We have identified and characterized an N-acetylgalactosamine-4-O-sulfotransferase designated chondroitin-4-sulfotransferase-3 (C4ST-3) (GenBank accession number AY120869) 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 341 amino acids with a 12-amino acid cytoplasmic domain and a 311-amino acid luminal domain containing a single potential N-linked glycosylation site. C4ST-3 has the greatest amino acid sequence identity when aligned with chondroitin-4-O-sulfotransferase 1 (C4ST-1) (45%) but also shows significant amino acid identity with chondroitin-4-O-sulfotransferase 2 (C4ST-2) (27%), dermatan-4-O-sulfotransferase 1 (29%), HNK-1 ST (26%), N-acetylgalactosamine-4-O-sulfotransferase 1 (26%), and N-acetylgalactosamine-4-O-sulfotransferase 2 (23%). C4ST-3 transfers sulfate to the C-4 hydroxyl of beta,1-linked GalNAc that is substituted with a beta-linked glucuronic acid at the C-3 hydroxyl. The open reading frame of C4ST-3 is encoded by three exons located on human chromosome 3q21.3. Northern blot analysis reveals a single 2.1-kilobase transcript. C4ST-3 message is expressed in adult liver and at lower levels in adult kidney, lymph nodes, and fetal liver. Although C4ST-3 and C4ST-1 have similar specificities, the highly restricted pattern of expression seen for C4ST-3 suggests that it has a different role than C4ST-1.

We and others have recently reported the cloning and functional characterization of members of the HNK-1 family of sulfotransferases. These include human HNK-1 sulfotransferase (HNK-1 ST) itself (1, 2), GalNAc-4-ST1 (3–5), GalNAc-4-ST2 (6), dermatan-4-sulfotransferase-1 (D4ST-1) (7), chondroitin-4-sulfotransferase-1 (C4ST-1) (8, 9), and chondroitin-4-sulfotransferase-2 (C4ST-2) (8). With the exception of HNK-1 ST itself, which transfers sulfate to the C-3 hydroxyl of terminal glucuronic acid in the sequence GlcUA(1,3)Gal(1,4)GlcNAc-R to produce the HNK-1 epitope SO4-3-GlcUA, C4ST-1 transfers sulfate to the C-4 hydroxyl of terminal beta,1-linked GalNAc on N-linked oligosaccharides such as those found on the glycoprotein hormones lutropin and thyrotropin (10, 11). Whereas C4ST-1, C4ST-2, and D4ST-1 are also GalNAc-4-O-sulfotransferases, they only transfer sulfate to the C-4 hydroxyl of internal beta,1-linked GalNAc moieties within the repeating disaccharide sequences found in chondroitin and dermatan (7–9). We now report the cloning of another member of this family of sulfotransferases. Like C4ST-1, this new sulfotransferase, chondroitin-4-sulfotransferase-3 (C4ST-3) transfers sulfate to the C-4 hydroxyl of beta,1-linked GalNAc that is flanked by GlcUA residues in chondroitin. In contrast to C4ST-1, C4ST-3 is labile at 37 °C and has a restricted distribution, suggesting that it may have a unique biological role in vivo.

EXPERIMENTAL PROCEDURES

Molecular Cloning of a cDNA Encoding Human Chondroitin-4-O-sulfotransferase-3—A human genomic BAC clone, RP11-390G14, derived from human chromosome 3 (GenBank accession number AC024558) was identified in TBLASTN searches (12) against the nonredundant data base at the NCBI using deduced protein sequences of human and rat HNK-1 sulfotransferases (1, 2) as query sequences. The putative open reading frame (ORF) encodes a protein that shows homology to the C4ST-1 family of sulfotransferases (8, 9). Subsequent BLASTN queries of the dbEST data set using this region of homology identified two matching EST sequences (GenBank accession numbers BF448098 and AI074148, respectively). The corresponding cDNA clones (IMAGp998H017603 and IMAGp998M214153, respectively) were obtained from the RZPD (Berlin, Germany) (13) and sequenced on both strands. The partial ORF of C4ST-3, encoding most of its luminal domain, was amplified by PCR. The 5′-specific primer, 5′-ctg gga acg cgg cca cca cta gga gaa gaa ggg ccc tct-3′, contains a HindIII restriction enzyme site, the Kozak consensus sequence GGGAC, and an artificial start codon. The 3′-specific primer, 5′-gct tca gac tag age cgg aag gcc gaa gac ggc ccc tct-3′, contains an XbaI site and a stop codon. The product was directionally subcloned into pcDNA3.1 (Invitrogen, Karlsruhe, Germany) and designated pcDNA3.1-C4ST-3-ORF309. Subsequent BLASTN queries of the dbEST using the partial ORF of C4ST-3 identified an EST (GenBank accession number B1908522) that overlaps the NH2 terminus of the sequence used to construct frame; PAPS, 3′-phosphoadenosine-5′-phosphosulfate; UTR, untranslated region; EST, expressed sequence tag; NTA, nitrilotriacetic acid; CHO, Chinese hamster ovary.
Chondroitin-4-sulfotransferase-3

**Fig. 1.** Nucleotide and deduced amino acid sequence of human C4ST-3 cDNA (GenBank™ accession number AY120869). A, the predicted amino acid sequence of C4ST-3 is denoted by capital letters below the nucleotide sequence. The single predicted membrane-spanning domain and a single potential N-linked glycosylation site are indicated by the *thick* underlines and by the *underlines* with a *black dot* below the glycosylated Asn, respectively. B, Kyte-Doolittle mean hydrophobicity plot for C4ST-3 (scan window size was 13 amino acids).

pcDNA3.1-C4ST-2-ORF309. B1908522 was used to evaluate a segment of the working draft of chromosome 3 (GenBank™ accession number NT_005586; also see below) for the presence of exons predicted by the program FGENES (version 1.6). Based on genomic sequence information, the start codon of C4ST-3 was predicted to be ~40 bp to the 5′-end of the genomic sequence represented by B1908522. The 5′-terminal part of the C4ST-3 ORF was amplified by PCR from human placenta first-strand cDNA (generated with Omniscript Reverse Transcriptase; Qiagen, Hilden, Germany) using the 5′-specific primer 5′-ttg atc tcg gcg ggg g-3′ and the 3′-specific primer 5′-act ccg ccc cct cct acc-3′. The 482-bp product was subcloned into the pGEM KS vector (Stratagene, Amsterdam, The Netherlands) by sequentially subcloning an 850-bp PCR fragment from the pGEM®-T Easy vector (Promega) and sequenced on both strands. Finally, the full-length ORF was assembled in the pBluescript® KS vector (Stratagene, Amsterdam, The Netherlands) by sequentially subcloning an 850-bp *Pst*I fragment derived from pcDNA3.1-C4ST-3-ORF310 and an additional 170-bp *HindIII*-Specific fragment from the pGEM®-T construct above to produce pBKS-C4ST-3-ORF341. The 1026-bp PCR fragment was directionally subcloned into the eukaryotic expression vector pcDNA3.1 (Invitrogen), completely sequenced on both strands, and designated pcDNA3.1-C4ST-3.

**Construction of pcDNA3.1-C4ST-3**—The ORF of C4ST-3 was amplified from pBKS-C4ST-3-ORF341 (see above) by PCR using 1) the 5′-specific primer 5′-atg ggg agg cgc-3′ and 2) the 3′-specific primer 5′-gct cta gac tag agc agc-3′. The 1026-bp PCR fragment was directionally subcloned into the eukaryotic expression vector pcDNA3.1 (Invitrogen), completely sequenced on both strands, and designated pcDNA3.1-C4ST-3.

**Genomic Organization and Chromosomal Localization of C4ST-3**—The cDNA sequences of C4ST-3 and C4ST-1 (GenBank™ accession numbers AA677272, A1041547, AA857547, and BI464064, respectively) using the Lasergene (DNASTAR Inc., Madison, WI) software suite. The deduced protein sequences of all members of the HNK-1 family of sulfotransferases were analyzed by multiple alignments using the ClustalW (version 1.4) algorithm (14) implemented in the Bioedit suite. The deduced protein sequences of all members of the HNK-1 family of sulfotransferases were analyzed by multiple alignments using the ClustalW (version 1.4) algorithm (14) implemented in the Bioedit suite. The deduced protein sequences of all members of the HNK-1 family of sulfotransferases were analyzed by multiple alignments using the ClustalW (version 1.4) algorithm (14) implemented in the Bioedit suite. The deduced protein sequences of all members of the HNK-1 family of sulfotransferases were analyzed by multiple alignments using the ClustalW (version 1.4) algorithm (14) implemented in the Bioedit suite.
Buffer, pH 7.4, 5 mM MgCl₂, 175 mM KCl, 2% Triton X-100 protease inhibitors (23 milli-trypsin inhibitor units of aprotinin and 4 milli-trypsin inhibitor units 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 g for 20 min. The supernatant was designated as the cell extract. The culture medium was pooled and sedimented at 12,000 g for 20 min. The culture supernatant was adjusted to a final concentration of 20 mM HEPES, pH 7.4, and protease inhibitors were added as noted above.

Sulfotransferase activities were assayed as described (3, 8). Each reaction (50 µl) contained 50 mM imidazole-HCl, pH 6.8, 2 mM dithiothreitol, 5 × 10⁶ cpm [³²P]PAPS, 2 µM PAPS, and protamine (0.005% (w/v) for dermatan), and protease inhibitors (see solubilization buffer above). Desulfated chondroitin (Seikagaku America, Inc.) and desulfated dermatan from porcine intestine (Sigma) (18) were used as acceptors. Incorporation was determined to be linear for 2–3 h under these conditions. Assays were carried out for 2 h at 25 °C and terminated by boiling for 3 min. The [³²P]SO₄-labeled chondroitin and dermatan products were precipitated by the addition of 0.1 volume of 4 M potassium acetate and 3 volumes of ethanol. After sedimentation, the precipitates were dissolved in 140 µl of H₂O and separated from any remaining [³²P]SOPs and degradation products by passage over Bio-Spin 6 columns (Bio-Rad). Transfer to GαNac1, GαNac2, 2Man, MGO (19) and GlcUAβ1,3Galβ1,4GlcNAcβ1,3Galβ1,4Glcβ1,C2H4NHC(O)F₃ was assayed as described (2, 3). When sulfotransferase bound to Ni²⁺-NTA-agarose was analyzed, the reaction was gently mixed every 5 min over 3 h under these conditions. Assays were carried out for 2–8 h with the remaining [³²P]SOPs and degradation products by passage over Bio-Spin columns (Bio-Rad). Transfer to GαNac1, GαNac2, 2Man, MGO (19) and GlcUAβ1,3Galβ1,4GlcNAcβ1,3Galβ1,4Glcβ1,C2H4NHC(O)F₃ was assayed as described (2, 3). When sulfotransferase bound to Ni²⁺-NTA-agarose was analyzed, the reaction was gently mixed every 30 min.

Product Characterization—[³²P]SO₄-chondroitin and [³²P]SO₄-dermatan products were digested with 50 units of chondroitinase AC.
I (Calbiochem) in 100 mM Tris acetate buffer, pH 7.3, for 16 h at 37°C or with 25 milliunits of chondroitinase B (Calbiochem) in 100 mM Tris acetate buffer, pH 8.0, for 16 h at 37°C. The digestion products were analyzed by high pressure liquid chromatography on a 4.6xH11003 250-mm MicroPak AX-5 column (Varian) using a linear gradient of 10–450 mM KH2PO4 over 40 min with a flow rate of 1.0 ml/min (20). Standards were detected by absorbance at 210 or 234 nm, and fractions were collected at 0.5-min intervals for determination of radioactivity. [35S]SO4-labeled products separated on a Micro Pak AX-5 column were pooled separately and further characterized by gel filtration on a 16x/11003 500-mm SuperdexTM 30 preparative grade gel filtration column (Amersham BioSciences) eluted at 1.0 ml/min in 100 mM NH4HCO3. The location of the sulfate on the disaccharide digestion products was confirmed by digestion with chondroitin-4-sulfatase (50 milliunits) and analysis on Micro Pak AX-5 columns as above.

RESULTS
Identification of a Human cDNA Related to HNK-1 ST—The nonredundant data base at the NCBI was probed with the deduced amino acid sequences of human and rat HNK-1 STs (GenBank accession number AF033827). A BAC clone RP11-390G14 derived from human chromosome 3 (GenBank accession number AC024558) that contained an ORF with a length of 843 bp in the region displaying homology was identified. This sequence was used for further BLASTN searches against dbEST. Retrieval of EST BI908522 that overlapped the 5'-region of the ORF in the BAC clone allowed us to evaluate FGENES-predicted exons on a corresponding working draft sequence segment (GenBank accession number NT_005588) of chromosome 3. An ORF with a length of 1026 bp that is encoded by three exons was identified and cloned (Fig. 1A).

T A B L E  I
Substrate specificities of C4ST-3, C4ST-1, C4ST-2, D4ST-1, GalNAc-4-ST-1, and GalNAc-4-ST2

| Transferase | Chondroitin | Dermatan | GGnMMCO |
|-------------|-------------|-----------|----------|
| pmol/h/plate | pmol/h/plate | pmol/h/plate | pmol/h/plate |
| C4ST-3 | 195.4 | 46.9 | 3.5 |
| C4ST-1 | 1053.9 | 143.4 | 3.2 |
| C4ST-2 | 86.6 | 128.6 | 3.8 |
| D4ST-1 | 89.1 | 1156.0 | 3.7 |
| GalNAc-4-ST1 | 73.3 | 51.8 | 1678.5 |
| GalNAc-4-ST2 | 64.3 | 51.7 | 1249.7 |
| Vector | 19.8 | 21.6 | 7.7 |

Chrondroitin-4-sulfotransferase-3
34769

FIG. 3. Structure of the human C4ST-3 and C4ST-1 genes. A, exonic sequences that contribute to coding regions of C4ST-3 and C4ST-1 are shaded in black, and untranslated regions are shaded in gray. The size of intronic regions not shown are indicated in parentheses. The locations of CpG islands are indicated by the solid bars. B, schematic of the aligned protein sequences of C4ST-3 and C4ST-1. The numbers and letters above each sequence indicate the first amino acid encoded by the particular exons (E1–E3). The location of the transmembrane domain (TM) and the conserved motifs/regions, 5’-phosphosulfonate binding site (5’-PSB), 3’-phosphate binding site (3’-PB), III, IV, and V, are also indicated.

FIG. 4. Expression of pSec-C4ST-3, pSec-C4ST-1, and pSec-D4ST-1 in CHO cells. CHO/Tag cells were transfected with pSec-C4ST-3, pSec-C4ST-1, pSec-C4ST-2, pSec-D4ST-1, pSec-GalNAc-4-ST1, pSec-GalNAc-4-ST2, or the pSec vector. After 60 h in culture, the medium was collected, and secreted sulfotransferases were incubated with Ni2+-NTA-agarose (Qiagen). The beads were washed, and sulfotransferase activity was determined as described under “Experimental Procedures” using fixed aliquots of beads. Results are expressed as pmol incorporated/h/100-mm plate.
Azole-HCl at the pH values indicated following incubation for 2 h at 37 °C of [35S]SO4 to chondroitin by pSec-C4ST-3 was monitored using imidazole-HCl at the pH values indicated following incubation for 2 h at 37 °C of [35S]SO4 to chondroitin by pSec-C4ST-3 was monitored using imidazole-HCl. The amount of [35S]SO4 incorporated was determined as described under “Experimental Procedures.”

B. The amount of [35S]SO4 transferred to chondroitin by immobilized pSec-C4ST-1 (■) and pSec-C4ST-3 (□) or dermatan by pSec-D4ST-1 (●) was determined following incubation for 2 h at the temperatures indicated with 50 μg of desulfated chondroitin or desulfated dermatan and 2 μM PAPS[S] in a 50-μl reaction. The transfer of [35S]SO4 to chondroitin by pSec-C4ST-3 was monitored using imidazole-HCl at the pH values indicated following incubation for 2 h at 28 °C.

C. The [35S]SO4-labeled products were developed with a linear gradient of 10–450 mM KH2PO4 over 40 min at a flow rate of 1.0 ml/min. (■). The product of the chondroitinase AC I digestion was further digested with chondroitin-4-sulfatase and analyzed on the MicroPak AX-5 column under identical conditions (○). The elution positions of standards are as follows: GalNAc-4-PSB (1), D-glucono-6-enepyranoside (2), D-gluco-4-enepyranoside (3), and SO4 free sulfate (4).

C4ST-3 is 45% identical to C4ST-1, 27% identical to C4ST-2, 26% identical to GalNAc-4-ST1, 23% identical to GalNAc-4-ST2, and 26% identical to HNK-1 ST (all of the protein sequences shown in Fig. 2 are of human origin). The regions with the highest degree of identity are the putative 5′-phosphosulfate binding site (5′-PSB), the putative 3′-phosphate binding site (3′-PB), and three regions of unknown function designated III, IV, and V that are carboxyl-terminal to the 3′-phosphate binding site (Fig. 2). Identical and similar amino acids are shaded if they occur at a specific position in at least five of the seven sequences shown in the multiple alignment in Fig. 2.

Genomic Organization and Chromosome Localization of C4ST-1—A BLAST analysis of the genomic sequence available through the NCBI Web site using C4ST-3 cDNA as a query sequence mapped the C4ST-3 gene to human chromosome 3q21.3. The ORF and the 3′-UTR of C4ST-3 are encoded by three exons (Fig. 3). The genomic sequence was examined for the presence of CpG islands as defined by Gardiner-Garden and Frommer (16). A CpG island was identified extending from 620 bp upstream to 880 bp downstream of the C4ST-3 start codon. Such CpG islands have been detected in 82% of analyzed genes that show widespread expression and are indicative of the presence of a promoter region.
**Fig. 8.** RNA dot blot analysis of C4ST-3 transcripts. The human multiple tissue expression (MTE™) array was hybridized with a 32P-labeled human C4ST-3-specific cDNA probe. Tissue sources for the RNA are indicated below the blot. C4ST-3 expression is detected in adult liver (A9) and at significantly lower levels in adult kidney (A7), lymph nodes (F7), and fetal liver (D11). *,** paracentral gyrus of cerebral cortex; ***,** peripheral blood leukocytes; ****,** Burkitt’s lymphoma Raji; *****,** Burkitt’s lymphoma Daudi; ******,** colorectal adenocarcinoma, SW280.

**C4ST-3** is a Chondroitin-4-sulfotransferase—Whereas C4ST-3 displayed the highest percentage of identical amino acids, 45.1%, when compared with C4ST-1, it also displayed significant homology with other members of the HNK-1 family of sulfotransferases. Initial experiments using C4ST-3 expressed by CHO/Tag cells revealed detectable levels of an activity in both cell extracts and medium able to transfer sulfate from [35S]PAPS to chondroitin (not shown). Due to the low levels of activity, a secreted form of C4ST-3 was prepared by substituting the cytosolic and transmembrane domains of C4ST-3 with the signal sequence of human IgG (Invitrogen) to produce pSec-C4ST-3. The identical constructs were prepared for C4ST-1, C4ST-2, D4ST-1, HNK-1 ST, GaINAc-4-ST1, and GaINAc-4-ST2 and designated pSec-C4ST-1, pSec-C4ST-2, pSec-D4ST-1, pSec-GaINAc-4-ST1, and pSec-GaINAc-4-ST2. The myc epitope followed by six histidines was located at the carboxyl terminus of each of these constructs. Following transfection into CHO/Tag cells, the secreted transferases were allowed to bind to Ni²⁺-NTA-agarose and assayed for transfer of sulfate to chondroitin, dermatan, and GGM-M-CO while bound to the agarose beads as summarized in Table I.

pSec-C4ST-3 transferred sulfate to chondroitin (195 pmol/µl/plate) but not to either dermatan or GGM-M-CO (Table I). Likewise, pSec-C4ST-1 was highly active with chondroitin but not dermatan or GGM-M-CO. pSec-D4ST-1 and pSec-GaINAc-4-ST1 and -ST2 were active with dermatan and GGM-M-CO, respectively. Whereas the transfer of sulfate to chondroitin was 10-fold higher than seen with mock-transfected cells for pSec-C4ST-3 (195 versus 20 pmol/µl/plate), the rate of transfer was one-tenth of that seen for pSec-C4ST-1. Since this may reflect differences in the level of expression, the relative amounts of pSec-C4ST-1 and pSec-C4ST3 were estimated by Western blot analysis using an anti-myc antibody following SDS-PAGE and electrophoretic transfer to polyvinylidene difluoride (Fig. 4). The level of expression was significantly lower for pSec-C4ST-3 than for either pSec-C4ST-1 or pSec-D4ST-1, accounting for much of the apparent lower level of activity.

In addition to lower levels of expression, pSec-C4ST-3 was found to be labile at temperatures above 28 °C (Fig. 5B). Whereas pSec-C4ST-1 and pSec-D4ST-1 displayed similar levels of activity at 37 and 28 °C, the transfer of sulfate by pSec-C4ST-3 at 37 °C was 35% of that seen at 28 °C (Fig. 5B).
As with pSec-C4ST-3, the transfer of sulfate by pSec-C4ST-1 was reduced at 45 and 55 °C, whereas the transfer of sulfate to dermatan by D4ST-1 was not markedly reduced at these temperatures. When transferase reactions were carried out at 28 °C, incorporation of sulfate by C4ST-3 into chondroitin remained linear for up to 2 h under the assay conditions utilized (Fig. 5A). As with C4ST-1 and pSec-C4ST-1, the incorporation of \( ^{35}\text{S}\)SO\(_4\) into chondroitin by pSec-C4ST-3 is saturated at an acceptor concentration of 200 \(\mu\)g/ml (Fig. 6). pSec-C4ST-3 has a pH optimum of 6.5 for transfer of sulfate to chondroitin (Fig. 5C).

Digestion of the \( ^{35}\text{S}\)-sulfated chondroitin product with chondroitinase AC I yielded a single peak that comigrated with \(\alpha\)-d-glucono-4-enepyranoside \(\beta\)-1,3-GalNAc-4-SO\(_4\) (Fig. 7). No product was obtained upon digestion with chondroitinase B (not shown). The location of the sulfate on the C-4 hydroxyl of the GalNAc was confirmed by the release of free sulfate upon digestion of this product with chondroitin-4-sulfatase (Fig. 7). Thus, C4ST-3 is a GalNAc-4-sulfotransferase that is specific for chondroitin sequences (i.e. GlcUA\(\beta\)1,4GalNAc\(\beta\)-1,3).

**Tissue Expression Pattern of C4ST-3—Macroarray and Northern blot analyses were used to determine the expression pattern for C4ST-3 in human tissues.** As evident by MTF™ macroarray analysis (Fig. 8), C4ST-3 transcripts are most highly represented in adult liver (A9). Significantly lower expression levels were detected in adult kidney (A7), lymph nodes (F7), and fetal liver (D11). By Northern blot analysis (not shown), a single 2.1-kb transcript was detected in adult liver (Fig. 8). The level of expression was not sufficient to detect a transcript in the kidney upon Northern blot analysis.

**DISCUSSION**

C4ST-3 represents the seventh member of the HNK-1 family of sulfotransferases. C4ST-3 has the highest percentage of identical amino acids when aligned with C4ST-1, 45.1% identical. Like C4ST-1, C4ST-3 transfers sulfate to the C-4 hydroxyl of GalNAc substituted at C-3 with IdoUA (Fig. 7). Neither sulfotransferase transfers sulfate to the C-4 hydroxyl of GalNAc substituted at C-3 with \(\text{Gal}^{\alpha\text{-3,3}}\text{GalNAc}\). Both C4ST-3 and C4ST-1 have similar, if not identical, specificities and are restricted in their biological role by C4ST-3. Since C4ST-3 highly represents in adult liver, the biological role of these complex sulfated structures.