Molecular Cloning and Expression of an N-Acetylgalactosamine-4-O-Sulfotransferase that Transfers Sulfate to Terminal and Non-Terminal β1,4-linked N-Acetylgalactosamine

Hyung-Gyoo Kang #*, Matthias R. Evers ‡*, Guoqing Xia ‡*, Jacques U. Baenziger #, and Melitta Schachner §§

From the ‡ Zentrum fuer Molekulare Neurobiologie, Universitaet Hamburg, Martinistr. 52, D-20246 Hamburg, Germany, and # Department of Pathology, Washington University School of Medicine, 660 S. Euclid Ave, St. Louis, MO 63110, USA

The nucleotide sequence reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession numbers AF332472 and AF332473.

§ To whom correspondence should be addressed: Zentrum fuer Molekulare Neurobiologie, Universitaet Hamburg, Martinistr. 52, 20246 Hamburg, Germany.
Tel + 49 - 40 - 42803 – 6246;
Fax + 49 - 40 - 42803 – 6248;
E-Mail: melitta.schachner@zmnh.uni-hamburg.de

Running title: Cloning of an N-Acetylgalactosamine-4-O-Sulfotransferase

* These three authors contributed equally to this work.
Abstract

We have identified and characterized an N-acetylgalactosamine 4-O-sulfotransferase designated GalNAc-4-ST2 (GenBank™ accession number AF332472) based on its homology to HNK-1 sulfotransferase (HNK-1 ST). The cDNA predicts an open reading frame encoding a type II membrane protein of 443 amino acids with a 12 amino acid cytoplasmic domain, a 23 amino acid transmembrane domain, and a 408 amino acid luminal domain containing 4 potential N-linked glycosylation sites. GalNAc-4-ST2 displays a high degree of amino acid sequence identity with GalNAc-4-ST1 (46%), HNK-1 ST (23%), chondroitin 4-O- sulfotransferase-1 (C4ST-1) (27%), and chondroitin 4-O-sulfotransferase-2 (C4ST-2) (24%). GalNAc-4-ST2 transfers sulfate to the C-4 hydroxyl of terminal β1,4-linked GalNAc in the sequence GalNAc-β1,4GlcNAcβ-R found on N-linked oligosaccharides and non-terminal β1,4-linked GalNAc in chondroitin and dermanan. The translated region of GalNAc-4-ST2 is encoded by five exons located on human chromosome 18q11.2. Northern blot analysis reveals a 2.1 kilobase transcript. GalNAc-4-ST2 message is most highly expressed in trachea and to a lesser extent in heart, liver, pancreas, salivary gland and testis. The I.M.A.G.E. cDNA clone 49547 contains a putative GalNAc-4-ST2 splice form with an open reading frame encoding a protein of 358 amino acids that lacks the transmembrane domain and the stem region. This form of GalNAc-4-ST2 is not retained by transfected cells and is active against chondroitin but not terminal β1,4-linked GalNAc. Thus, as with GalNAc-4-ST1, sequences amino terminal to the catalytic domain contribute to the specificity of GalNAc-4-ST2 toward terminal β1,4-linked GalNAc.
Introduction

Sulfated carbohydrate structures have been shown to play important roles in a large number of different molecular interactions including: symbiotic interactions between plants and nitrogen-fixing bacteria, homing of lymphocytes to lymph nodes, control of the circulatory half-life of the glycoprotein hormones LH and TSH, binding of growth factors by proteoglycans, and triggering preferential neurite outgrowth. HNK-1, an example of such a sulfated glycan, has the structure $\text{SO}_4^{-3-}\text{GlcA}\beta_1,3\text{Gal}\beta_1,4\text{GlcNAc-R}$ and is found at the non-reducing termini of glycoprotein and glycolipid oligosaccharides. The structure was first identified with monoclonal antibodies as an epitope on human natural killer cells. The HNK-1 epitope is a hallmark of many neural recognition molecules and displays phylogenetic conservation, highlighting its functional importance. It is expressed in the central and peripheral nervous systems during development and regeneration, and has recently been recognized to be a crucial player in synaptic plasticity involving inhibitory interneurons in the hippocampus. HNK-1 is also a predominant autoantigen in demyelinating diseases of the peripheral nervous system.

Expression cloning of the rat and human sulfotransferases responsible for synthesis of the HNK-1 epitope revealed the presence of two sequence motifs that are associated with 3’-phospho-adenosyl-5’-phosphosulfonate (PAPS) binding and are common among all sulfotransferases cloned to date. Despite the presence of these motifs, HNK-1 sulfotransferase shows only limited similarity to other sulfotransferases outside of the PAPS binding regions. Nonetheless, three additional sulfotransferases, chondroitin 4-O-sulfotransferase 1 (C4ST-1), chondroitin 4-O-sulfotransferase 2 (C4ST-2), and N-acetylgalactosamine-4-O-sulfotransferase 1 (GalNAc-4-ST1) were identified based on their homology to the catalytic C-
terminal domain of HNK-1 ST. C4ST-1 and C4ST-2 (Fig. 2) show 25% and 22% identity to HNK-1 ST, respectively, while GalNAc-4-ST1 is 23% identical to HNK-1 ST. In addition to the motifs associated with binding PAPS, HNK-1 ST, C4ST-1, C4ST-2, and GalNAc-4-ST1 have three additional motifs of unknown function that are located carboxy terminal to the PAPS binding motifs. While homologous, these sulfotransferases have distinct saccharide specificities. HNK-1 ST transfers sulfate to the C-3 hydroxyl of terminal β1,3-linked GlcA in the sequence GlcAβ1,3Galβ1,4GlcNAcβ-17;18. C4ST-1 and C4ST-2 transfer sulfate to the C-4 hydroxyl of non-terminal β1,4-linked GalNAc in chondroitin and dermatan 22;23. GalNAc-4-ST1 transfers sulfate to the C-4 hydroxyl of terminal β1,4-linked GalNAc in the sequence GalNAc-β1,4GlcNAcβ-R that is found on N-linked oligosaccharides 24;25.

We have identified and characterized a fourth sulfotransferase that is homologous to HNK-1 ST with 23% identical amino acids. N-Acetylgalactosamine-4-O-sulfotransferase 2 (GalNAc-4-ST2) transfers sulfate to the C-4 hydroxyl of terminal β1,4-linked GalNAc in the sequence GalNAc-β1,4GlcNAcβ-R and to the C-4 hydroxyl of internal β1,4-linked GalNAc moieties in chondroitin and dermatan. We describe the properties of this novel sulfotransferase that differs in its specificity and expression from other GalNAc-specific sulfotransferases.

**Experimental Procedures**

**Materials.** α-[32P]dCTP and Megaprime labelling kit were purchased from Amersham Pharmacia Biotech, human Multiple Tissue Northern blots (MTN®) and
human Multiple Tissue Expression (MTE™) arrays from Clontech Laboratories, Inc. Bovine LH was provided by the NIDDK National Hormone and Pituitary Program, Dr. A.F. Parlow.

**Molecular Cloning of a human cDNA homologous to HNK-1 sulfotransferase.** A human EST derived from human infant brain (GenBank™ accession no. H15485) with high sequence similarity to rat HNK-1 ST\(^{17}\) was identified using BLASTN and TBLASTN algorithms\(^{26}\) against the databases of GenBank™, EMBL and DDBJ EST Divisions at the National Center for Biotechnology Information (National Institutes of Health, Bethesda, MD). The insert of the corresponding I.M.A.G.E.Consortium [LLNL] cDNA clone, ID 49547 (IMAGp951B2016 obtained from RZPD, Berlin, Germany)\(^{27}\) was sequenced and found to contain a complete open reading frame encoding for a 358 amino acid protein showing significant sequence similarity to rat and human HNK-1 sulfotransferases. This sequence was designated as GalNAc-4-ST2(Met86) (GenBank™ accession no. AF332473).

The forward primer 5’-GGG AGA GTG GAG AAG AGA AGA GAA C-3’ and the reverse primer 5’-AAG CCA ATC CAT TTA GTA CCA TCA GA-3’ were used to amplify a 595 bp fragment (nts at position 211-805 in the GalNAc-4-ST2 cDNA) including the putative start codon from first-strand cDNA produced from 100 ng of total RNA obtained from a human glioma and the human kidney 293 cell line using Omniscript™ RT (Qiagen). An additional 38 bps (designated exon 4 in Fig. 3) were present in this product when compared to GalNAc-4-ST2(Met86) and shifted the stop codons located 5’ out of the reading frame of the putative translation start site. Similarity searches identified a bovine EST (GenBank™ accession no. BE724107) with sequence homology to the 5’ end of the amplified fragment, allowing us to
predict a chimeric sequence from the human and bovine cDNAs that would contain an ORF encoding for a putative 443 amino acid protein. Human genomic clones (GenBank™ accession nos. AC009872, AC023575, AC010854) were analyzed with FGENESH and the predicted exons evaluated using the chimeric GalNAc-4-ST2 cDNA. Primer pairs derived from exonic sequences were designed and used to amplify the full-length GalNAc-4 ST2 sequence shown in Figure 1A (GenBank™ accession. No. AF332472) from both human embryonic kidney 293 cell line and glioma cDNA.

**In-silico analysis of the human GalNAc-4-ST2 locus.** A BLASTN search with the GalNAc-4-ST2 and GalNAc-4-ST2(Met86) sequences against the High Throughput Genomic Sequences (hgts) data base retrieved several entries (GenBank™ accession nos. AC012269, AP001272, AC032032, AC009872, AC023575) annotated to be derived from chromosome 18q11.2. The genomic clone with the GenBank™ accession number AC010854 was retrieved with a BLASTN search against the non-redundant database. The organization of the GalNAc-4-ST2 gene shown in Figure 3 was deduced by comparing the cDNA of GalNAc-4-ST2 with these genomic sequences.

**Construction of pcDNA3.1-GalNAc-4-ST2(Met86) and pcDNA3.1-GalNAc-4-ST2.** The complete open reading frame of GalNAc-4-ST2 was amplified from cDNA prepared from human kidney 293 cells by polymerase chain reaction (PCR) using: 1) The 5´-specific primer 5´-ATA TGG ATC CGC CAC CAT GCA GCC ATC TGA AAT G-3´ containing a BamHI site, the consensus Kozak sequence GCCACC and a start codon; and 2) The 3´-specific primer 5´-ATA TTC TAG ACT ACA AAA ATG GAG TTG TAT AAT TAA-3´ containing a stop codon and an Xba I site. The open reading frame of GalNAc-4-ST2(Met86) was amplified by PCR from
the human I.M.A.G.E. cDNA clone 49547 using the 5'-specific primer 5'-GCG GAT CCG CCA CCA TGC CTG AGG ATG TAC GAG AA-3' (contains an *Bam*HI site, consensus Kozak sequence GCCACC, and a start codon) and the same 3' primer that was used to obtain GalNAc-4-ST2, above. PCR reactions were carried out using PfuTurbo® polymerase (Stratagene) with 35 cycles of a reaction consisting of 45 s denaturation at 95°C, 60 s annealing at 55°C, and 60 s of elongation at 72°C. The PCR fragments had the expected lengths of 1332 bp and 1077 bp, respectively, and were directionally subcloned into the eukaryotic expression vector pcDNA3.1 (Invitrogen).

**Transient Expression of GalNAc-4-ST2 and GalNAc-4-ST2(MET86).**

CHO/Tag cells were transfected with 13 µg of pcDNA3.1-GalNAc-4-ST2, pcDNA3.1-C4ST-1, pcDNA3.1- GalNAc-4-ST2(Met86) or pcDNA3.1 and 35 µg of Lipofectamine (GIBCO/BRL) in serum-free medium for 6 h per the manufacturer's protocol. Sixty hours after transfection, the cells and medium were collected separately for analysis. Cells were lysed with 200 µl of 20 mM HEPES buffer, pH 7.4, 5 mM MgCl₂, 175 mM KCl, 2% Triton X-100, protease inhibitors (23 millitrypsin inhibitor units of aprotinin and 4 µg each of leupeptin, antipain, pepstatin, and chymostatin) per 100 mm-diameter culture plate. The homogenate was mixed by rotation for 1 h and sedimented at 12,000 x g for 20 min. The supernatant was designated as the cell extract. The culture medium was pooled, sedimented at 12,000 x g for 20 min, and the supernatant adjusted to a final concentration of 20 mM HEPES, pH 7.4 and protease inhibitors added as noted above.

**Sulfotransferase Assays.** GalNAc-4-sulfotransferase reactions (50 µl) were carried out as described at 28°C for 15 h. Each reaction contained 15 mM HEPES, pH 7.4, 1% Triton X-100, 40 mM 2-mercaptoethanol, 10 mM NaF, 1 mM ATP, 4
mM magnesium acetate, 13% glycerol, protease inhibitors, 2 µM unlabelled PAPS, 1x 10^6 cpm [^{35}S]PAPS, 20 µM GalNAcβ1,4 GlcNAcβ1,2Manα-MCO, and enzyme. [^{35}S]SO_4-GalNAcβ1,4 GlcNAcβ1,2Manα-MCO was separated from [^{35}S]PAPS and labelled endogenous acceptors by passage over a Sep-Pak C_{18} cartridge (Waters)^{30}. Control reactions were done in the absence of substrate or enzyme. Chondroitin and dermatan GalNAc-4-sulfotransferase activities were determined using a previously described assay^{23} and 50 µg of desulfated chondroitin (Seikagaku America, Inc.) or desulfated dermatan^{31} as acceptor. HNK-1-sulfotransferase activity was assayed using the acceptor GlcAβ1,3Galβ1,4GlcNAcβ1,3Galβ1,4Glcβ1,C_2H_4NHCOCF_3 as described previously^{17} except that the reactions were carried out for 15 h.

**Product Characterization.** [^{35}S]SO_4-GalNAcβ1,4 GlcNAcβ1,2Manα-MCO obtained by incubation with GalNAc-4-ST2 or GalNAc-4-ST2(Met86) for 16 h at 28 °C was isolated on a Sep-Pak C_{18} (Waters) and characterized as described previously with a CarboPak PA1 column (Dionex)^{30}. Sulfated monosaccharides were released from the [^{35}S]SO_4-GalNAcβ1,4 GlcNAcβ1,2Manα-MCO product by mild acid hydrolysis and separated from intact oligosaccharides and other degradation products by gel filtration on Sephadex G-10 in 100 mM NH_4HCO_3 prior to analysis by HPLC on CarboPak PA1 column (Dionex) as described^{30}.

The [^{35}S]SO_4-labelled chondroitin and dermatan products were digested with 30 milliunits of chondroitinase ABC (Seikagaku America, Inc.) in 100 mM Tris-HCl, pH 8.0 for 16 h at 37°C. The completeness of digestion was determined by gel filtration on Sephadex G-25 in 100 mM NH_4HCO_3. The released [^{35}S]SO_4-labelled saccharides were characterized by HPLC using a 4.6 x 250 mm MicroPak AX-5 column (Varian Associates). Saccharides were separated using a linear gradient of 10 mM KH_2PO_4 to 450 mM KH_2PO_4 over 40 min. at 1.0 ml per min.^{32} Standards
were detected by their absorbance at 210 or 234 nm and fractions were collected at
0.5 min intervals for determination of radioactivity. In excess of 90% of the
\([^{35}S]SO_4\) label was recovered.

**Sulfation of glycoprotein acceptors.** Glycoproteins were tested as
substrates for GalNAc-4-ST2 using the conditions described above for
oligosaccharide acceptors with the following modifications. No unlabelled PAPS
was added to the \([^{35}S]PAPS\). Each 50 µl reaction contained 3 µg of purified bovine
parotid carbonic anhydrase VI (CA-VI)\(^{33}\), bovine LH, enzymatically desulfated
bovine LH, or asialo human chorionic gonadotrophin (hCG). After 16 h at 28°C,
duplicate reactions were either stopped by addition of an equal volume of sample
buffer (10% glycerol, 5% 2-mercaptoethanol, 2% SDS, 0.003% bromophenol blue,
and 62.5 mM Tris, pH 6.8) or were digested with 34 microunits of peptide:N-
glycosidase F (PNGase F) as described\(^3\) prior to addition of sample buffer. The
\([^{35}S]SO_4\)-labelled proteins were separated by SDS-PAGE (12% acrylamide) and
detected by autoradiography.

**Northern Blot and Expression Array Analysis.** Human Multiple Tissue
Northern (MTN\(^\text{®}\)) blots and Human Multiple Tissue Expression (MTE\(^\text{TM}\)) arrays
were purchased from Clontech Laboratories, Inc. They were hybridized with
5 - 15 \(10^6\) cpm of a GalNAc-4-ST2 specific α-[\(^{32}\)P]-labelled cDNA probe and
washed according to the manufacturer’s specifications. The membranes were
exposed to Biomax MS films (Kodak) for 2 to 6 days at –80°C with intensifying
screens. The 503 bp probe used for the labeling reactions corresponds to nt 745 to
1247 of the GalNac-4-ST2 cDNA (GenBank\(^\text{TM}\) accession No. AF332472).
RESULTS

Identification of a human cDNA related to HNK-1 ST and GalNAc-4-ST1.

We recently identified and characterized a human GalNAc-4-O-sulfotransferase, GalNAc-4-ST1, based on its homology to the HNK-1 ST that transfers sulfate to GlcAβ1,3Galβ1,4GlcNAc-R to produce the HNK-1 epitope SO₄³⁻-GlcAβ1,3Galβ1,4GlcNAc-R. In contrast to HNK-1 ST, GalNAc-4-ST1 transfers sulfate to the C4 hydroxyl of β1,4-linked GalNAc on N-linked oligosaccharides such as those found on the glycoprotein hormones LH and TSH. Further screening of the EST database at GenBank, National Center for Biotechnology Information, using the deduced amino acid sequence of the catalytic domain of rat HNK-1 ST (GenBank accession no. AF022729) identified the human EST H15485. The insert of the corresponding I.M.A.G.E. Consortium [LLNL] cDNA clone ID 49547 contains an open reading frame of 1077 bp with a single in-frame ATG codon at the 5'-end preceded by two in-frame stop codons. The open-reading frame predicts a protein of 358 amino acid residues with a molecular mass of 42.1 kDa and four potential N-linked glycosylation sites. The predicted protein sequence is closely related to GalNAc-4-ST1 with 44% identical residues. However, no membrane-spanning regions are predicted to be present by hidden Markov model based algorithm TMHMM (Sonnhammer et al. 1998). This suggests the presumptive sulfotransferase, designated as GalNAc-4-ST2(Met86) (GenBank accession No. AF332473), is synthesized as a soluble glycoprotein.

Transcripts derived from a human embryonic kidney 293 cell line and glioma cDNA contained an additional 38 bp sequence, later identified as exon 4 (see below), that was not present in GalNAc-4-ST2(Met86). A virtual cDNA consisting of this
amplified sequence and an EST (GenBank™ accession no. BE724107) from the 5’ end of the coding region of the bovine orthologue was deduced and used to analyze genomic clones representing the human GalNAc-4-ST2 locus (see below). Primer pairs were designed for the predicted open reading frame (ORF) and used to amplify the sequence designated GalNAc-4-ST2 in Figure 1A (GenBank™ accession No. AF332472) from first strand cDNAs derived from both 293 cells and a glioma. The ORF of GalNAc-4-ST2 is preceded at its 5’ end by a stop codon in all three reading frames and predicts a protein of 443 amino acids with four potential glycosylation sites and a molecular mass of 52.1 kDa. GalNAc-4-ST2 is predicted to have a membrane spanning domain by the TMHMM algorithm and a type II topology. If exon 4 is absent due to alternative splicing, GalNAc-4-ST2(Met86) would be initiated from the alternative ATG start codon located at the position of the Met that is equivalent to amino acid 86 in full-length GalNAc-4-ST2 (Fig.1).

Multiple alignment of the protein sequence of GalNAc-4-ST2 with other members of the HNK-1 sulfotransferase family using the ClustalW algorithm implemented in the BIOEDIT suite (Fig. 2) revealed that GalNAc-4-ST2 is 46% identical to GalNAc-4-ST1. GalNAc-4-ST2 also has a high percentage of identical amino acids when compared to other HNK-1 ST related sulfotransferases; 23% to human HNK-1 ST, 27% to human chondroitin 4-O-sulfotransferase-1 (C4ST-1), and 24% to human chondroitin 4-O-sulfotransferase-2 (C4ST-2) (Fig. 2). Regions with the highest degree of identity include the putative 5’-phosphosulfonate binding site (5’-PSB), the putative 3’-phosphate binding site (3’-PB), and three regions designated III, IV, & V that do not have an identified function.
Genomic organization and chromosome localization of GalNAc-4-ST2.
BLAST similarity searches of the High Throughput Genomic Sequences (htgs) and the non-redundant data set were performed using the GalNAc-4-ST2 cDNA sequence and retrieved multiple matching sequences (see “Experimental Procedures”). Among the genomic clones retrieved GenBank™ accession no. AP001272 and AP001087 were both annotated to be mapped on chromosome 18q11.2. Comparisons between the cDNA and the genomic sequences AC009872, AC023575 and AC010854 showed that the coding region of GalNAc-4-ST2 is distributed over five exons (Fig. 3A). The size of the untranslated regions of exon 1 and exon 5 have not yet been determined. As noted above, GalNAc-4-ST2(Met86) does not contain exon 4. Synthesis of GalNAc-4-ST(Met86) is therefore initiated from the alternative start codon located at amino acid position 86 of GalNAc-4-ST2 producing a protein that does not have a transmembrane domain (Fig. 3B). Each of the exons identified, including exon 4, was found to have conserved donor and acceptor splice sites (Fig. 3C).

GalNAc-4-ST2 transfers sulfate to terminal β1,4-linked GalNAc on N-linked oligosaccharides. Since GalNAc-4-ST2 is closely related to GalNAc-4-ST1, the specificity of GalNAc-4-ST2 for saccharides terminating with β1,4-linked GalNAc was examined using the substrates shown in Table I. pcDNA3.1-GalNAc-4-ST2 was transfected into CHO/Tag cells and cell extracts prepared for analysis. Like GalNAc-4-ST1,GalNAc-4-ST2 transfers sulfate to GalNAcβ1,4 GlcNAcβ1,2Manα-MCO and to GalNAcβ1,4 GlcNAcβ-MCO. GalNAc-4-ST2 does not transfer sulfate to either Galβ1,4GlcNAcβ-MCO or GlcNAcβ1,2Manα-MCO (Table I) indicating that the β1,4-linked GalNAc is required and likely the sugar modified with sulfate. The oligosaccharide acceptors in Table I were tested at 20 µM and 100 µM; concentrations that are equal to and five-fold greater than the $K_m$ of 15 µM that we
previously reported for GalNAcβ1,4 GlcNAcβ1,2Manα-MCO with the GalNAc-4-sulfotransferase found in the pituitary gland. As we have observed with the GalNAc-4-sulfotransferase from the pituitary gland and with GalNAc-4-ST1, GalNAc-4-ST2 transfers sulfate to GalNAcβ1,4GlcNAcβ1,2Manα-MCO more efficiently than to GalNAcβ1,4GlcNAcβ-MCO.

The sulfated product was characterized as described previously to establish the location of the sulfate. The [35S]SO₄ labelled GalNAcβ1,4 GlcNAcβ1,2Manα-MCO product co-migrated with authentic SO₄-4-GalNAcβ1,4 GlcNAcβ1,2Manα-MCO when analyzed by HPLC (Fig. 4A). Following mild acid hydrolysis, GalNAc-4-SO₄ was the only product obtained (Fig. 4B). Thus, GalNAc-4-ST2 transfers sulfate exclusively to the C4 hydroxyl of the terminal GalNAc in GalNAcβ1,4 GlcNAcβ1,2Manα-MCO.

We also examined GalNAc-4-ST2 for its ability to transfer sulfate to glycoproteins bearing N-linked oligosaccharides terminating with the sequence GalNAcβ1,4 GlcNAcβ1,2Manα such as are found on the glycoprotein hormones LH and TSH and carbonic anhydrase VI (CA-VI). GalNAc-4-ST2 transfers sulfate to bovine CA-VI isolated from parotid gland and to bovine LH (Fig. 5). The N-linked oligosaccharides on CA-VI isolated from parotid glands bear terminal β1,4-linked GalNAc due to the absence of sulfotransferase in the parotid gland. In contrast, >90% of the GalNAc on the N-linked oligosaccharides of LH is substituted with sulfate. Following removal of terminal sulfate moieties from the N-linked oligosaccharides on LH by digestion with GalNAc-4-sulfatase, there is a marked increase in the amount of sulfate transferred to LH. This indicates the sulfate is added to the terminal β1,4-linked GalNAc that has been exposed by the digestion (Fig. 5). The sulfate label is released from CA-VI and LH by digestion with PNGase F further
confirming its location on N-linked oligosaccharides. Glycoproteins bearing oligosaccharides with terminal β1,4-linked Gal, such as asialo-hCG, were not modified with sulfate (not shown). Thus, GalNAc-4-ST2, like GalNAc-4-ST1, is capable of specifically modifying oligosaccharides on glycoproteins bearing terminal β1,4-linked GalNAc. We did note that GalNAc-4-ST1 shows a preference for LH while GalNAc-4-ST2 prefers CA-VI, suggesting the peptide portion of these glycoproteins may contribute to recognition by one or the other GalNAc-4-ST.

**GalNAc-4-ST2 transfers sulfate to internal β1,4-linked GalNAc moieties in chondroitin and dermatan.** Even though GalNAc-4-ST2 displays the highest degree of identity with GalNAc-4-ST1, it is also homologous to C4ST-1, C4ST-2, and HNK-1 ST. We therefore examined extracts and medium from CHO/Tag cells transfected with pcDNA3.1-GalNAc-4-ST2 and pcDNA3.1-GalNAc-4-ST2(Met86) for transfer of sulfate to chondroitin, dermatan, and the HNK-1 precursor GlcAβ1,3Galβ1,4GlcNAcβ1,3Galβ1,4Glcβ1,C2H4NHCOCF3. No evidence of transfer to GlcAβ1,3Galβ1,4GlcNAcβ1,3Galβ1,4Glcβ1,C2H4NHCOCF3 was seen (not shown). Following transfection of pcDNA3.1-GalNAc-4-ST2 into CHO/Tag cells, 15% of the sulfotransferase activity directed at GalNAcβ1,4 GlcNAcβ1,2Manα-MCO was found in the cell extract and 85% in the medium (Fig.6). Release of GalNAc-4-ST2 into the medium, most likely results from proteolytic cleavage as has been seen for a number of transferases including GalNAc-4-ST1. While GalNAc-4-ST2 in the cell extract is able to transfer sulfate to chondroitin, there is a 3-fold increase in the relative rate of transfer of sulfate to chondroitin versus GalNAcβ1,4 GlcNAcβ1,2Manα-MCO for GalNAc-4-ST2 that has been released into the medium as compared to GalNAc-4-ST2 retained by the cell (Fig. 6)
When pcDNA3.1-GalNAc-4-ST2(Met86) is expressed in CHO/Tag cells virtually all of the activity directed at either GalNAcβ1,4 GlcNAcβ1,2Manα-MCO or chondroitin is present in the medium indicating little or no retention in the Golgi in the absence of the N-terminal transmembrane domain and stem region. Notably the GalNAc-4-ST2(Met86) released into the medium transfers sulfate to chondroitin while showing little transfer to GalNAcβ1,4 GlcNAcβ1,2Manα-MCO (Fig. 6). Furthermore, GalNAc-4-ST2 and GalNAc-4- ST2(Met86) that have been released into the medium transfer sulfate to dermatan as well as to chondroitin. Thus the relative rates of transfer to GalNAcβ1,4 GlcNAcβ1,2Manα-MCO, chondroitin, and dermatan differ significantly for GalNAc-4-ST2 in the cell (0.4:0.1:0), GalNAc-4-ST2 released into the medium (2.3:1.7:1.0), and GalNAc-4-ST2(Met86) released into the medium (0.2:4.5:1.0). This suggests that the transmembrane domain and stem region dictates the specificity for terminal GalNAc and also has an impact on recognition of chondroitin versus dermatan.

The location of the sulfate in the \[^{35}\text{S}]\text{SO}_4^-\text{chondroitin} and in \[^{35}\text{S}]\text{SO}_4^-\text{dermatan} products was determined by digestion with chondroitinase ABC and HPLC analysis (Fig. 7). Both chondroitin and dermatan yielded predominantly the sulfated disaccharide \(\Delta\text{Di-4S (D-gluco-4-enePyranosideβ1,3GalNAc-4-SO}_4\)). The same result was obtained with the product produced by human C4ST-1 expressed in CHO/Tag cells (not shown). Thus GalNAc-4-ST2 transfers sulfate predominantly to the C-4 hydroxyl of internal GalNAc moieties in both chondroitin and dermatan.

**Expression Pattern of GalNAc-4-ST2.** Array and Northern blot analyses were used to determine the expression pattern for GalNAc-4-ST2 in human tissues (Fig. 8 and Fig. 9, respectively). A strong signal was obtained for the trachea with the human Multiple Tissue Expression (MTE\(^\text{TM}\)) array system when probing with
radiolabelled GalNAc-4-ST2 cDNA (Fig. 8, 7H). Significantly weaker signals were detected in the following tissues listed in order of decreasing intensity of hybridization signal: fetal lung (G11), adult pancreas (B9), testis (F8), and salivary gland (E9) stronger than pituitary gland (D3), apex of the heart (H4), lung (A8), prostate (E8), and mammary gland (F9) stronger than heart (A4), liver (A9) and the spinal cord (E3). Even though the signals did not reproduce well and could not be quantitated by densitometry, they are considered specific since the negative controls showed no visual signal (Fig. 8, 12A-H). In addition, a specific species of mRNA could be detected by Northern blot analysis in a number of the tissues that provided a weak signal by array analysis (Fig. 9). A transcript of approximately 2.1 kb was detected by Northern blot analysis in heart, liver, and pancreas, with a significantly lower signal in lung.

**Discussion**

We and others recently reported the cloning and characterization of a GalNAc-4-O-sulfotransferase, GalNAc-4-ST1, based on its homology to HNK-1 ST and C4ST, respectively. GalNAc-4-ST1 is highly expressed in the pituitary and other regions of the brain. GalNAc-4-ST1 accounts for the addition of sulfate to terminal β1,4-linked GalNAc on the N-linked oligosaccharides of LH and other pituitary glycoproteins. We have now cloned a second, novel GalNAc-4-O-sulfotransferase, GalNAc-4-ST2, utilizing the same strategy. Differences in the specificity and expression GalNAc-4-ST2 as compared to GalNAc-4-ST1 indicate it has a distinct biologic role *in vivo*.

GalNAc-4-ST2 is the fifth member of a family of structurally related sulfotransferases that thus far include HNK-1 ST, C4ST-1, C4ST-2, GalNAc-4-ST1
and GalNAc-4-ST2. Since HNK-1 ST was the first member to be cloned we and others have used the term HNK-1 ST family to describe this family. Like all other sulfotransferases the members of HNK-1 family of sulfotransferases have two motifs that are hypothesized to mediate binding of the 5’-phosphosulfonate (5’-PSB in Fig. 2) and 3’-phosphate (3’ PB in Fig. 2) group of the high-energy donor PAPS (3’-phospho-adenosine-5’-phosphosulfonate). Three additional regions (III, IV, & V in Fig. 2) located carboxy terminal to the 5’-phosphosulfonate and 3’-phosphate binding regions also have a high percentage of identical amino acids whereas the cytosolic, transmembrane, and stem regions have few identical amino acids. GalNAc-4-ST2 and GalNAc-4-ST1 have the highest percentage of identical amino acids with 46%. In contrast, C4ST-1 and C4ST-2 have 29% identical amino acid residues. The multiple sequence alignment in Figure 2 indicates that GalNAc-4-ST2 is 23% identical to HNK-1 ST, 27% to C4ST-1, and 24% to C4ST-2.

The specificities of the members of the HNK-1 sulfotransferase family are summarized in Table II. HNK-1 ST is the only family member that transfers sulfate to the C-3 hydroxyl of terminal β1,3-linked glucuronic acid. Each of the other family members transfers sulfate to the C-4 hydroxyl of β1,4-linked GalNAc. In the case of C4ST-1 and C4ST-2 the GalNAc is found within the repeating disaccharide sequences of chondroitin and dermatan; i.e., it is not located at the non-reducing terminus. In contrast, GalNAc-4-ST1 and GalNAc-4-ST2 transfer sulfate to GalNAc moieties located at the non-reducing termini of oligosaccharide acceptors. Thus, like GalNAc-4-ST1 and GalNAc-4-ST2, HNK-1 ST transfers sulfate to a terminal β-linked sugar. The specificities of GalNAc-4-ST1 and GalNAc-4-ST2 are indicative of a structural relationship with C4ST-1 and C4ST-2. While the native, membrane-associated form of GalNAc-4-ST1 does not transfer sulfate to chondroitin, the native,
membrane-associated form of GalNAc-4-ST2 transfers sulfate to chondroitin as well as to GalNAcβ1,4 GlcNAcβ1,2Manα-MCO but at a lower rate. The truncated forms of GalNAc-4-ST1 and GalNAc-4-ST2; i.e., GalNAc-4-ST1(Met119) and GalNAc-4-ST2(Met86), both transfer sulfate to non-terminal β1,4-linked GalNAc residues in chondroitin. GalNAc-4-ST1(Met119) retains its ability to transfer sulfate to the terminal GalNAc of GalNAcβ1,4 GlcNAcβ1,2Manα-MCO whereas GalNAc-4-ST2(Met86) no longer transfers sulfate to GalNAcβ1,4 GlcNAcβ1,2Manα-MCO.

Thus, for both GalNAc-4-ST1 and GalNAc-4-ST2 the transmembrane and stem region have a major impact on the specificity for terminal versus internal β1,4-linked GalNAc. The presence of the stem and transmembrane domains markedly reduce but do not abolish transfer of sulfate to chondroitin and dermatan for GalNAc-4-ST2, whereas these regions completely abolish sulfate transfer to chondroitin by GalNAc-4-ST1. The different rates of sulfate transfer to chondroitin and dermatan seen with GalNAc-4-ST2 proteolytically released into the medium and truncated GalNAc-4-ST2(Met86) suggest that the stem region may also have an impact on the specificity for the β1,4-linked GalNAc and the adjacent uronic acid. However, more detailed analyses will be required to address these issues fully.

The five members for the HNK-1 family of sulfotransferases thus far defined are localized to five different chromosomes: HNK-1 ST to chromosome 2 (GenBank™ accession no. AC012493), C4ST-1 to chromosome 12q23, C4ST-2 to chromosome 7p22,23 and GalNAc-4-ST1 to chromosome 19q13.1,24 and GalNAc-4-ST2 to chromosome 18q11.2. The coding sequence for C4ST-2 (GenBank™ accession no. AC004840) is found within a single exon. In contrast the coding sequence of GalNAc-4-ST1 is found in 3 exons while that for GalNAc-4-ST2 is found in 5 exons. For both GalNAc-4-ST1 and GalNAc-4-ST2 the entire catalytic domain
including the 5’-phosphosulfonate binding site, the 3’-phosphate binding site, and regions III, IV and V are encoded by a single exon. The additional exons encode the cytosolic domain, the transmembrane domain, and the majority of the stem region. It is these regions that contribute to shifting the specificity away from chondroitin and dermatan to terminal GalNAcβ1,4GlcNAcβ. Thus, GalNAc-4-ST1 and GalNAc-4-ST2 may have evolved from C4ST-1 and C4ST-2 by the introduction of these addition exons which may have markedly altered the properties of the stem region. The presence of these five homologous sulfotransferases on five different chromosomes suggests that they diverged quite some time ago.

Even though GalNAc-4-ST1 and GalNAc-4-ST2 are the most closely related HNK-1 ST family members in terms of their genomic organization and protein sequence identity, the differences in their enzymatic properties and their patterns of expression indicate they have distinct functions in vivo. The high level of expression of GalNAc-4-ST1 in pituitary is consistent with its role in adding sulfate to the GalNAcβ1,4GlcNAcβ termini found on N-linked oligosaccharides of LH, TSH, POMC and other hormones. Expression of GalNAc-4-ST1 in other regions of the brain indicates that the same structures are present on other glycoproteins produced in these regions. We have, for example, found that GalNAc-4-ST1 and glycoproteins bearing sulfated N-linked oligosaccharides are abundant in the cerebellum. In contrast, Northern blots indicate that GalNAc-4-ST2 is not highly expressed in brain but is highly expressed in trachea and to lesser extent in heart, liver, pancreas, salivary gland, testis, and lung. Since GalNAc-4-ST2 can add sulfate to chondroitin as well as to GalNAcβ1,4GlcNAcβ, the actual products produced in these tissues remain to be established. In addition, it is possible that the truncated form of GalNAc-4-ST2, GalNAc-4-ST2(Met86), is expressed in specific cells and/or tissues. GalNAc-4-
ST2(Met86) is missing exon 4 (see Fig. 3) and does not contain a transmembrane
domain or the stem region due to initiation from the alternative start codon that
encodes Met86 in GalNAc-4-ST2. As a result GalNAc-4-ST2(Met86) is synthesized
and released as a soluble chondroitin and dermatan specific GalNAc-4-
sulfotransferase. Whether the truncated form of GalNAc-4-ST2 is synthesized in vivo
remains to be established; however, GalNAc-4-ST2(Met86) and GalNAc-4-ST2
would clearly have different functions.

GalNAc-4-ST1 and GalNAc-4-ST2 appear to have different expression
patterns in vivo; however, we have found that two cell lines known to produce N-
linked oligosaccharides terminating with β1,4-linked GalNAc-4-SO₄, human kidney
293 cells and human SH-SY5Y neuroblastoma cells, contain message for both
sulfotransferases. Clearly understanding the relationship of these two closely related
sulfotransferases will be important for assessing the role of the sulfated saccharides
they produce in vivo. The cloning and characterization of a second GalNAc-4-
sulfotransferase, GalNAc-4-ST2, that is closely related to GalNAc-4-ST1 but clearly
has distinct properties adds further strength to the view that the sulfated saccharides
produced have critical biologic roles.

Acknowledgements

This work was supported by National Institutes of Health Grant R01-
DK41738 to J.U.B., by Deutsche Forschungsgemeinschaft Grant SFB 470 to M.S.,
and by a German Academic Exchange Service (DAAD) postdoctoral fellowship to
G.X. The authors are grateful to Dr. M. Westphal for human glioma tissue and to Dr.
J.J. Hopwood, Women’s and Children’s Hospital, Adelaide, Australia for purified,
recombinant GalNAc-4-sulphatase, and to Jeremy Keusch for helpful suggestions.
Abbreviations

Glucuronic acid, GlcA; Iduronic acid, IdA; lutropin, LH; pro-opiomelanocortin, POMC; thyrotropin, TSH; HNK-1 sulfotransferase, HNK-1 ST; chondroitin 4-O-sulfotransferase, C4ST; N-acetylgalactosamine 4-O-sulfotransferase, GalNAc-4-ST; (CH$_2$)$_8$-COOCH$_3$. MCO; bp, base pair; PAPS, 3’-phospho-adenosine-5’-phosphosulfonate; GlcNAcβ-R., R = underlying saccharide structure.

References

11. Kruse, J., Mailhammer, R., Wernecke, H., Faissner, A., Sommer, I., Goridis, C.,


42. Woodworth, A., Mi, Y. L., and Baenziger, J. U. Unpublished observation.

Table I. The GalNAc-4-ST2 transfers sulfate to terminal $\beta$1,4-linked GalNAc. CHO/Tag cells were transfected with pcDNA3.1-GalNAc-4-ST2. After 60 h in culture cells were harvested and solubilized in 200 $\mu$l of 1% Triton X-100. The cell extract (10 $\mu$l) was tested for transfer of $[^{35}\text{S}]\text{SO}_4$ ($10^6$ cpm/nmol $[^{35}\text{S}]\text{PAPS}$) to the acceptors shown. Results are expressed as pmol incorporated/h/100 mm plate.

<table>
<thead>
<tr>
<th>Acceptor</th>
<th>GalNAc-4-ST2 (pmol/h/plate)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>20$\mu$M</td>
</tr>
<tr>
<td>GlcNAc$\beta$1,2Man$\alpha$-MCO</td>
<td>0.11</td>
</tr>
<tr>
<td>GalNAc$\beta$1,4 GlcNAc$\beta$1,2Man$\alpha$-MCO</td>
<td>10.47</td>
</tr>
<tr>
<td>GalNAc$\beta$1,4 GlcNAc$\beta$-MCO</td>
<td>5.46</td>
</tr>
<tr>
<td>Gal$\beta$1,4 GlcNAc$\beta$1,2Man$\alpha$1,6Man$\beta$-MCO</td>
<td>0.18</td>
</tr>
</tbody>
</table>
Table II. Specificities of the HNK-1 ST family members. The ability of the various members of the HNK-1 family of sulfotransferases to modify the acceptors shown is indicated by a +, - , or +/- for active, inactive, or weakly active, respectively. The saccharide that is modified with sulfate is in each case shown in bold letters. GalNAc-4-ST1 and GalNAc-4-ST2 indicate the native, membrane associated, intracellular forms.

<table>
<thead>
<tr>
<th>Sulfotransferases</th>
<th>Acceptors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(GlcAβ1,3GalNAcβ1,4)</td>
</tr>
<tr>
<td></td>
<td>(IdAα1,3GalNAcβ1,4)</td>
</tr>
<tr>
<td>C4ST-1</td>
<td>+</td>
</tr>
<tr>
<td>C4ST-2</td>
<td>+</td>
</tr>
<tr>
<td>GalNAc-4-ST1</td>
<td>-</td>
</tr>
<tr>
<td>GalNAc-4-ST1(Met119)</td>
<td>+</td>
</tr>
<tr>
<td>GalNAc-4-ST2</td>
<td>+/-</td>
</tr>
<tr>
<td>GalNAc-4-ST2(Met86)</td>
<td>+</td>
</tr>
<tr>
<td>HNK-1 ST</td>
<td>-</td>
</tr>
</tbody>
</table>
Figure Legends

Figure 1. **Nucleotide and deduced amino acid sequence of human GalNAc-4-ST2 cDNA (Genbank™ accession No. AF332472).** Panel A, the predicted amino acid sequence of GalNAc-4-ST2 is denoted by capital letters below the nucleotide sequence. The single membrane spanning domain predicted by the hidden Markov model based algorithm TMHMM \(^{34}\) and four potential N-linked glycosylation sites are indicated by the bold underline and by the underline with a ● below the glycosylated Asn, respectively. The truncated form of GalNAc-4-ST2, GalNAc-4-ST2(Met86), is initiated at the AUG (double underline) encoding Met86 (double underline) of GalNAc-4-ST2. The nucleotide sequence derived from exon 4 is shown in lower case letters. **Panel B,** shows the hydrophilicity plot for GalNAc-4-ST2.

Figure 2. **Comparison of amino acid sequences of human GalNAc-4-ST2, GalNAc-4-ST1, C4ST-1, C4ST-2, and HNK-1 ST.** Alignment was performed using the ClustalW program implemented in the BioEdit suite \(^{35}\). Introduced gaps are shown as hyphens and aligned amino acids are boxed (black for identical residues and dark grey for similar residues). Putative binding sites for the 5’-phosphosulfonate group (5’-PSB) and 3’-phosphate group (3’-PB) of PAPS, and three additional highly conserved domains (III, IV, & V) are marked.

Figure 3. **Structure of the human GalNAc-4 ST2 gene.** Panel A, exons that contribute to the 5’ and 3’ UTR are boxed. Exonic sequences that are transcribed into the coding sequence are boxed and shaded in black. Exons are numbered from 1 (E1)
to 5 (E5). Relevant genomic clones are denoted with their GenBank™ accession numbers below the schematic. The splicing pattern for GalNAc-4-ST2 is shown above and for GalNAc-4-ST2(Met86) below the schematic. Note that I.M.A.G.E. cDNA clone 49547 only contains exonic sequences from the position marked “a” to the position marked “b”. **Panel B**, mRNA and the translated GalNAc-4-ST2 protein are shown. Nucleotide positions are denoted below the mRNA scheme. TM denotes the transmembrane domain. The AUG in brackets marks the start codon for GalNAc-4-ST2(Met86) that would be generated if exon 4 is absent. **Panel C**, the intron-exon boundaries are shown for the human GalNAc-4-ST2 gene. Exon derived sequences are capitalized. Note that all boundaries have an invariant intron derived GT and AG at the donor and acceptor splice sites.

Figure 4. **GalNAc-4-ST2 transfers sulfate exclusively to C-4 of the β1,4-linked GalNAc.** GalNAcβ1,4 GlcNAcβ1,2Manα-MCO was incubated with [35S]PAPS and GalNAc-4-ST2 and the [35S]SO₄-labelled product isolated by passage over a Sep-Pak C₁₈. **Panel A**, analysis of the [35S]SO₄-GalNAcβ1,4 GlcNAcβ1,2Manα-MCO product on a Carbopac PA1 column. **Panel B**, analysis of the [35S]SO₄- GalNAcβ1,4 GlcNAcβ1,2Manα-MCO product on Carbopac PA1 following mild acid hydrolysis. The elution positions of standards are indicated by the numbers: 1, GlcNAc-3-SO₄; 2, SO₄; 3, GalNAc-3-SO₄; 4, SO₄-3-GalNAcβ1,4 GlcNAcβ1,2Manα-MCO; 5, GalNAc-4-SO₄; 6, SO₄-4-GalNAcβ1,4 GlcNAcβ1,2Manα-MCO; 7, GalNAc-6-SO₄
Figure 5. **GalNAc-4-ST2 transfers sulfate to N-linked oligosaccharides terminating with the sequence GalNAcβ 1,4GlcNAcβ 1,2Manα on glycoproteins.**

GalNAc-4-ST2 was incubated with 3 µg each of bovine CA-VI (from parotid gland), LH, or desulfated LH (treated with GalNAc-4-sulfatase) and [35S]PAPS. An equal amount of each reaction was digested with PNGase F to release N-linked oligosaccharides. No sulfated products were seen in control lanes containing hCG, asialo-hCG, or media from cells transfected with pcDNA3.1 with no cDNA (not shown). Lane 1, CA-VI; Lane 2, CA-VI + PNGase F; Lane 3, LH; Lane 4, LH + PNGase F; Lane 5, desulfated LH; Lane 6, desulfated LH + PNGase F.

Figure 6. **GalNAc-4-ST2 transfers sulfate to GalNAcβ 1,4GlcNAcβ 1,2Manα, chondroitin, and dermatan.** CHO/Tag cells were transfected with pcDNA3.1-GalNAc-4-ST2, pcDNA3.1-GalNAc-4-ST2(Met86), or pcDNA3.1. After 60 h in culture cells were harvested and solublized in 200 µl of 1% Triton X-100. The medium was collected and concentrated 10 fold to 2.0 ml by ultrafiltration. Each assay utilized 10 µl of either the cell extract or concentrated medium. Incorporation is reported as pmol/h/100-mm plate. Basal levels of incorporation have been substracted. The transferase reactions were carried out with 20µM GalNAcβ1,4GlcNAcβ 1,2Manα-MCO (Grey bar), 50 µg chondroitin (striped bar), or 50 µg dermatan (Dark bar) in a 50µl reaction containing [35S]PAPS (1 x 10^6 cpm/nmole) for 16 h. the products were separated as described in the experimental procedures.

Figure 7. **GalNAc-4-ST2 transfers sulfate to the C-4 hydroxyl of internal β1,4 linked GalNAc residues in chondroitin.** Panel A, Chondroitin (50 µg) and
Panel B, dermatan (50 µg) were incubated with GalNAc-4-ST2 and [35S]PAPS. The labelled product was separated from free label by gel filtration and digested with 30 milliunits of Chondroitinase ABC in 100 mM Tris-Acetate buffer, pH 8.0 for 6 h at 37°C. The product was analyzed on a Micropak AX-5 column (Varian Assoc.) developed with a gradient of 10-450 mM KH₂PO₄ over 40 min. at a flow rate of 1.0 ml/min. The elution times for for authentic standards are indicated by the numbers: 1. GalNAc-4-SO₄, 2. ΔDi6 (D-gluco-4-enepyranosideβ1,3GalNAc-6-SO₄), 3. ΔDi4 (D-gluco-4-enepyranosideβ1,3GalNAc-4-SO₄), 4. ΔDi-diSₐE (D-gluco-4-enepyranosideβ1,3GalNAc-4,6-diSO₄) and 5. free SO₄.

Figure 8. RNA dot blot analysis of GalNAc-4-ST2 transcripts. The human Multiple Tissue Expression (MTE™) array shown was hybridized with a 32P-labelled human GalNAc-4-ST2 specific cDNA probe (see Experimental Procedures). Tissue sources for the RNA are indicated below the blot. Asterisks indicate: * - paracentral gyrus of cerebral cortex, ** - peripheral blood leukocytes, *** - Burkitt’s lymphoma Raji, **** - Burkitt’s lymphoma Daudi, ***** - colorectal adenocarcinoma, SW280.

Figure 9. Northern blot analysis of GalNAc-4-ST2 transcripts. Each lane of the MTN® Northern blot contains 2 µg poly (A)⁺ RNA and was hybridized with a 32P-labelled cDNA probe specific for human GalNAc-4-ST2 (see Experimental Procedures). Tissues used to prepare the RNA are indicated above each lane. Migration positions of standards are indicated at the left.
Figure 1

<table>
<thead>
<tr>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>TATGGAATTGGGCTCTTAGAGAACAAGAAAAGACTGAAGTTTTACGGGAAAACAAATCATGTGGTCTTCAGATTCTGAAATAAGGAGAAATGCAGCCATCTGAAATGGTCATGAACCCCAAACAAGTCTTCCTCTCTGTGCTGATATTT</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>MM MM</td>
</tr>
<tr>
<td>Q  P  S  E  M  V  M  N  P  K  Q  V  F  L  S  V  L  I  F</td>
</tr>
<tr>
<td>20</td>
</tr>
<tr>
<td>GGAGTAGCTGGGCTACTCCTCTTCATGTATTTGCAAGTCTGGATTGAAGAACAACATACAGGGAGAGTGGAGAAGAGAAGAGAACAAAAAGTAACTTCAGGATGGGGACCAGTGAAGTACTTGCGGCCTGTACCCAGAATCATGAATCAGTACA</td>
</tr>
<tr>
<td>150</td>
</tr>
<tr>
<td>300</td>
</tr>
<tr>
<td>gaaaaaatccaggaacatatcaccaaccagAACCCCAAGTTTCACATGACCTGAGGATGTACGAGAAAAAAAGGAAAATCTTCTACTCAATTCTGAGAGATCTACTAGGCTCTTAACAAAGACCAGTCATTCACAAGGAGGGGATCAAGCT</td>
</tr>
<tr>
<td>450</td>
</tr>
<tr>
<td>E  K  I  Q  E  H  I  T  N  Q  N  P  K  F  H  M  P  E  D  V  R  E  K  K  E  N  L  L  N  S  E  R  S  T  R  L  L  T  K  T  S  H  S  Q  G  G  D  Q  A</td>
</tr>
<tr>
<td>600</td>
</tr>
<tr>
<td>AATACTTACACCAAAGCTGTGTTTGTTCGTGATCCCATGGAAAGATTAGTATCAGCCTTTAGGGACAAATTTGAACACCCCAATAGTTATTACCATCCAGTATTCGGAAAGGCAATTATCAAGAAATATCGACCAAATGCCTGTGAAGAA</td>
</tr>
<tr>
<td>900</td>
</tr>
<tr>
<td>1200</td>
</tr>
<tr>
<td>GCATTAATTAATGGATCTGGAGTCAAGTTCAAAGAGTTTATCCACTACTTGCTGGATTCCCACCGTCCAGTAGGAATGGACATTCACTGGGAAAAGGTCAGCAAACTCTGCTATCCGTGTTTGATCAACTATGATTTTGTAGGGAAATTT</td>
</tr>
<tr>
<td>1500</td>
</tr>
<tr>
<td>1800</td>
</tr>
<tr>
<td>TGTATGACAGAAATTTAACCAAGTGCAGTTGTCTTGATTTAATGTAGATTTTTACCAAATAGTATGACACCAATTGGCACAAAGTTATAGGAAAATCACCTACAGGAGATGTAAACAACTTGAGTTGCTCTAAAATGTTTGGAAAAGAGC</td>
</tr>
<tr>
<td>2100</td>
</tr>
<tr>
<td>TGCTTTTGCATTATGAATTATATTGTTGAAGCAATAACCTAGCCAGCTGTTGCATTAGCTAAAGCAGCCTCTTGCAATGGTAGGAAAAAAGGATCTCAAATAGCATGAGTGTATGTCTATATCCTGAAATTTATTGTCTAAAATGCATGA</td>
</tr>
<tr>
<td>2400</td>
</tr>
<tr>
<td>...</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Hydrophobicity</td>
</tr>
<tr>
<td>Position</td>
</tr>
<tr>
<td>20</td>
</tr>
<tr>
<td>30</td>
</tr>
<tr>
<td>40</td>
</tr>
<tr>
<td>50</td>
</tr>
<tr>
<td>60</td>
</tr>
<tr>
<td>70</td>
</tr>
<tr>
<td>80</td>
</tr>
<tr>
<td>90</td>
</tr>
<tr>
<td>100</td>
</tr>
<tr>
<td>110</td>
</tr>
<tr>
<td>120</td>
</tr>
<tr>
<td>130</td>
</tr>
<tr>
<td>140</td>
</tr>
<tr>
<td>150</td>
</tr>
</tbody>
</table>

[Graph showing mean hydrophobicity over position]
Figure 2

GalNAc-4 ST 2
1. MOPSE-MYNPKQVSLVHCMAGLIELELYMYO-VWIEEWHATGKVEKREQYKTSVGSKEKLYRPRPRIMSTEIKGIEHITNONPK-FHHPDVKYK 96
2. MTLRP-GTNRLACGGSLSAGAALLESLDPOETAPQPGKFYRPROPHODPLGGSDODGKLEPTVERTLSAGQPRGRNPAPDOPOP 98
3. MSHQALLACCGTVEIKHVASLTTLFQATF 43
4. MKPALLEVMRNRMCRMPATELSFLTYF 45
5. FARKALQLWEGSVFMIVYWDSDGAAGAHFHYLTSFSPHTGPPLPT 72

GalNAc-4 ST 1
6. LLLNSERSRTLLSHSQQDOSSKSTSGPTKLIEKRGQAKTVFNKFSNMWVPDSLNLKSLKKNWKKEETDEKRSFLGEG KYG 191
7. LO-GLLKLRRR-RRLL1KKMPAAATIPANSDDAFPRPSLTLCAWSLHRDNGVRVMDCAK KYRA 169
8. LTMPRVRKLP-EEEKHIPELKTQQ-KEI-TLELQVPLYMEKELIWNDDAL 97
9. FGQVDDCRKGS-RS-PQDPYMNRS-IO-LELSNTAIVLHIGNQ10-D440 94
10. FLGAVKGDHIKETEOP-PAPGMSYESV-YDGK-RPSPDGRGMAECQVGLHNNSLAF 141

HNK-1 ST
11. MHHQWLLLAACFWVIFMFMVASKFITLTFK-------------------------DP----------------------------DVYSAKQEFLF 43
12. ------MHHQWLLLAACFWVIFMFMVASKFITLTFK-------------------------DP----------------------------DVYSAKQEFLF 43
13. ------LTTMPEVRKLP------------------EEK---HIPEELKPTG------KELP-------DSQLVQPLVYMER--LELIRNVCR----DDAL 97
14. ------LTTMPEVRKLP------------------EEK---HIPEELKPTG------KELP-------DSQLVQPLVYMER--LELIRNVCR----DDAL 97
15. ------LTTMPEVRKLP------------------EEK---HIPEELKPTG------KELP-------DSQLVQPLVYMER--LELIRNVCR----DDAL 97
16. ------LTTMPEVRKLP------------------EEK---HIPEELKPTG------KELP-------DSQLVQPLVYMER--LELIRNVCR----DDAL 97

C4ST-1
17. MKPALLEVMRMNRICRMVLATCLGSFILVIFYF------------------------------------------------------QSMLHPVMRNP 45
18. ------FMGIDCRKGS-RR----PLQELYNPS-IO-LELSNTAVLHIGNQ10-D440 94
19. ------FMGIDCRKGS-RR----PLQELYNPS-IO-LELSNTAVLHIGNQ10-D440 94
20. ------FMGIDCRKGS-RR----PLQELYNPS-IO-LELSNTAVLHIGNQ10-D440 94

C4ST-2
21. --------MTKARLFRLWLVLGSVFMILLIIVYWDSAGAAHFHYLTSFSPHTGPPLPT 72
22. FLSAGVKGDFPKETE-OP-PAPGMSYESV-YDGK-RPSPDGRGMAECQVGLHNNSLAF 141
23. FLSAGVKGDFPKETE-OP-PAPGMSYESV-YDGK-RPSPDGRGMAECQVGLHNNSLAF 141
24. FLSAGVKGDFPKETE-OP-PAPGMSYESV-YDGK-RPSPDGRGMAECQVGLHNNSLAF 141

5'-PSB
25. LLYDFVGKFETLEDDANYFLQMSAGKELKFPNFKDRHSDERTNAQVVRQYLKDLTRTERQLIYDFYYLDYLMFNYTTP-FL--- 443
26. LIDYDFVGKFESMEDDANFFLSLIRAPRNLTFPRFKDRHSAEARTTARIAHQYFAQLSALQRQRTYDFYYMDYLMFNYSKP-FADLY 424
27. EIMYSVIGHHETLEDDAPYILKEAGIDHLVSYP---TIPPGITVYNRTKVEHYFLGISKRDIRRLYARFEGDFKLFGYQKPDFLLN- 356
28. HIHYDLVGKYETLEEDSNYFLQLAGVGSYLKFP----TYAKSTRTTDEMTTEFFQNISSEHQTQLYEVYKLDFLMFNYSVPSYLKLE 352
29. QIDYDFVGKLETLDEDAAQLLQLLQVDRQLP----PSYRRTASSWEEDWFAKIPLAWRQQLYKLYEADFVLFGYPKVAEMLDCW 333

3'-PB
30. KAVFVRDPMERLVSAFRDKFEHP---NSYYHPVFGKAILARYR-----ANASREALRTGSGVRFPEFVQYLLDVH------RPVGMDTHWDHVSRLCSPC 338
31. KMLFVREPFERLVSAYRNKFTQKY--NISFHKRYGTKIIKRQR-----KNATQEALRKGDDVKFEEFVAYLIDPHT----QREEPFNEHWQTVYSLCHPC 269
32. KFLFVREPFERLVSAYRNKFTQKY--NISFHKRYGTKIIKRQR-----KNATQEALRKGDDVKFEEFVAYLIDPHT----QREEPFNEHWQTVYSLCHPC 269
33. KFLFVREPFERLVSAYRNKFTQKY--NISFHKRYGTKIIKRQR-----KNATQEALRKGDDVKFEEFVAYLIDPHT----QREEPFNEHWQTVYSLCHPC 269

11
34. LSYDGVKGDFELEDFYTOMIGAKPELFFNFKDHRSSQDNQNYVRDDKDLTRTERLOFDLYLSFP--- 443
35. LSYDGVKGDFELEDFYTOMIGAKPELFFNFKDHRSSQDNQNYVRDDKDLTRTERLOFDLYLSFP--- 443
36. LSYDGVKGDFELEDFYTOMIGAKPELFFNFKDHRSSQDNQNYVRDDKDLTRTERLOFDLYLSFP--- 443
37. LSYDGVKGDFELEDFYTOMIGAKPELFFNFKDHRSSQDNQNYVRDDKDLTRTERLOFDLYLSFP--- 443
38. LSYDGVKGDFELEDFYTOMIGAKPELFFNFKDHRSSQDNQNYVRDDKDLTRTERLOFDLYLSFP--- 443
39. LSYDGVKGDFELEDFYTOMIGAKPELFFNFKDHRSSQDNQNYVRDDKDLTRTERLOFDLYLSFP--- 443
40. LSYDGVKGDFELEDFYTOMIGAKPELFFNFKDHRSSQDNQNYVRDDKDLTRTERLOFDLYLSFP--- 443
41. LSYDGVKGDFELEDFYTOMIGAKPELFFNFKDHRSSQDNQNYVRDDKDLTRTERLOFDLYLSFP--- 443
42. LSYDGVKGDFELEDFYTOMIGAKPELFFNFKDHRSSQDNQNYVRDDKDLTRTERLOFDLYLSFP--- 443
43. LSYDGVKGDFELEDFYTOMIGAKPELFFNFKDHRSSQDNQNYVRDDKDLTRTERLOFDLYLSFP--- 443
Figure 3

A

B

C

<table>
<thead>
<tr>
<th>intron</th>
<th>exon</th>
<th>3' splice site</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>...AAGAACAACATACAG gtaacacatgtgctt</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>...GGAGGTTGGAAAGA ...AAAAAGTAACTTCAG gtaagttttgtcca</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>...GTGGGAGCCAGAAGG ...CTGTACCCAGAATCA gtaglaacclccc</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>...TGATACAGAAAAAA ...CATATACAAAAGC gttttctctctcttccc</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>...AACCCAAAGTTTCAC ... gttttatttttacag</td>
</tr>
</tbody>
</table>

5' splice site

3' splice site

100 bp
Figure 4

Retention Time (min.)

Radioactivity (cpm)

A

B

Retention Time (min.)
Figure 6

![Graph](image-url)
<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 3 4 5 6 7 8 9 10 11 12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>H</td>
</tr>
<tr>
<td>1 brain, whole</td>
<td>2 cerebral cortex</td>
<td>3 frontal lobe</td>
<td>4 parietal lobe</td>
<td>5 occipital lobe</td>
<td>6 temporal lobe</td>
<td>7 pons</td>
<td>8 pons</td>
</tr>
<tr>
<td>2 cerebellum, left</td>
<td>3 cerebellum, right</td>
<td>4 thalamus</td>
<td>5 amygdala</td>
<td>6 medulla oblongata</td>
<td>7 medulla oblongata</td>
<td>8 medulla oblongata</td>
<td>9 medulla oblongata</td>
</tr>
<tr>
<td>3 substantia nigra</td>
<td>4 substantia nigra</td>
<td>5 substantia nigra</td>
<td>6 substantia nigra</td>
<td>7 substantia nigra</td>
<td>8 substantia nigra</td>
<td>9 substantia nigra</td>
<td>10 substantia nigra</td>
</tr>
<tr>
<td>4 heart</td>
<td>5 heart</td>
<td>6 heart</td>
<td>7 heart</td>
<td>8 heart</td>
<td>9 heart</td>
<td>10 heart</td>
<td>11 heart</td>
</tr>
<tr>
<td>5 esophagus</td>
<td>6 esophagus</td>
<td>7 esophagus</td>
<td>8 esophagus</td>
<td>9 esophagus</td>
<td>10 esophagus</td>
<td>11 esophagus</td>
<td>12 esophagus</td>
</tr>
<tr>
<td>6 colon, descending</td>
<td>7 colon, descending</td>
<td>8 colon, descending</td>
<td>9 colon, descending</td>
<td>10 colon, descending</td>
<td>11 colon, descending</td>
<td>12 colon, descending</td>
<td>13 colon, descending</td>
</tr>
<tr>
<td>7 kidney</td>
<td>8 kidney</td>
<td>9 kidney</td>
<td>10 kidney</td>
<td>11 kidney</td>
<td>12 kidney</td>
<td>13 kidney</td>
<td>14 kidney</td>
</tr>
<tr>
<td>8 liver</td>
<td>9 liver</td>
<td>10 liver</td>
<td>11 liver</td>
<td>12 liver</td>
<td>13 liver</td>
<td>14 liver</td>
<td>15 liver</td>
</tr>
<tr>
<td>A1</td>
<td>B1</td>
<td>C1</td>
<td>D1</td>
<td>E1</td>
<td>F1</td>
<td>G1</td>
<td>H1</td>
</tr>
<tr>
<td>1 brain, whole</td>
<td>2 cerebral cortex</td>
<td>3 frontal lobe</td>
<td>4 parietal lobe</td>
<td>5 occipital lobe</td>
<td>6 temporal lobe</td>
<td>7 pons</td>
<td>8 pons</td>
</tr>
<tr>
<td>2 cerebellum, left</td>
<td>3 cerebellum, right</td>
<td>4 thalamus</td>
<td>5 amygdala</td>
<td>6 medulla oblongata</td>
<td>7 medulla oblongata</td>
<td>8 medulla oblongata</td>
<td>9 medulla oblongata</td>
</tr>
<tr>
<td>3 substantia nigra</td>
<td>4 substantia nigra</td>
<td>5 substantia nigra</td>
<td>6 substantia nigra</td>
<td>7 substantia nigra</td>
<td>8 substantia nigra</td>
<td>9 substantia nigra</td>
<td>10 substantia nigra</td>
</tr>
<tr>
<td>4 heart</td>
<td>5 heart</td>
<td>6 heart</td>
<td>7 heart</td>
<td>8 heart</td>
<td>9 heart</td>
<td>10 heart</td>
<td>11 heart</td>
</tr>
<tr>
<td>5 esophagus</td>
<td>6 esophagus</td>
<td>7 esophagus</td>
<td>8 esophagus</td>
<td>9 esophagus</td>
<td>10 esophagus</td>
<td>11 esophagus</td>
<td>12 esophagus</td>
</tr>
<tr>
<td>6 colon, descending</td>
<td>7 colon, descending</td>
<td>8 colon, descending</td>
<td>9 colon, descending</td>
<td>10 colon, descending</td>
<td>11 colon, descending</td>
<td>12 colon, descending</td>
<td>13 colon, descending</td>
</tr>
<tr>
<td>7 kidney</td>
<td>8 kidney</td>
<td>9 kidney</td>
<td>10 kidney</td>
<td>11 kidney</td>
<td>12 kidney</td>
<td>13 kidney</td>
<td>14 kidney</td>
</tr>
<tr>
<td>8 liver</td>
<td>9 liver</td>
<td>10 liver</td>
<td>11 liver</td>
<td>12 liver</td>
<td>13 liver</td>
<td>14 liver</td>
<td>15 liver</td>
</tr>
</tbody>
</table>

**Figure 8**
Molecular cloning and expression of an N-Acetylgalactosamine-4-O-Sulfotransferase that transfers Sulfate to Terminal and Non-Terminal beta1,4-linked N-Acetylgalactosamine

Hyung-Gyoo Kang, Matthias R. Evers, Guoqing Xia, Jacques U. Baenziger and Melitta Schachner

*J. Biol. Chem.* published online January 3, 2001

Access the most updated version of this article at doi: 10.1074/jbc.M011560200

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 0 references, 0 of which can be accessed free at http://www.jbc.org/content/early/2001/01/03/jbc.M011560200.citation.full.html#ref-list-1