Molecular Cloning and Characterization of a Dermatan-specific N-Acetylgalactosamine 4-O-Sulfotransferase

DOI: 10.1074/jbc.M105848200

Document Version
Final published version

Link to publication record in Manchester Research Explorer

Citation for published version (APA):
Evers, M. R., Xia, G., Kang, H. G., Schachner, M., & Baenziger, J. U. (2001). Molecular Cloning and Characterization of a Dermatan-specific N-Acetylgalactosamine 4-O-Sulfotransferase. Journal of Biological Chemistry, 276(39), 36344-36353. https://doi.org/10.1074/jbc.M105848200

Published in: Journal of Biological Chemistry

Citing this paper
Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights
Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy
If you believe that this document breaches copyright please refer to the University of Manchester’s Takedown Procedures [http://man.ac.uk/04Y6Bo] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.

Download date: 20. Jun. 2020
Molecular Cloning and Characterization of a Dermatan-specific N-Acetylgalactosamine 4-O-Sulfotransferase*

Received for publication, June 22, 2001, and in revised form, July 23, 2001
Published, JBC Papers in Press, July 24, 2001, DOI 10.1074/jbc.M105848200

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

From the Department of Pathology, Washington University School of Medicine, St. Louis, Missouri 63110
and Zentrum fuer Molekulare Neurobiologie, Universitaet Hamburg, Martinistr. 52, D-20246 Hamburg, Germany

We have identified and characterized an N-acetylgalactosamine-4-O-sulfotransferase designated dermatan-4-sulfotransferase-1 (D4ST-1) (GenBank™ accession number AF401222) based on its homology to HNK-1 sulfotransferase. The cDNA predicts an open reading frame encoding a type II membrane protein of 376 amino acids with a 43-amino acid cytoplasmic domain and a 316-amino acid luminal domain containing two potential N-linked glycosylation sites. D4ST-1 has significant amino acid identity with HNK-1 sulfotransferase (21.4%), N-acetylgalactosamine-4-O-sulfotransferase-1 (GalNAc-4-ST1) (24.7%), N-acetylgalactosamine-4-O-sulfotransferase-2 (GalNAc-4-ST2) (21.0%), chondroitin-4-O-sulfotransferase 1 (27.3%), and chondroitin-4-O-sulfotransferase 2 (22.8%). D4ST-1 transfers sulfate to the C-4 hydroxyl of β1,4-linked GalNAc that is substituted with an O-linked iduronic acid (IdoUA) at the C-3 hydroxyl. D4ST-1 shows a strong preference in vitro for sulfate transfer to IdoUA1,3GalNAcβ1,4 that is flanked by GlcUAβ1,3GalNAcβ1,4 as compared with IdoUAα1,3GalNAcβ1,4 flanked by IdoUAα1,3GalNAcβ1,4. The specificity of D4ST-1 when assayed in vitro suggests that the addition of sulfate to GalNAc occurs immediately after epimerization of GlcUA to IdoUA. The open reading frame of D4ST-1 is encoded by a single exon located on human chromosome 15q14. Northern blot analysis reveals a single 2.4-kilobase transcript. D4ST-1 message is expressed in virtually all tissues at some level but is most highly expressed in pituitary, placenta, uterus, and thyroid. The properties of D4ST-1 indicate that sulfation of the GalNAc moieties in dermatan is mediated by a distinct GalNAc-4-O-sulfotransferase and occurs following epimerization of GlcUA to IdoUA.

The oligosaccharide moieties on glycoproteins, glycolipids, and proteoglycans are frequently selectively modified with sulfate. The addition of sulfate is temporally and spatially regulated, indicating that the presence of sulfate in many cases confers biological function to these saccharides (1–3). Chondroitin sulfate proteoglycans consist of repeating units of (GlcUAβ1,3GalNAcβ1,4)n that are sulfated on the 4-hydroxyl of GalNAc (chondroitin sulfate A), the 6-hydroxyl of GalNAc (chondroitin sulfate C), or both the 4- and 6-hydroxyls of GalNAc (chondroitin sulfate E). Variable amounts of GlcUA in chondroitin are converted to IdoUA by glucuronolactone-5-epimerase to produce dermanan sulfate, also referred to as chondroitin sulfate B, a glycan consisting of IdoUA1,3GalNAcβ1,4n and (GlcUAβ1,3GalNAcβ1,4)n disaccharides (4, 5). The proportion of IdoUA and GlcUA in dermnan varies considerably in tissues and is influenced by the level of free sulfate in the medium of cultured cells (6). In addition to sulfation of the GalNAc on the C-4 hydroxyl, the GalNAc in dermatan can also be sulfated at the C-6 hydroxyl (7), and the IdoUA can be sulfated at the C-2 hydroxyl (8). While chondroitin and dermatan glycans are selectively expressed on a number of proteoglycans, many aspects of the regulation of their synthesis remain to be elucidated. In light of the many functions attributed to proteoglycans bearing dermatan and chondroitin oligosaccharides (9, 10), developing an understanding of the regulation of their synthesis will yield information about the biologic significance of these complex sulfated structures.

Two groups independently reported the cloning of a GalNAc-4-sulfotransferase (C4ST-1) that selectively adds sulfate to the C-4 hydroxyl of GalNAc when adjacent to GalUA in both chondroitin and dermatan (11, 12). A second GalNAc-4-sulfotransferase (C4ST-2), which is 36% identical to C4ST-1, was also cloned (12). The preferential addition by C4ST-2 of sulfate to GlcUA-GalNAc in both chondroitin and dermatan led to the speculation that sulfation may precede the epimerization of GlcUA to IdoUA in dermatan (12). C4ST-1 and C4ST-2 are both homologous to the HNK-1 sulfotransferase (HNK-1 ST), a 3-O-glucuronosylsulfotransferase that is responsible for synthesis of the HNK-1 epitope SO4-3-GlcUA1,3Galβ1,4GlcNAcR (13, 14). Two additional GalNAc-4-O-sulfotransferases, GalNAc-4-ST1 and GalNAc-4-ST2, that add sulfate to terminal β1,4-linked GalNAc in the sequence GalNAcβ1,4GlcNAcR on N-linked oligosaccharides such as are present on the glycoprotein hormones lutropin and thyrotropin were cloned and also found to be homologous to HNK-1 ST. The intracellular forms of

* This work was supported by National Institutes of Health Grant R01-DR41738 (to J. U. B.), by Deutsche Forschungsgemeinschaft Grant SCHAI185/15-1 (to M. S.), and by a German Academic Exchange Service postdoctoral fellowship (to G. X.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AF401222.

† These three authors contributed equally to this work.

§ To whom correspondence should be addressed: Washington University School of Medicine, Dept. of Pathology, 660 S. Euclid Ave., St. Louis, MO 63110. Tel.: 314-362-8730; Fax: 314-362-8888; E-mail: Baenziger@Pathology.wustl.edu.

¶ The abbreviations used are: GlcUA, glucuronic acid; IdoUA, iduronic acid; HNK-1 ST, HNK-1 sulfotransferase; C4ST, chondroitin 4-O-sulfotransferase; GalNAc-4-ST, N-acetylgalactosamine 4-O-sulfotransferase; bp, base pair(s); ORF, open reading frame; nt, nucleotide; PAPS, 3’-phosphoadenosine-5’-phosphosulfate; UTR, untranslated region; PCR, polymerase chain reaction; HPLC, high pressure liquid chromatography; D4ST, dermatan-4-sulfotransferase-1; ΔDi-4S, Δ-gluco-4-enepyranoside β1,3-GalNAc-4-SO42; ΔDi-6S, Δ-gluco-4-enepyranoside β1,3-GalNAc-6-SO42; ΔUA, unsaturated uronic acid; ΔDi-di Sβ, Δ-gluco-4-enepyranoside β1,3-GalNAc-4,6-diSO42.
GalNAc-4-ST1 and GalNAc-4-ST2 that are located in the Golgi do not add sulfate to the nonterminal GalNAc residues in either chondroitin or dermatan (15, 16).

We now report the cloning of an additional GalNAc-4-sulfotransferase, dermatan-4-sulfotransferase (D4ST-1) that adds sulfate to the C-4 hydroxyl of GalNAc in the sequence IdOA-GalNAc but not GlcUA-GalNAc. Like C4ST-1, C4ST-2, GalNAc-transferase, dermatan-4-sulfotransferase (D4ST-1) that adds either chondroitin or dermatan (15, 16).

EXPERIMENTAL PROCEDURES

Isolation of a cDNA Encoding Human Dermatan 4-O-Sulfotransferase 1—A 1311-bp open reading frame (ORF) encoding a novel sulfotransferase, designated D4ST-1, was identified using FGENES (version 1.6) (17). A single exon gene on the genomic clone RP11-64K12 with the GenBank™ accession number AC013356 (see “Results”) was predicted. PCR was used to amplify this 1311-bp ORF from human brain cDNA (see below). Filtering of repetitive elements (Repeat Masker, version 5/3) searches against the databases of GenBank™, EMBL, and DBJ expressed sequence tag divisions (dbEST) at the forward primer 5′-ATG CTG ATG TTG TCG GGT ATC GTG-3′ and the reverse primer 5′-GCC ACC GCT GCT GGA AG-3′, located on this genomic clone to verify exonic sequences of D4ST-1 and to locate exons of neighboring genes. CpG islands as defined by Gardiner-Garden and Frommer (18) were located in genomic sequences 5′ of the D4ST-1 exon using the WWWCPG program (19).

5′- and 3′-rapid amplification of cDNA ends using Marathon Ready™ cDNA (adult kidney; BD Biosciences-CINTOCHEN, Heidelberg, Germany) as a template was performed to obtain a full-length sequence. The 3′-region was obtained by amping the cDNA using a primer containing a nucleotides 130–153 (D4ST-1-GSP1; 5′-ATG CTG ATG TTG TCG GGT ATC GTG-3′) and the AP1 anchor primer (CLONTECH), followed by a nested PCR (30 cycles) using a primer consisting of nucleotides 473–499 (D4ST-1-NGSP1; 5′-CCT GCT CTA ACT GGA AGC GGG TGA TGA-3′) and the AP2 anchor primer (CLONTECH). The DNA products were gel-purified (Life Technologies, Inc.), cloned into pCR-TOPO 2.1 (Invitrogen, Karlsruhe, Germany), and sequenced on both strands. Amplified sequences were assembled using the Lasergene (DNASTAR Inc., Madison, WI) software suite. The 5′-UTR could not be amplified using a number of different primer combinations.

The protein sequences of HNK-1 sulfotransferase family members were analyzed using the ClustalW (version 1.4) algorithm (20) implemented in the Bioedit suite to perform multiple alignments. Pairwise alignments were executed using the Smith-Waterman algorithm (21) implemented in the European Molecular Biology Open Software Suite at the European Bioinformatics Institute.

Amplification of ORF of D4ST-1 and Construction of pcDNA3.1-D4ST-1—The ORF of D4ST-1 was amplified from a SuperScript™ human fetal brain cDNA library (Life Technologies) by PCR using 1) the 5′-specific primer 5′-ATA TGA ATT CGC CAC CAT GTT CCC CCG CCC GAT GAC CC-3′ containing an EcoRI site, the consensus Kozak sequence GCCACC (22), and a start codon and 2) the 3′-specific primer 5′-ATG CTG ATG TTG TCG GGT ATC GTG-3′ containing a stop codon and an XhoI site. PCRs were carried out using Fis/Turbo® HotStart polymerase (Stratagene, Amsterdam, The Netherlands) with an initial denaturation for 2 min at 96°C, followed by 35 cycles of a reaction consisting of 30 s denaturation at 96°C, 60 s annealing at 60°C, and 120 s of elongation at 72°C. The PCR fragment containing the ORF with the expected length of 1131 bp was directionally subcloned into the eukaryotic expression vector pcDNA3.1 (Invitrogen), completely sequenced on both strands, and designated pcDNA3.1-D4ST-1.

Construction of pcDNA3.1-C4ST-1 and pcDNA3.1-C4ST-2—The ORF of human C4ST-1 was amplified from the I.M.A.G.E. Consortium (LLNL) cDNA clone, ID 2388199 (GenBank™ accession number AJ289133; IMAGp998J085931.1 obtained from RZPD, Berlin, Germany) (23). PCRs were carried out with Fis/Turbo® HotStart polymerase (Stratagene, Amsterdam, The Netherlands) with an initial denaturation for 2 min at 96°C, followed by 35 cycles of a reaction consisting of 30 s denaturation at 96°C, 60 s annealing at 60°C, and 120 s of elongation at 72°C. The PCR fragment containing the ORF with the expected length of 1131 bp was directionally subcloned into the eukaryotic expression vector pcDNA3.1 (Invitrogen), completely sequenced on both strands, and designated pcDNA3.1-C4ST-1.

Based on the sequence of GenBank™ entry AJ289131, the primers 5′-CCG CCC GCA GTA GCA CCA-3′ and 5′-TCC AGG CAC ACC GGA AAA AG-3′ were used to amplify (40 cycles) the complete ORF of human C4ST-2 including adjacent untranslated regions from Marathon Ready™ cDNA clone, ID 2388199 (GenBank™ accession number AJ289133; IMAGp998J085931.1 obtained from RZPD, Berlin, Germany) (23). PCRs were carried out with Fis/Turbo® HotStart polymerase (Stratagene, Amsterdam, The Netherlands) with an initial denaturation for 2 min at 96°C, followed by 35 cycles of a reaction consisting of 30 s denaturation at 96°C, 60 s annealing at 60°C, and 120 s of elongation at 72°C. The PCR fragment containing the ORF with the expected length of 1131 bp was directionally subcloned into the eukaryotic expression vector pcDNA3.1 (Invitrogen), completely sequenced on both strands, and designated pcDNA3.1-C4ST-2.
Dermatan-specific GaINAc-4-sulfotransferase

FIG. 2. Multiple amino acid sequence alignment of human D4ST-1, GaINAc-4-ST1, GalNAc-4-ST2, CaST-1, CaST-2, and HKN-1 ST. Alignment was performed using the ClustalW algorithm. Introduced gaps are shown as hyphens, and aligned amino acids are boxed (black for identical residues and dark gray 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, and V) are indicated.

Transient Expression of Human D4ST-1, CaST-1, and CaST-2—CHO/Tag cells were transfected with 13 μg of pcDNA3.1-D4ST-1, pcDNA3.1-CaST-1, pcDNA3.1-CaST-2, or pcDNA3.1 using 35 μg of LipofectAMINE (Life Technologies) in serum-free medium for 6 h according to 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, and protease inhibitors (25 μl 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 × 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 Assays—Chondroitin- and Dermatan-4-sulfotransferase activities were assayed as described (12, 15). Each reaction (50 μl) contained 50 mM imidazole-HCl, pH 6.5, 2 mM dithiothreitol, 5 × 10⁵ cpm of [35S]-labeled 3′-phosphoadenosine-5′-phosphosulfate (PAPS), 2 μM PAPS, protamine (0.005% [w/v]) for chondroitin or 0.05% (w/v) for dermatan), and protease inhibitors (see solubilization buffer above). Desulfated chondroitin (Seikagaku America, Inc.) and desulfated dermatan from porcine intestine (Sigma) (24) were used as acceptors.

Further screening of the nonredundant data base at the NCBI using the deduced amino acid sequence of human HNK-1 ST produced a new human ST with 72% identity to residues 219–308 of HNK-1 ST. This cDNA corresponded to an HNK-1-like ST that was cloned into pcDNA3.1 and transiently expressed in CHO cells.

RESULTS

Identification of a Human cDNA Related to HKN-1 ST—We recently identified and characterized two human GaINAc-4-O-sulfotransferases, GaINAc-4-ST1 (15) and GaINAc-4-ST2 (16), based on their homology to the HKN-1 ST. HKN-1 ST produces the HNK-1 epitope SO₄-3-GlcUA by transferring sulfate to the C-4 hydroxyl of terminal β1,4-linked GaINAc present on N-linked oligosaccharides such as those found on the glycans of the thyroid hormones and thyrotrpin (28). CaST-1 and CaST-2 also GalNAc-4-O-sulfotransferases that are homologous to HKN-1 ST, and CaST-4, CaST-1, and CaST-2 transfer sulfate to the C-4 hydroxyl of internal β1,4-linked GalNAc moieties within the repeating disaccharide sequences found in chondroitin and dermatan (11, 12).

Further screening of the nonredundant data base at the NCBI using the deduced amino acid sequence of human HKN-1...
ST (GenBank™ accession no. AF033827) identified a 177-kb genomic clone, GenBank™ accession number AC013356, that included a sequence related to HNK-1 ST (E value 1e-23). We identified an ORF with a length of 1131 bp encoded by a single predicted exon in the region displaying homology. This ORF was amplified from human brain cDNA and sequenced (Fig. 1A). 5' and 3'-rapid amplification of cDNA ends was carried out to obtain a full-length sequence of D4ST-1 but only succeeded in elongation of the 3'-end (see "Experimental Procedures"). The amplified sequences contained the 3'-UTR with a polyadenylation signal and a poly(A) tail. Fig. 1A shows the 1960-bp cDNA that includes the 1131-bp ORF encoding a protein of 376 amino acid residues with two potential N-glycosylation sites and a calculated molecular mass of 43 kDa. The deduced protein, designated D4ST-1 (GenBank™ accession number AF401222) is a typical type II transmembrane protein with a transmembrane region close to the N-terminus (see the Kyte-Doolittle hydrophobicity profile (29) in Fig. 1B).

Multiple alignment of the protein sequence of D4ST-1 with other members of the HNK-1 sulfotransferase family was performed using the ClustalW algorithm (Fig. 2). The alignment indicates that D4ST-1 is 27.3% identical to C4ST-1, 22.8% identical to C4ST-2, 24.7% identical to GalNAc-4-ST1, 21.0% identical to GalNAc-4-ST2, and 21.4% identical to HNK-1 ST (all of the protein sequences shown 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 six sequences shown in the multiple sequence alignment in Fig. 2. While this emphasizes key features of a protein family, relationships between individual members are under-

---

**Table I**

| Acceptor   | Chondroitin  | Dermatan    |
|------------|--------------|-------------|
|            | Cell/Medium  | Cell/Medium |
| D4ST-1     | 38/346       | 80/3396     |
| C4ST-1     | 702/2266     | 154/448     |
| C4ST-2     | 42/352       | 44/438      |
| Mock       | 12/178       | 26/192      |

---

**Fig. 3. Structure of the human D4ST-1 gene.** Exonic sequences that contribute to coding regions are boxed and shaded in black, while untranslated regions are only boxed. The intronless D4ST-1 gene is found between the genes encoding the KIAA0945 protein (the 3' terminal part of the KIAA0944 gene is shown) and a pseudogene copy of the ribosomal L9 protein (termed L9v). Transcriptional orientations are shown by arrows below the gene identifiers. The underlying genomic clone is denoted with its GenBank™ accession number below the schematic. nt, nucleotide.

**Fig. 4. Concentration dependence for transfer of sulfate to chondroitin and dermatan by D4ST-1, C4ST-1, and C4ST-2.**

CHO-Tag cells were transfected with pcDNA3.1-D4ST-1, pcDNA3.1-C4ST-1, pcDNA3.1-C4ST-2, or pcDNA3.1. After 60 h in culture, cells were harvested and solubilized in 200 μl of 1% Triton X-100, and 25 μl of cell extract was utilized for a 2-h assay. The medium was collected and concentrated to 2 ml by ultrafiltration, and 25 μl was utilized. Results are expressed as pmol incorporated/hr/100-mm plate.

---

**Table II**

| Acceptor | Chondroitin | Dermatan |
|----------|-------------|----------|
|          | Cell/Medium | Cell/Medium |
| D4ST-1   | 38/346      | 80/3396  |
| C4ST-1   | 702/2266    | 154/448  |
| C4ST-2   | 42/352      | 44/438   |
| Mock     | 12/178      | 26/192   |
that D4ST-1 is more closely related to C4ST-1 (32.2% identical amino acids) and GalNAc-4-ST1 (31.1% identical amino acids) than to C4ST-2 (28.7% identical amino acids), HNK-1 ST (26.1% identical amino acids), and GalNAc-4-ST2 (24.6% identical amino acids).

**Identification of a Murine ORF Encoding the Mouse Ortholog of D4ST-1**

The human D4ST-1 cDNA sequence was used to query the nonredundant database at the NCBI. A mouse cDNA clone 2600016L03 (GenBank™ accession number AK011230) was identified, and the predicted ORF with the identical length as the ORF of the human counterpart of 1131 bp (nt 93–1223) was amplified by PCR from adult mouse lung and kidney cDNA (data not shown). The mouse ortholog of D4ST-1 is 92.7% identical to human D4ST-1 at the protein level (89% identical at the nucleotide level). In contrast to the coding regions, the 3′-UTRs of human and mouse D4ST-1 show a decreased level of identity (59%).

**Genomic Organization and Chromosome Localization of D4ST-1**

The genomic BAC clone RP11-64K12 (GenBank™ accession number AC013356) initially identified as encoding D4ST-1 is annotated to be mapped on human chromosome 15q14. Comparison of the genomic sequence with the amplified D4ST-1 cDNA indicates that the ORF and the 3′-UTR of D4ST-1 are encoded by a single exon (Fig. 3). Only 2.96 kilobase pairs of sequence separate the 3′-UTR of the gene encoding the KIAA0945 protein (GenBank™ accession number AB023162) (30) from the start codon of the D4ST-1 gene. Downstream of the D4ST-1 locus, we identified a copy of the gene encoding the ribosomal protein L9 (GenBank™ accession number BC004156) that probably represents a pseudogene (31). An analysis of splice sites directly upstream of the start codon with a minimum score of 0.9 determined the nearest 3′-acceptor splice site to be 532 bp upstream (Fig. 3). The start codon of the 1131-bp D4ST-1 ORF in Fig. 1 is the first potential site of initiation downstream of the nearest potential splice site in the genomic sequence that is in the correct reading frame and preceded by a stop codon. We were not, however, able to amplify the predicted sequences upstream of the start codon from any human cDNA.

The genomic sequence was also examined for the presence of CpG islands as defined by Gardiner-Garden and Frommer (18).
Dermatan-specific GalNAc-4-sulfotransferase

A CpG island was identified extending from 620 bp upstream to 880 bp downstream of the D4ST-1 start codon. Such CpG islands have been detected in 82% of analyzed genes showing widespread expression and are indicative of the presence of a promoter region (32, 33). The high G + C content of this region may have contributed to the difficulty of amplifying the complete cDNA.

**D4ST-1 Is a Dermatan-specific GalNAc-4-O-sulfotransferase**—While D4ST-1 is homologous to HNK-1 ST, C4ST-1, C4ST-2, GalNAc4-ST1, and GalNAc4-ST2 and the highest percentage of identical residues are located within the regions that are characteristic of the HNK-1 family of sulfotransferases, it was not possible to predict the saccharide specificity of D4ST-1 based on its peptide sequence. Initial experiments using D4ST-1 expressed by CHO/Tag cells demonstrated that D4ST-1 was not able to transfer sulfate to either terminal glucuronic acid in GlcUAβ1,3Galβ1,4GlcNAc-R or terminal GalNAc in GalNAcβ1,4GlcNAcβ1,2Man-R but did transfer sulfate to dermatan and chondroitin (not shown). The specificity of D4ST-1 was examined in greater detail and compared with that of C4ST-1 and C4ST-2 as summarized in Table I. CHO/Tag cells were transfected with pcDNA3.1-D4ST-1, pcDNA3.1-C4ST-1, pcDNA3.1-C4ST-2, or pcDNA3.1. After 60 h in culture, both the culture medium and a Triton X-100 extract of the cells were tested for sulfotransferase activity.

The results summarized in Table I indicate that D4ST-1, C4ST-1, and C4ST-2 have distinct specificities for chondroitin and dermatan. Although D4ST-1 has a hydrophobic sequence near its amino terminus that is predicted to act as a transmembrane domain, only 3% of the activity is retained by the transfected cells as compared with that released into the culture medium. The D4ST-1 released into the medium may reflect a proteolytic clip, as has been seen for a number of sulfotransferases including GalNAc4-ST1 and GalNAc4-ST2 (15, 16, 34, 35); however, this possibility has not been formally addressed for D4ST-1. While D4ST-1 is able to transfer sulfate to both dermatan and chondroitin, the rate of transfer is nearly 10-fold greater to dermatan than to chondroitin when the same substrate concentrations are used (Table I). Like D4ST-1, C4ST-1 transfers sulfate to both chondroitin and dermatan; however, C4ST-1 transfers 5-fold more sulfate to chondroitin than to dermatan at the same substrate concentrations. While a large percentage of C4ST-1 activity is also found in the medium, roughly 25% of the sulfotransferase activity produced is retained by the transfected cells. C4ST-2 is also found predominately in the medium following transfection. In contrast to either D4ST-1 or C4ST-1, C4ST-2 transfers sulfate to dermatan and chondroitin at similar rates; however, the rates are one-tenth of that obtained with either D4ST-1 or C4ST-1.

The relative specificities of D4ST-1, C4ST-1, and C4ST-2 for desulfated chondroitin and dermatan were further examined by comparing their rates of transfer at multiple concentrations of chondroitin and dermatan as shown in Fig. 4, A and B, respectively. D4ST-1 only displays a significant transfer of sulfate to chondroitin at relatively high substrate concentrations, raising the possibility that the chondroitin is contaminated with small amounts of dermatan and/or that occasional iduronic acid residues are present in the chondroitin and utilized by the D4ST-1. C4ST-1 transfers sulfate to GlcUAβ1,3GalNAcβ1,4α sequences (11, 12). Since dermatan is known to contain GlcUAβ1,3GalNAcβ1,4 sequences interspersed with the IdoUAα1,3GalNAcβ1,4 sequences (36), the transfer of sulfate to both chondroitin and dermatan by C4ST-1 is expected. However, the rate of transfer to dermatan from porcine intestine (90% IdoUA, 10% GlcUA) should be lower than to chondroitin due to the low content of GlcUA (Fig. 4B). C4ST-2 also transfers sulfate to both chondroitin and dermatan (12); however, it does not demonstrate a marked preference for chondroitin or dermatan and requires relatively high concentrations of acceptor to obtain significant levels of sulfate transfer (Fig. 4).

The [35S]-sulfated chondroitin and dermatan products were digested with chondroitinase AC I and chondroitinase B. Digestion of both the [35S]SO4-chondroitin and [35S]SO4-dermatan products obtained with C4ST-1 yielded a single peak that comigrated with a-glucurono-4-enepyranoside β-1,3-GalNAc4- SO4 (ΔDi-4S) (Fig. 5, A and B). 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 (not shown). Chondroitinase AC I digestion of the [35S]SO4-labeled chondroitin and dermatan products produced by D4ST-1 produced a minor peak (peak a) that comigrated with ΔDi-4S and was sensitive to digestion with chondroitin-4-sulfatase and a major peak (peak b) that is more acidic when examined by HPLC on a MicroPak AX-5 column (Fig. 5, C and D). The D4ST-1 products released by chondroitinase AC I digestion were isolated and analyzed by gel filtration. In both instances peak a comigrated with authentic disaccharide, and peak b comigrated with authentic tetrasaccharide on Superdex 30 (Fig. 6, A and B).

The product obtained upon digestion of either [35S]SO4-chondroitin or [35S]SO4-dermatan with chondroitinase AC I reflects the sulfotransferase utilized rather than the substrate. The characteristic product obtained upon chondroitinase AC I digestion in conjunction with the strong preference of D4ST-1 for dermatan suggested that D4ST-1 adds sulfate to a GalNAc that...
is adjacent to an IdoUA. Further, since chondroitinase AC I cleaves the GalNAcβ1,4GlcUA but not GalNAcβ1,4IdoUA, the tetrasaccharide was predicted to arise from an IdoUA-GalNAc flank by GlcUA-GalNAc as shown in Fig. 7 and to have the sequence ΔUAβ1,3GalNAcβ1,4IdoUAα1,3GalNAc. The structure of the tetrasaccharide and the location of the sulfate were established by the experiments shown in Fig. 8. The tetrasaccharide was treated with Hg(OAc)₂ to remove the unsaturated uronic acid (ΔUA). When analyzed on a MicroPak AX-5 column, the product eluted at an earlier time (Fig. 8A) due to the reduction in negative charge associated with loss of the unsaturated uronic acid. Following digestion with β-hexosaminidase, the sulfated product migrated just ahead of ΔDi-4S (Fig. 8B) consistent with the structure IdoUAα1,3GalNAcβ1,4SO₄. Digestion with α-1-iduronidase generated a product that migrated at the position of SO₄-4-GalNAc (Fig. 8C).

Digestion of the D4ST-1 [⁴³⁵S]SO₄ dermatan product with chondroitinase B yielded two [⁴³⁵S]SO₄-labeled products when analyzed by HPLC on MicroPak AX-5 column (Fig. 5F). The minor product, peak a, migrated at the location of ΔDi-4S, whereas the major peak, peak c, was more acidic than the tetrasaccharide produced by digestion with chondroitinase AC I and migrated at the position of a hexasaccharide when examined on Superdex 30 (Fig. 9B). No product was obtained when the D4ST-1 [⁴³⁵S]SO₄-chondroitin product was digested with chondroitinase B (Fig. 5E). This suggests that the major product of chondroitinase B digestion is ΔUAβ1,3GalNAc-4-SO₄β1,4GlcUAβ1,3GalNAcβ1,4GlcUAβ1,3GalNAc as shown in Fig. 7. The absence of either ΔDi-4S or this hexasaccharide in the chondroitinase B digestion product of [⁴³⁵S]SO₄-chondroitin produced by D4ST-1 indicates that the chondroitin is not contaminated with dermatan but does have an occasional IdoUAα1,3GalNAcβ1,4 sequence that is a substrate of the D4ST-1. Digestion of the hexasaccharide with chondroitinase AC I provided the [⁴³⁵S]SO₄-hexasaccharide (Fig. 9A, peak b, and Fig. 9B) to a [⁴³⁵S]SO₄-tetrasaccharide when analyzed by gel filtration on Superdex 30 (Fig. 9D). While less acidic than the hexasaccharide, the tetrasaccharide did not comigrate with the tetrasaccharide produced by chondroitinase AC I digestion (compare Fig. 9C with peak b of Fig. 5D). The difference in elution time reflects the difference in the location of the sulfate and the presence of GlcUA as opposed to IdoUA (Fig. 7). As was seen with the tetrasaccharide product of chondroitinase AC I digestion of the [⁴³⁵S]SO₄ dermatan, treatment of the [⁴³⁵S]SO₄-hexasaccharide produced by chondroitinase B digestion with Hg(Ac)₂ to remove the unsaturated uronic acid shifted the [⁴³⁵S]SO₄-labeled material (peak a) to a less acidic position (peak b) when examined by anion exchange on MicroPak AX-5 (Fig. 10A). The shift in elution time was similar to that seen upon removal of the unsaturated uronic acid from the tetrasaccharide released by chondroitinase AC I digestion (Fig. 6A). Digestion of the [⁴³⁵S]SO₄-labeled product obtained by removal of the unsaturated uronic acid (Fig. 10A) with β-hexosaminidase did not alter its elution time (Fig. 10B, peak d), indicating that the GalNAc was not released. In contrast, digestion with N-acetylgalactosamine-4-O-sulfatase (aryl.sulfa-

tase B) resulted in the appearance of free [⁴³⁵S]SO₄ (Fig. 10B, peak c) confirming the location of the sulfate on the GalNAc exposed by the removal of the unsaturated uronic acid.

The high yield of [⁴³⁵S]SO₄-labeled products from both the chondroitinase AC I and the chondroitinase B digestions indicates that the structure shown in Fig. 7 is the predominant sequence modified by D4ST-1 in the dermatan substrate. The origin of the ΔDi-4-S, peak a, following digestion of the D4ST-1 product with chondroitinase AC I is not at present clear. The ratio of peak b to peak a was typically higher upon chondroitinase AC I digestion of the dermatan than chondroitin product; however, there was considerable variation in this ratio. The presence of ΔDi-4S in the chondroitinase AC I digestes may indicate that D4ST-1 can transfer sulfate to a GalNAc flanked by GlcUA on both sides; however, further analyses with more highly defined substrates are in progress to more fully address this issue.

Expression Pattern of D4ST-1—Northern blot and array analyses were used to determine the expression pattern for D4ST-1 in human tissues (Fig. 11 and 12, respectively). A single 2.4-kilobase pair transcript was detected by Northern blot analysis in heart, placenta, liver, and pancreas with significantly lower levels in skeletal muscle and kidney (Fig. 11). Array analysis indicated that nearly every tissue represented expressed at least modest levels of the D4ST-1 message. The
strongest signals were seen in pituitary (D3), placenta (B8), uterus (D8), thyroid (D9), fetal lung (G11), and the colorectal adenocarcinoma cell line SW480 (G10) (Fig. 12).

**FIG. 9.** Digestion of chondroitinase B [\(^{35}\)S]SO\(_4\)-labeled product from dermanan with chondroitinase AC I. A, [\(^{35}\)S]SO\(_4\)-labeled dermanan produced by D4ST-1, digested with chondroitinase B, and fractionated on a Micro Pak AX-5 column. B, gel filtration on Superdex 30 of the major peak eluting at 31 min (peak b) in A. C, chromatography of peak b in A on Micro Pak AX-5 following digestion with chondroitinase AC I. D, gel filtration chromatography of peak b in A on Superdex 30 following digestion with chondroitinase AC I. The positions of standards are as follows. A and C, GalNAc-4-SO\(_4\) (1), \(\Delta\)Di-6S (2), \(\Delta\)Di-4S (3), \(\Delta\)Di-diSO\(_4\) (4), SO\(_4\), free sulfate (5). B and D, chondroitin hexasaccharide (1), chondroitin tetrasaccharide (2), \(\Delta\)Di-4S (3), SO\(_4\), free sulfate (4).

**DISCUSSION**

Dermatan-specific GalNAc-4-sulfotransferase represents the sixth member of the HNK-1 family of sulfotransferases. D4ST-1 transfers sulfate to the C-4 hydroxyl of GalNAc substituted at C-3 with \(\alpha\)-linked IdoUA (i.e. IdoUA-GalNAc). As a result D4ST-1 is only active on dermanan type structures (i.e. chondroitins that contain IdoUA). IdoUA makes up 90% of the uronic acid in the mature dermanan isolated from porcine intestine that was used for these studies. While small amounts of \(\Delta\)UA-GalNAc-4-SO\(_4\) were identified, it is not clear whether sulfate addition occurs at both the C-4 and C-3 hydroxyls.

**FIG. 10.** Characterization of the [\(^{35}\)S]SO\(_4\)-hexasaccharide product obtained by chondroitinase B digestion of the D4ST-1 dermanan product. A, peak c in Fig. 5 was isolated and examined by chromatography on Micro Pak AX-5 prior to (solid line, peak a) and following treatment with merceric acetate to remove the unsaturated uronic acid (broken line, peak b). B, following removal of the unsaturated uronic acid, peak b in A was isolated and digested with \(\beta\)-N-acetyethylhexosaminidase (solid line, peak d) or N-acetylgalactosamine-4-O-sulfatase (aryl sulfatase B) (broken line, peak c) and analyzed on the same MicroPak AX-5 column. The elution positions of standards are as follows. GalNAc-4-SO\(_4\) (1), \(\Delta\)Di-6S (2), \(\Delta\)Di-4S (3), \(\Delta\)Di-diSO\(_4\) (4), SO\(_4\), free sulfate (5).

**FIG. 11.** Northern blot analysis of D4ST-1 transcripts. Each lane of the MTN® Northern blot contains 2 \(\mu\)g poly(A)+ RNA and was hybridized with a \(^{32}\)P-labeled cDNA probe specific for human D4ST-1 (see “Experimental Procedures”). Tissues used to prepare the RNA are indicated above each lane. Migration positions of standards are indicated at the left.
configuration is favored over the L-configuration at equilibrium in the presence of the epimerase by a factor of 85 to 15 (39), selective sulfation of the GalNAc attached to the reducing terminus of the IdoUA by D4ST-1 would have the effect of shifting the equilibrium toward IdoUA, since it could not longer equilibrate back to GlcUA. Such a mechanism would also result in very efficient sulfation of IdoUA-GalNAc disaccharides in dermatan. If this is the case, the levels of PAPS, D4ST-1, and glucuronyl C5-epimerase would each have the potential to determine the extent to which GlcUA-GalNAc disaccharides are converted to IdoUA-GalNAc in a particular tissue or cultured cell.

Proteoglycans that bear dermatan sequences are widely expressed by a variety of different cells in the body (42). Thus, the nearly ubiquitous presence of message in tissues is expected. It is apparent from both the Northern blot and array analyses that the levels of D4ST-1 message expressed vary widely. Tissues such as heart, placenta, liver, pancreas, uterus, pituitary, and thyroid gland express much higher levels than other tissues. Since D4ST-1 and glucuronyl C5-epimerase may individually or collectively determine the IdoUA to GlcUA ratio for dermatan, it will be of interest to determine how expression of D4ST-1 is regulated and whether the levels of D4ST-1 expression correlate with the extent of GlcUA to IdoUA conversion on dermatan chains present on proteoglycans.

D4ST-1 is structurally related to the other members of the HNK-1 family of sulfotransferases. The highest percentage of sequence identity to other family members is seen in the regions hypothesized to bind the 5′-phosphosulfate and the 3′-phosphate of PAPS as well as the three regions of unknown function. The regions with the lowest percentage of identical amino acids among the HNK-1 sulfotransferase family include the stem region, the transmembrane domain, and the cytosolic domain. D4ST-1 has the highest percentage of identical amino acids when compared with C4ST-1 (27.3%) in the multiple sequence alignment. D4ST-1 and C4ST-1 transfer sulfate to the C-4 hydroxyl of GalNAc internal IdoUA-GalNAc and GlcUA-GalNAc disaccharide sequences, respectively. The specificity of C4ST-2, which has 22.8% identical amino acid residues with D4ST-1 on the multiple sequence alignment, differs from that of either D4ST-1 or C4ST-1. The relatively high concentrations of desulfated dermatan and chondroitin that are required to detect activity raise the possibility that the native substrate has not yet been identified. More extensive characterization of the product may reveal if this is the case. C4ST-2 is, however, reported to be a GalNAc-4-sulfotransferase that transfers sulfate to internal UA-GalNAc sequences (12). In contrast to D4ST-1, C4ST-1, and C4ST-2, GaINAc-4-ST1 and GaINAc-4-ST2 both transfer sulfate to GaINAcβ1,4GaINAcβ sequences found at the termini of N- and O-linked oligosaccharides (15, 16). Since deletion of the stem region from both GaINAc-4-ST1 and GaINAc-4-ST2 converts these terminal sulfotransferases to forms that will transfer sulfate to internal GlcUA-GalNAc sequences as well as to terminal GaINAcβ1,4GaINAcβ sequences, the high percentage of identical amino acids, 24.7 and 21.0% respectively, is not surprising. HNK-1 ST has 21.4% identical amino acids yet transfers sulfate to terminal H9252-UA-GalNAc sequences, whereas the highest percentage of identical amino acids among the HNK-1 sulfotransferases and how these relate to the structural features of the sulfotransferases.
D4ST-1 is located on human chromosome 15q14, while HNK-1 ST is on chromosome 2 (GenBank™ accession number AC012493). GalNAc-4-ST1 is on chromosome 19q13.1 (15), GalNAc-4-ST2 is on chromosome 18q11.2 (16), C4ST-1 is on chromosome 12q23 (12), and C4ST-2 is on chromosome 7p22 (12). D4ST-1, like C4ST-2 (12), appears to be encoded by a single exon. In contrast, GalNAc-4-ST1 and GalNAc-4-ST2 are encoded by multiple exons. The additional exons in GalNAc-4-ST1 and GalNAc-4-ST2 encode the cytosolic domain, transmembrane domain, and much of the stem region. The stem region appears to confer the specificity for terminal rather than internal GalNAc on GalNAc-4-ST1 and GalNAc-4-ST2 (15, 16). While 5'-rapid amplification of cDNA ends and attempts to amplify predicted sequences did not yield the entire 5'-UTR for the D4ST-1 cDNA, there are a number of indications that the entire coding sequence for the ORF is contained in the cDNA sequence we have obtained. This conclusion is based on the following. 1) There is a hydrophobic sequence near the N terminus that is predicted to function as a transmembrane domain and would make D4ST-1 a type II transmembrane protein as has been seen for both saccharide specific sulfotransferases and for glycosyltransferases. 2) When expressed in CHO cells, D4ST-1 is appropriately directed to the lumen of the endoplasmic reticulum. 3) Despite the high percentage of sequence identity in the coding regions of mouse and human D4ST-1 and the lower but still significant percentage of sequence identity in the 3'-UTRs of the human and mouse sequences, searches of available data bases with the putative mouse 5'-UTR did not reveal any human target sequence with significant homology. The orthologous mouse and human ORFs are of identical size, and all of the encoded protein domains of D4ST-1 have been conserved. 4) There is a CpG island flanking the start codon of the D4ST-1 ORF we have described, indicative of a promoter region. This finding further suggests that D4ST-1 has a short 5'-UTR with a complex structure due to the high G + C content of this region.

The HNK-1 family of sulfotransferases is proving to be extensive. By examining structurally, the various members of this family display significant differences in their specificities, in the structures that are synthesized, in the regulation of their expression, and in their biological functions. D4ST-1 and C4ST-1 play key roles in the sulfation of dermatan and chondroitin, respectively. The role of C4ST-2 in sulfation of dermatan and chondroitin should become clear as a more detailed understanding of its specificity is developed. In contrast to D4ST-1, C4ST-1, and C4ST-2, HNK-1 ST, GalNAc-4-ST1, and GalNAc-4-ST2 are more highly restricted in their expression, and the structures they produce are only found on limited numbers of glycoproteins. HNK-1-bearing structures are involved in cell interactions in the nervous system (43, 44), whereas N-linked oligosaccharides with terminal β1,4-linked GalNAc-4- SO₄, the product of GalNAc-4-ST1, on glycoprotein hormones such as luteinizing hormone play a critical role in vitro by regulating its circulatory half-life (45–47). Future structure function studies will provide insights into the biology as well as the specificity of this novel family of sulfotransferases.

Acknowledgments—We are grateful to Dr. J. J. Hopwood (Women's and Children's Hospital, Adelaide, Australia) for purified, recombinant GalNAc-4-sulfatase, to Drs. Kerstin Kutsche and Andres Veske (University Hospital Eppendorf, Hamburg, Germany) for cDNA pools, and to Dr. Christian Kubisch (University of Bonn) for helpful suggestions.