiology* **3,** 97–130.
3. Bhavanadan, V. P. & Furukawa, K. (1995) in *Biology of the Sialic Acids*, ed. Rosenberg, A. (Plenum, New York), pp. 145–196.
4. Kunemund, V., Jungalwala, F. B., Fischer, G., Chou, D. K. H., Keilhauer, G. & Schachner, M. (1988) *J. Cell Biol.* **106,** 213–223.
5. Schachter, H. (1991) *Glycobiology* **1,** 453–461.
6. 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.
7. Brew, K., Vanaman, T. C. & Hill, R. L. (1968) *Proc. Natl. Acad. Sci. USA* **59,** 491–497.
8. D'Agostaro, G., Bendiak, B. & Tropak, M. (1989) *Eur. J. Biochem.* **183,** 211–217.
9. Rajput, B., Shaper, N. L. & Shaper, J. H. (1996) *J. Biol. Chem.* **271,** 5131–5142.
10. Ioffe, E. & Stanley, P. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 728–732.
11. Metzler, M., Gertz, A., Sarkar, M., Schachter, H., Schrader, J. W. & Marth, J. D. (1994) *EMBO J.* **13,** 2056–2065.
12. Lu, Q., Hasty, P. & Shur, B. D. (1997) *Dev. Biol.* **181,** 257–267.
13. Asano, M., Furukawa, K., Kido, M., Matsumoto, S., Umesaki, Y., Kochibe, N. & Iwakura, Y. (1997) *EMBO J.* **16,** 1850–1857.
14. Morrison, J. F. & Ebner, K. E. (1971) *J. Biol. Chem.* **246,** 3992–3998.
15. Hill, R. L. & Brew, K. (1975) *Adv. Enzymol. Rel. Areas Mol. Biol.* **43,** 411–490.
16. Sheares, B, T. & Carlson, D. M. (1984) *J. Biol. Chem.* **259,** 8045–8047.
17. Rivera-Marrero, C. A. & Cummings, R. D. (1990) *Mol. Biochem. Parasitol.* **43,** 59–68.
18. Bakker, H., Agterberg, M., Van Tetering, A., Koeleman, C. A. M., Van den Eijnden, D. H. & Van Die, I. (1994) *J. Biol. Chem.* **269,** 30326–30333.
19. Bakker, H., Schoenmakers, P. S., Koeleman, C. A. M., Joziasse, D. H., Van Die, I. & Van den Eijnden, D. H. (1997) *Glycobiology* **7,** 539–548.
20. Masri, K. A., Appert, H. E. & Fukuda, M. N. (1988) *Biochem. Biophys. Res. Comm.* **157,** 657–663.
21. Shaper, N. L., Hollis, G. F., Douglas, J. G., Kirsch, I. R. & Shaper, J. H. (1988) *J. Biol. Chem.* **263,** 10420–10428.
22. Shaper, J. H., Joziasse, D. H., Meurer, J. A., Chou, T. D., Schnaar, R. L. & Shaper, N. L. (1995) *Glycoconjugate J.* **12,** 477–478.
23. Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., *et al.* (1994) *Nature (London)* **368,** 32–38.
24. Van Die, I., Bakker, H. & Van den Eijnden, D. H. (1997) *Glycobiology* **7,** v–ix.
25. Furukawa, K., Matsuta, K., Takeuchi, F., Kosuge, E., Miyamoto, T. & Kobata, A. (1990) *Intern. Immunol.* **2,** 105–112.
26. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
27. Uhlen, M., Guss, B., Nilsson, B., Gatenbeck, S., Philipson, L. & Lindberg, M. (1984) *J. Biol. Chem.* **259,** 1695–1702.
28. Sanchez-Lopez, R., Nicholson, R., Gesnel, M.-C., Matrisian, L. M. & Breathnach, R. (1988) *J. Biol. Chem.* **263,** 11892–11899.
29. Chu, G., Hayakawa, H. & Berg, P. (1987) *Nucleic Acids Res.* **15,** 1311–1326.
30. Palcic, M. M., Heerze, L. D., Pierce, M. & Hindsgaul, O. (1990) *Glycoconjugate J.* **5,** 49–63.
31. Glasgow, L. R., Paulson, J. C. & Hill, R. L. (1977) *J. Biol. Chem.* **252,** 8615–8623.
32. Merkle, R. K. & Cummings, R. D. (1987) *Methods Enzymol.* **138,** 232–259.
33. Furukawa, K., Roberts, D. D., Endo, T. & Kobata, A. (1989) *Arch. Biochem. Biophys.* **270,** 302–312.
34. Basu, M., De, T., Das, K. K., Kyle, J. W., Chon, H.-C., Schaeper, R. J. & Basu, S. (1987) *Methods Enzymol.* **138,** 575–607.
35. Ebner, K. E., Mawal, R., Fitzgerald, D. K. & Colvin, B. (1972) *Methods Enzymol.* **28B,** 507–510.
36. Nakazawa, K., Furukawa, K., Narimatsu, H. & Kobata, A. (1993) *J. Biochem.* **113,** 747–753.
37. Furukawa, K. & Roth, S. (1985) *Biochem. J.* **227,** 573–582.
38. Kyte, J. & Doolittle, R. F. (1982) *J. Mol. Biol.* **157,** 105–132.
39. Paulson, J. C. & Colley, K. J. (1989) *J. Biol. Chem.* **264,** 17615–17618.
40. Yadav, S. P. & Brew, K. (1991) *J. Biol. Chem.* **266,** 698–703.
41. Kurosawa, N., Inoue, M., Yoshida, Y. & Tsuji, S. (1996) *J. Biol. Chem.* **271,** 15109–15116.
42. Bierhuizen, M. A. & Fukuda, M. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 9326–9330.
43. Bakker, H., Van Tetering, A., Agterberg, M., Smit, A. B., Van den Eijnden, D. H. & Van Die, I. (1997) *J. Biol. Chem.* **272,** 18580–18585.
44. Baenziger, J. U. & Fiete, D. (1979) *J. Biol. Chem.* **254,** 9795–9799.
45. Rost, B., Casadio, R., Fariselli, P. & Sander, C. (1995) *Protein Sci.* **4,** 521–533.
46. Harduin-Lepers, A., Shaper, J. H. & Shaper, N. L. (1993) *J. Biol. Chem.* **268,** 14348–14359.
47. Furukawa, K. & Kobata, A. (1991) *Mol. Immunol.* **28,** 1333–1340.
48. Filley, E., Andreoli, A., Steele, J., Waters, M., Wagner, D., Nelson, D., Tung, K., Rademacher, T., Dwek, R. & Rook, G. A. W. (1989) *Clin. Exp. Immunol.* **76,** 343–347.
49. Almeida, R., Amado, M., David, L., Levery, S. B., Holmes, E. H., Merkx, G., van Kessel, A. G., Rygaard, E., Hassan, H., Bennett, E. & Clausen, H. (1997) *J. Biol. Chem.* **272,** 31979–31991.