cDNA cloning, genomic organization and chromosomal localization of human heparan glucosaminyl N-deacetylase/N-sulphotransferase-2

Donald E. HUMPHRIES*†[1], Julia LANCIOTTI* and Joel B. KARLINSKY*

*Department of Veterans Affairs Medical Center and the Department of Medicine, Boston University School of Medicine, Research 151, 150 S. Huntington Ave., Boston, MA 02130, U.S.A., and the †Department of Medicine, Harvard Medical School, Boston, MA 02130, U.S.A.

The cDNA and gene encoding human heparan glucosaminyl N-deacetylase/N-sulphotransferase-2 have been cloned. The cDNA encoded a protein of 883 amino acids that was 94% similar to heparan N-sulphotransferase-2 from mouse mast cells. Comparison of the deduced amino acid sequences of human heparan N-sulphotransferase-1 and -2 showed that the enzymes were 70% similar; greater than 90% of the amino acids between residues 418 and 543 were identical. The least conserved amino acids were found in the N-terminus/putative transmembrane regions of the two enzymes. The human heparan N-sulphotransferase-2 gene was localized to chromosome arm 10q (band 10q22) by *in situ* fluorescent hybridization. The gene contains 13 exons spanning 6.5 kb, ranging in size from 88 bp (exon 2) to > 1 kb (exon 1), and 12 introns, which were found to occur at similar sites within the coding sequence of the human heparan N-sulphotransferase-1 gene. The structure of the two genes differed in that the heparan N-sulphotransferase-1 gene contained one additional intron. The similarity of the heparan N-sulphotransferase-1 and -2 proteins and their similar exon-intron organization suggest that they derive from a common ancestral gene.

INTRODUCTION

The initial step in the processing of heparan polymers to heparan sulphate or heparin is N-deacetylation/N-sulphation. After N-sulphation of heparan, the polymer may undergo further modification, including epimerization of some of the glucuronates to iduronate and O-sulphation (reviewed in [1,2]). These later modifications result in a large number of possible polymer structures, and in production of biologically active oligosaccharide motifs contained within the polymer.

Since processing of heparan polymers is dependent on N-sulphation, the enzymes that catalyse this initial reaction are important in determining the final structures of heparin and heparan sulphate. Two heparan N-deacetylase/N-sulphotransferase enzymes have been cloned. The first heparan N-sulphotransferase (referred to here as heparan N-sulphotransferase-1) was cloned from rat liver and is believed to be primarily involved in heparan sulphate biosynthesis [3]. We and others have cloned the human equivalent of this enzyme [4,5]. The second heparan N-sulphotransferase (referred to here as heparan N-sulphotransferase-2) was cloned from a heparin-producing mouse mastocytoma cell line and is believed to be important for the synthesis of heparin [6,7].

The present work describes the cloning of the cDNA and gene for human heparan glucosaminyl N-deacetylase/N-sulphotransferase-2. Although the heparan N-sulphotransferase-2 gene was mapped to a different chromosome than heparan N-sulphotransferase-1, the similarity of the proteins and their similar genetic exon–intron organization suggest that they derive from a common ancestral gene.

MATERIALS AND METHODS

Amplification of human heparan N-sulphotransferase-2 cDNA

Human umbilical-vein endothelial cells (HUVEC) RNA (5 μg) was reversed transcribed using a specific oligonucleotide [5′-GTGIT(G/C)(A/G)(A/T)ACCAI(C/G)A-3′]. The 20 μl reaction contained 50 mM Tris/HCl, pH 8.3, containing 75 mM KCl, 3 mM $MgCl_2$, 10 mM DTT, 0.5 mM of each dNTP, 20 pmol primer and Moloney murine leukaemia virus reverse transcriptase (200 units of Superscript II; Gibco-BRL) for 2 h at 42 °C. PCR reactions (GeneAmp Kit; Cetus) contained 10 mM Tris/HCl, pH 8.3, containing 50 mM KCl, 1.5 mM $MgCl_2$, 0.01% (w/v) gelatin, 4 mM dNTPs, 2 μl of the prepared HUVEC cDNA, 50 pmol of each oligonucleotide primer [5′-AA(A/G)GA(A/G)-TT(C/T)TGGTGGTT(C/T)CC-3′] and [5′-AA(A/G)AA(C/T)TCCAT(A/G)TACCA(A/G)TC-3′], and 2.5 units of Taq polymerase for 35 cycles. Each cycle consisted of 30 s at 94 °C, 45 s at 50 °C and 2 min at 72 °C, and was followed by a final extension of 5 min at 72 °C. The sequences of these oligonucleotides were based on the published cDNA sequence for rat liver heparan N-sulphotransferase [3]. Amplified DNA was subcloned into pCRII (Invitrogen) and sequenced.

Hybridizations

HUVEC and IMR-90 cDNA libraries were from Clontech. A human genomic lambda Fix library was purchased from Stratagene. Nylon membranes (MSI) were prehybridized for 2 h at 42 °C in 50% (v/v) formamide/0.75 M NaCl/0.05 M sodium phosphate/5 mM EDTA/5 × Denhardt's solution/0.2% (w/v) SDS/100 μg/ml denatured salmon sperm DNA before adding ^{32}P[dCTP]- (DuPont–New England Nuclear) labelled random-primed probes. After 20 h, membranes were washed three times with 2 × SSC (1 × SSC is 0.15 M NaCl/0.015 M sodium citrate) containing 0.1% (w/v) SDS for 10 min at 42 °C. This was followed by a single stringency wash with 0.2 × SSC containing 0.1% SDS for 60 min at 60 °C. After purification of the clones, phage was isolated and inserts were subcloned into pcDNA3 (cDNA clones) or pBluescript SK− (genomic clones).

A tissue blot containing poly(A)$^+$ mRNA from several human tissues was purchased from Clontech. The blot was prehybridized for 30 min in Quickhyb (Stratagene) containing 100 μg/ml denatured salmon sperm DNA at 68 °C before hybridization with

Abbreviations used: HUVEC, human umbilical-vein endothelial cells; SSC, standard saline citrate; UTR, untranslated region.

[1] To whom correspondence should be addressed (e-mail humphries.donald_e@boston.va.gov).

a ^{32}P 3′-untranslated region (UTR) DNA fragment (bp 2787–3204) for 90 min. The blot was washed several times in 2 × SSC at room temperature and finally with 0.1 × SSC containing 0.1 % SDS for 30 min at 60 °C.

DNA sequencing

Sequencing was done by the dideoxy chain termination method [8,9] with Sequenase (Amersham) and ^{35}S[dATP] (DuPont-New England Nuclear). Compressions were resolved by repeating the reactions with deaza-7 dGTP (Amersham) or use of thermal cycling (New England Biolabs). Both strands of the cDNA were sequenced. Sequence alignments were performed with Geneworks software (Intelligenetics).

Chromosomal localization

The chromosomal localization of the heparan N-sulphotransferase-2 gene was determined by *in situ* fluorescent hybridization (Genome Systems). DNA was labelled with digoxigenin dUTP by nick translation and hybridized to normal metaphase chromosomes derived from phytohaemagglutinin-stimulated peripheral blood lymphocytes in 50 % (v/v) formamide, 10 % (w/v) dextran sulphate and 2 × SSC. Specific hybridization signals were detected by fluoresceinated anti-digoxigenin antibodies.

RESULTS

Cloning of human heparan N-sulphotransferase-2 cDNA

PCR amplification of HUVEC cDNA using primers based on the sequence of rat liver heparan N-sulphotransferase resulted in an 842 bp cDNA encoding human heparan N-sulphotransferase-2. Screening a human endothelial cell cDNA library with this DNA fragment resulted in the isolation of a single cDNA that encoded a protein whose sequence was 92 % similar to the N-terminal region of mouse heparan N-sulphotransferase-2, but which did not contain sequence encoding the C-terminal end. Screening of an oligo(dT)-primed human fibroblast cDNA library resulted in the isolation of two additional cDNAs that overlapped the original clone and contained the missing 3′ sequence. When taken together, the combined cDNA contained a short 5′-UTR of 27 bp; an open reading frame of 2649 bp encoding a protein of 883 amino acids; and a 3′-UTR of 545 bp including a polyadenylation signal (human heparan N-sulphotransferase-2; GeneBank accession number U26601).

Characterization of human heparan N-sulphotransferase-2

Comparison of the deduced amino sequence of human heparan N-sulphotransferase-2 with that from mouse mast cells showed that 94 % of the amino acids were identical. Comparison of the deduced amino acid sequences of human heparan N-sulphotransferase-1 and -2 showed that the enzymes were 70 % similar and contained many conserved regions, which are presumably important for binding of substrates and catalysis (Figure 1). A region in the middle of the enzymes was the most conserved, in that > 90 % of the amino acids between residues 418 and 543 were identical. The least conserved amino acids were found in the N-terminus/putative transmembrane regions of the two enzymes.

Tissue expression of heparan N-sulphotransferase-2 mRNA

Expression of heparan N-sulphotransferase-2 mRNA was evaluated in human tissues (Figure 2). A 3′-UTR DNA fragment of heparan N-sulphotransferase-2 (bp 2787–3204) hybridized most strongly to a ≈ 4 kb transcript and less strongly to a 4.4 kb transcript in the tissues examined. Heparan N-sulphotransferase-2 mRNA levels were highest in pancreas, with intermediate amounts in heart and placenta. The lowest levels were found in

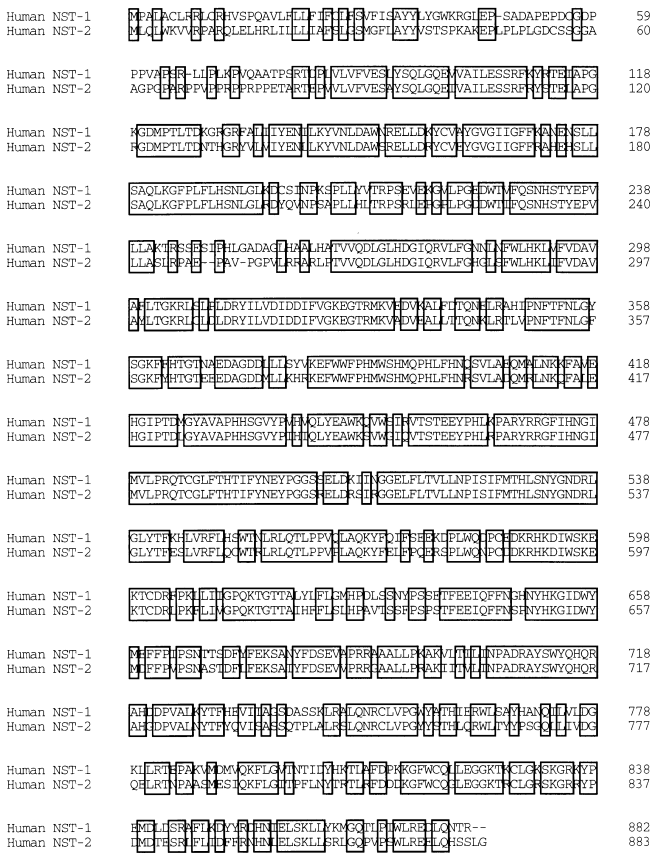

Figure 1 Comparison of deduced amino acid sequences of human heparan N-sulphotransferase-1 and human heparan N-sulphotransferase-2

The deduced amino acid sequences of human heparan N-sulphotransferase-1 (NST-1) and human heparan N-sulphotransferase-2 (NST-2) are aligned. Identical residues are boxed.

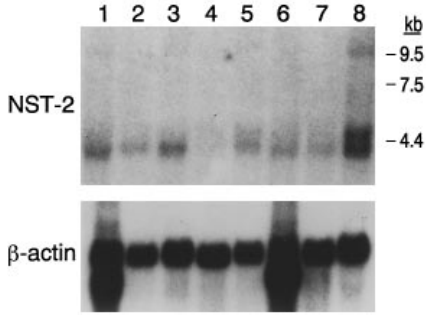

Figure 2 Expression of heparan N-sulphotransferase-2 (NST-2) mRNA in human tissues

A human multiple-tissue blot containing mRNA (2 μg) from heart, brain, placenta, lung, liver, skeletal muscle, kidney and pancreas (lanes 1–8 respectively) was hybridized with a ^{32}P-labelled human heparan N-sulphotransferase-2 3′-UTR DNA fragment obtained by PCR (nt 2787–3204) and with β-actin. Migrations of RNA standards are indicated on the right.

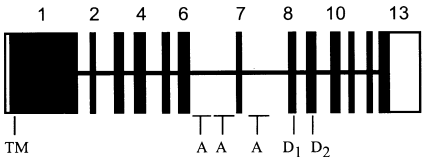

Figure 3 Genomic organization of the human heparan N-sulphotransferase-2 gene

Filled boxes represent coding exons; open boxes represent UTRs. The exon number is indicated at the top. The location of the transmembrane domain (TM), Alu repetitive sequences (A), and the two conserved sequences shown in Figure 6 (D_1 and D_2) are indicated at the bottom.

lung. Heparan N-sulphotransferase-1 mRNA expression was previously found to be high in heart and pancreas and lowest in lung [4]. The major differences between N-sulphotransferase-1 and -2 expression were found in placenta and liver. Placenta contained relatively high amounts of heparan N-sulphotransferase-2 mRNA but only intermediate amounts of heparan N-sulphotransferase-1 mRNA. Heparan N-sulphotransferase-1 mRNA expression was high in human liver, whereas only intermediate levels of heparan N-sulphotransferase-2 mRNA were expressed.

Structure of the human heparan N-sulphotransferase-2 gene

Four unique overlapping phage clones were isolated from a human lung genomic library. Three overlapping DNA fragments that contained the entire gene were subcloned into pBluescript and partially sequenced to determine the intron/exon organization of the coding region of the heparan N-sulphotransferase-2 gene (GeneBank accession number AF042084).

The coding region of the human heparan N-sulphotransferase-2 gene was found to contain 13 exons spanning 6.5 kb (Figure 3). Exons ranged in size from 88 bp (exon 2) to >1 kb (exon 1). Exon 1 contained the translational start methionine, and included nucleotides encoding amino acids comprising the cytoplasmic, transmembrane and proline-rich regions of the heparan N-sulphotransferase-2 enzyme. Exons 2–12 were found to be relatively small (88–186 bp). Exon 13 is the last and second-largest exon; it contained the translational stop site and all of the 3′-UTR of human heparan N-sulphotransferase-2 mRNA.

Intron–exon junction sequences within the human heparan N-sulphotransferase gene were noted to be consistent with the 5′-donor (GT) and 3′-acceptor (AG) consensus sequences of Breathnach and Chambon [10]. Intron sizes ranged from 97 bp (intron 12) to 760 bp (intron 7). Intron 7 contained one Alu repetitive element, whereas intron 6 is essentially comprised of two Alu elements (Figure 3).

Chromosomal localization of the human heparan N-sulphotransferase-2 gene

Initial *in situ* hybridization experiments resulted in specific labelling of the mid-long arm of a group C chromosome, which was believed to be chromosome 10 on the basis of size, morphology and banding pattern (results not shown). A second experiment was conducted in which a biotin-labelled probe specific for the centromere of chromosome 10 (D10Z1) was co-hybridized with digoxigenin dUTP-labelled human heparan N-sulphotransferase-2 DNA. A total of 80 metaphase cells were analysed, with 65 exhibiting specific labelling. Measurements of 10 specifically labelled chromosome 10 centromeres (red in Figure 4) demonstrated that human heparan N-sulphotransferase-2 (green) is located at a position that is 39 % of the distance from the centromere to the telomere of chromosome arm 10q, in an area corresponding to band 10q22.

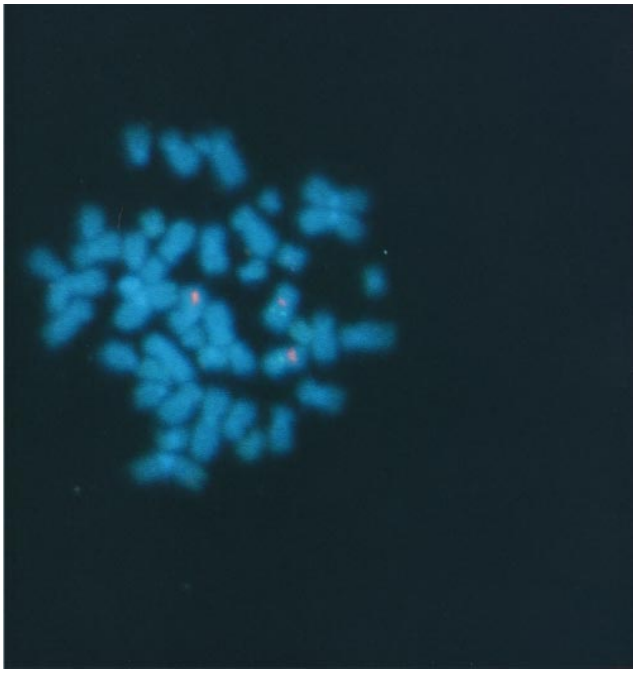

Figure 4 Localization of the human heparan N-sulphotransferase-2 gene by fluorescence *in situ* hybridization

A biotin-labelled probe specific for the centromere of chromosome 10 (D10Z1) was co-hybridized with digoxigenin dUTP-labelled human heparan N-sulphotransferase-2 DNA. Specific hybridization signals were detected by fluoresceinated anti-digoxigenin antibodies for heparan N-sulphotransferase-2 (green) and Texas red avidin (red).

Comparison of the human heparan N-sulphotransferase-1 and -2 genes

As demonstrated in Figure 5, all 12 introns that interrupt the coding region of the human heparan N-sulphotransferase-2 gene were found to occur at similar sites within the coding sequence of the human heparan N-sulphotransferase-1 gene [11]. The two genes differed in that the heparan N-sulphotransferase-1 gene contained one additional intron. Thus exons 1A and 1B of the

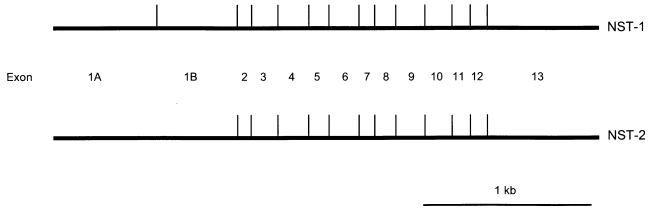

Figure 5 Comparison of the structure of the human heparan N-sulphotransferase-1 and -2 genes

The cDNAs are represented by horizontal lines, with vertical lines showing the positions of the introns that interupt the coding region. The exons are numbered such that the similar exons have the same number.

Table 1 Comparison of the exons in the human heparan N-sulphotransferase (NST)-1 and -2 genes

Comparison of the size and similarity between exons of human heparan NST-1 and NST-2. Note that exon 1 of the heparan NST-2 gene encodes 335 amino acids. To compare exons 1A and 1B, the exon sizes for NST-2 exons 1A and 1B were determined assuming this gene contained an intron in a position similar to intron 1 in the NST-1 gene.

	Amino acids		
Exon	NST-1	NST-2	Homology (%)
1A	171	173 } 335	53
1B	165	162 }	74
2	29	29	72
3	52	52	79
4	62	62	89
5	43	43	91
6	61	61	80
7	32	32	84
8	42	42	67
9	59	59	76
10	56	56	63
11	37	37	51
12	34	34	73
13	39	41	57
Total	882	883	70

human heparan N-sulphotransferase-1 gene encode amino acids similar to amino acids encoded by the first exon of the human heparan N-sulphotransferase-2 gene (Figure 5). Comparison of the corresponding exons of the two heparan N-sulphotransferases (Table 1) showed that exon 5 was the most similar (91 %).

DISCUSSION

Isolation and sequencing of human heparan glucosaminyl N-deacetylase/N-sulphotransferase-2 cDNA has permitted comparison of the deduced amino acid sequences of human heparan N-sulphotransferase-1 and -2 from a single species for the first time. The overall nucleotide similarity of these enzymes was 70 %; numerous regions of similarity or identity were found.

We also report the first description of the intron/exon organization of the heparan N-sulphotransferase-2 gene and found the structure of the gene to be very similar to the structure of the human heparan N-sulphotransferase-1 gene [11]. All 12 introns that interrupt the coding region of the heparan N-sulphotransferase-2 gene occurred at similar sites within the coding sequence of the heparan N-sulphotransferase-1 gene. The heparan N-sulphotransferase-2 gene contained one less intron than the heparan N-sulphotransferase-1 gene (Figure 5), such that the structural organization of exons 2–13 of the heparan N-sulphotransferase-2 gene was exactly the same as that of exons 3–14 of the heparan N-sulphotransferase-1 gene. In contrast with the similar exon sizes in the two heparan N-sulphotransferase genes, introns in the human heparan N-sulphotransferase-1 gene were longer than those in the heparan N-sulphotransferase-2 gene, resulting in the coding sequences within the human heparan N-sulphotransferase-1 spanning approx. 35 kb compared with 6.5 kb in the human heparan N-sulphotransferase-2 gene.

Comparison of the amino acids encoded by the similar exons of the heparan N-sulphotransferase-1 and -2 genes showed the least conserved exons were 1A (53 %), 11 (51 %) and 13 (57 %). Although the overall amino acid similarity of exon 1A was 53 %,

	Motif 1	Motif 2
NST	**YHK-GIDW**	**Y-SWYQH**
STa	**YPKSGTNW**	**YGSWFEH**
Aryl	**YPKSGTTW**	**YGSWYQH**
Est	**YPKSGTTW**	**YGSWFEH**
Fla	**YPKSGTTW**	**FGPYWDH**

Figure 6 Conserved amino acid sequences in sulphotransferases

The amino acid sequences of heparan N-sulphotransferases (NST), mouse hydroxysteroid sulphotransferase (Sta; Genebank accession number X84816), rat aryl sulphotransferase (Aryl; accession number X68640), bovine oestrogen sulphotransferase (Est; accession number M54942) and plant flavonol sulphotransferase (Fla; accession number M84136) were aligned using Geneworks and edited manually. Conserved amino acids in the consensus sequences are shown in bold.

the first 86 amino acids were not similar whereas the remaining 87 amino acids were 82 % similar. This non-homologous region of exon 1A contains the amino acids comprising the cytosolic, transmembrane and spacer regions of the heparan N-sulphotransferase-2 enzyme. Since this region is thought to be important for targeting and retention of many Golgi enzymes [12–16], the two enzymes may localize to different areas of the Golgi complex. Utilizing full-length and truncated epitope-tagged heparan N-sulphotransferase-1 cDNA constructs, we have already shown that the N-terminal region of this enzyme is necessary for retention and localization in the trans-Golgi network [4].

Other functional domains within the heparan N-sulphotransferases have not yet been defined, and thus correlation of exon sequences with catalytic or substrate-binding domains is not possible at this time. However, comparison of the amino acid sequences of heparan N-sulphotransferase-1 and -2 with the amino acid sequences of 17 other cloned soluble sulphotransferases has revealed a short domain (amino acids 640–720) that may be important for sulphotransferase activity. Two conserved motifs were found within this domain: YXKGXXW (amino acids 651–657 in heparan N-sulphotransferase-1) and (F/Y)XW(F/Y)XH (amino acids 710–716 in heparan N-sulphotransferase-1) (Figure 6). Although the function of these sequences is not yet known, they may form part of the 3′-phosphoadenosine 5′-phosphosulphate-binding or active sites of these sulphotransferases; the degree of conservation suggests that this region of the N-sulphotransferase is important for sulphotransferase activity. Analysis of the amino acid sequence of chondroitin 6-sulphotransferase, recently cloned from chicken cartilage [17], failed to reveal either of these motifs, suggesting the existence of at least two functionally distinct classes of sulphotransferase enzymes.

The two heparan N-sulphotransferase genes map to different human chromosomes, with heparan N-sulphotransferase-1 at 5q32-q33.1 and heparan N-sulphotransferase-2 at 10q22. The similarity of the cDNAs and the genomic organization of the two heparan N-sulphotransferase genes suggests that they arose by gene duplication and are evolutionarily related. Duplication was followed by either the addition of an intron in the heparan N-sulphotransferase-1 gene or the loss of an intron in the heparan N-sulphotransferase-2 gene respectively. On the basis of reports describing intron loss in the rat insulin-1 gene [18,19], the chloroplast *rpoC1* gene [20] and the rhodopsin gene in fish [21], intron loss appears to be the more likely event. This probably occurs by reverse transcription of mRNA and homologous recombination of the resulting DNA into the genome [22,23]. If

loss is the mechanism, then the heparan N-sulphotransferase-1 gene is similar to the ancestral gene. Analysis of heparan N-sulphotransferase gene structures of other species may confirm this speculation.

The full genomic structure of the two heparan N-sulphotransferase genes remains to be clarified, since the transcriptional start sites of both heparan N-sulphotransferases have yet to be defined. The human heparan N-sulphotransferase-2 cDNA clones that were isolated contained only 27 bp of the 5′-UTR. An additional ≈ 500 bp of the 5′-UTR was obtained by rapid amplification of cDNA ends-PCR. Sequencing these PCR products revealed a high GC content ($> 65\%$), which probably accounts for the difficulty in defining this region. Sequencing of the isolated genomic clones revealed that the 5′-UTR determined thus far is contained within three exons (results not shown).

Although both heparan N-deacetylase/N-sulphotransferases appear to be expressed in all tissues, the functional relationship of these two enzymes is not yet clear. The enzymes differ most noticeably in their N-terminal segments, which suggests they may be localized in different subcompartments of the Golgi, and may therefore subsume different functions. Alternatively, the enzymes may interact with, or function together with, different enzymes involved in the biosynthesis of heparin and heparan sulphate.

This work was supported by the Medical Research Service of the Department of Veterans Affairs.

REFERENCES

1 Lindahl, U., Lidholt, K., Spillmann, D. and Kjellén, L. (1994) Thrombosis Res. **75**, 1–32
2 Lindahl, U. (1989) in Heparin. Chemical and Biological Properties, Clinical Applications (Lane, D. A. and Lindahl, U., eds.), pp. 159–191, Edward Arnold, London
3 Hashimoto, Y., Orellana, A., Gil, G. and Hirschberg, C. B. (1992) J. Biol. Chem. **267**, 15744–15750
4 Humphries, D. E., Sullivan, B. M., Aleixo, M. D. and Stow, J. L. (1997) Biochem. J. **325**, 351–357
5 Dixon, J., Loftus, S. K., Gladwin, A. J., Scambler, P. J., Wasmuth, J. and Dixon, M. J. (1995) Genomics **26**, 239–244
6 Orellana, A., Hirschberg, C. B., Wei, Z., Swiedler, S. J. and Ishihara, M. (1994) J. Biol. Chem. **269**, 2270–2276
7 Eriksson, I., Sandback, D., Ek, B., Lindahl, U. and Kjellén, L. (1994) J. Biol. Chem. **269**, 10438–10443
8 Sanger, F., Nicklen, S. and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 5463–5467
9 Zhang, H., Scholl, R., Browse, J. and Somerville, C. (1988) Nucleic Acids Res. **16**, 1220
10 Breathnach, R. and Chambon, P. (1981) Annu. Rev. Biochem. **50**, 349–383
11 Gladwin, A. J., Dixon, J., Loftus, S. K., Wasmuth, J. J. and Dixon, M. J. (1996) Genomics **32**, 471–473
12 Munro, S. (1991) EMBO J. **10**, 3577–3588
13 Teasdale, R. D., D'Agostaro, G. and Gleeson, P. A. (1992) J. Biol. Chem. **267**, 4084–4096
14 Russo, R. N., Shaper, N. L., Taatjes, D. J. and Shaper, J. H. (1992) J. Biol. Chem. **267**, 9241–9247
15 Burke, J., Pettit, J. M., Schachter, H., Sarkar, M. and Gleeson, P. A. (1992) J. Biol. Chem. **267**, 24433–24440
16 Aoki, D., Lee, N., Yamaguchi, N., Dubois, C. and Fukuda, M. N. (1992) Proc. Natl. Acad. Sci. U.S.A. **89**, 4319–4323
17 Fukuta, M., Uchimura, K., Nakashima, K., Kato, M., Kimata, K., Shinomura, T. and Habuchi, O. (1995) J. Biol. Chem. **270**, 18575–18580
18 Lomedico, P., Rosenthal, N., Efstratiadis, A., Gilbert, W., Kolodner, R. and Tizard, R. (1979) Cell **18**, 545–558
19 Perler, F., Efstratiadis, A., Lomedico, P., Gilbert, W., Kolodner, R. and Dodgson, J. (1980) Cell **20**, 555–566
20 Wallace, R. S. and Cota, J. H. (1996) Curr. Genet. **29**, 275–281
21 Fitzgibbon, J., Hope, A., Slobodyanyuk, S. J., Bellingham, J., Bowmaker, J. K. and Hunt, D. M. (1995) Gene **164**, 273–277
22 Baltimore, D. (1985) Cell **40**, 481–482
23 Boeke, J. D., Garfinkel, D. J., Styles, C. A. and Fink, G. R. (1985) Cell **40**, 491–500

Received 24 October 1997/5 February 1998; accepted 20 February 1998