Molecular cloning and expression of a human hST8Sia VI (α2,8-sialyltransferase) responsible for the synthesis of the diSia motif on O-glycosylproteins

Mélanie TEINTENIER-LELIÈVRE*, Sylvain JULIEN*[1], Sylvie JULIANT†, Yann GUERARDEL*, Martine DUONOR-CÉRUTTI†, Philippe DELANNOY* and Anne HARDUIN-LEPERS*[2]

*Unité de Glycobiologie Structurale et Fonctionnelle, CNRS UMR 8576, IFR 118, GDR CNRS 2590, Université des Sciences et Technologies de Lille, F-59655 Villeneuve d'Ascq, France, and †Centre de Pharmacologie et de Biotechnologie pour la Santé, CNRS UMR 5160, GDR CNRS 2590, 2352, F-30380 Saint Christol-lès-Alès, France

Based on BLAST analysis of the human and mouse genome databases using the human CMP sialic acid; α2,8-sialyltransferase cDNA (hST8Sia I; EC 2.4.99.8), a putative sialyltransferase gene, was identified on human chromosome 10. The genomic organization was found to be similar to that of hST8Sia I and hST8Sia V. Transcriptional expression analysis showed that the newly identified gene was constitutively expressed at low levels in various human tissues and cell lines. We have isolated a full-length cDNA clone from the breast cancer cell line MCF-7 that encoded a type II membrane protein of 398 amino acid residues with the conserved motifs of sialyltransferases. We have established a mammary cell line (MDA-MB-231) stably transfected with the full-length hST8Sia VI and the analysis of sialylated carbohydrate structures expressed at the cell surface clearly indicated the disappearance of Neu5Acα2-3-sialylated structures. The transient expression of a truncated soluble form of the enzyme in either COS-7 cells or insect *Sf*-9 cells led to the production of an active enzyme in which substrate specificity was determined. Detailed substrate specificity analysis of the hST8Sia VI recombinant enzyme *in vitro*, revealed that this enzyme required the trisaccharide Neu5Acα2-3Galβ1-3GalNAc (where Neu5Ac is *N*-acetylneuraminic acid and GalNAc is *N*-acetylgalactosamine) to generate diSia (disialic acid) motifs specifically on *O*-glycans.

Key words: α2,8-sialyltransferase (ST8Sia), *O*-glycosylproteins, disialic acid (diSia), enzymatic characterization, MDA-MB-231 cells, *Sf*-9.

INTRODUCTION

Sialic acids are negatively charged monosaccharides found at the non-reductive terminal position of the carbohydrate groups of glycoconjugates or oligosaccharides. In mammals they are α2,3- or α2,6-linked to a β-D-galactopyranose (Gal) residue or α2,6-linked to a β-D-*N*-acetylglucosamine (GlcNAc) or a β-D-*N*-acetylgalactosamine (GalNAc) residue. Sialic acid residues can also form linear homopolymers of α2,8-linked sialic acids that are named according to their DP (degree of polymerization) diSia (disialic acid) for DP = 2, oligosialic acid (oligosia) for 3 ⩾ DP ⩽ 7 and PSA [polySia (polysialic acid)] for DP ⩾ 8 [2]. Until now, PSA has been described on a small group of *N*-glycosylproteins including the extensively studied NCAM (neural cell adhesion molecule), the α-subunit of the voltage-sensitive sodium channels from *Electrophorus electricus* and most recently the *O*-glycosylprotein CD-36 of human milk [3–7]. PSAs confer anti-adhesive properties to the polysialylated glycoproteins that attenuate cell–cell interactions [8]. Associated with NCAM in the developing nervous system, PSA can be extended to as many

as 60 sialic acid residues and is implicated in various biological processes such as neural cell migration, axonal growth or synaptogenesis [6,9,10]. diSia and oligoSia chains are commonly found on gangliosides of the b and c series that are known to play important roles in differentiation, signal transduction and cell adhesion [11–13]. More recently, these shorter sialic acid chains were also described on both *N*- and *O*-glycosylproteins. diSia motifs were found on the *N*-glycans of Band 3 of human erythrocytes, fetuin and α2-macroglobulin of calf serum, and on the *O*-glycans of human glycophorin, bovine adipo-Q, CD-166 of Neuro2A cells and the bovine chromogranins [14–19]. OligoSia motifs were described on the *N*-glycans of ceruloplasmin and integrin α5 subunit [20,21] and on the *O*-glycans of megalin [22,23].

Until now, six members of the ST8Sia (α2,8-sialyltransferase) family have been described and cloned from various animal species. However, the role of each enzyme in the biosynthesis of diSia, oligoSia and PSA *in vivo* remains unclear. ST8Sia I and ST8Sia V are both involved in the α2,8-sialylation of gangliosides. ST8Sia I is responsible for the biosynthesis of G_{D3}

from G_{M3} whereas ST8Sia V synthesizes mainly G_{T3} [24,25]. ST8Sia II and ST8Sia IV are polysialyltransferases catalysing the polymerization of $\alpha2,8$-sialic acid residues onto the Neu5Ac$\alpha2$-3(6)Gal$\beta1$-4GlcNAc structures found on the N-glycans of NCAM [26]. A third $\alpha2,8$-sialyltransferase, ST8Sia III, transfers sialic acid residues to a terminal $\alpha2,3$-linked sialic acid of glycoproteins and glycolipids [27]. Interestingly, h (human) ST8Sia III as well as ST8Sia II and ST8Sia IV are capable of autopolysialylation [28,29]. This autopolysialylation has been suggested to be implicated in the formation of active polysialyltransferases *in vivo* and *in vitro* [29–31]. However, nothing is known about the enzymes driving the synthesis of diSia, oligoSia or PSA chains on mammalian O-glycans.

In the present study, we describe the molecular cloning, genomic organization and the enzymatic characterization of the hST8Sia VI, which is orthologous to the recently cloned mouse enzyme [32]. The expression of an active, soluble recombinant enzyme clearly shows that the hST8Sia VI synthesizes mainly diSia motifs on the $\alpha2,3$-sialylated core 1 structure (Neu5Ac$\alpha2$-3Gal$\beta1$-3GalNAc$\alpha1$-O-Ser) found in O-glycosylproteins. In contrast with the mouse enzyme, hST8Sia VI exhibits almost no activity towards either gangliosides or N-glycans.

EXPERIMENTAL

Materials

CMP-[^{14}C]Neu5Ac (CMP-[^{14}C]N-acetylneuraminic acid, 10.7 GBq/mmol) and First Strand cDNA Synthesis kit were from Amersham Pharmacia Biotech (Little Chalfont, U.K.). NucleoSpin® RNA II kit was from Macherey-Nagel (Düren, Germany). Oligo nts were synthesized and purified by Eurogentec (Seraing, Belgium) and dNTP were from Promega Biosciences (Son Luis Obispo, CA, U.S.A.). EXTRA-POL I DNA polymerase and DyNazyme Ext DNA polymerase were from Eurobio (Courtaboeuf, France) and Ozyme (Saint-Quentin-en-Yvelines, France). DMEM (Dulbecco's modified Eagle's medium) containing 4.5 g/l glucose without glutamine was from BioWhittaker Europe. TC100 medium, MEM (minimal essential medium), L-glutamine, antibiotics, geneticin G418, FCS (foetal calf serum) used in cell culture, LIPOFECTAMINE™ Plus reagent and TOPO TA-cloning kit were from Invitrogen (Cergy-Pontoise, France). Foetal bovine serum was from D. Deutscher (Issy-les-Moulineaux, France). DOTAP transfection reagent, MAA (digoxigenin-labelled *Maackia amurensis* agglutinin), SNA (digoxigenin-labelled *Sambucus nigra* agglutinin), PGNase F (peptide N-glycosidase F) and anti-digoxigenin fluorescein Fab fragments were from Roche (Meylan, France). HiSpeed Plasmid Midi kit was from Qiagen (Courtaboeuf, France). CMP-Neu5Ac, α_1-acid glycoprotein, fetuin, pFLAG-CMV-1, DMB (1,2-diamino-4,5-methylene dioxybenzene) dihydrochloride, aryl glycosides and Triton CF-54 were from Sigma-Aldrich (St Louis, MO, U.S.A.). Glyco® Sialidase S, Glyco® Sialidase C and Glyco® Sialidase A™ were from Glyko Inc. (Novato, CA, U.S.A.). Colon cancer cell lines HT-29 and Caco-2 were generously given by Dr G. Huet (INSERM U560, Lille, France) and NBEC (normal breast epithelial cells) were kindly provided by Dr Xuefen Le Bourhis (UPRES-EA 1033, Université des Sciences et Technologies de Lille, France) [33]. mRNA of lung cancer cell line NCI and the cDNA of Dami megakaryocyte cell line were given by Dr I. Van Seuningen (INSERM U560) and Dr A. Pierce (CNRS UMR 8576, Université des Sciences et Technologies de Lille, France) respectively. G_{T1b}, G_{D1b} and G_{M3} were a generous gift from J.-P. Zanetta (CNRS UMR 8576).

Cell culture

MCF-7 and MDA-MB-231 cells were cultured in MEM medium supplemented with 5% or 10% (v/v) FCS, 100 units/ml penicillin, 100 μg/ml streptomycin and 45 μg/ml gentamycin. COS-7 cells were grown in DMEM supplemented with 10% (v/v) FCS, L-glutamine 20 mM, penicillin and streptomycin at 37°C under 5% CO$_2$. *Sf*-9 cells (ATCC, CRL-1711) were cultured in TC100 medium supplemented with 5% (v/v) FCS.

Bio-informatic analysis

BLAST algorithm (http://www.ncbi.nlm.nih.gov/genome/seq/HsBlast.html) [34] was used to retrieve nt sequences highly similar to hST8Sia I (GenBank® accession number D26 360), from the human HTG (High-Throughput Genomic) sequences division or the human and mouse EST (expressed sequence tag) divisions of the GenBank®/EBI databases at the NCBI (National Center for Biotechnology Information). The exon-intron junctions were analysed with the internet splice program (http://www.fruitfly.org/cgi-bin/seq_tools/splice.pl). The amino acid sequence analysis was performed using the program of translation for publication (http://www.infobiogen.fr/services/analyseq/cgi-bin/forpub; Infobiogen, France). Hydropathy analyses and determination of potential N-glycosylation sites were performed using the servers TM-Pred Prediction of Transmembrane Regions and orientation (http://www.ch.embnet.org/software/TMPRED_form) and the NetNGlyc 1.0 program (http://www.cbs.dtu.dk/services/NetNGlyc/) from Expert Protein Analysis System (Swiss Institute of Bioinformatics, Switzerland). Sequence alignments were performed using the LFASTAp program (http://www.infobiogen.fr/services/analyseq/cgi-bin/lfastap) and clustalW algorithms (http://www.infobiogen.fr/services/analyseq/cgi-bin/clustalw).

To examine hST8Sia VI gene expression in various human cell lines total RNA from various cell lines (MCF-7, MDA-MB-231, T47-D and NBEC) was extracted using the NucleoSpin® RNA II kit and cellular RNA was quantified by spectrophotometry at 260 nm. For subsequent PCR amplifications, first-strand cDNA was synthesized from total RNA using the First Strand cDNA Synthesis kit according to the manufacturer's instructions.

A specific hST8Sia VI fragment of 834 bp was obtained after RT-PCR (reverse transcription-PCR) of RNAs isolated from various cultured cells, using 200 nM of sense (5'$_{376}$ AGAGCCAAAC-TTGCTTCCTGC 3') and antisense (5'$_{1194}$ TTAGGCGACTTCA-CATTTGCT 3') primers, 200 μM dNTP and 1 unit of EXTRA-POL I DNA polymerase using the following conditions: 95°C for 1 min, 38 cycles of 1 min at 95°C, 1 min at 58°C and 2 min at 72°C, and 10 min at 72°C.

Isolation of ST8Sia VI cDNA and construction of expression vectors

For subsequent plasmid constructions, restriction digestion and DNA sequencing (GenoScreen, France) confirmed the insert junctions and the total conservation between the inserted sequence and hST8Sia VI sequence (EMBL accession number: AJ621583). To isolate cDNA clones, RT-PCR was performed using 150 ng of first-strand cDNA synthesized from MCF-7 total RNA. A first cDNA amplification was performed by PCR using a sense (5'$_{-45}$ GCTGTGCTTCGCCCCGGCAGCAGC 3') and an antisense (5'$_{1388}$ TATGAGTCAGATATGGTGTCCATG 3') primer at 94°C for 2 min, followed by 40 cycles (94°C for 1 min, 62°C for 1 min and 72°C for 2 min) and an extension step of 10 min at 72°C. To obtain a cDNA encoding the full-length protein, the amplified products were used as template for a nested PCR using a sense primer containing the restriction site NotI (5'$_{-14}$ GCCTGGCGGCCGCGATGCGGCCGGGGGGCGCAC 3') and

an antisense primer containing the restriction site XbaI ($5'_{1233}$ CCTTTGGTGTTTGGAGACATCTAGAATCACCTACTGC 3′). The RT-PCR amplified fragments (1285 bp) were subcloned into PCR2.1 TOPO TA cloning vector giving the TOPO-L-ST8Sia VI plasmid. The inserted fragment was cut out by digestion with NotI and XbaI and inserted into the NotI and XbaI sites of pRc-CMV expression vector. The resulting plasmid (pRc-CMV-L-ST8Sia VI) was further purified and used for stable transfection of MDA-MB-231 cells, as described below.

An expression vector was also prepared for subsequent transient transfection in animal cells. A cDNA encoding a truncated form of hST8Sia VI lacking the first 27 amino acids of the open reading frame was obtained by PCR amplification using the TOPO-L-ST8Sia VI as DNA template, the same antisense primer containing the restriction site XbaI ($5'_{1233}$ CCTTTGGTGTTT-GGAGACATCTAGAATCACCTACTGC 3′) and a sense primer containing the restriction site NotI ($5'_{82}$ GACGCGCCCGCGG-CCGCCAGGATTCTGGTGGAG 3′). Reactions were run for 2 min at 94 °C, 1 min at 62 °C and 2 min at 72 °C for 37 cycles. The resulting 1197 bp fragment was subcloned into the pCR2.1 vector of TOPO TA-cloning kit, was cut out by digestion with NotI and XbaI, and inserted into the NotI and XbaI sites of the pFLAG expression vector. The resulting plasmid encoded a fusion protein with a signal peptide sequence, the FLAG sequence and a truncated form of hST8Sia VI (pFLAG–S-ST8Sia VI). Baculovirus-mediated insect cell expression was used to produce a large amount of a soluble form of hST8Sia VI. For this purpose, a truncated form of hST8Sia VI, lacking the first 24 amino acids, fused to the signal peptide sequence of a viral gene EGT (ecdysone-S-glycosyltransferase), and His$_6$-tagged in N-terminal was expressed into Sf-9 cells. A 1169 bp PCR fragment corresponding to the hST8Sia VI amino acids 25–398 was amplified using TOPO-L-ST8Sia-VI plasmid as DNA template (0.6 ng), a sense primer containing the last codon of the EGT signal, the restriction site HpaI and 6 histidine codons $5'_{72}$ GATGTTAACGCTCACCATCACCATCACCATTGCCCGGC-AGACGCG 3′ and an antisense primer containing the restriction sites HindIII and BamHI $5'_{1170}$ GGATCCAAGCTTTGTTTA-GGCGACTTCACATTTGCTAAA 3′. Reactions were run for 45 s at 94 °C, 1 min at 65 °C and 1 min 30 s at 72 °C for 10 cycles, followed by 45 s at 94 °C and 2 min at 72 °C for 25 cycles. The amplified DNA fragment was subcloned into the TOPO pCR2.1 vector for further sequencing, digested with HpaI and BamHI and the resulting digested fragment was inserted into the HpaI-BglII sites of the pUC-PSEGT plasmid [35]. This construction, named pUC-PSEGT-ST8Sia-VI was then digested with BamHI and HindIII and the fragment was inserted at the BglII-HindIII sites of the p119 transfer vector designed for recombination into the p10 locus of the AcMNPV baculovirus (clone 1.2) giving rise to the p119-ST8Sia-VI construct [36].

Production of soluble forms of hST8Sia VI

Qiagen-purified pFlag-S-ST8Sia-VI (5 μg) or pFlag plasmids (5 μg) were transiently transfected into COS-7 cells in 100 mm diameter dishes using LIPOFECTAMINE™ Plus reagent, according to the manufacturer's instructions. The medium was harvested 48 h after transfection and the recombinant protein expressed in the medium was used as the enzyme source for sialyltransferase assays. Sf-9 cells were co-transfected by lipofection using DOTAP with the p119-ST8Sia-VI transfer vector and purified viral DNA [37]. The recombinant baculoviruses were plaque purified and viral clones were tested for sialyltransferase activity as described below. The recombinant protein was used as the enzyme source for sialyltransferase assays.

Sialyltransferase assays and sialylated product characterization

Sialyltransferase assays were performed in 100 mM cacodylate buffer (pH 6.2) containing 10 mM MnCl$_2$, 0.2 % Triton CF-54, 40 μM CMP-[^{14}C]Neu5Ac (1.94 KBq) and one of the acceptor substrates (2 mg/ml for glycoproteins, 1 mM for arylglycosides or glycolipids) and 23 μl of the enzyme source in a total volume of 50 μl. Unless stated otherwise, the reactions were performed at 32 °C for 4 h. For glycoproteins, the reactions were stopped by addition of 2.5 × SDS/sample buffer and the reaction products were separated by SDS/PAGE. After transfer onto nitrocellulose membrane (Biotrace, Pall Corporation) the radioactive products were detected and quantified by radio-imaging using a Personal Molecular Imager FX (Bio-Rad, France). For glycolipids and arylglycosides, the incubation reaction was stopped with 1 ml of H$_2$O and products were applied onto C$_{18}$ Sep-Pak cartridges (Millipore Corp., Milford, MA, U.S.A.), eluted with 30 % CH$_3$OH and processed for scintillation counting.

[^{14}C]Neu5Ac-labelled fetuin was produced using soluble recombinant hST8Sia VI in the incubation conditions described above. For linkage analysis of sialic acids, sialylated fetuin was treated with specific sialidases at 37 °C for 1 h, according to the manufacturer's instructions (Glyko Inc.). For analysis of the glycan acceptor, sialylated fetuin was incubated with PGNase F at 37 °C for 30–120 min according to the manufacturer's instructions. In a second experiment, native fetuin was desialylated by Glyco® Sialidase S and was resialylated by the soluble recombinant hST8Sia VI. The resulting products were also resolved by SDS/PAGE and detected by radio-imaging. For analysis of the DP, sialylated fetuin and colominic acid were separately incubated for 2.5 h at 50 °C in 50 μl of DMB reagent (2.7 mM DMB, 9 mM sodium hydrosulphite, 0.5 mM 2-mercaptoethanol in 20 mM trifluoroacetic acid) [38]. A volume (one-fifth) of 0.1 M NaOH was added and incubated for 1 h at room temperature to terminate the reaction and to eventually hydrolyse lactones [39]. DMB-tagged oligosialylated sequences from fetuin and colominic acid were co-injected onto a CarboPac PA-100 column (Dionex). Elution was performed at 0.8 ml/min with a concentration gradient of 0–32 % of 1 M KNO$_3$ in water. Elution was monitored by an on-line fluorescence detector set to 372 nm for excitation and 456 nm for emission. The samples were automatically collected and subsequently counted for radioactivity. The concentration of fetuin and asialofetuin were calculated on the basis of the number of O-glycosidically linked Galβ1-3GalNAc residues, determined by gas chromatography (61 nmol/mg).

Stable transfections

MDA-MB-231 cells were transfected with 20 μg of pRc-CMV or pRc-CMV-L-ST8SiaVI plasmids using LIPOFECTAMINE™. After transfections (2 days), the dishes were split at a 1:10 ratio and cells were cultured in MEM containing 5 % (v/v) FCS and geneticin G418 (1 mg/ml). MEM and geneticin were replaced every 2 days. After 15 days, the resistant colonies were well established, the cells selected and maintained with 1 mg/ml geneticin and 10 % (v/v) FCS and used for flow cytometric analysis experiments.

Detection of α2,6-linked and α2,3-linked sialic acids in MDA-MB-231 transfected cells

Cell surface expression of sialic acids in two different types of linkages was quantified using SNA specific for α2,6-linked sialic acids and MAA specific for α2,3-linked sialic acids. Stably transfected cells were fixed with 4 % paraformaldehyde (30 min at 4 °C in the dark), quenched for 30 min with 50 mM NH$_4$Cl in PBS
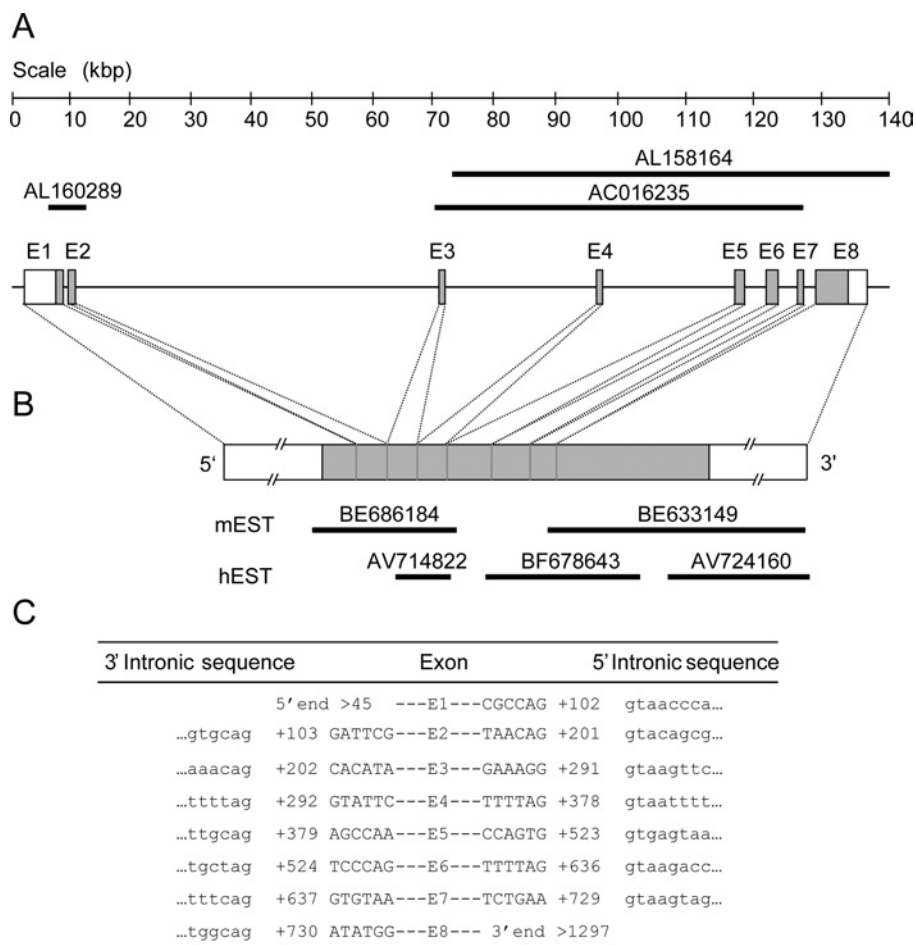
Figure 1 Genomic organization and exon/intron junctions of the hST8Sia VI gene

(**A**) The hST8Sia VI gene located on chromosome 10p12.31, spans 140 kb and contains 8 exons (labelled E1–E8). Black lines represent the DNA sequences identified in the human genomic databases. The entire genomic sequence of hST8Sia VI is included in the human contig: NT 008682.3. (**B**) Schematic representation of the hST8Sia VI mRNA. The grey boxes represent the open reading frame and the open boxes the 5′ and 3′ untranslated regions respectively. The black lines represent the mouse (m) and human (h) EST identified in the public databases. (**C**) The nt sequences at the intron (lowercase letters) and exon (uppercase letters) junctions are shown. Exons are numbered from the 5′ end with the initiator methionine denoted as +1.

and incubated overnight at 4°C with SNA or MAA [1:50 in PBS, 1% (v/v) BSA]. After washing with PBS containing 1% BSA the cells were incubated with anti-digoxigenin-fluorescein Fab fragments (1:500 in PBS, 1% BSA) for 2 h on ice. Flow cytometric analysis of the cells stained with lectins was performed using a FACScan fluorospectrometer (Becton Dickinson).

RESULTS

Identification of hST8Sia VI gene

Similarity searches using the BLAST algorithms (BLASTn and tBLASTn) were conducted with the hST8Sia I cDNA (GenBank® accession number: D26360) used as a bioinformatics probe in the human genomic databases (HTG division of NCBI). This study led us to the identification of two genomic clones AC016235 and AL158164 (Figure 1A), which partially aligned with the hST8Sia I cDNA and encompassed the signature motifs of mammalian sialyltransferases. In addition, these nt sequences were localized on human chromosome 10 where no sialyltransferase gene was described. Upon screening the human EST division of the NCBI databases, three ESTs, AV714822, BF678643 and AV724160 were retrieved, which corresponded to the 3′ region of the putative

gene whereas no human EST could be found to correspond with the 5′ end of the gene (Figure 1B). However, BLAST analysis conducted in the mouse EST division of NCBI revealed the existence of two ESTs (BE686184 and BE633149) belonging to the same Integrated Molecular Analysis of Genome Expression (I.M.A.G.E.) clone (ID 34128403) that represented the mouse counterpart of the human sequence used for the search (Figure 1B) and which allowed us to identify the missing upstream sequences of the putative sialyltransferase in the human genomic clone, AL160289. The gene identification was completed at the Ensembl site (http://www.ensembl.org). As shown in Figure 1(A) the putative sialyltransferase gene spans over 140 kb of human genomic sequence (NT 008682) on human chromosome 10 (10p12.31). This nt sequence contains an open reading frame of 1197 bp (Figure 1B). In order to gain insights into the evolutional relationship of the hST8Sia VI with other sialyltransferase genes, the hST8Sia VI genomic organization was determined (Figure 1C) through the use of a specialized internet site (http://www.fruitfly.org/seq_tools/splice.html). The sequence of the splice junctions of the hST8Sia VI gene obey the AG/GT rule and the gene splits into eight coding exons in a genomic organization closely related to that of hST8Sia I and hST8Sia V (results not shown, see [40]).

```
                              ctgtgcttcgccccggcagcagcggtggcggcggcgcctgtggctcagg    -1

  M   R   P   G   G   A   L   L   A   L   L   A   S   L   L   L   L   L   L   L   R   L   L   W   C    25
 ATG CGG CCG GGG GGC GCA CTG CTC GCC CTG CTC GCC AGC CTG CTG CTG CTG CTG CTG CTG CGC CTG CTC TGG TGC    75

  P   A   D   A   P   G   R   A   R   I   L   V   E   E   S   R   E   A   T   H   G   T   P   A   A    50
 CCG GCA GAC GCG CCC GGC CGC GCC AGG ATT CTG GTG GAG GAA AGC AGG GAG GCC ACC CAC GGC ACC CCC GCA GCG   150

  L   R   T   L   R   S   P   A   T   A   V   P   R   A   T  (N)  S   T   Y   L   N   E   K   S   L    75
 CTG AGG ACG CTC CGG AGC CCG GCG ACC GCG GTA CCG CGC GCC ACT AAC AGC ACA TAT CTG AAT GAG AAG TCG CTC   225

  Q   L   T   E   K   C   K   N   L   Q   Y   G   I   E   S   F   S  (N)  K   T   K   R   Y   S   E   100
 CAA CTG ACG GAG AAA TGC AAA AAT CTG CAA TAT GGC ATT GAG TCT TTC TCT AAC AAA ACG AAA AGG TAT TCA GAG   300

  N   D   Y   L   Q   I   I   T   D   I   Q   S   C   P   W   K   R   Q   A   E   E   Y   A   N   F   125
 AAC GAC TAC CTT CAG ATT ATC ACA GAT ATA CAG AGT TGT CCA TGG AAA CGG CAA GCA GAA GAA TAT GCA AAT TTT   375

  R   A   K   L   A   S   C   C   D   A   V   Q   N   F   V   V   S   Q  (N)  N   T   P   V   G   T   150
 AGA GCC AAA CTT GCT TCC TGC TGT GAT GCT GTT CAA AAC TTT GTT GTT TCT CAG AAT AAC ACT CCA GTT GGG ACT   450

 (N)  M   S   Y   E   V   E   S   K   K   E   I   P   I   K   K   N   I   F   H   M   F   P   V   S   175
 AAT ATG AGT TAC GAG GTG GAA AGC AAA AAA GAA ATC CCA ATT AAG AAG AAC ATT TTT CAT ATG TTT CCA GTG TCC   525

  Q   P   F   V   D   Y   P   Y   N   Q   C   A   V   V   G   N   G   G   I   L  (N)  K   S   L   C   200
 CAG CCT TTT GTG GAC TAC CCT TAT AAT CAG TGT GCA GTG GTC GGA AAT GGG GGA ATT CTG AAT AAG TCT CTC TGT   600

  G   T   E   I   D   K   S   D   F   V   F   R   C   N   L   P   P   T   T   G   D   V   S   K   D   225
 GGA ACT GAA ATA GAT AAA TCC GAC TTC GTT TTT AGG TGT AAC CTA CCC CCA ACC ACA GGA GAT GTT AGT AAA GAT   675

  V   G   S   K   T   N   L   V   T   I   N   P   S   I   I   T   L   K   Y   G   N   L   K   E   K   250
 GTT GGC AGT AAA ACA AAT CTT GTG ACT ATA AAT CCA AGC ATC ATA ACT CTG AAA TAT GGG AAC TTA AAG GAA AAA   750

  K   A   L   F   L   E   D   I   A   T   Y   G   D   A   F   F   L   L   P   A   F   S   F   R   A   275
 AAA GCC CTA TTT CTG GAG GAC ATT GCA ACC TAT GGA GAT GCA TTT TTT CTT CTG CCA GCA TTT TCC TTC AGG GCC   825

  N   T   G   T   S   F   K   V   Y   Y   T   L   E   E   S   K   A   R   Q   K   V   L   F   F   H   300
 AAC ACG GGT ACC TCT TTC AAA GTA TAC TAC ACG CTC GAA GAG TCT AAA GCA AGA CAA AAG GTT CTA TTT TTC CAT   900

  P   K   Y   L   K   D   L   A   L   F   W   R   T   K   G   V   T   A   Y   R   L   S   T   G   L   325
 CCC AAG TAC CTG AAA GAT CTG GCC CTT TTC TGG AGA ACT AAA GGT GTG ACT GCA TAC CGC TTG TCC ACC GGC TTG   975

  M   I   T   S   V   A   V   E   L   C   K   N   V   K   L   Y   G   F   W   P   F   S   K   T   V   350
 ATG ATC ACA AGT GTT GCA GTG GAA CTG TGT AAA AAT GTG AAG CTG TAT GGA TTC TGG CCC TTC TCT AAA ACT GTA  1050

  E   D   I   P   V   S   H   H   Y   Y   D   N   K   L   P   K   H   G   F  [H]  Q   M   P   K  [E]  375
 GAA GAC ATA CCT GTC AGC CAT CAC TAT TAT GAC AAC AAG CTA CCT AAA CAT GGT TTC CAT CAG ATG CCC AAA GAA  1125

  Y   S   Q   I   L   Q   L   H   M   K   G   I   L   K   L   Q   F   S   K   C   E   V   A   *        398
 TAC AGC CAG ATC CTC CAA CTT CAC ATG AAA GGA ATC CTC AAA CTG CAA TTT AGC AAA TGT GAA GTC GCC TAA       1197

 acaaagtatcttaaaatgggaataattttaatataa                                                                1230
```

Figure 2   nt and predicted amino acid sequences of hST8Sia VI

Numbering of the cDNA begins with the initiation codon. The amino acid sequence is shown in single-letter code. The putative 19 amino acid N-terminal transmembrane domain is underlined. Putative *N*-glycosylation sites are circled. Dashed and dotted lines underline the sialyl motifs L and S respectively. The conserved His and Glu residues in sialyl motif VS are boxed.

Cloning and nt sequencing of hST8Sia VI

The nt sequences described above were used to design oligo nt primers for PCR amplification with MCF-7 first strand cDNA as template. However, probably because of the low level of expression of a long transcript and strong secondary RNA structures, we could not amplify the target sequence at once, so we amplified a full-length cDNA by nested PCR. Upon DNA sequencing of several of these MCF-7 PCR-amplified cDNA clones, we noticed various splicing variants of hST8Sia VI gene lacking 1 or 2 exons (results not shown). Indeed, theoretically, exons numbered 2, 3, 4 and 7 can be deleted without frame-shift in the open reading frame. The nt and deduced amino acid sequences of the full-length cDNA sequence obtained are shown in Figure 2. Hydropathy analysis of this protein indicated the presence of a hydrophobic sequence of 19 amino acids in the NH$_2$-terminal region, corresponding to the transmembrane domain characteristic of all sialyltransferases cloned to date (results not shown). The predicted protein consisted of 398 amino acids with a calculated molecular mass of 44 835 Da,

contained in the catalytic domain the four sialylmotifs L, S, III and VS characteristic of all the animal sialyltransferases, and showed five potential *N*-glycosylation sites (Figure 2). The position of the initiation codon was estimated according to the Kozak consensus sequence [41] (http://www.hri.co.jp/atgpr/). In addition, the protein showed the (I/L)(F/Y)GFW(P/A)F sequence at the 3′ end of the sialylmotifs, which is conserved among the members of the α2,8-sialyltransferase family. Besides, comparison of the amino acid sequence with those of the other human sialyltransferases indicated a phylogenetic linkage with the α2,8-sialyltransferases and particularly the greatest sequence identities with hST8Sia I (35 %) and hST8Sia V (32 %). Finally, hST8Sia VI is the human counterpart of a mouse sequence (GenBank® accession number AB059554) cloned during the time-course of our study [32]. The two sequences share 82.7 % sequence identity. Orthologues of the hST8Sia VI gene are also detected in several other vertebrate genomes such as *Rattus norvegicus* (AJ699423), *Bos taurus* (AJ868432), *Gallus gallus* (AJ699424), *Takifugu rubripes* (AJ715549, AJ715550) and *Danio*
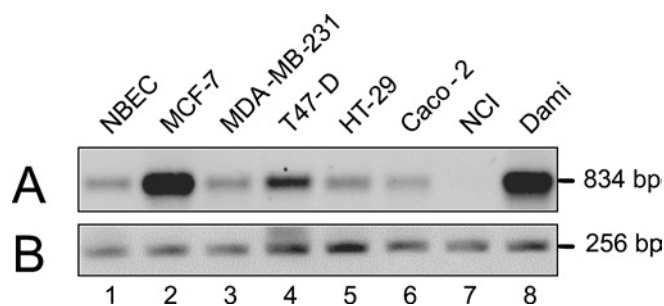
Figure 3   RT-PCR analysis of the expression of hST8Sia VI gene in various human cell lines

(**A**) mRNAs levels of hST8Sia VI were evaluated by RT-PCR, as described in the Experimental section, in various cell lines; lane 1, NBEC; lane 2, MCF-7; lane 3, MDA-MB-231; lane 4, T47-D; lane 5, HT-29; lane 6, Caco-2; lane 7, NCI; lane 8, Dami. (**B**) Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used as standard as previously described [50]. The sizes of the amplified fragments are indicated on the right side of the gel.

rerio (AJ715551) [40]. The deduced amino acid sequences of these genes display 81.4 %, 82.2 %, 59.4 %, 35.2 %, 34.4 % and 36.9 % sequence identity respectively with the hST8Sia VI gene.

hST8Sia VI gene expression in human tissues and cell lines

In order to determine the expression pattern of hST8Sia VI gene and the size of hST8Sia VI mRNA, hybridization of a Northern blot of 12 human tissues and of an expression array of 72 different human tissues was performed with a full-length hST8Sia VI cDNA (1.27 kb). However, the expression levels of the ST8Sia VI gene were very low in all the tissues examined (results not shown). Thus we performed semi-quantitative RT-PCR analysis in several human cancer and normal cell lines. As shown in Figure 3, higher levels of expression of the hST8Sia VI gene were observed in breast cancer cell line MCF-7 and in Dami megakaryocyte cell line whereas very low levels of expression were observed in NBEC, in breast cancer cell line MDA-MB-231 and T47-D and in colon cancer cell lines HT-29 and Caco-2. No expression could be detected in the lung cancer cell line NCI.

Decreased expression of MAA-binding sites in the MDA-MB-231 cell line stably transfected with full-length hST8Sia VI

In order to assess the function of hST8Sia *in vivo*, a MDA-MB-231 cell line was stably transfected with the expression vector pRc-CMV-L-hST8Sia VI (full-length) as described in the Experimental section. Cell clones were selected by mRNA expression (Figure 4A) and by ST8Sia activity towards fetuin used as an acceptor substrate for *in vitro* assays (Figure 4B). Characterization of the sialylated carbohydrate profile of MDA-ST8Sia VI, as well as of the mock-transfected cells was performed by flow cytometry analysis of these cells using MAA and SNA lectins that are specific for $\alpha2,3$- and $\alpha2,6$-sialylated oligo-saccharides respectively. As shown in Figure 4(C), *in vivo* expression of hST8Sia VI led to a 4-fold decrease in MAA-binding sites in MDA-MB-231 transfected cells whereas very limited changes were observed for the SNA-binding sites in MDA-MB-231 transfected cells (Figure 4C). Our results suggest that *in vivo* hST8Sia VI catalyses the transfer of the sialic acid residue onto the disaccharide structure Neu5Acα2-3Gal, but had very limited activity towards the disaccharide structures Neu5Acα2-6GalNAc.
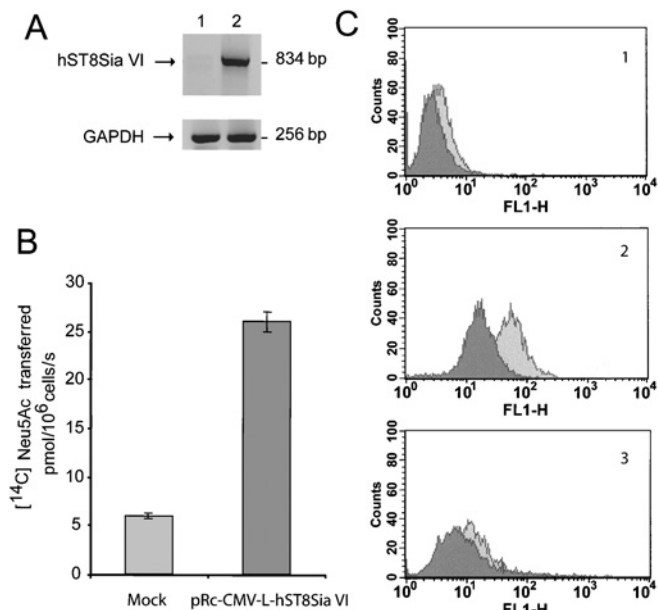
Figure 4   Expression of full-length human ST8Sia VI in MDA-MB-231 breast cancer cells

(**A**) The expression of human ST8Sia VI cDNA in MDA-MB-231 breast cancer cells was determined by RT-PCR. Lane 1, mock transfected cells; lane 2, pRc-CMV-L-hST8Sia VI transfected cells. The size of the amplified fragments are indicated on the right side. (**B**) hST8Sia VI enzymatic activity measured in selected cell clones. Enzymatic activity was measured on fetuin and results are expressed in pmol of transferred Neu5Ac residues per 10^6 cells/s. (**C**) Flow cytometry analysis of pRc-CMV-L-hST8Sia VI transfected cells. MDA-MB-231 expressing (pRc-CMV-L-ST8Sia VI transfected cells; dark grey peaks) or not expressing hST8Sia VI (mock transfected cells; light grey peaks), were subjected to MAA and SNA labelling. MAA revealed the presence of a terminal α3-linked sialic acid and SNA the presence of a terminal α6-linked sialic acid. Panel 1, negative control (secondary antibodies only); panel 2, MAA staining; panel 3, SNA staining.

Sialyltransferase activity of the hST8Sia VI

In order to conclusively verify that the cloned DNA which we had obtained from MCF-7 cells indeed represented human ST8Sia VI, and to facilitate functional analysis of the enzyme, we produced a soluble form of the enzyme with deleted cytoplasmic and transmembrane domains, using the baculovirus/*Sf*-9 system. Thus a truncated cDNA lacking the first 24 amino acid residues of the N-terminus region was generated by PCR and used for the construction of a p119-hST8Sia VI vector, as described in the Experimental section. The recombinant hST8Sia VI enzyme produced in insect cells was used for *in vitro* sialyltransferase assays with various glycoconjugate acceptors (Table 1). This recombinant hST8Sia VI enzyme exhibited highly decreased sialyltransferase activity towards asialoglycoproteins compared with the corresponding sialylated substrates indicating that the enzyme required a sialylated substrate. hST8Sia VI exhibited strong sialyltransferase activity towards BSM (bovine submaxillary mucin) and OSM (ovine submaxillary mucin), which contains only *O*-glycans and towards fetuin, which contains both *N*- and *O*-glycans. On the contrary, it showed very low activity towards α1-acid glycoprotein, which contains only *N*-glycans. This set of results clearly showed that the recombinant hST8Sia VI preferred sialylated *O*-glycans such as Neu5Acα2-3Galβ1-3GalNAc-*O*-Ser/Thr or Neu5Acα2-6GalNAc-*O*-Ser/Thr found on fetuin, to sialylated *N*-glycans as acceptor substrates. We next examined hST8Sia VI activity towards various glycolipids such as G_{M3}, G_{D3}, or G_{D1b} containing a sialylated lactosylceramide fraction (Table 1). hST8SiaVI showed almost no activity (< 11 %) towards these

Table 1 Comparison of the acceptor substrate specificity of hST8Sia VI and m (mouse) ST8Sia VI

Various acceptor substrates were incubated in the standard assay mixture using transfected *Sf*-9 cells as an enzyme source. Each substrate was used at a concentration of 1 mM for arylglycosides, 2 mg/ml for glycoproteins and 1 mM for glycolipids. The specific [^{14}C]Neu5Ac incorporation resulting from recombinant hST8Sia VI activity was estimated by subtracting the background sialyltransferase activities of mock transfected cells from the radioactivity in the presence of exogenous acceptor. Relative rates were calculated as a percentage of the incorporation of sialic acid onto fetuin. A value of 0 indicates < 1 %; nd, not determined; Cer, ceramide.

| Acceptor	Structures	Relative rate (%)	
		hST8Sia VI	mST8Sia VI *
Glycoproteins			
Fetuin	Neu5Acα2-3Galβ1-3GalNAcα1-O-Ser/Thr		
	Neu5Acα2-3Galβ1-3[Neu5Acα2-6]GalNAcα1-O-Ser/Thr		
	Neu5Acα2-6(3)Galβ1-4GlcNAc-R	100 (0.15)†	100 (0.206)†
Asialofetuin	Galβ1-3GalNAcα1-O-Ser/Thr and Galβ1-4GlcNAc-R	10	0
α_1-Acid glycoprotein	Neu5Acα2-6Galβ1-4GlcNAc-R	15	0
Asialo-α_1-acid glycoprotein	Galβ1-4GlcNAc-R	6	0
BSM	GlcNAcβ1-3[Neu5Acα2-6]GalNAcα1-O-Ser/Thr		
	Neu5Acα2-6GalNAcα1-O-Ser/Thr	350	375
Asialo-BSM	GlcNAcβ1-3GalNAcα1-O-Ser/Thr		
	GalNAcα1-O-Ser/Thr	15	0
OSM	Neu5Acα2-6GalNAcα1-O-Ser/Thr	334	nd
Asialo-OSM	GalNAcα1-O-Ser/Thr	12	nd
Glycolipids			
G$_{M3}$	Neu5Acα2-3Galβ1-4Glcβ1-1'Cer	3	13
G$_{D3}$	Neu5Acα2-8Neu5Acα2-3Galβ1-4Glcβ1-1'Cer	8	0
G$_{D1}$	Neu5Acα2-3Galβ1-3GalNAcβ1-4[Neu5Acα2-3]Galβ1-4Glcβ1-1'Cer	12	6
G$_{D1}$	Galβ1-3GalNAcβ1-4[Neu5Acα2-8Neu5Acα2-3]Galβ1-4Glcβ1-1'Cer	11	0
G$_{T1}$	Neu5Acα2-3Galβ1-3GalNAcβ1-4[Neu5Acα2-8Neu5Acα2-3]Galβ1-4Glcβ1-1'Cer	25	1.1
Arylglycosides			
	Galβ1-3GlcNAcβ-octyl	11	nd
	Neu5Acα2-6Galβ1-4GlcNAcβ1-2Manα-octyl	3	nd
	Neu5Gcα2-3Galβ1-3GalNAcβ-octyl	3	nd

* Data are from [32].
† Actual activities are shown in brackets in pmol/h per μl.

compounds in contrast with the mouse ST8SiaVI enzyme, which showed noticeable activity towards G$_{M3}$. However, we observed higher sialyltransferase activity of hST8Sia VI towards the gangliosides G$_{T1b}$ and G$_{D1a}$, (25 % and 12 % respectively) which both contain a terminal Neu5Acα2-3Galβ1-3GalNAc sequence whereas the mouse enzyme showed almost no activity towards these compounds. Finally, hST8Sia VI exhibited almost no activity towards arylglycosides (< 11 %) (Table 1). These results suggest that hST8Sia VI prefers Neu5Acα2-3Galβ1-3GalNAc-R (where R is the remainder of the N-linked oligosaccharide chain) or Neu5Acα2-6GalNAc-R found in the O-glycans and also is highly sensitive to the underlying sequences.

In order to confirm that hST8Sia VI preferentially sialylates O-glycans, α2,8-[^{14}C]Neu5Ac-labelled sialylated fetuin was produced using a soluble recombinant hST8Sia VI (see the Experimental section) and was subjected to PNGase F, which hydrolyses N-glycans (Figure 5). After 30 min treatment, we observed a shift of mass corresponding to the loss of N-glycans but the intensity of the signal remained unchanged, suggesting that the sialic acid residue was transferred onto O-glycans of fetuin rather than onto N-glycans.

To characterize the linkage catalysed by hST8Sia VI, α2,8-[^{14}C]Neu5Ac-labelled fetuin was produced using soluble recombinant hST8Sia VI. It was then subjected to linkage specific sialidases, Glyco® Sialidase S (specific for α2,3-linked sialic acid), Glyco® Sialidase C (specific for α2,3/6-linked sialic acid) and Glyco® Sialidase A™ (specific for α2,3/6/8-linked sialic acid) (Figure 6A). We observed that the treatment with the Glyco® Sialidase S and C, which hydrolyse the α2,3- and α2,6-linked sialic acid is ineffective on the signal. Conversely, this signal disappeared when labelled fetuin was treated with Glyco®
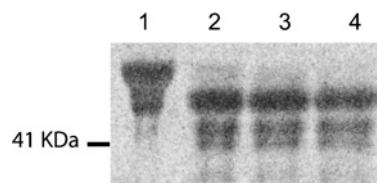
Figure 5 *N*-deglycosylation of sialylated fetuin

[^{14}C]Neu5Ac-labelled fetuin was produced using soluble recombinant hST8Sia VI and was subjected to PNGase F treatment for 0 min (lane 1), 30 min (lane 2), 60 min (lane 3) or 120 min (lane 4). The resulting products were separated by SDS/PAGE and detected by phosphorimaging. The molecular mass marker is indicated on the left side of the gel.

Sialidase A™, which is specific for α2,3/6/8-linked sialic acid, indicating that the incorporated sialic acid was α2,8-linked. These results confirmed that the cloned hST8Sia VI belonged to the ST8Sia subfamily.

We next aimed to determine whether hST8Sia VI preferentially sialylated Neu5Acα2-3Galβ1-3GalNAc-R acceptor substrate or the Neu5Acα2-6GalNAc-R acceptor substrate that are both present on native bovine fetuin. Thus either native fetuin or α2,3-desialylated fetuin was incubated with recombinant hST8Sia VI. Indicative of the efficiency of the desialylation, a shift of mass was observed after Ponceau Red staining of the Western blot, when α2,3-sialic acid was removed before incubation (results not shown). As shown in Figure 6(B), hST8Sia VI did not mediate the efficient transfer of sialic acid residue onto α2,3-desialylated fetuin. This result suggested that the recombinant hST8Sia VI preferentially catalysed the transfer of a sialic acid residue to an α2,3-linked Neu5Ac rather than to an α2,6-linked Neu5Ac and
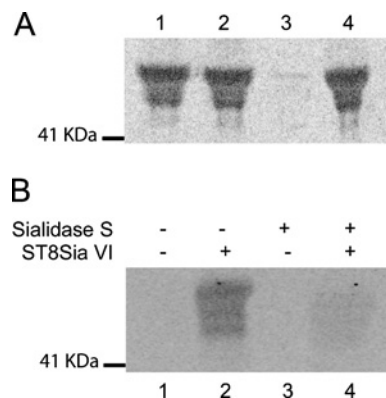
Figure 6   Linkage analysis of incorporated sialic acid

(**A**) [^{14}C]Neu5Ac-labelled fetuin was produced using soluble recombinant hST8Sia VI and was subjected to sialidase treatment with Glyco® Sialidase S (specific for α2,3-linked sialic acid, lane 1), Glyco® Sialidase C (specific for α2,3-and α2,6-linked sialic acids, lane 2), Glyco® Sialidase A^{TM} (specific for α2,3-α2,6 and α2,8/9-linked sialic acids, lane 3) or none (lane 4). The resulting products were separated by SDS/PAGE and detected by phosphorimaging. (**B**) Native fetuin (lanes 1 and 2) or α3-desialylated fetuin (lanes 3 and 4) were [^{14}C]Neu5Ac-sialylated using soluble recombinant hST8Sia VI (lanes 2 and 4) or using mock-transfected cells conditioned medium (lanes 1 and 3). The resulting products were separated by SDS/PAGE and detected by phosphorimaging. The molecular mass marker is indicated on the left side.

confirmed the flow cytometric analysis of the transfected MDA-ST8Sia VI cells with the MAA lectin.

DP analysis

Based on the above results, we next addressed the question of whether hST8Sia VI may synthesize di-, oligo-, or PSA chains onto fetuin. For this purpose, native bovine fetuin was first incubated with hST8Sia VI and CMP-[^{14}C]Neu5Ac. Sialylated sequences were then liberated from glycoproteins by mild acid hydrolysis, tagged with DMB and separated by HPLC on a CarboPac PA-100 column. Retention times for [Neu5Acα2-8]$_n$-Neu5Ac-DMB sequences were determined by injecting tagged total colominic acid standard, as well as di-, tri- and tetra-sialylated sequences purified from a mild hydrolysate of colominic acid. As shown in Figure 7, in the hST8Sia VI incubation mixture, we clearly observed a radioactive peak with a retention time similar to Neu5Acα2-8Neu5Ac-DMB, strongly suggesting that synthesis of the di-sialylated epitope occurred from sialylated fetuin. Accordingly, such a peak was not observed in either the mock-transfected cell incubation mixture or native fetuin, indicating that hST8Sia VI catalysed the transfer of a single sialic acid residue onto native fetuin. Furthermore, a time course analysis of the sialylation of fetuin clearly confirmed the specificity of this enzyme for biosynthesis of diSia motifs.

DISCUSSION

In the present study, we report on the molecular cloning and enzymatic characterization of a human sialyltransferase gene that encodes an ST8Sia with preference for *O*-glycans. This gene spans over 140 kb of human chromosome 10 (10p12.31) and splits into eight coding exons with an organization closely related to that of the hST8Sia I and hST8Sia V genes. Phylogenetic analyses have shown that the ST8Sia gene family can be divided into two groups of genes, which evolved from a common ancestor and it was observed that the human sialyltransferases of each group showed distinct enzymatic properties [40]. In the first group, ST8Sia II, III and IV are known as oligo- or poly-sialyltransferases involved in the elongation of linear chains of sialic
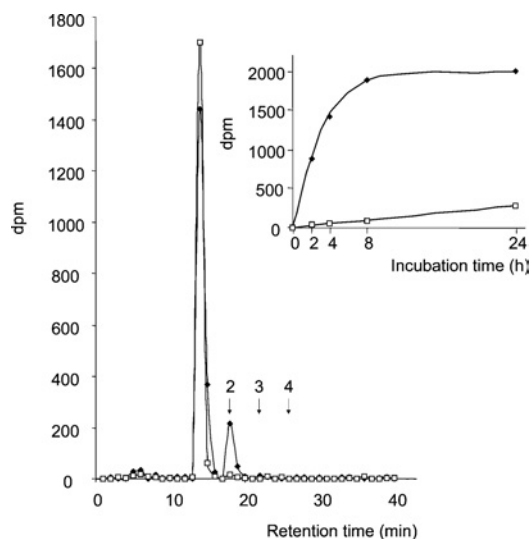
Figure 7   HPLC/DMB analysis of α2,8-linked-[^{14}C]-Neu5Ac on fetuin

Native fetuin was [^{14}C]Neu5Ac-sialylated for 4 h using soluble recombinant hST8Sia VI or using mock-transfected cells conditioned medium and subjected to mild acid hydrolysis followed by DMB derivation. Oligosialylated derivatives were fractionated by HLPC on CarboPac PA-100 column according to the DP. Arrows indicate the retention time of standard di-, tri- and tetrasialo-DMB derivatives prepared from partially hydrolysed colominic acid. Inset: time course analysis of [^{14}C]-labelled disialo-DMB derivatives synthesis. (◆) hST8Sia VI; (□) mock.

acids found mainly in glycoproteins. ST8Sia II and IV were clearly shown to be responsible for the elongation of polysialylated chains of NCAM [26] and even if the substrate specificity of ST8Sia III is not clearly demonstrated, this latter enzyme is capable of autopolysialylation *in vitro* [29]. In the second group encompassing ST8Sia I, ST8Sia V and ST8Sia VI, ST8Sia I (G$_{D3}$ synthase) and ST8Sia V (G$_{T3}$ synthase) are acting as di-sialyltransferases, transferring only one sialic acid residue onto G$_{M3}$ or G$_{D3}$ to convert these gangliosides into G$_{D3}$ or G$_{T3}$ respectively. It could be predicted that the enzymatic properties of hST8Sia VI would be those of a di-sialyltransferase. As hypothesized, we have clearly demonstrated that the human ST8Sia VI transferred only one sialic acid residue in α2,8-linkage. However, in contrast with its recently characterized mouse counterpart [32], which also acts on the Neu5Acα2-3Galβ1-4Glc sequence of G$_{M3}$ and on the Neu5Acα2-6Galβ sequences such as are found in *N*-glycosyl-proteins, the hST8Sia VI works almost exclusively on *O*-glycans. Stable transfections of hST8Sia VI in the MDA-MB-231 cell line have shown that hST8Sia VI works preferentially on cell surface α2,3-sialylated structures. Our *in vitro* assays using a soluble form of the enzyme have demonstrated that hST8Sia VI preferentially uses the trisaccharide Neu5Acα2-3Galβ1-3GalNAcα1-*O*-Ser/Thr found in bovine fetuin, and to a lesser extent, to that found in gangliosides (G$_{D1a}$). It shows also an activity towards the disaccharide Neu5Acα2-6GalNAcα1-*O*-Ser/Thr found on *O*-glycans of BSM and OSM. This discrepancy between the specific enzymatic activity of the mouse and the human enzymes could be related to inter-species enzymatic differences which indeed lead to species-specific glycosylation. Currently work is in progress to determine which *O*-glycosylprotein(s) might be disialylated in these stably-transfected MDA cells using specific monoclonal antibodies towards diSia, oligoSia or PSA [42].

The occurrence of diSia motifs in animal tissues was suggested a long time ago by Finne et al. [43] but indeed, Neu5Acα2-8Neu5Acα2-3Galβ1-3GalNAc-α1-*O*-Ser/Thr or Galβ1-3[Neu5Acα2-8Neu5Acα2-6]GalNAc-α1-*O*-Ser/Thr oligosaccharides

are not frequently described in mammalian glycoproteins. However, these diSia motifs were shown previously to occur in several embryonic and adult pig brain glycoproteins [44], in the murine CD-166 [18], on the *N*-glycans of Band 3 of human erythrocytes [14], on the *N*-glycans of fetuin and α2-macroglobulin of calf serum [17], on the *O*-glycans of human erythrocyte glycophorin [15], bovine adipo-Q [19] and chromogranin [16]. It has been suggested that the biosynthesis of these diSia residues on glycoproteins might be catalysed by an α2,8-sialyltransferase distinct from the known polysialyltransferases ST8Sia II and ST8Sia IV and it was shown that a recombinant ST8Sia III could sialylate purified bovine adipo-Q [19]. We have clearly demonstrated in *in vitro* assays that hST8Sia VI is highly specific for the synthesis of the diSia epitope found on *O*-glycosylproteins although the presence of this structure remains to be demonstrated *in vivo*. However, diSia epitopes on *O*-glycosylproteins might represent biosynthetic intermediates, thus generating a multiplicity of di-, oligo- and poly-Sia structures that have recently been described in mammalian cell lines and tissues [44,45]. The development of chemical detection tools [46] or biochemical tools [42] should help us to solve these biosynthetic pathways.

Whereas the importance of the diSia epitope on glycolipids is well documented, the role of the diSia epitopes found on glycoproteins is not known. As suggested by the fact that the anti-diSia antibody specifically blocked neurite extension in the Neuro2A mouse neuroblastoma cell line, it has been proposed that the diSia epitope on glycoproteins could be involved in neurite formation [18]. During the adipocyte differentiation of murine fibroblastic 3T3-L1 preadipocytes, significant increases in diSia structures on glycoproteins were observed in fully differentiated adipocyte cells and among them the serum adipo-Q glycoprotein has been shown to have the diSia epitope [19]. Adipo-Q is considered to play an important role in energy homoeostasis and it is believed that the diSia epitope could be functionally involved in these physiological phenomena in a way that remains to be determined [19]. Of the siglecs cloned so far, it is known that siglec-7 and siglec-11 have a high affinity for the diSia epitopes found on glycoproteins [47]. Siglec-11 shows a unique expression pattern on tissue macrophages and infiltrating mononuclear leucocytes in inflammatory tissues suggesting that this molecule and the diSia epitopes might be involved in the regulation of innate immune responses. Siglec-7 is the major siglec expressed by natural killer cells [48] and displays unique ligand binding properties, different from other members of the siglec family, towards α2,8-linked sialic acids over α2,6- and α2,3-linked sialic acids [49]. One can assume that glycan recognition by siglec-7 is likely to be directly linked to its function in modulating the activation of natural killer cells. The characterization of this new ST8Sia VI might help to determine the function of the diSia epitopes found on *O*-glycosylated proteins, although the presence of this structure remains to be demonstrated *in vivo*.

REFERENCES

1 Tsuji, S., Datta, A. K. and Paulson, J. C. (1996) Systematic nomenclature for sialyltransferases. Glycobiology **6**, 5–7

2 Troy, F. A. I. (1996) Sialobiology and the polysialic acid glycotope. In Biology of the Sialic Acids (Rosenberg, A., ed.), pp. 95–144, Plenum Press, New York and London

3 Finne, J. (1982) Occurrence of unique polysialosyl carbohydrate units in glycoproteins of developing brain. J. Biol. Chem. **257**, 11966–11970

4 Livingston, B. D., Jacobs, J. L., Glick, M. C. and Troy, F. A. (1988) Extended polysialic acid chains (n > 55) in glycoproteins from human neuroblastoma cells. J. Biol. Chem. **263**, 9443–9448

5 Yabe, U., Sato, C., Matsuda, T. and Kitajima, K. (2003) Polysialic acid in human milk. CD36 is a new member of mammalian polysialic acid-containing glycoprotein. J. Biol. Chem. **278**, 13875–13880

6 Yoshida, K., Rutishauser, U., Crandall, J. E. and Schwarting, G. A. (1999) Polysialic acid facilitates migration of luteinizing hormone-releasing hormone neurons on vomeronasal axons. J. Neurosci. **19**, 794–801

7 Zuber, C., Lackie, P. M., Catterall, W. A. and Roth, J. (1992) Polysialic acid is associated with sodium channels and the neural cell adhesion molecule N-CAM in adult rat brain. J. Biol. Chem. **267**, 9965–9971

8 Rutishauser, U. (1996) Polysialic acid and the regulation of cell interactions. Curr. Opin. Cell Biol. **8**, 679–684

9 Cremer, H., Chazal, G., Goridis, C. and Represa, A. (1997) NCAM is essential for axonal growth and fasciculation in the hippocampus. Mol. Cell Neurosci. **8**, 323–335

10 Muller, D., Wang, C., Skibo, G., Toni, N., Cremer, H., Calaora, V., Rougon, G. and Kiss, J. Z. (1996) PSA-NCAM is required for activity-induced synaptic plasticity. Neuron **17**, 413–422

11 Cheresh, D. A., Pierschbacher, M. D., Herzig, M. A. and Mujoo, K. (1986) Disialogangliosides GD$_2$ and GD$_3$ are involved in the attachment of human melanoma and neuroblastoma cells to extracellular matrix proteins. J. Cell Biol. **102**, 688–696

12 Kasahara, K., Watanabe, Y., Yamamoto, T. and Sanai, Y. (1997) Association of Src family tyrosine kinase Lyn with ganglioside GD$_3$ in rat brain. Possible regulation of Lyn by glycosphingolipid in caveolae-like domains. J. Biol. Chem. **272**, 29947–29953

13 Liu, H., Kojima, N., Kurosawa, N. and Tsuji, S. (1997) Regulated expression system for GD$_3$ synthase cDNA and induction of differentiation in Neuro2a cells. Glycobiology **7**, 1067–1076

14 Fukuda, M., Dell, A. and Fukuda, M. N. (1984) Structure of fetal lactosaminoglycan. The carbohydrate moiety of Band 3 isolated from human umbilical cord erythrocytes. J. Biol. Chem. **259**, 4782–4791

15 Fukuda, M., Lauffenburger, M., Sasaki, H., Rogers, M. E. and Dell, A. (1987) Structures of novel sialylated O-linked oligosaccharides isolated from human erythrocyte glycophorins. J. Biol. Chem. **262**, 11952–11957

16 Kiang, W. L., Krusius, T., Finne, J., Margolis, R. U. and Margolis, R. K. (1982) Glycoproteins and proteoglycans of the chromaffin granule matrix. J. Biol. Chem. **257**, 1651–1659

17 Kitajima, K., Sato, C., Honda, N., Matsuda, T., Yokoyama, M.-H., Close, B. and Colley, K. (1999) Occurence of α2,8-linked oligosialic acid residues in mammalian glycoproteins. In Sialobiology and Other Novel Forms of Glycosylation (Inoue, Y., Lee, Y. C. and II, F. A. T., eds.), pp. 69–76, Gakushin Publishing Co., Osaka, Japan

18 Sato, C., Matsuda, T. and Kitajima, K. (2002) Neuronal differentiation-dependent expression of the disialic acid epitope on CD166 and its involvement in neurite formation in Neuro2A cells. J. Biol. Chem. **277**, 45299–45305

19 Sato, C., Yasukawa, Z., Honda, N., Matsuda, T. and Kitajima, K. (2001) Identification and adipocyte differentiation-dependent expression of the unique disialic acid residue in an adipose tissue-specific glycoprotein, adipo Q. J. Biol. Chem. **276**, 28849–28856

20 Nadanaka, S., Sato, C., Kitajima, K., Katagiri, K., Irie, S. and Yamagata, T. (2001) Occurrence of oligosialic acids on integrin α5 subunit and their involvement in cell adhesion to fibronectin. J. Biol. Chem. **276**, 33657–33664

21 Ziak, M., Meier, M., Novak-Hofer, I. and Roth, J. (2002) Ceruloplasmin carries the anionic glycan oligo/poly α2,8 deaminoneuraminic acid. Biochem. Biophys. Res. Commun. **295**, 597–602

22 Ziak, M., Kerjaschki, D., Farquhar, M. G. and Roth, J. (1999) Identification of megalin as the sole rat kidney sialoglycoprotein containing poly α2,8 deaminoneuraminic acid. J. Am. Soc. Nephrol. **10**, 203–209

23 Ziak, M., Meier, M. and Roth, J. (1999) Megalin in normal tissues and carcinoma cells carries oligo/poly α2,8 deaminoneuraminic acid as a unique posttranslational modification. Glycoconj. J. **16**, 185–188

24 Kono, M., Yoshida, Y., Kojima, N. and Tsuji, S. (1996) Molecular cloning and expression of a fifth type of α2,8-sialyltransferase (ST8Sia V). Its substrate specificity is similar to that of SAT-V/III, which synthesize GD1c, GT1a, GQ1b and GT3. J. Biol. Chem. **271**, 29366–29371

25 Sasaki, K., Kurata, K., Kojima, N., Kurosawa, N., Ohta, S., Hanai, N., Tsuji, S. and Nishi, T. (1994) Expression cloning of a GM$_3$-specific α-2,8-sialyltransferase (GD$_3$ synthase). J. Biol. Chem. **269**, 15950–15956

26 Angata, K., Suzuki, M. and Fukuda, M. (1998) Differential and cooperative polysialylation of the neural cell adhesion molecule by two polysialyltransferases, PST and STX. J. Biol. Chem. **273**, 28524–28532

27   Lee, Y. C., Kim, Y. J., Lee, K. Y., Kim, K. S., Kim, B. U., Kim, H. N., Kim, C. H. and Do, S. I. (1998) Cloning and expression of cDNA for a human Sia α2,3Galβ1, 4GlcNA:α2,8-sialyltransferase (hST8Sia III). Arch. Biochem. Biophys. **360**, 41–46

28   Angata, K., Suzuki, M., McAuliffe, J., Ding, Y., Hindsgaul, O. and Fukuda, M. (2000) Differential biosynthesis of polysialic acid on neural cell adhesion molecule (NCAM) and oligosaccharide acceptors by three distinct α2,8-sialyltransferases, ST8Sia IV (PST), ST8Sia II (STX), and ST8Sia III. J. Biol. Chem. **275**, 18594–18601

29   Muhlenhoff, M., Eckhardt, M., Bethe, A., Frosch, M. and Gerardy-Schahn, R. (1996) Autocatalytic polysialylation of polysialyltransferase-1. EMBO **15**, 6943–6950

30   Close, B. E. and Colley, K. J. (1998) *In vivo* autopolysialylation and localization of the polysialyltransferases PST and STX. J. Biol. Chem. **273**, 34586–34593

31   Muhlenhoff, M., Manegold, A., Windfuhr, M., Gotza, B. and Gerardy-Schahn, R. (2001) The impact of N-glycosylation on the functions of polysialyltransferases. J. Biol. Chem. **276**, 34066–34073

32   Takashima, S., Ishida, H. K., Inazu, T., Ando, T., Ishida, H., Kiso, M., Tsuji, S. and Tsujimoto, M. (2002) Molecular cloning and expression of a sixth type of α2,8-sialyltransferase (ST8Sia VI) that sialylates O-glycans. J. Biol. Chem. **277**, 24030–24038

33   Toillon, R. A., Chopin, V., Jouy, N., Fauquette, W., Boilly, B. and Le Bourhis, X. (2002) Normal breast epithelial cells induce p53-dependent apoptosis and p53-independent cell cycle arrest of breast cancer cells. Breast Cancer Res. Treat. **71**, 269–280

34   Altschul, S., Madden, T., Schaffer, A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25**, 3389–3402

35   O'Reilly, D. R. and Miller, L. K. (1989) A baculovirus blocks insect molting by producing ecdysteroid UDP-glucosyl transferase. Science (Washington, DC) **245**, 1110–1112

36   Croizier, G. and Quiot, J. M. (1975) [Electron microscopic study of the reticulated cellular structures induced in *Lepidoptera* by baculovirus]. Acad. Sci. Hebd. Seances Acad. Sci. D **281**, 1055–1057

37   Felgner, P. L. and Ringold, G. M. (1989) Cationic liposome-mediated transfection. Nature (London) **337**, 387–388

38   Sato, C., Inoue, S., Matsuda, T. and Kitajima, K. (1999) Fluorescent-assisted detection of oligosialyl units in glycoconjugates. Anal. Biochem. **266**, 102–109

39   Cheng, M. C., Lin, S. L., Wu, S. H., Inoue, S. and Inoue, Y. (1998) High-performance capillary electrophoretic characterization of different types of oligo- and polysialic acid chains. Anal. Biochem. **260**, 154–159

40   Harduin-Lepers, A., Mollicone, R., Delannoy, P. and Oriol, R. (2005) The animal sialyltransferases and sialyltransferase-related genes: a phylogenetic approach. Glycobiology **15**, 805–817

41   Kozak, M. (1992) Regulation of translation in eukaryotic systems. Annu. Rev. Cell Biol. **8**, 197–225

42   Sato, C., Kitajima, K., Inoue, S., Seki, T., Troy, 2nd, F. A. and Inoue, Y. (1995) Characterization of the antigenic specificity of four different anti-(α2,8-linked polysialic acid) antibodies using lipid-conjugated oligo/polysialic acids. J. Biol. Chem. **270**, 18923–18928

43   Finne, J., Krusius, T., Rauvala, H. and Hemminki, K. (1977) The disialosyl group of glycoproteins. Occurrence in different tissues and cellular membranes. Eur. J. Biochem. **77**, 319–323

44   Sato, C., Fukuoka, H., Ohta, K., Matsuda, T., Koshino, R., Kobayashi, K., Troy, 2nd, F. A. and Kitajima, K. (2000) Frequent occurrence of pre-existing α2,8-linked disialic and oligosialic acids with chain lengths up to 7 Sia residues in mammalian brain glycoproteins. Prevalence revealed by highly sensitive chemical methods and anti-di-, oligo-, and poly-Sia antibodies specific for defined chain lengths. J. Biol. Chem. **275**, 15422–15431

45   Martersteck, C. M., Kedersha, N. L., Drapp, D. A., Tsui, T. G. and Colley, K. J. (1996) Unique α2,8-polysialylated glycoproteins in breast cancer and leukemia cells. Glycobiology **6**, 289–301

46   Sato, C., Inoue, S., Matsuda, T. and Kitajima, K. (1998) Development of a highly sensitive chemical method for detecting α2,8-linked oligo/polysialic acid residues in glycoproteins blotted on the membrane. Anal. Biochem. **261**, 191–197

47   Angata, T., Kerr, S. C., Greaves, D. R., Varki, N. M., Crocker, P. R. and Varki, A. (2002) Cloning and characterization of human Siglec-11. A recently evolved signaling that can interact with SHP-1 and SHP-2 and is expressed by tissue macrophages, including brain microglia. J. Biol. Chem. **277**, 24466–24474

48   Nicoll, G., Ni, J., Liu, D., Klenerman, P., Munday, J., Dubock, S., Mattei, M. G. and Crocker, P. R. (1999) Identification and characterization of a novel siglec, siglec-7, expressed by human natural killer cells and monocytes. J. Biol. Chem. **274**, 34089–34095

49   Yamaji, T., Teranishi, T., Alphey, M. S., Crocker, P. R. and Hashimoto, Y. (2002) A small region of the natural killer cell receptor, Siglec-7, is responsible for its preferred binding to α2,8-disialyl and branched α2,6-sialyl residues. A comparison with Siglec-9. J. Biol. Chem. **277**, 6324–6332

50   Recchi, M. A., Hebbar, M., Hornez, L., Harduin-Lepers, A., Peyrat, J. P. and Delannoy, P. (1998) Multiplex reverse transcription polymerase chain reaction assessment of sialyltransferase expression in human breast cancer. Cancer Res. **58**, 4066–4070