

Functional expression and genomic structure of human chondroitin 6-sulfotransferase

Kae Tsutsumi, Hiromi Shimakawa, Hiroshi Kitagawa, Kazuyuki Sugahara*

Department of Biochemistry, Kobe Pharmaceutical University, 4-19-1 Motoyamakita-machi, Higashinada-ku, Kobe 658-8558, Japan

Received 13 October 1998; received in revised form 2 November 1998

Abstract The cDNA and gene encoding human chondroitin 6-sulfotransferase (C6ST) have been cloned. The expression of a soluble recombinant form of the protein in COS-1 cells produced an active sulfotransferase, which used as acceptor substrates polymer chondroitin, various chondroitin sulfate isoforms and chondroitin sulfate tetrasaccharides. The identification of the reaction products demonstrated that the enzyme transferred sulfate to position 6 of GalNAc in the GlcA β 1-3GalNAc but not the IdoA α 1-3GalNAc nor the GlcA β 1-3GalNAc(4-*O*-sulfate) sequences. The human C6ST gene spans more than 20 kb and consists of three exons. The protein-coding domain of the C6ST gene is divided into two discrete exons.

© 1998 Federation of European Biochemical Societies.

Key words: Chondroitin sulfate; Gene structure; Glycosaminoglycan; Proteoglycan; Sulfotransferase

1. Introduction

Chondroitin sulfate proteoglycans, which consist of a core protein with at least one covalently attached glycosaminoglycan (GAG) chain, are distributed on the surfaces of most cells and the extracellular matrix in virtually every tissue (for reviews, see [1,2]). Despite the ubiquity of this family of molecules, a wide variety of chondroitin sulfate proteoglycans with characteristic sulfated GAG chains exhibit tissue-specific and developmentally regulated expression [3], and have been implicated in the regulation and maintenance of cell proliferation, cytodifferentiation, and tissue morphogenesis [4]. In neural development, chondroitin sulfate governs developmentally significant events such as cellular adhesion, migration, and neurite outgrowth [5–8]. The molecular basis for the developmentally regulated and tissue-specific synthesis of chondroitin sulfate, as well as other GAGs, has yet to be clarified.

Chondroitin sulfate has a linear polymer structure that possesses repetitive, sulfated disaccharide units containing glucuronic acid (GlcA) and *N*-acetylgalactosamine (GalNAc) [1,2].

*Corresponding author. Fax: (81) (78) 441-7571.
E-mail: k-sugar@kobepharm-u.ac.jp

Abbreviations: C6ST, chondroitin 6-sulfotransferase; GalNAc, *N*-acetyl-D-galactosamine; GlcA, D-glucuronic acid; GAG, glycosaminoglycan; Δ HexA, 4,5-unsaturated hexuronic acid or 4-deoxy- α -L-threo-hex-4-ene-pyranosyluronic acid; HPLC, high-performance liquid chromatography; KSGal6ST, keratan sulfate Gal-6-sulfotransferase; PAPS, 3'-phosphoadenosine 5'-phosphosulfate; PCR, polymerase chain reaction

The nucleotide sequence reported in this paper has been submitted to the DDBJ/GenBank/EBI data banks with accession number AB017915.

Since GAG structures are largely determined by the specificities of the sulfotransferases responsible for their synthesis, it is presumed that the differential expression of the enzymes is the key for the controlled synthesis of GAGs. Despite the growing number of sulfotransferase cDNAs which have been cloned, limited information is available about the organization and regulation of the expression of sulfotransferase genes.

The major chondroitin sulfate found in the mammalian tissues bears sulfate groups at position 4 or 6 of GalNAc residues. It was reported that the ratio of 4-sulfation/6-sulfation changed during the development of chicken and human epiphyseal cartilage and rat skin [4,9–11]. We recently demonstrated that the ratio of 4-sulfation/6-sulfation in the embryonic chick brain changes with development and that the relative levels of the specific sulfotransferase activities are closely coordinated with the relative changing levels of the specific chondroitin sulfate structures [12]. These findings suggested that the expression of sulfotransferases is a predominant factor regulating the sulfation profile of chondroitin sulfate structures. Among the sulfotransferases involved in the biosynthesis of chondroitin sulfate, the cDNAs encoding chondroitin 6-sulfotransferase (C6ST), which catalyzes the transfer of sulfate from PAPS to position 6 of the GalNAc residue, were recently cloned from chicken and mouse [13,14]. In addition, during the preparation of this article, a report of the human C6ST cDNA sequence appeared [15]. However, the precise specificities of these recombinant enzymes and these gene organizations have not been described. In this report, we have characterized the human recombinant C6ST and determined the genomic organization.

2. Materials and methods

2.1. Materials

[³⁵S]PAPS and unlabeled PAPS were purchased from New England Nuclear (Boston, MA, USA) and Sigma Chemicals (St. Louis, MO, USA), respectively. Chondroitin (a chemically desulfated derivative of whale cartilage chondroitin sulfate A), various chondroitin sulfate isoforms and four unsaturated standard disaccharides derived from chondroitin sulfate, i.e. $\Delta^{4,5}$ HexA α 1-3GalNAc(4-*O*-sulfate), $\Delta^{4,5}$ HexA α 1-3GalNAc(6-*O*-sulfate), $\Delta^{4,5}$ HexA α 1-3GalNAc(4,6-*O*-disulfate), and $\Delta^{4,5}$ HexA(2-*O*-sulfate) α 1-3GalNAc(6-*O*-sulfate), chondroitinase ABC (EC 4.2.2.4), chondro-4-*O*-sulfatase (EC 3.1.6.9), and chondro-6-*O*-sulfatase (EC 3.1.6.10) were purchased from Seikagaku Corp. (Tokyo, Japan). Tetrasaccharides used as acceptors (see Table 2) were isolated by high-performance liquid chromatography (HPLC) after the hyaluronidase digestion of chondroitin, king crab cartilage chondroitin sulfate K and bovine bronchial cartilage chondroitin sulfate as described [16,17]. They were structurally characterized enzymatically and also by 500-MHz ¹H NMR spectroscopy when required. HITRAP Desalting columns were obtained from Amersham Pharmacia Biotech (Uppsala, Sweden). All other reagents and chemicals were of the highest quality available.

-440	gacgccttggacgcggttcaggaatccgccgcgctagccggctcgggcctgagcggggaggggcgccaggcaggacctct	-361
-360	cgaggctcgtgcgcaggacggcgcccgctggcgccgcttccttcagcgtgccgacggcccccgagcgccctccatccctccgccc	-271
-270	cgccccggagaagacgcacagctcgggcgccgcgggcgccggggcgcggaaccgcttctgcccggatcttccaggaggaaagcgaagt	-181
-180	gcgagcggatgctgcccgcgcccagcccgagcgaggggtcgggcgccgggtgagctctcgccggcggggacaaaggggtgtcccca	-91
-90	cctgaagacggcaagctgggtcctgagtgatgccctcagctgagtgccaaggctggcccgaggagccccacggccccacctttcccc	-1
1	ATGGAGAAAGGACTCACTTTGCCCCAGGACTGCCGGGACTTTGTGCACAGCCTGAAGATGAGAAGCAAATACGCCCTTTTCTTGTTT	90
1	M E K G L T L P Q D C R D F V H S L K M R S K <u>Y A L F L V F</u>	30
91	GTGGTGATAGTTTTGTCTTCATCGAAAAGGAAAATAAAATCATATCAAGGGTCTCAGACAAGCTGAAGCAGATTCCCCAAGCTCTAGCA	180
31	<u>V V I V F V F I</u> E K E N K I I S R V S D K L K Q I P Q A L A	60
181	GATGCCAACAGCACCAGCCAGCCCTGATCTTAGCTGAGAACGCATCTCTCTGTGCCCTGAGCGAGCTCGATTACAGCCTTCTCCAGCTT	270
61	D A N S T D P A L I L A E N A S L L S L S E L D S A F S Q L	90
271	CAGAGCCGCTCTCCGCAACCTCAGCTTGCGAGCTGGGCGTGGAGCCAGCCATGGAGGCCGAGGGGAGGAAGAGCAGAGAAAGGAG	360
91	Q S R L R N L S L Q L G V E P A M E A A G E E E E E Q R K E	120
361	GAGGAGCCGCCCAGACCGGCGTGGCGGGGCCCCGGCGCCACGTGCTGCTCATGGCCACCACGCGCACCGGCTCCTCGTTCTGTTGGGCGAG	450
121	E E P P R P A V A G P R R H V L L M A T T R T G S S S F V G E	150
451	TTCTTCAACCAGCAGGGCAACATCTTCTACCTCTTCGAGCCGCTGTGGCACATCGAGCGCACAGTGTCTTCGAGCCGGGGGCGCCAAC	540
151	F F N Q Q G N I F Y L F E P L W H I E R T V S F E P G G A N	180
541	GCCGCGGGCTCGGCCCTGGTGATCCGCGACGTGCTCAAGCAGCTCTTCCTGTGCGACCTGTACGTGCTGGAGCATTTCATCAGCCGCTG	630
181	A A G S A L V Y R D V L K Q L F L C D L Y V L E H F I T P L	210
631	CCCGAGGACCACCTGACTCAGTTTATGTTCCGCGGGGCTCCAGCCGCTCCCTGTGCGAGGACCCCGTCTGTACGCCCTTCGTCAAGAAG	720
211	P E D H L T Q F M F R R G S S R S L C E D P V C T P F V K K	240
721	GTCTTCGAGAAGTACCAGTCAAGAACCAGCGCTGCGGCCCTCAACGTGACGCTGGCCGAGAGGCTGCCGCCGCAAGGAGCACATG	810
241	V F E K Y H C K N R R C G P L N V T L A A E A C R R K E H M	270
811	GCCCTCAAGGCGGTGCGCATCCGGCAGCTGGAGTTCTGTCAGCCGCTGGCCGAGGACCCCGCTGGACCTGCGCGTTCATCCAGCTGGTG	900
271	A L K A V R I R Q L E F L Q P L A E D P R L D L R V I Q L V	300
901	CGCGACCCCCGGGCGTGTGCGCTCGCGCATGGTGGCCTTCGCCGCAAGTATAAGACCTGGAAGAAGTGGCTGGACGACGAGGGCCAG	990
301	R D P R A V L A S R M V A F A G K Y K T W K K W L D D E G Q	330
991	GACGGCTGAGGGAAGAGGAGGTGCAGCGGCTGCGGGGCAACTGCGAGAGCATCCGCTGTCCGCGGAGCTGGGGCTGCGGCAGCCCGCC	1080
331	D G L R E E E V Q R L R G N C E S I R L S A E L G L R Q P A	360
1081	TGGCTGCGGGGCGCTACATGCTGGTGCCTACGAGGACGTGGCACGCGGGCGCTGCAGAAAGGCCCGGAGATGTACCGCTTCGCCGGC	1170
361	W L R G R Y M L V R Y E D V A R G P L Q K A R E M Y R F A G	390

Fig. 1. Nucleotide and deduced amino acid sequences of the putative human recombinant C6ST cDNA and the positions of introns. The isolated cDNA contained a single open reading frame that predicts a protein composed of 479 amino acid residues. The hydropathy analysis revealed the putative transmembrane hydrophobic domain which is underlined in the amino-terminal region. The presumptive polyadenylation signal ATTTAA is boxed. The sequences are numbered relative to the translation initiation site, which begins at the first in-frame ATG codon. The locations of introns are indicated by arrowheads.

2.2. Cloning of the human C6ST cDNA

Chick C6ST cDNA was first amplified by polymerase chain reaction (PCR) using a 5' primer, 5'-CGAGAAGGAAAACAACCTTCA-3' (the nucleotide sequence corresponding to 321–340 of the cDNA for chick C6ST [13]), and the 3' primer, 5'-CTCGGGCGCTGGTGA-GAT-3' (the nucleotide sequence corresponding to 769–786), using chick embryo brain cDNA as a template as described [12]. A human placenta cDNA library in a λ gt 10 phage vector (Clontech, Palo Alto, CA, USA) was screened with the cloned chick C6ST cDNA fragment. Three positive clones were plaque-purified, and then the insert cDNA fragments were subcloned into Bluescript plasmid vectors (Stratagene, La Jolla, CA, USA) and sequenced using a 377 DNA sequencer (PE Applied Biosystems, Foster City, CA, USA). To obtain the 5' and 3' ends of the C6ST cDNA, a cloning strategy of rapid amplification of cDNA ends (RACE) was employed according to the manufacturer's instructions (5'/3' RACE kit, Boehringer Mannheim, Mannheim, Germany), using 0.2 μ g of human placenta poly(A)⁺ RNA as a template. The nucleotide sequence of the amplified cDNA was determined in the DNA sequencer. Several clones were sequenced to compensate for misreading by *EX Taq* polymerase (Takara Shuzo Co., Kyoto, Japan).

2.3. Construction of a soluble form of the human C6ST

A truncated form of C6ST, lacking the first *N*-terminal 48 amino acids of the C6ST, was amplified by PCR using a 5' primer (5'-GGAAGATCTTCAGACAAGCTGAAGCAGAT-3') containing an in-frame *Bgl*II site and a 3' primer (5'-GGAAGATCTCTACGT-GACCCAGAAGGTG-3') containing an in-frame *Bgl*II site located 40 bp downstream of the stop codon. PCR reactions were carried out with *Pfu* polymerase (Stratagene, La Jolla, CA, USA) by 30 cycles of 95°C for 30 s, 60°C for 45 s, and 72°C for 90 s. The PCR fragment was subcloned into the *Bam*HI site of pGIR201protA [18] resulting in the fusion of C6ST to the insulin signal sequence and the protein A sequence present in the vector. An *Nhe*I fragment containing the fusion protein was inserted into the *Xba*I site of the expression vector pSVL (Amersham Pharmacia Biotech).

2.4. Expression of the soluble form of the C6ST and enzyme assay

The expression plasmid (11 μ g) was transfected into COS-1 cells on 100-mm plates using Lipofectamine (Life Technologies, Gaithersburg, MD, USA) according to the instructions provided by the manufacturer. Two days after transfection, 1 ml of the culture medium was collected and incubated with 10 μ l of IgG-Sepharose (Amersham

1171	ATCCCCCTGACCCCGCAGGTGGAAGACTGGATCCAAAAGAACACGCAGGCGGCCACGACGGCAGCGGCATCTACTCCACGCAGAAGAAC	1260
391	I P L T P Q V E D W I Q K N T Q A A H D G S G I Y S T Q K N	420
1261	TCCTCGGAGCAGTTCGAGAAGTGGCGCTTCAGCATGCCCTTCAAGCTGGCCCCAGGTGGTGCAGGCCCGCTGCGGCCCTGCCATGCGCCTC	1350
421	S S E Q F E K W R F S M P F K L A Q V V Q A A C G P A M R L	450
1351	TTCGGGTACAAACTGGCGCGGGACGCCGCCGCCCTCACCAACCGCTCAGTCAGCCTGCCTGGAGGAGAGGGGCACCTTCTGGGTACGCTAG	1440
451	F G Y K L A R D A A A L T N R S V S L L E E R G T F W V T *	479
1441	gggggccccgggccccgtagtccccctcctcgtgaaaggcctgccccgttttctgcgcagccctcgagagggcggggtgcacagcgccat	1530
1531	gagcgggcgagcgctcctgtagcagtagggccccagccagcgctccagcacaagcgggccccagggtaattgcgagagaacaggaca	1620
1621	gtgccccgtcccccttaggggccccacacccagaccccaacgggttgcaagcctcctgagcagggccttaggcagggccccgctgtggcaag	1710
1711	cttcgatctcacacacacagaaacatacattcgtgcctggagaccctgcaggccagagtcacaaatatttaacaatcagaaggggcaaggc	1800
1801	tctgaccagtgacagtcagacctcctcgtttatttgggtgtaaacggtttcctgtctggtggtgaagtcgtggaatctgggtgggtccttg	1890
1891	aggaggggcttaggacagccgtgggtgtcaaaggtggcatttgaggctcgtttgaggtgacagtggtgtttaccaactagtagacagcga	1980
1981	ttcagcgttttcatgaattgggttggtgtgctggtttactgatgaatatgggcccccttatagagctgcaaacacacacacatgctatagac	2070
2071	atacatatcatgtacacacacatagacaagcatcataaaggcacaagtcgacacacatctatgcagacaagcttccctcgtcgtttatgcc	2160
2161	cacagggtttttctgtatgacacacccagaggagcctggtgcttaacattttaggattatttgcagggcgagggcgaggggaaagaaacgc	2250
2251	gttaaaaggggccatgacatgacacagttccctggccgggtttggaactgaacagtggtattcagaactgcagcgttcaaagccccctgccccgc	2340
2341	ttagattctcatggctacacctcctcgtgacccatccacatccttgccggtggcagggcaggtggtttgaaatgggtgcagcacaaggccccgt	2430
2431	ctagcttggtcgtggtccccgaacatgctcatatttgaaggctgctcagggcctggcagggcagagcagcagctggttccctggagggcaaaagg	2520
2521	cagctcacggggccttccatttccagccaggcctcatctgggttagggccaagggaaagtacagagtgcccggtgggggagaagtccac	2610
2611	caggatgccccccctccctgagaagcccgcttcttccctcgtggacaggagtcagccagttggttgcgtgtggttacagccagtgccctg	2700
2701	aattcctccttagggccctgggaagagtagtcttaacgcaggtatgtgctgggtgtttgttccgggcttttattatggcttggtgtct	2790
2791	ttcttgtttcatggctgtgttttgcctttgtttctgtaggagcgctccttccaagcagtgaggcgctgaggccttaggggaggtcgggg	2880
2881	aaaaccagcaggggagcacttggtcctcgtggcagggagggagggagggagggccttccccagggaggaacagggagaggtggccacgtg	2970
2971	gcaggccacagtgccctatgggttttctgtttgaaccccatgtgggatgaatctcaaagccaccagcgttcttgcctagtgtccagagc	3060
3061	agcagagatttgcagccctcgccccctccttgagaagaagctggatttgaataccctcacagggccctaacatgttctcgtgggtgggcaggt	3150
3151	ggcctcatgacaagcgacaagtgtaccagagtagaaggccttccccactccacctcaaaccagggagcagaactcagtagtccagagct	3240
3241	tacaaggagctgcaagggtttgcccagggtcgtcccagctcgtccttctggtttcctggacaatttctctgtcagatacggcccatgtgaaa	3330
3331	cccagaggggtgcattttgggttgatgtaatatggacaccatggaagtacaggatctgagtcacaattggcatgggggaggggaagatgtg	3420
3421	tcctggcaggaagccaccattggagaaggtagggccagtaaccatctcatgggattcgtataatcagcctatttgagctgctgggtctc	3510
3511	tgtoactgtggaactccagtcattcaaggaggtctccactcctgggttt	3600
3601	acactcttgggttttagagtccaagaataagttgtgcagaacatggttctaatttggagttcaaaggaaacaccagctctgaaacatgac	3690
3691	ctgtggggccaggtatttctgtggcagggccagagtgctggccgttcttccccagcctccccactgctccctcctctgttccccatgtgga	3780
3781	ggtcggggggggtgggactggggagggggcagccacctccatccaaggctcttccctaggggccccgtgtaattgtggacagcagactttatc	3870
3871	cctccttcttactctggggccaagacctccactccgcctcctgaggtggttacaagaatgggccttgggagcttctcaggacattgaca	3960
3961	atgtgcacactgcaggttgacttaattttattt	4050
4051	cccaaatcaagaagacaaaataacaatagaaaccttcaacacactccgttcatgtagcagcagcagcagcagcagcagcagcagcagcagc	4140
4141	agaaatgagctggagccagccagc	4230
4231	ctacatcagggcccttgcgtgctgtattacctcgtcccgggtgtttgagacaggtcacaagcgctggtggttccctggaagtgtattccc	4320
4321	ctggggctcagcgatgctcgtggaaaccttgggggtttggggaaggacagtggtgtgactattggcttgaacggttatggaatgtgttg	4410
4411	gagtcacccatgtggatatgattttcagcttggagctgggagtggaagcagtggtggcgccggcctctttttccagcgaggcc	4500
4501	agccctggcttctcctcgttccagcttgcctgcagaagccttccctgcaaggtgcccacctgcccggagacagagggcagcagctgag	4590
4591	gcctctctccactggcccgccagcctccttagcacaaggggaccattcagcaatgagtggttactcatgttcttcttctgttgatgagttct	4680
4681	gccccagcatatagaaatggctcagaacggtccaggccttggtgagacacagcagcgtccaatgctcacaggctgtccccactcctgcc	4770
4771	tgttcccatcttccccctcgggacgtgtttataaattgaggaatggatgagggccctggagggtccttgggagtagtttagtggagggcagg	4860
4861	attctcaggtccccatagagagggaaagagagaagcagatgtgcataagcacaatctctgtcacatacacccgcagtggtgctaccaggtt	4950
4951	attttccaaaagctgaggaacataaagcaaattaggcttttgcctctcgtcaatacatgcacttgaataaacaagaaaagagatgttt	5040
5041	aataacaagtttacttccggtcgtgtagcaccctcagcctgggagctactgccagaggctggcgtagtggggcagctgctgaccaaacc	5130
5131	cacagaacagaggggcaccagggcatcacacgacatcttccctcccatctcgtccatctgtctgtctggcaatggaaggagcttcagcagg	5220
5221	agatccttccagaaggttgatttcttggcctgggtggttagaggagatatcctgacccataaagtcttccagcccccgggtcattcctcctta	5310
5311	tactaggccctattgaccgtagagggcaggtgcggtggtcctacacctgtaatccccagcatttggggagggccaaggtgggaggtattgcttg	5400
5401	agcccaagagttggagaccagcctgggcaacatggtgagaccaattgactgtagggtgttcccttgctaagacagggagcagaagactgg	5490
5491	atggctgtgtcctcaaggcagtcctccctcccatccccattcccaaagtgcagaagcgaaagccagatctcaagggtgtatcctgaggcag	5580
5581	gagagctaaggggagagaaattggggctgagtggaagcagatgcctgcacaagcgaagtgtgtcttattgatctgtacaagctgatagaa	5670
5671	ggaccatctgtgaaatccagggtcctgagttggtgggtttgtctgtgaaagcttctaggaacaggcgagcagatggaagggagaggg	5760
5761	tccagggtcctgtaaatgcagcccgacacgtccttctgctgagttcagagacagcagcagcagcagcagcagcagcagcagcagcagcagc	5850
5851	gcagaattggcaacttcacatgtctcgagagctccaagatcctttggtctcgtgtccttcggcaccocgtaactggaactggggacaaatt	5940
5941	tgttacgtgtttccaaggctacagacatggtgccatcctcagggcctagctcatgtcatggtcaggtgctttgctgggggatacaaat	6030
6031	gacccatgtggtcgtcgtcttctgcatggtgactgggacccccctgaagtctccatctcagctgtatattctagtccaaattttc	6120
6121	ctggctccctgctccctccaggtcctgctcactgtccttcttctgctgagttcagagacagcagcagcagcagcagcagcagcagcagcagc	6210
6211	acaccccttcacagcatacactgcatggtatgtacatatgcattccagctgtgtgtatgtgtagcatgtagctcattctgctcagcctct	6300
6301	ggcgccgttccacattgctcccagccccattgctgtcactgtcgtcccagatgtccttgccatggccacagacaatctgccgtctcctgg	6390
6391	aaacgctggggctgcccttcagagagctggcagccccagcagtcagggcctgctttgcagaatagaatgtgaacccaactcctgatggcc	6480
6481	tacttgacttatttaatttaagatgaaatcatgcaatgagcaa	6552

Fig. 1 (continued).

Pharmacia Biotech) for 1 h at 4°C. The beads recovered by centrifugation were washed with and then resuspended in the assay buffer, and tested for a sulfotransferase activity using polymer chondroitin, various chondroitin sulfate isoforms and tetrasaccharides prepared from chondroitin and chondroitin sulfate as sulfate acceptor substrates and [³⁵S]PAPS as a sulfate donor as described [12].

2.5. Characterization of the reaction products

The reaction products were isolated by gel filtration on a column of HITRAP desalting equilibrated with 0.25 M NH₄HCO₃/7% 1-propanol. The radioactive peak containing the product was pooled and evaporated to dryness. The isolated products were digested with 20 mIU of chondroitinase ABC or 5 mIU of chondroitinase B. Each digest was analyzed by an anion-exchange HPLC on an amine-bound silica PA03 column (YMC Co., Kyoto, Japan) as described [19]. To confirm the disaccharide structure, chondro-4-sulfatase or 6-sulfatase digestion of chondroitinase ABC digest was conducted with the remainder of each ³⁵S-labeled isolated product as described [12].

2.6. Isolation of human C6ST genomic clones

A human placenta genomic library in a λFIX II phage vector (Stratagene, La Jolla, CA, USA) was screened with the cloned chick C6ST cDNA fragment described above. Three independent clones were isolated and characterized. Insert DNA fragments were initially characterized by restriction digestion and Southern blot analysis. Human genomic DNA fragments that hybridized to the C6ST cDNA probes were subcloned into Bluescript plasmid vectors and sequenced.

3. Results and discussion

3.1. Cloning of human C6ST cDNA

The screening of the human placenta cDNA library with the cloned chick C6ST cDNA fragment [12] resulted in the isolation of three independent clones that overlapped with each other. When taken together, the combined cDNA contained a long 5'-untranslated region of 435 bp, a single open reading frame of 1437 bp encoding a protein of 479 amino acids, and a 3'-untranslated region of 1421 bp. To clone the full-length cDNA, the RACE cloning strategy was employed [20]. The revealed sequence of the overlapping cDNA fragments indicated that the cDNA consisted of a long 5'-untranslated region of 440 bp, a single open reading frame of 1437 bp encoding a protein of 479 amino acids, and a very long 3'-untranslated region of 5115 bp, including a poly(A)⁺ tail (Fig. 1). The 5'-untranslated region had a high G+C content of 75%, which probably accounts for the difficulty in defining this region. The Northern blot analysis indicated that the human C6ST mRNA was about 7.5 kb in length in human placenta [14], suggesting that the cDNA was nearly full-length. Database searches indicated that the amino acid sequence displayed 74, 85 and 36% identity to the chick C6ST, mouse C6ST and human keratan sulfate Gal-6-sulfotransferase (KSGal6ST) [21], respectively, and the highest sequence

identity was found in the COOH-terminal catalytic domain (Fig. 2). There were four regions in which more than six consecutive amino acid clusters were identical among the four sequences.

3.2. Expression and characterization of a soluble form of the human C6ST

To facilitate the functional analysis of the human C6ST, a soluble form of the protein was generated by replacing the first 48 amino acids of the C6ST with a cleavable insulin signal sequence and a protein A IgG binding domain as described in Section 2, and then the soluble C6ST was expressed in COS-1 cells as a recombinant enzyme fused with the protein A IgG binding domain. The fused enzyme expressed in the medium was absorbed on IgG-Sepharose beads to eliminate endogenous C6ST, and then the enzyme-bound beads were used as an enzyme source. The bound fusion protein was assayed for sulfotransferase activity using a variety of acceptor substrates. As shown in Table 1, activity was detected with polymer chondroitin and every chondroitin sulfate isoform tested. However, the incorporated sulfate level of the chondroitin sulfate isoforms was significantly less than that into chondroitin. For example, the sulfate incorporation into chondroitin sulfate D was only approximately 40% of that incorporated into chondroitin, indicating that chondroitin is a better substrate than chondroitin sulfate isoforms. Besides, only a small amount of incorporation into chondroitin sulfate E was observed since chondroitin sulfate E is the most highly sulfated chondroitin sulfate containing the least acceptor site. In addition, no detectable sulfotransferase activity was recovered by the affinity purification from a control pSVL transfection sample.

To identify the sulfotransferase reaction products, polymer chondroitin and chondroitin sulfates A, B, C and D were labeled with ³⁵S-sulfate by incubation with [³⁵S]PAPS as a sulfate donor and the enzyme-bound beads as an enzyme source, and the products from each reaction were isolated by gel filtration, then digested with chondroitinase ABC. The digest was analyzed by an anion-exchange HPLC on an amine-bound silica column as described in Section 2. As shown in Fig. 3, the sulfate was incorporated exclusively into the GalNAc C6 position in a non-sulfated disaccharide unit, GlcA-GalNAc (Fig. 3A–E). Notably, no disulfated disaccharides were detected, suggesting that the enzyme did not transfer sulfate to the C6 position of a GalNAc residue in a monosulfated disaccharide unit, GlcA-GalNAc(4-sulfate). Moreover, no sulfated di- or tetrasaccharides were observed for the chondroitinase B digest of the products obtained with chondroitin sulfate B (Fig. 3F), indicating that the enzyme did

Table 1
Sulfotransferase activities of the recombinant C6ST towards polymer chondroitin and chondroitin sulfate isoforms

Acceptor	Activity ^a (pmol/ml medium/h)	Relative rate (%)
Chondroitin	992	100
Chondroitin sulfate A	288	29
Chondroitin sulfate B	227	23
Chondroitin sulfate C	259	26
Chondroitin sulfate D	397	40
Chondroitin sulfate E	11	1

The recombinant C6ST was incubated for 1 h with 100 µg each of the various acceptor substrates listed under the conditions described in Section 2. The reaction products were separated from [³⁵S]PAPS and its degradation products by gel filtration chromatography on a column of HITRAP desalting, and were quantified by liquid scintillation counting.

^aThe values represent the averages of two independent experiments.

hc6ST	1	MEKGLTLPQDCDFVHSLKMSKXALFLVFWVVFTEKENKISRVSDKIKQIFQALA	60
cC6ST	1	MERRSALPODFNEVLHCLMRSKXAVLLVFWV-GLVITEKENNFISRVSDKIKQSQVLV	59
mC6ST	1	MEKGLALPODFRDLVHSLKIRGVLFLLAFVFWVVFTEKENKISRVSDKIKQIFHFVA	60
hKSGal6ST	1	-----	1
hc6ST	61	DANSTDPALILAHNASLISLSELSAFLSOLSRIRNLSLOLGVEPAMEAAGEEEEEERKE	120
cC6ST	60	EANETEASPVQAENGSLASLROLDTAFSOLRTFRNVTIQLAGE--L--G-----	105
mC6ST	61	DANSTDPALLLSFNASLISLSELSLSTFSLRSRIRNLSLOLGVEPAM-----ESQEAGA	114
hKSGal6ST	1	-----MQCSW-KAVLLALASIAIQYTAIRT-FTAKSFH-TC-GLAEAG-LAERLCEE	49
hc6ST	121	EEPPRPAVAGPRRHVLLMATTRGSSFVGEEFNQGNIFYLFEPLWHIERT-V-SFEPG-	177
cC6ST	106	-IAA-P---EPRRHVLMMATTRGSSFVGEEFNQGNIFYLFEPLWHIERT-V-TFEPG-	157
mC6ST	115	EKPSQQAGAGTRRHVLLMATTRGSSFVGEEFNQGNIFYLFEPLWHIERT-V-FTQOR-	171
hKSGal6ST	50	SPTFAYNLS-RKTHILLIATTRGSSFVGOLFNOHLDVLYLFEPLVHVNQTLIPRFTQSK	108
hc6ST	178	GANAAGSAL-VVRDYLKCLFCDLYVLEHETITLPEDHITQFMFRRCSSRLCEDPVC-T	235
cC6ST	158	GANAAGSAL-VVRDYLKCLFCDLYVLEHETITLPEDHITQFMFRRCSSRLCEDPVC-T	215
mC6ST	172	GANAAGSAL-VVRDYLKCLFCDLYVLEHETITLPEDHITQFLFRRCSSRLCEDPVC-T	229
hKSGal6ST	109	SPARRVMI GASRDILRSLYDCDLYLENYIKPPVNHITDRIFRRCASVLCSPVCDP	168
hc6ST	236	PFVKK-VFEKYECRRRCGELNVTVAEACRRKEMAKAVRIRGLEELQELAEDPRLEI	294
cC6ST	216	PSLKK-VFEKYECRRRCGELNVTVAEACRRKEMAKAVRIRGLEELQELAEDPRLEI	274
mC6ST	230	PFVKK-VFEKYECRRRCGELNVTVAEACRRKDEVAKAVRIRGLEELQELVEDPRLEI	288
hKSGal6ST	169	PGPADIVLEEGDC-VRKCGLLNLTVAEACRRSEVAIKTVRVPEVNDLRALVEDPRLEI	227
hc6ST	295	KVIQLVRDPRAVLASRMVAFAGYKTIWAKWLDDGQDGLREEVGLRGNCESTRISALL	354
cC6ST	275	KVIQLVRDPRAVLASRMVAFAGYKTIWAKWLDDGQDGLREEVGLRGNCESTRISALL	333
mC6ST	289	KVIQLVRDPRAVLASRMVAFAGYKTIWAKWLDDGQDGLREEVGLRGNCESTRISALL	347
hKSGal6ST	228	KVIQLVRDPRAVLASRMVAFAGYKTIWAKWLDDGQDGLREEVGLRGNCESTRISALL	285
hc6ST	355	GLRCPAWLFGRYMLVRYEDVARPPKAREMYRFAGIHPTECVDFEWIRANTQAPQD-SNG	413
cC6ST	334	GLRCPAWLFGRYMLVRYEDVARPPKAREMYRFAGIHPTECVDFEWIRANTQAPQD-SNG	392
mC6ST	348	GLRCPAWLFGRYMLVRYEDVARPPKAREMYRFAGIHPTECVDFEWIRANTQAPQD-SNG	406
hKSGal6ST	286	GLMRPFWLKGKMYLVRYEDLARNPMKITEEYIGFLGIPDSHVARWLNNTNRGDPTLGKH	345
hc6ST	414	IYSTQKNSSEQFEKWRFSMEFKLAQVVOACGFAMRLGYKLARDAALNRSVSLVEER	473
cC6ST	393	IYSTQKNSSEQFEKWRFSMEFKLAQVVOACGFAMRLGYKLARDAALNRSVSLVEER	452
mC6ST	407	IYSTQKNSSEQFEKWRFSMEFKLAQVVOACGFAMRLGYKLARDAALNRSVSLVEER	466
hKSGal6ST	346	KVGVTRNSAATAEKWRERLSYDIVAFACNACQOVLAQLGYKIAASEEELKNPSVSLVEER	405
hc6ST	474	GTFWVT	479
cC6ST	453	PPTRIT	458
mC6ST	467	GTFWVT	472
hKSGal6ST	406	DFRPFS	411

Fig. 2. Sequence comparison of human, chick and mouse C6STs and human KSGal6ST. The predicted amino acid sequences were aligned using the GENETYX-MAC (Ver. 9) computer program. Closed and shaded boxes indicate that the predicted amino acid in the alignment is identical among all four sequences and any three sequences, respectively. There is 74, 85 or 36% identity at the amino acid level between human and chick C6STs, between human and mouse C6STs or between human C6ST and KSGal6ST, respectively. There are four regions in which more than six consecutive amino acid clusters were identical among the four sequences.

not transfer sulfate to the GalNAc residue in the sequence -IdoA-GalNAc-IdoA- in chondroitin sulfate B [22]. These re-

sults together indicate that the expressed protein is C6ST with marked specificity for a GlcA-GalNAc sequence. These obser-

Table 2
Sulfotransferase activities of the recombinant C6ST towards tetrasaccharides isolated from chondroitin sulfate

Acceptor	Activity ^a (pmol/ml medium/h)
GlcAβ1-3GalNAcβ1-4GlcAβ1-3GalNAc	10.6
GlcAβ1-3GalNAc(4-O-sulfate)β1-4GlcAβ1-3GalNAc	1.7
GlcAβ1-3GalNAc(6-O-sulfate)β1-4GlcAβ1-3GalNAc	1.2
GlcAβ1-3GalNAc(4-O-sulfate)β1-4GlcAβ1-3GalNAc(4-O-sulfate)	ND ^b

The recombinant C6ST was incubated for 3 h with 1 nmol each of the various individual tetrasaccharides listed under the conditions described in Section 2. The reaction products were separated from [³⁵S]PAPS and its degradation products by gel filtration chromatography on a column of Superdex 30, and were quantified by liquid scintillation counting.

^aThe values represent the averages of two independent experiments.

^bND, not detected (<0.01 pmol/ml medium/h).

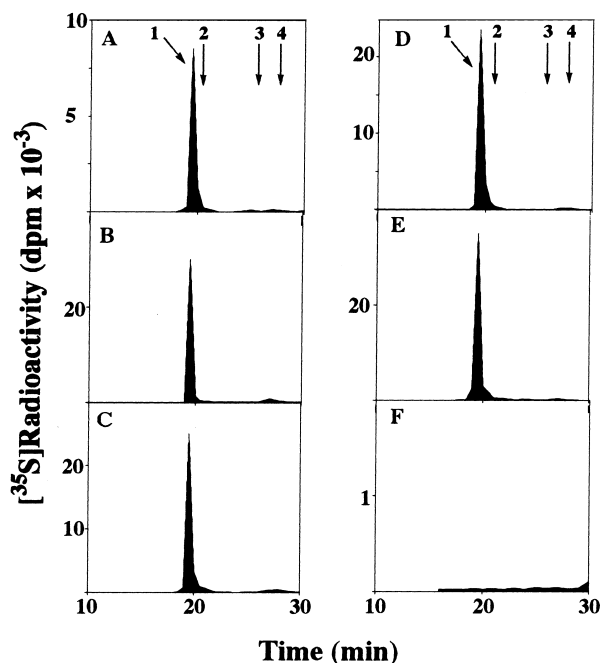


Fig. 3. Identification of the C6ST reaction products. Enzymatic reactions were carried out using polymer chondroitin (A), whale cartilage chondroitin sulfate A (B), pig skin chondroitin sulfate B (C), shark cartilage chondroitin sulfate C (D), shark cartilage chondroitin sulfate D (E), or pig skin chondroitin sulfate B (F) as acceptor substrates under the incubation conditions described in Section 2, and the reaction products were isolated by gel filtration on a column of HTRAP desalting, and digested with chondroitinase ABC (A–E) or chondroitinase B (F). The digests were analyzed by anion-exchange HPLC on an amine-bound silica PA03 column as described in Section 2. The eluate was collected at 30-s intervals for radioactivity measurement by liquid scintillation counting. Arrows indicate the elution positions of the authentic unsaturated disaccharides: 1, $\Delta^{4,5}\text{HexA}\alpha 1\text{-3GalNAc(6-O-sulfate)}$; 2, $\Delta^{4,5}\text{HexA}\alpha 1\text{-3GalNAc(4-O-sulfate)}$; 3, $\Delta^{4,5}\text{HexA(2-O-sulfate)}\alpha 1\text{-3GalNAc(6-O-sulfate)}$; 4, $\Delta^{4,5}\text{HexA}\alpha 1\text{-3GalNAc(4,6-O-disulfate)}$.

variations together with the finding that the -IdoA-GalNAc(6-sulfate)-IdoA- sequence was present in dermatan sulfate isolated from human umbilical cord [23] suggested that another as yet unidentified sulfotransferase that catalyzes the transfer of sulfate to the C6 position of the GalNAc in the -IdoA-GalNAc-IdoA- sequence probably exists. In fact, we have found a sulfotransferase in fetal bovine serum that catalyzed sulfate transfer from PAPS to a GalNAc residue in the -IdoA-GalNAc-IdoA- sequence [24].

To further characterize the substrate specificity of the recombinant C6ST, we examined four structurally defined tetrasaccharides. These tetrasaccharides included those isolated

from commercial chondroitin, king crab cartilage chondroitin sulfate K and bovine tracheal cartilage chondroitin sulfate using testicular hyaluronidase. As shown in Table 2, a non-sulfated compound, GlcA β 1-3GalNAc β 1-4GlcA β 1-3GalNAc, was a better substrate than the monosulfated compounds GlcA β 1-3GalNAc(4-O-sulfate) β 1-4GlcA β 1-3GalNAc and GlcA β 1-3GalNAc(6-O-sulfate) β 1-4GlcA β 1-3GalNAc, whereas a disulfated tetrasaccharide, GlcA β 1-3GalNAc(4-O-sulfate) β 1-4GlcA β 1-3GalNAc(4-O-sulfate) showed no acceptor activity. These ^{35}S -labeled products obtained with the former three tetrasaccharide acceptors were completely digested with chondroitinase ABC, yielding quantitatively a ^{35}S -labeled peak at the position of $\Delta^{4,5}\text{HexA}\alpha 1\text{-3GalNAc(6-O-sulfate)}$ (data not shown). The results confirmed that the sulfate was incorporated exclusively into the GalNAc C6 position in a non-sulfated disaccharide unit, GlcA-GalNAc, as above. Thus, the prior sulfation of the acceptor substrates has some influence upon further sulfation. It will be interesting to determine whether a preceding sulfation of saccharide residues on the reducing side of oligosaccharides has a stimulatory effect on the sulfation of the GalNAc at the non-reducing terminus using sulfated oligosaccharides such as GalNAc β 1-4GlcA β 1-3GalNAc(4-O- or 6-O-sulfate) when such oligosaccharides become available.

We previously reported that strikingly similar patterns of developmental change were observed for bovine serum C6ST and keratan sulfate Gal-6-sulfotransferase activities, suggesting that the two enzyme activities might be attributable to a single enzyme molecule [25]. In fact, the recombinant C6ST efficiently utilized corneal keratan sulfate as an acceptor (data not shown), confirming an earlier proposal that the single enzyme C6ST catalyzes the sulfation of keratan sulfate as well as that of chondroitin sulfate, as has been reported for the chick recombinant C6ST [13].

3.3. Cloning of human C6ST gene

Our screening of a human placenta genomic library with the cloned chick C6ST cDNA fragment resulted in the isolation of three independent clones. Insert DNA fragments were characterized by restriction digestion and Southern blot analysis. Sequencing revealed that these three clones overlapped with each other and covered the entire coding region of the C6ST cDNA. As summarized in Fig. 4, the human C6ST gene spans more than 20 kb and consists of three exons, the coding region of the gene being divided into two exons. The nucleotide sequences of the exons were found to be completely identical to those of the cDNA shown in Fig. 1. The intron/exon junctions were found to follow the GT/AG rule and were flanked by the conserved sequences [26].

To date, although at least eight distinct sulfotransferase

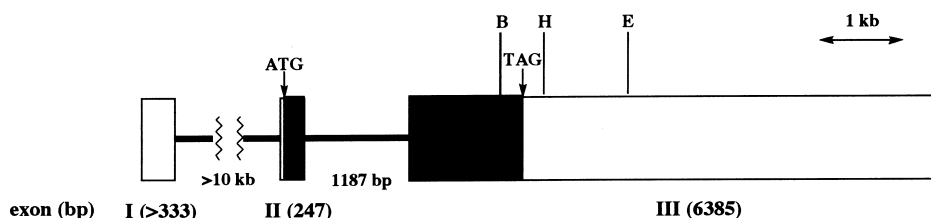


Fig. 4. Genomic organization of the human C6ST gene. *HindIII* (H), *EcoRI* (E) and *BamHI* (B) restriction sites are shown as hash marks. Exon regions are denoted by boxes. Closed boxes represent the coding sequence and open boxes denote the 5'- and 3'-untranslated sequences. The translation initiation codon (ATG) and the termination codon (TAG) are also shown. Black horizontal bars denote the introns.

cDNAs involved in GAG formation have been isolated [13–15,21,27–32], the genomic organization has been elucidated only for the human heparan sulfate *N*-deacetylase/*N*-sulfotransferase-1 and -2 genes [33,34]. The protein-coding sequence of the human heparan sulfate *N*-deacetylase/*N*-sulfotransferase-1 gene is distributed over 14 exons that span approximately 35 kb in length [33]. Similarly, the protein-coding sequence of the human heparan sulfate *N*-deacetylase/*N*-sulfotransferase-2 gene is distributed over 13 exons that span approximately 6.5 kb in length [34]. Comparison of the genomic organization of the human heparan sulfate *N*-deacetylase/*N*-sulfotransferase-1 and -2 genes shows a similar genetic exon-intron organization within the coding sequences [34]. In contrast, the genomic organization of the human C6ST gene is relatively simple, its protein-coding sequence being divided into only two exons that span approximately 8 kb of the genomic sequence (Fig. 4). In addition, the intron insertion site within the protein-coding sequence is identical to that of the chick C6ST gene (data not shown). These findings suggest the existence of at least two evolutionarily distinct classes of sulfotransferases involved in GAG biosynthesis.

In summary, we have characterized the human recombinant C6ST and determined the genomic organization. Since a deficiency in C6ST activity has been reported to be associated with a heritable form of spondyloepiphyseal dysplasia [35,36], the present findings provide molecular tools to study the function and the regulated expression of chondroitin 6-sulfate structure as well as the molecular mechanisms of such diseases.

Acknowledgements: This work was supported in part by the Science Research Promotion Fund of the Japan Private School Promotion Foundation and Grant-in-Aid for Scientific Research on Priority Areas 10178102 from the Ministry of Education, Science, Sports, and Culture of Japan.

References

- [1] Rodén, L. (1980) in: *The Biochemistry of Glycoproteins and Proteoglycans* (Lennarz, W.J., ed), pp. 267–371, Plenum, New York.
- [2] Hascall, V.C. and Hascall, G.K. (1981) in: *Cell Biology of Extracellular Matrix* (Hay, E.D., Ed), pp. 39–63, Plenum, New York.
- [3] Poole, A.R. (1986) *Biochem. J.* 236, 1–14.
- [4] Kimata, K., Okayama, M., Oohira, A. and Suzuki, S. (1973) *Mol. Cell. Biochem.* 1, 211–228.
- [5] Herndon, M.E. and Lander, A.D. (1990) *Neuron* 4, 949–961.
- [6] Faissner, A., Clement, A., Lochter, A., Streit, A., Mandl, C. and Schachner, M. (1994) *J. Cell Biol.* 126, 783–799.
- [7] Nadanaka, S., Clement, A., Masayama, K., Faissner, A. and Sugahara, K. (1998) *J. Biol. Chem.* 273, 3296–3307.
- [8] Clement, A., Nadanaka, S., Masayama, K., Mandl, C., Sugahara, K. and Faissner, A. (1998) *J. Biol. Chem.* 273, 28444–28453.
- [9] Mathews, M.B. and Glagov, S. (1966) *J. Clin. Invest.* 45, 1103–1111.
- [10] Robinson, H.C. and Dorfman, A. (1969) *J. Biol. Chem.* 244, 348–352.
- [11] Habuchi, H., Kimata, K. and Suzuki, S. (1986) *J. Biol. Chem.* 261, 1031–1040.
- [12] Kitagawa, H., Tsutsumi, K., Tone, Y. and Sugahara, K. (1997) *J. Biol. Chem.* 272, 31377–31381.
- [13] Fukuta, M., Uchimura, K., Nakashima, K., Kato, M., Kimata, K., Shinomura, T. and Habuchi, O. (1995) *J. Biol. Chem.* 270, 18575–18580.
- [14] Uchimura, K., Kadomatsu, K., Fan, Q.-W., Muramatsu, H., Kurosawa, N., Kaname, T., Yamamura, K., Fukuta, M., Habuchi, O. and Muramatsu, T. (1998) *Glycobiology* 8, 489–496.
- [15] Fukuta, M., Kobayashi, Y., Uchimura, K., Kimata, K. and Habuchi, O. (1998) *Biochim. Biophys. Acta* 1399, 57–61.
- [16] Kitagawa, H., Tsutsumi, K., Ujikawa, M., Goto, F., Tamura, J., Neumann, K.W., Ogawa, T. and Sugahara, K. (1997) *Glycobiology* 7, 531–537.
- [17] Sugahara, K., Tanaka, Y., Yamada, S., Seno, N., Kitagawa, H., Haslam, S.M., Morris, H.R. and Dell, A. (1996) *J. Biol. Chem.* 271, 26745–26754.
- [18] Kitagawa, H. and Paulson, J.C. (1994) *J. Biol. Chem.* 269, 1394–1401.
- [19] Sugahara, K., Okumura, Y. and Yamashina, I. (1989) *Biochem. Biophys. Res. Commun.* 162, 189–197.
- [20] Frohman, M.A., Dush, M.K. and Martin, G.R. (1988) *Proc. Natl. Acad. Sci. USA* 85, 8998–9002.
- [21] Fukuta, M., Inazawa, J., Torii, T., Tsuzuki, K., Shimada, E. and Habuchi, O. (1997) *J. Biol. Chem.* 272, 32321–32328.
- [22] Malmström, A., Fransson, L.-Å., Höök, M. and Lindahl, U. (1975) *J. Biol. Chem.* 250, 3419–3425.
- [23] Fransson, L.-Å. (1968) *J. Biol. Chem.* 243, 1504–1510.
- [24] Nadanaka, S. and Sugahara, K. (1996) *Seikagaku* 68, 895.
- [25] Sugahara, K., Okamoto, H., Nakamura, M., Shibamoto, S. and Yamashina, I. (1987) *Arch. Biochem. Biophys.* 258, 391–403.
- [26] Breathnach, R. and Chambon, P. (1981) *Annu. Rev. Biochem.* 50, 349–383.
- [27] Hashimoto, Y., Orellana, A., Gil, G. and Hirschberg, C.B. (1992) *J. Biol. Chem.* 267, 15744–15750.
- [28] Orellana, A., Hirschberg, C.B., Wei, Z., Swiedler, S.J. and Ishihara, M. (1994) *J. Biol. Chem.* 269, 2270–2276.
- [29] Shworak, N.W., Liu, J., Fritze, L.M.S., Schwartz, J.J., Zhang, L., Logeart, D. and Rosenberg, R.D. (1997) *J. Biol. Chem.* 272, 28008–28019.
- [30] Kobayashi, M., Habuchi, H., Yoneda, M., Habuchi, O. and Kimata, K. (1997) *J. Biol. Chem.* 272, 13980–13985.
- [31] Habuchi, H., Kobayashi, M. and Kimata, K. (1998) *J. Biol. Chem.* 273, 9208–9213.
- [32] Uchimura, K., Muramatsu, H., Kadomatsu, K., Fan, Q.-W., Kurosawa, N., Mitsuoka, C., Kannagi, R., Habuchi, O. and Muramatsu, T. (1998) *J. Biol. Chem.* 273, 22577–22583.
- [33] Gladwin, A.J., Dixon, J., Loftus, S.K., Wasmuth, J.J. and Dixon, M.J. (1996) *Genomics* 32, 471–473.
- [34] Humphries, D.E., Lanciotti, J. and Karlinsky, J.B. (1998) *Biochem. J.* 332, 303–307.
- [35] Toledo, S.P., Mouraõ, P.A., Lamego, C., Alves, C.A., Dietrich, C.P., Assis, L.M. and Mattar, E. (1978) *Am. J. Med. Genet.* 2, 385–395.
- [36] Mouraõ, P.A., Kato, S. and Donnelly, P.V. (1981) *Biochem. Biophys. Res. Commun.* 98, 388–396.