



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

DOI:

[10.1074/jbc.M105848200](https://doi.org/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.



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 β 1,4-linked GalNAc that is substituted with an α -linked iduronic acid (IdoUA) at the C-3 hydroxyl. D4ST-1 shows a strong preference *in vitro* for sulfate transfer to IdoUA α 1,3GalNAc β 1,4 that is flanked by GlcUA β 1,3GalNAc β 1,4 as compared with IdoUA α 1,3GalNAc β 1,4 flanked by IdoUA α 1,3GalNAc β 1,4. 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.

confers biological function to these saccharides (1–3). Chondroitin sulfate proteoglycans consist of repeating units of (GlcUA β 1,3GalNAc β 1,4)_n¹ 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 glucuronylC5-epimerase to produce dermatan sulfate, also referred to as chondroitin sulfate B, a glycan consisting of (IdoUA α 1,3GalNAc β 1,4)_n and (GlcUA β 1,3GalNAc β 1,4)_n disaccharides (4, 5). The proportion of IdoUA and GlcUA in dermatan 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 GlcUA 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-glucuronylsulfotransferase that is responsible for synthesis of the HNK-1 epitope SO₄-3-GlcUA β 1,3Gal β 1,4GlcNAc-R (13, 14). Two additional GalNAc-4-O-sulfotransferases, GalNAc-4-ST1 and GalNAc-4-ST2, that add sulfate to terminal β 1,4-linked GalNAc in the sequence GalNAc β 1,4GlcNAc-R 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

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

* This work was supported by National Institutes of Health Grant R01-DK41738 (to J. U. B.), by Deutsche Forschungsgemeinschaft Grant SCHA185/15-1 (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.

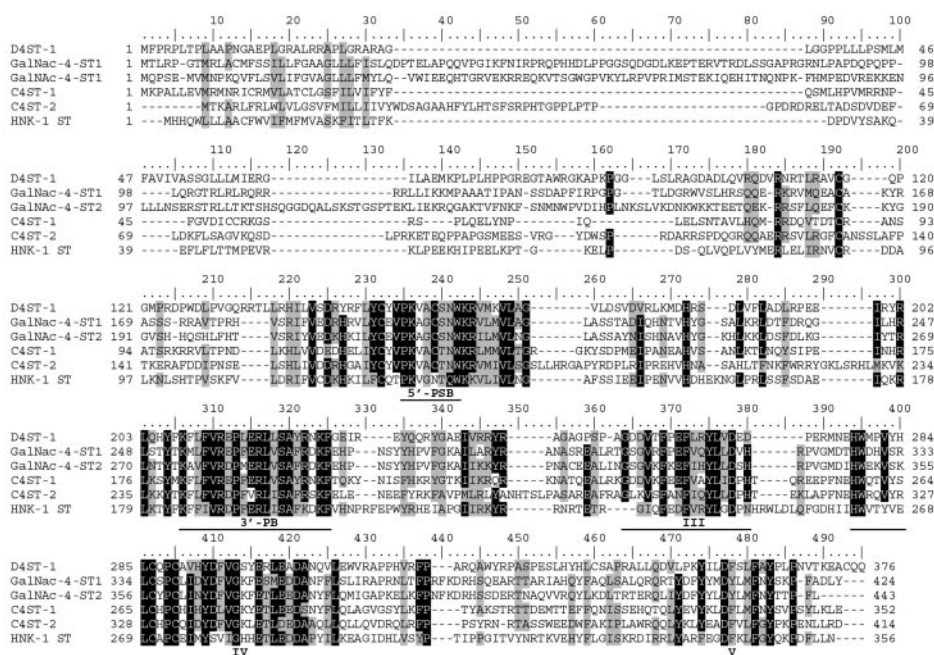
The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AF401222.

§ 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.

¹ 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'-phosphosulfonate; UTR, untranslated region; PCR, polymerase chain reaction; HPLC, high pressure liquid chromatography; D4ST, dermatan-4-sulfotransferase-1; Δ Di-4S, D-glucono-4-eneopyranoside β -1,3-GalNAc-4-SO₄; Δ Di-6S, D-glucono-4-eneopyranoside β -1,3-GalNAc-6-SO₄; Δ UA, unsaturated uronic acid; Δ Di-di S_{EP}, D-gluco-4-eneopyranoside β -1,3-GalNAc-4,6-diSO₄.

FIG. 2. Multiple amino acid sequence alignment of human D4ST-1, GalNAc-4-ST1, GalNAc-4-ST2, C4ST-1, C4ST-2, and HNK-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, C4ST-1, and C4ST-2—CHO/Tag cells were transfected with 13 μ g of pcDNA3.1-D4ST-1, pcDNA3.1-C4ST-1, pcDNA3.1-C4ST-2, or pcDNA3.1 using 35 μ g 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 μ l of 20 mM HEPES buffer, pH 7.4, 5 mM MgCl₂, 175 mM KCl, 2% Triton X-100, and protease inhibitors (23 millitrypsin inhibitor units of aprotinin and 4 μ g 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 \times g for 20 min. The supernatant was designated as the cell extract. The culture medium was pooled and sedimented at 12,000 \times 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 μ l) contained 50 mM imidazole-HCl, pH 6.8, 2 mM dithiothreitol, 5 \times 10⁵ cpm of ³⁵S-labeled 3'-phosphoadenosine-5'-phosphosulfate (PAPS), 2 μ M 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. Incorporation was determined to be linear for 3 h under these conditions. Assays were carried out for 2 h at 37 $^{\circ}$ C and terminated by boiling for 3 min. The [³⁵S]SO₄-labeled chondroitin and dermatan product was 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 μ l of H₂O and separated from any remaining [³⁵S]PAPS and degradation products by passage over Bio-Spin 6 columns (Bio-Rad). Transfer to GalNAc β 1,4GlcNAc β 1,2Man α -(CH₂)₃-COOCH₃ (25) and GlcUA β 1,3Gal β 1,4GlcNAc β 1,3Gal β 1,4Glc β 1,2H₄NHCOCF₃ was assayed as described (14, 15).

Product Characterization—³⁵S]SO₄-dermatan and [³⁵S]SO₄-chondroitin products were digested with 50 milliunits of chondroitinase AC I (Calbiochem) in 100 mM Tris acetate buffer, pH 7.3, for 16 h at 37 $^{\circ}$ C or 25 milliunits of chondroitinase B (Calbiochem) in 100 mM Tris acetate buffer, pH 8.0, for 16 h at 37 $^{\circ}$ C. The digestion products were analyzed by HPLC on a 4.6 \times 250-mm MicroPak AX-5 column (Varian) using a linear gradient of 10 mM KH₂PO₄ to 450 mM KH₂PO₄ over 40 min with a flow rate of 1.0 ml/min (26). 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 16 \times 500-mm SuperdexTM 30 preparative grade gel filtration column (Amersham Pharmacia Biotech) eluted at 1.0 ml/min in 100 mM NH₄HCO₃.

Oligosaccharides were reduced as indicated by the addition of 50 mM NaBH₄, 50 mM sodium acetate, pH 7.0, and incubation for 30 min at 37 $^{\circ}$ C. Reductions were terminated by placing the samples on ice and adding acetic acid until the borohydride was discharged. The reduced samples were dried under vacuum, resuspended in 1 ml of methanol, and dried under vacuum until excess borate was removed. Terminal unsaturated uronic acid was removed by treatment with 35 mM mercuric acetate, pH 5.0, for 30 min at 25 $^{\circ}$ C (27). Residual mercury was removed by passage over a Dowex 50W-X8 H⁺ column and washing with H₂O. Digestion with 30 milliunits of β -N-acetylhexosaminidase_r (New England BioLabs) was carried out in 50 mM sodium acetate buffer, pH 4.5. Digestion with 0.2 milliunits of α -L-iduronidase (Glyko) was carried out in 50 mM sodium acetate buffer, pH 5.0. All products were analyzed by HPLC on a MicroPak AX-5 column as described above.

Northern Blot and Expression Array Analysis—Human Multiple Tissue Northern (MTN®) blots and Human Multiple Tissue Expression (MTETM) arrays were purchased from CLONTECH. They were hybridized with 5–15 \times 10⁶ cpm of a specific α -³²P-labeled cDNA probe (α -³²P]dCTP and MegaprimeTM labeling kit purchased from Amersham Pharmacia Biotech) and washed according to the manufacturer's specifications. The membranes were exposed to Biomax MS films (Eastman Kodak Co.) for 2–5 days at –80 $^{\circ}$ C with intensifying screens. The 425-bp probe used for the labeling reactions corresponds to nt 673–1097 of the D4ST-1 cDNA (GenBankTM accession number AF401222).

RESULTS

Identification of a Human cDNA Related to HNK-1 ST—We recently identified and characterized two human GalNAc-4-O-sulfotransferases, GalNAc-4-ST1 (15) and GalNAc-4-ST2 (16), based on their homology to the HNK-1 ST. HNK-1 ST produces the HNK-1 epitope SO₄-3-GlcUA β 1,3Gal β 1,4GlcNAc-R by transferring sulfate to the acceptor GlcUA β 1,3Gal β 1,4GlcNAc-R (13, 14). In contrast, GalNAc-4-ST1 and GalNAc-4-ST2 both transfer sulfate to the C-4 hydroxyl of terminal β 1,4-linked GalNAc present on N-linked oligosaccharides such as those found on the glycoprotein hormones lutropin and thyrotropin (28). C4ST-1 and C4ST-2 are also GalNAc-4-O-sulfotransferases that are homologous to HNK-1 ST; however, C4ST-1 and C4ST-2 transfer sulfate to the C-4 hydroxyl of internal β 1,4-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 HNK-1

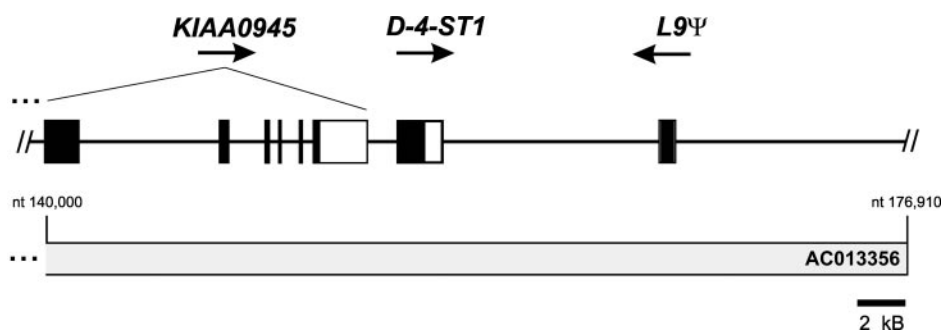


FIG. 3. **Structure of the human D4ST-1 gene.** Exonic sequences that contribute to coding regions are boxed and shaded in *black*, while untranslated regions are only *boxed*. The intronless D4ST-1 gene is found between the genes encoding the KIAA0945 protein (the 3'-terminal part of the *KIAA094* gene is shown) and a pseudogene copy of the ribosomal L9 protein (termed *L9Ψ*). 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.

TABLE I

D4ST-1, *C4ST-1*, and *C4ST-2* differ in their rates of sulfate transfer to chondroitin and dermatan

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 μ l of 1% Triton X-100, and 25 μ l of cell extract was utilized for a 2-h assay. The medium was collected and concentrated to 2 ml by ultrafiltration, and 25 μ l was utilized. Results are expressed as pmol incorporated/h/100-mm plate.

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

ST (GenBank™ accession no. AF033827) identified a 177-kilobase pair 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-

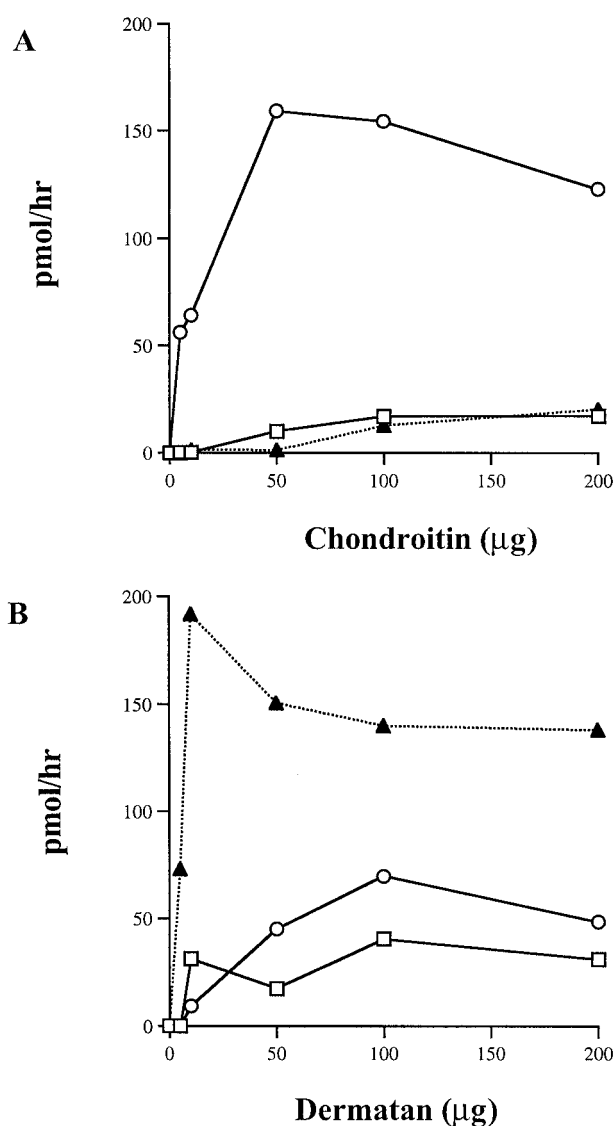


FIG. 4. **Concentration dependence for transfer of sulfate to chondroitin and dermatan by D4ST-1, C4ST-1, and C4ST-2.** D4ST-1 (▲), C4ST-1 (○), and C4ST-2 (□) were incubated with increasing concentrations (0–200 μ g per 50 μ l assay) of desulfated chondroitin (A) or desulfated dermatan (B) for 2 h at 37 °C, and the amount of [35 S]SO $_4$ incorporated was determined. The amount of [35 S]SO $_4$ incorporated by an equal amount of extract from cells transfected with pcDNA3.1 has been subtracted.

emphasized, especially in the highly variable N-terminal regions of this particular enzyme family. Pairwise alignments using the Smith-Waterman algorithm, however, confirmed

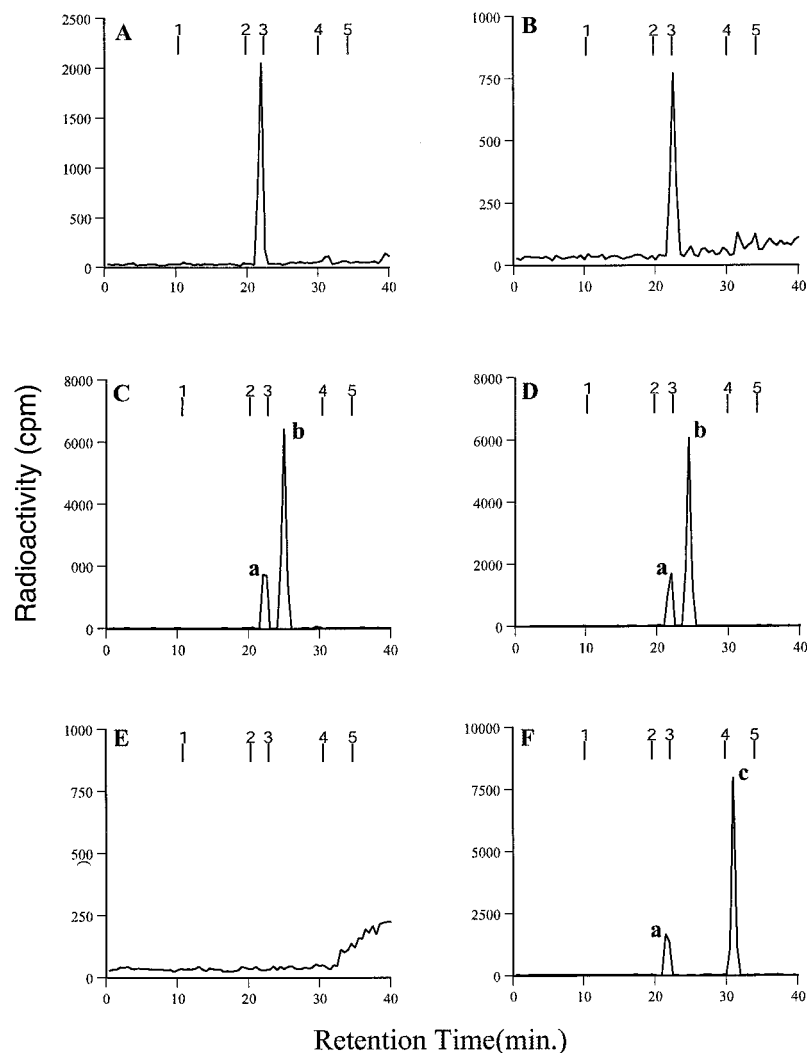


FIG. 5. Characterization of the $[^{35}\text{S}]\text{SO}_4$ -chondroitin and $[^{35}\text{S}]\text{SO}_4$ -dermatan products produced by D4ST-1 and C4ST-1. Desulfated chondroitin (50 μg) (A, C, and E) and desulfated dermatan (50 μg) (B, D, and F) were incubated with C4ST-1 (A and B) or D4ST-1 (C–F). The $[^{35}\text{S}]\text{SO}_4$ -labeled products were isolated by gel filtration and digested with chondroitinase AC I (A–D) or chondroitinase B (E and F). The digestion products were analyzed on a MicroPak AX-5 column (Varian) developed with a linear gradient of 10–450 mM KH_2PO_4 over 40 min at a flow rate of 1.0 ml/min. The elution positions of standards are as follows: GalNAc-4- SO_4 (1), $\Delta\text{Di-6S}$ (2), $\Delta\text{Di-4S}$ (3), D-gluco-4-enepyranoside β -1,3 GalNAc-4,6-di SO_4 ($\Delta\text{Di-diS}_E$) (4), SO_4 , free sulfate (5).

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 data base at the NCBI. A mouse cDNA clone 2600016L03 (GenBankTM 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%). The 92 nts upstream of the start codon of the mouse D4ST-1 transcript, probably a part of the 5'-UTR, were used as a query sequence to search the nonredundant and High Throughput Genomic Sequences data bases at the NCBI. Regions showing significant homology to this stretch could not be detected in human genomic DNA, although most of the sequences of chromosome 15 have been released and the complete sequence of the gap between the D4ST-1 and the *KIAA0945* (see below) locus is known. Since coding regions are typically more highly conserved between orthologs than their UTRs, the absence of human genomic sequences homologous to the 5'-end of the mouse cDNA clone indicates that the query sequence

represents the murine 5'-UTR and that we have cloned the complete human ORF.

Genomic Organization and Chromosome Localization of D4ST-1—The genomic BAC clone RP11-64K12 (GenBankTM 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 (GenBankTM 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 (GenBankTM accession number BC004156) that probably represents a pseudogene (31). An analysis of splice sites directly upstream of the start codon of D4ST-1 with a minimum score of 0.9 determined the nearest 3'/acceptor splice site to be 532 bp upstream (analysis performed with Splice Site Prediction as available at the Berkeley Drosophila Genome project Web site). 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).

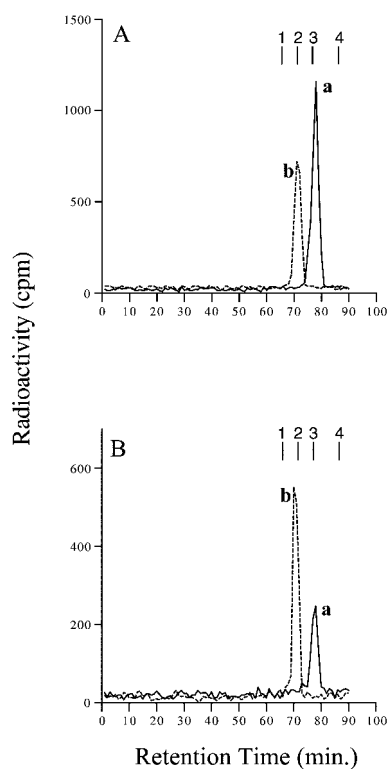


FIG. 6. Characterization of the [^{35}S]SO $_4$ -labeled products produced by digestion with chondroitinase AC I by gel filtration on Superdex 30. The [^{35}S]SO $_4$ -labeled products separated by chromatography on Micro Pak AX-5 (Fig. 5) following incubation of desulfated chondroitin and desulfated dermatan with D4ST-1 and digestion with chondroitinase AC I were subjected to gel filtration of Superdex 30. A, analysis of peak a (solid line) and peak b (broken line) obtained in Fig. 5C. B, analysis of peak a (solid line) and peak b (broken line) obtained in Fig. 5D. The positions of standards are as follows: chondroitin hexasaccharide (1), chondroitin tetrasaccharide (2), $\Delta\text{Di-4S}$ (3), SO $_4$, free sulfate (5).

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, GalNAc-4-ST1, and GalNAc-4-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 GlcUA β 1,3Gal β 1,4GlcNAc-R or terminal GalNAc in GalNAc β 1,4GlcNAc β 1,2Man-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 transmem-

brane 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 transferases including GalNAc-4-ST1 and GalNAc-4-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 predominantly 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 (GlcUA β 1,3GalNAc β 1,4) $_n$ sequences (11, 12). Since dermatan is known to contain GlcUA β 1,3GalNAc β 1,4 sequences interspersed with the IdoUA α 1,3GalNAc β 1,4 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 ^{35}S -sulfated chondroitin and dermatan products were digested with chondroitinase AC I and chondroitinase B. Digestion of both the [^{35}S]SO $_4$ -chondroitin and [^{35}S]SO $_4$ -dermatan products obtained with C4ST-1 yielded a single peak that comigrated with D-glucono-4-enepyranoside β -1,3-GalNAc-4-SO $_4$ ($\Delta\text{Di-4S}$) (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 [^{35}S]SO $_4$ -labeled chondroitin and dermatan products produced by D4ST-1 produced a minor peak (peak a) that comigrated with $\Delta\text{Di-4S}$ 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 [^{35}S]SO $_4$ -chondroitin or [^{35}S]SO $_4$ -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

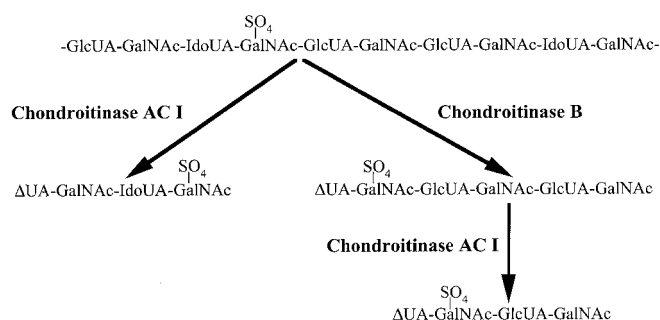


FIG. 7. Schematic illustrating the location of the sulfate incorporated into dermatan by D4ST-1 and the products that are produced by digestion with chondroitinase AC I and chondroitinase B.

is adjacent to an IdoUA. Further, since chondroitinase AC I cleaves the GalNAc β 1,4GlcUA but not GalNAc β 1,4IdoUA, the tetrasaccharide was predicted to arise from an [IdoUA-GalNAc] flanked by [GlcUA-GalNAc] as shown in Fig. 7 and to have the sequence Δ UA β 1,3GalNAc β 1,4IdoUA α 1,3GalNAc. 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 (Δ UA). 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 Δ Di-4S (Fig. 8B) consistent with the structure IdoUA α 1,3GalNAc α 1,4-SO $_4$. Digestion with α -L-iduronidase generated a product that migrated at the position of SO $_4$ -4-GalNAc (Fig. 8C).

Digestion of the D4ST-1 [35 S]SO $_4$ -dermatan product with chondroitinase B yielded two [35 S]SO $_4$ -labeled products when analyzed by HPLC on MicroPak AX-5 column (Fig. 5F). The minor product, *peak a*, migrated at the location of Δ Di-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 [35 S]SO $_4$ -chondroitin product was digested with chondroitinase B (Fig. 5E). This suggests that the major product of chondroitinase B digestion is Δ UA β 1,3GalNAc-4-SO $_4$ - β 1,4GlcUA β 1,3GalNAc β 1,4GlcUA β 1,3GalNAc as shown in Fig. 7. The absence of either Δ Di-4S or this hexasaccharide in the chondroitinase B digestion product of [35 S]SO $_4$ -chondroitin produced by D4ST-1 indicates that the chondroitin is not contaminated with dermatan but does have an occasional IdoUA α 1,3GalNAc β 1,4 sequence that is a substrate of the D4ST-1. Digestion of the hexasaccharide with chondroitinase AC I converted the [35 S]SO $_4$ -hexasaccharide (Fig. 9A, *peak b*, and Fig. 9B) to a [35 S]SO $_4$ -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 (See Fig. 7).

As was seen with the tetrasaccharide product of chondroitinase AC I digestion of the [35 S]SO $_4$ -dermatan, treatment of the [35 S]SO $_4$ -hexasaccharide produced by chondroitinase B digestion with Hg(Ac) $_2$ to remove the unsaturated uronic acid shifted the [35 S]SO $_4$ -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.

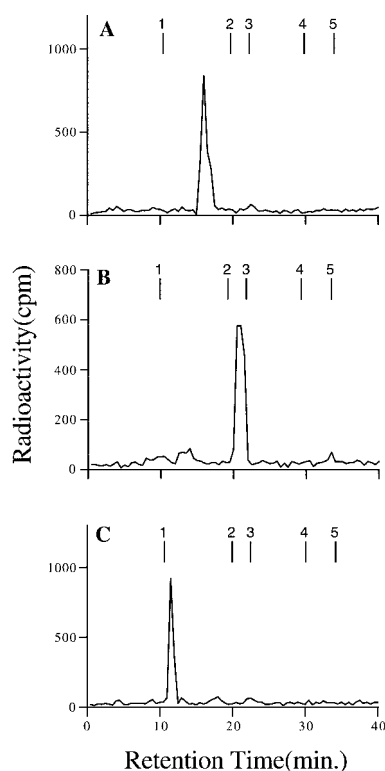


FIG. 8. Characterization of the [35 S]SO $_4$ -tetrasaccharide product obtained by chondroitinase AC I digestion of the D4ST-1 dermatan product. *Peak b* in Fig. 5D was isolated and examined by chromatography on Micro Pak AX-5 following reduction and treatment with mercuric acetate to remove the unsaturated uronic acid (A), digestion of the material shown in A with β -N-acetylhexosaminidase (B), or digestion of the material shown in B with α -L-iduronidase (C). The elution positions of standards are as follows: GalNAc-4-SO $_4$ (1), Δ Di-6S (2), Δ Di-4S (3), Δ Di-diS $_E$ (4), SO $_4$, free sulfate (5).

6A). Digestion of the [35 S]SO $_4$ -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 (arylsulfatase B) resulted in the appearance of free [35 S]SO $_4$ (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 [35 S]SO $_4$ -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 Δ Di-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 Δ Di-4S in the chondroitinase AC I digests 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

FIG. 9. Digestion of chondroitinase B $[^{35}\text{S}]\text{SO}_4$ -labeled product from dermatan with chondroitinase AC I. A, $[^{35}\text{S}]\text{SO}_4$ -labeled dermatan 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-SO₄ (1); $\Delta\text{Di-6S}$ (2), $\Delta\text{Di-4S}$ (3), $\Delta\text{Di-diS}_E$ (4), SO₄, free sulfate (5). B and D, chondroitin hexasaccharide (1), chondroitin tetrasaccharide (2), $\Delta\text{Di-4S}$ (3), SO₄, free sulfate (4).

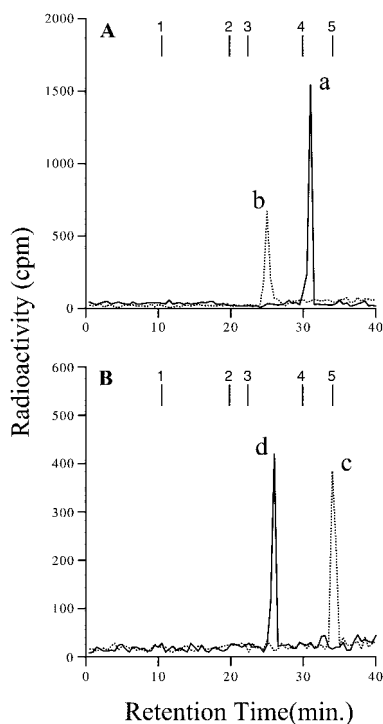
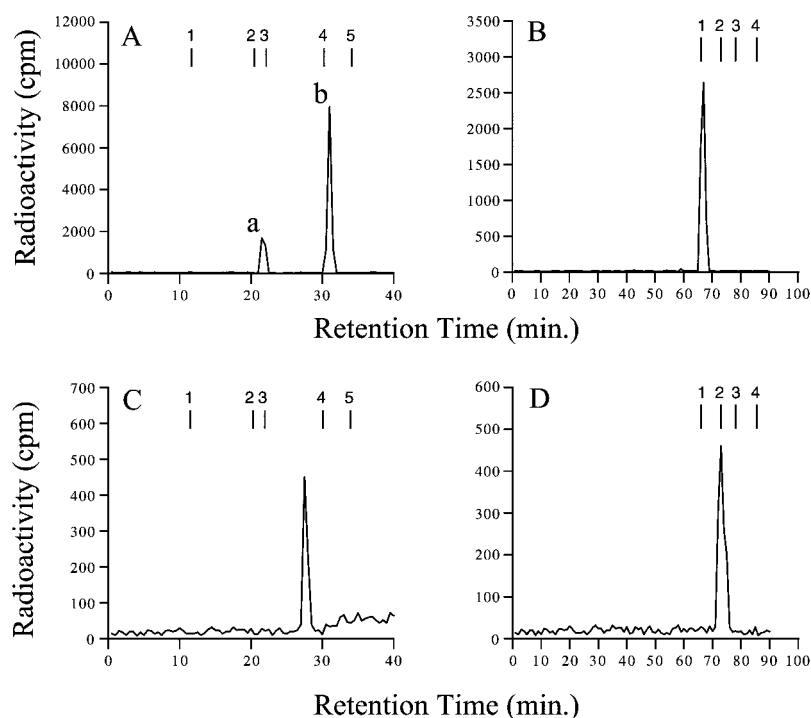


FIG. 10. Characterization of the $[^{35}\text{S}]\text{SO}_4$ -hexasaccharide product obtained by chondroitinase B digestion of the D4ST-1 dermatan product. A, *peak c* in Fig. 5F was isolated and examined by chromatography on Micro Pak AX-5 prior to (*solid line, peak a*) and following treatment with mercuric 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 β -N-acetylgalactosaminidase (*solid line, peak d*) or N-acetylgalactosamine-4-O-sulfatase (arylsulfatase B) (*broken line, peak c*) and analyzed on the same MicroPak AX-5 column. The elution positions of standards are as follows: GalNAc-4-SO₄ (1), $\Delta\text{Di-6S}$ (2), $\Delta\text{Di-4S}$ (3), $\Delta\text{Di-diS}_E$ (4), SO₄, free sulfate (5).

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).

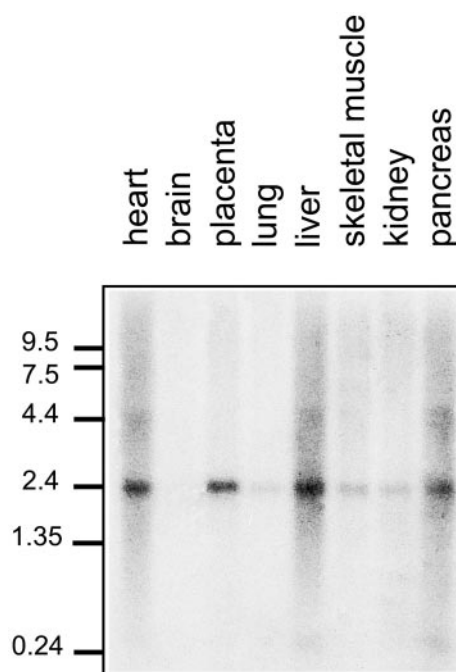
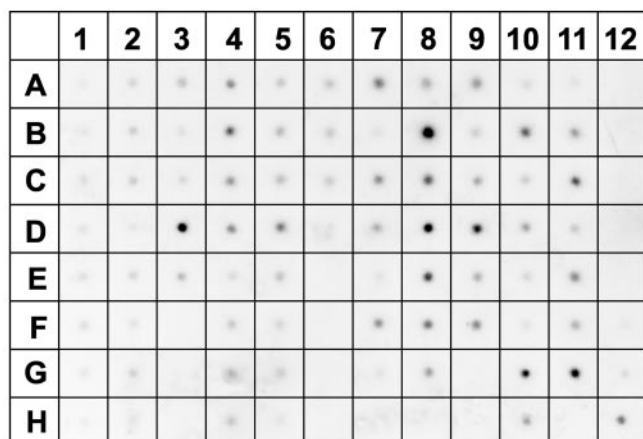


FIG. 11. Northern blot analysis of D4ST-1 transcripts. Each lane of the MTN® Northern blot contains 2 μg poly(A)⁺ RNA and was hybridized with a ^{32}P -labeled cDNA probe specific for human D4ST-1 (see "Experimental Procedures"). Tissues used to prepare the RNA are indicated above each lane. Migration positions of standards are indicated at the left.

DISCUSSION

Dermatan-4-O-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 α -linked IdoUA (*i.e.* IdoUA-GalNAc). As a result D4ST-1 is only active on dermatan type structures (*i.e.* chondroitins that contain IdoUA). IdoUA makes up 90% of the uronic acid in the mature dermatan isolated from porcine intestine that was used for these studies. While small amounts of $\Delta\text{UA-GalNAc-4-SO}_4$



A1 brain, whole	B1 cerebral cortex	C1 frontal lobe	D1 parietal lobe
A2 cerebellum, left	B2 cerebellum, right	C2 corpus callosum	D2 amygdala
A3 substantia nigra	B3 accumbens nucleus	C3 thalamus	D3 pituitary gland
A4 heart	B4 aorta	C4 atrium, left	D4 atrium, right
A5 esophagus	B5 stomach	C5 duodenum	D5 jejunum
A6 colon, transverse	B6 colon, descending	C6 rectum	D6 empty
A7 kidney	B7 skeletal muscle	C7 spleen	D7 thymus
A8 lung	B8 placenta	C8 bladder	D8 uterus
A9 liver	B9 pancreas	C9 adrenal gland	D9 thyroid gland
A10 leukemia, HL-60	B10 HELA S3	C10 leukemia, K-562	D10 leukemia, MOLT-4
A11 brain, fetal	B11 heart, fetal	C11 kidney, fetal	D11 liver, fetal
A12 yeast total RNA	B12 yeast tRNA	C12 E.coli rRNA	D12 E. coli DNA
E1 occipital lobe	F1 temporal lobe	G1 p.g. of cerebral cortex*	H1 pons
E2 caudate nucleus	F2 hippocampus	G2 medulla oblongata	H2 putamen
E3 spinal cord	F3 empty	G3 empty	H3 empty
E4 ventricle, left	F4 ventricle, right	G4 interventricular space	H4 apex of the heart
E5 ileum	F5 ilocecum	G5 appendix	H5 colon, ascending
E6 empty	F6 empty	G6 empty	H6 empty
E7 periph. blood leuk.**	F7 lymph node	G7 bone marrow	H7 trachea
E8 prostate	F8 testis	G8 ovary	H8 empty
E9 salivary gland	F9 mammary gland	G9 empty	H9 empty
E10 Burkitt's ***	F10 Burkitt's ****	G10 SW480 *****	H10 lung carcinoma, A549
E11 spleen, fetal	F11 thymus, fetal	G11 lung, fetal	H11 empty
E12 Poly r(A)	F12 human C _α -1 DNA	G12 human DNA, 100 ng	H12 human DNA, 500 ng

FIG. 12. RNA dot blot analysis of D4ST-1 transcripts. The human Multiple Tissue Expression (MTETM) array shown was hybridized with a ³²P-labeled human D4ST-1-specific cDNA probe (see "Experimental Procedures"). Tissue sources for the RNA are indicated below the blot. *, paracentral gyrus of cerebral cortex; **, peripheral blood leukocytes; ***, Burkitt's lymphoma Raji; ****, Burkitt's lymphoma Daudi; *****, colorectal adenocarcinoma, SW280.

were released from dermatan by chondroitinase B indicating that they originated from IdoUA-GalNAc-4-SO₄-IdoUA-GalNAc, the vast majority of IdoUA-GalNAc sequences that were sulfated by D4ST-1 *in vitro* were flanked on both sides by GlcUA-GalNAc (see Fig. 7). While such sequences have been described in dermatan (37, 38), they represent minor components in the mature proteoglycan. The selective sulfation of IdoUA-GalNAc that is flanked by GlcUA-GalNAc suggests that sulfation of GalNAc in dermatan is immediately followed by the epimerization of GlcUA to IdoUA. Following sulfate addition to GalNAc, the glucuronyl C5-epimerase is no longer able to convert IdoUA back to GlcUA (39). This raises the possibility that chondroitin sequences are converted to dermatan by the processive action of the glucuronyl C5-epimerase and D4ST-1.

Previous studies by a number of investigators would support such a mechanism. For example, the amount of IdoUA in dermatan produced by cultured fibroblasts is enhanced by increasing the levels of sulfate in the culture medium (6). This supported the hypothesis that there is a relationship between epimerization and sulfation. Furthermore, the addition of PAPS to microsomes from fibroblasts enhances the *in vitro* epimerization of GlcUA to IdoUA (40, 41), again suggesting that epimerization and sulfation are linked. Since the *D*-gluco configuration is favored over the *L*-ido 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, GalNAc-4-ST1 and GalNAc-4-ST2 both transfer sulfate to GalNAc β 1,4GlcNAc β -sequences found at the termini of *N*- and *O*-linked oligosaccharides (15, 16). Since deletion of the stem region from both GalNAc-4-ST1 and GalNAc-4-ST2 converts these terminal sulfotransferases to forms that will transfer sulfate to internal GlcUA-GalNAc sequences as well as to terminal GalNAc c β 1,4GlcNAc β -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 the C-3 hydroxyl of terminal GlcUA β 1,3Gal β 1,4GlcNAc β - (13, 14). The percentage of identical amino acids does not allow a prediction as to the specificity for internal *versus* terminal β 1,4-linked GalNAc or terminal β 1,3-linked GlcUA.

While D4ST-1 requires an IdoUA-GalNAc for transfer it also appears to be sensitive to the character of the flanking uronic acid moieties favoring surrounding sequences that contain GlcUA-GalNAc over those with IdoUA-GalNAc. This suggests that the binding site may accommodate more than just the IdoUA-GalNAc. It will be of interest to determine the minimum features of the saccharide sequence that are required for recognition by the various members of the HNK-1 family of 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. While related 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 β 1,4-linked GalNAc-4-SO₄, the product of GalNAc-4-ST1, on glycoprotein hormones such as luteinizing hormone play a critical role *in vivo* 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.

REFERENCES

- Hooper, L. V., Manzella, S. M., and Baenziger, J. U. (1996) *FASEB J.* **10**, 1137–1146
- Bowman, K. G., and Bertozzi, C. R. (1999) *Chem. Biol.* **6**, R9–R22
- Hemmerich, S., and Rosen, S. D. (2000) *Glycobiology* **10**, 849–856
- Kjell n, L., and Lindahl, U. (1991) *Annu. Rev. Biochem.* **60**, 443–475
- Prydz, K., and Dalen, K. T. (2000) *J. Cell Sci.* **113**, 193–205
- Silbert, J. E., Palmer, M. E., Humphries, D. E., and Silbert, C. K. (1986) *J. Biol. Chem.* **261**, 13397–13400
- Nadanaka, S., Fujita, M., and Sugahara, K. (1999) *FEBS Lett.* **452**, 185–189
- Cheng, F., Heinegard, D., Malmstrom, A., Schmidtchen, A., Yoshida, K., and Fransson, L. A. (1994) *Glycobiology* **4**, 685–696
- Iozzo, R. V. (1998) *Annu. Rev. Biochem.* **67**, 609–652
- Bovolenta, P., and Feraud-Espinosa, I. (2000) *Prog. Neurobiol.* **61**, 113–132
- Yamauchi, S., Mita, S., Matsubara, T., Fukuta, M., Habuchi, H., Kimata, K., and Habuchi, O. (2000) *J. Biol. Chem.* **275**, 8975–8981
- Hiraoka, N., Nakagawa, H., Ong, E., Akama, T. O., Fukuda, M. N., and Fukuda, M. (2000) *J. Biol. Chem.* **275**, 20188–20196
- Ong, E., Yeh, J. C., Ding, Y., Hindsgaul, O., and Fukuda, M. (1998) *J. Biol. Chem.* **273**, 5190–5195
- Bakker, H., Friedmann, I., Oka, S., Kawasaki, T., Nifant'ev, N., Schachner, M., and Mantei, N. (1997) *J. Biol. Chem.* **272**, 29942–29946
- Xia, G., Evers, M. R., Kang, H.-G., Schachner, M., and Baenziger, J. U. (2000) *J. Biol. Chem.* **275**, 38402–38409
- Kang, H. G., Evers, M. R., Xia, G., Baenziger, J. U., and Schachner, M. (2001) *J. Biol. Chem.* **276**, 10861–10869
- Salamov, A. A., and Solov'yev, V. V. (2000) *Genome Res.* **10**, 516–522
- Gardiner-Garden, M., and Frommer, M. (1987) *J. Mol. Biol.* **196**, 261–282
- Milanesi, L., D'Angelo, D., and Rogozin, I. B. (1999) *Bioinformatics* **15**, 612–621
- Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* **22**, 4673–4680
- Smith, T. F., and Waterman, M. S. (1981) *J. Mol. Biol.* **147**, 195–197
- Kozak, M. (1991) *J. Biol. Chem.* **266**, 19867–19870
- Lennon, G., Auffray, C., Polymeropoulos, M., and Soares, M. B. (1996) *Genomics* **33**, 151–152
- Nagasawa, K., Inoue, Y., and Tokuyasu, T. (1979) *J. Biochem. (Tokyo)* **86**, 1323–1329
- Skelton, T. P., Hooper, L. V., Srivastava, V., Hindsgaul, O., and Baenziger, J. U. (1991) *J. Biol. Chem.* **266**, 17142–17150
- Baenziger, J. U. (1994) *Methods Enzymol.* **230**, 237–249
- Yang, H. O., Gunay, N. S., Toida, T., Kuberan, B., Yu, G., Kim, Y. S., and Linhardt, R. J. (2000) *Glycobiology* **10**, 1033–1039
- Baenziger, J. U., and Green, E. D. (1991) in *Biology of Carbohydrates* (Ginsberg, V., and Robbins, P. W., eds) Vol. 3, pp. 1–46, JAI Press Ltd., London
- Kyte, J., and Doolittle, R. F. (1982) *J. Mol. Biol.* **157**, 105–322
- Nagase, T., Ishikawa, K., Suyama, M., Kikuno, R., Hirose, M., Miyajima, N., Tanaka, A., Kotani, H., Nomura, N., and Ohara, O. (1999) *DNA Res.* **6**, 63–70
- Mazuruk, K., Schoen, T. J., Chader, G. J., Iwata, T., and Rodriguez, I. R. (1996) *Biochim. Biophys. Acta* **1305**, 151–162
- Jones, P. A. (1999) *Trends Genet.* **15**, 34–37
- Larsen, F., Gundersen, G., Lopez, R., and Prydz, H. (1992) *Genomics* **13**, 1095–1107
- Paulson, J. C., and Colley, K. J. (1989) *J. Biol. Chem.* **264**, 17615–17618
- Colley, K. J. (1997) *Glycobiology* **7**, 1–13
- Gu, K., Liu, J., Pervin, A., and Linhardt, R. J. (1993) *Carbohydr. Res.* **244**, 369–377
- Fransson, L. A., Coster, L., Havasmark, B., Malmstrom, A., and Sjoberg, I. (1974) *Biochem. J.* **143**, 379–389
- Coster, L., Malmstrom, A., Sjoberg, I., and Fransson, L. (1975) *Biochem. J.* **145**, 379–389
- Malmstrom, A. (1984) *J. Biol. Chem.* **259**, 161–165
- Eklund, E., Roden, L., Malmstrom, M., and Malmstrom, A. (2000) *Arch. Biochem. Biophys.* **383**, 171–177
- Malmstrom, A., and Fransson, L. A. (1975) *J. Biol. Chem.* **250**, 3419–3425
- Rosenberg, L. C., Choi, H. U., Tang, L. H., Johnson, T. L., Pal, S., Webber, C., Reiner, A., and Poole, A. R. (1985) *J. Biol. Chem.* **260**, 6304–6313
- Jungalwala, F. B. (1994) *Neurochem. Res.* **19**, 945–957
- Schachner, M., and Martini, R. (1995) *Trends Neurosci.* **18**, 183–191
- Woodworth, A., and Baenziger, J. U. (2001) *Results Probl. Cell Differ.* **33**, 123–138
- Manzella, S. M., Hooper, L. V., and Baenziger, J. U. (1996) *J. Biol. Chem.* **271**, 12117–12120
- Baenziger, J. U. (1996) *Endocrinology* **137**, 1520–1522