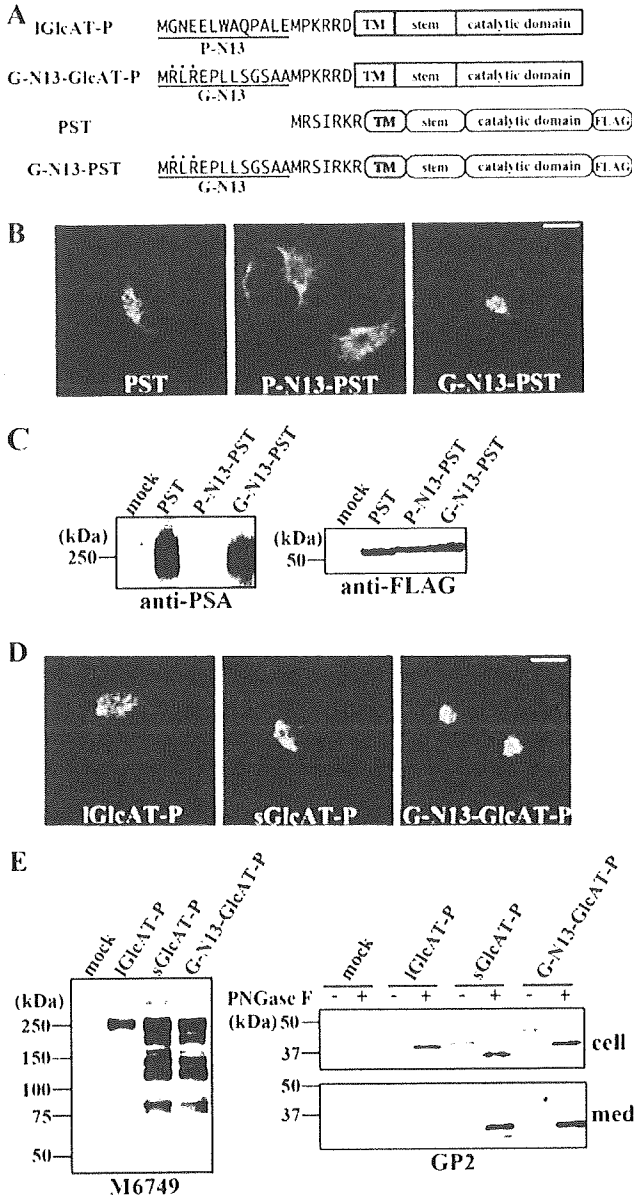

FIGURE 7. **Intracellular activities and localization of chimeric GlcAT-P and PST.** *A*, schematic diagrams of chimeric GlcAT-P and PST, which were fused with the N-terminal 13 amino acids from long form B4GalT-I (G-N13) at their N termini, are shown. The N-terminal cytoplasmic tails are depicted as the amino acid sequences. A dibasic motif in G-N13 near the N terminus is indicated by *dots*. *TM*, transmembrane domain. *B*, Neuro2A cells expressing wild-type PST, P-N13-PST, or G-N13-PST were immunostained with anti-FLAG mAb. Similar Golgi accumulation of PST and G-N13-PST was observed. *Bar*, 10 μm. *C*, Neuro2A cells were transfected with the PST, P-N13-PST, or G-N13-PST expression plasmid or the empty vector (*mock*). Cells were lysed and Western blotted with anti-PSA mAb (12E3) (*left*) or anti-FLAG mAb (*right*). *D*, Neuro2A cells expressing lGlcAT-P, sGlcAT-P, or G-N13-GlcAT-P were immunostained with GP2 pAb. sGlcAT-P and G-N13-GlcAT-P were similarly localized in the Golgi apparatus strictly. *Bar*, 10 μm. *E*, Neuro2A cells were transfected with the lGlcAT-P, sGlcAT-P, or G-N13-GlcAT-P expression plasmid or the empty vector (*mock*). Cells were lysed and Western blotted with M6749 mAb (*left*). Cellular proteins and proteins secreted into the medium (*med*) were treated with or without PNGase F, followed by Western blotting with GP2 pAb (*right*).

process mediated by the dibasic motif in the cytoplasmic tail. Consistent with a previous report (16), the dibasic motif in GlcAT-P near P-N13 was also essential for both Golgi localization and interaction with Sar1 (Figs. 4 and 5). In addition, we revealed, using a co-precipitation assay, that the interaction of sGlcAT-P with Sar1 was stronger than that of lGlcAT-P (Fig. 5*C*), probably because the existence of P-N13 interfered with this binding. Sar1 recruitment initiates the budding of COPII vesicles from ER (31), and COPII vesicles selectively incorporate and convey diverse cargoes, including glycosyltransferases from ER (32). In fact, the expression of a GTP-restricted or GDP-restricted form of Sar1 blocked the ER exit process, leading to the disruption of proper glycosyltransferase localization (33). Moreover, the direct binding of a cargo protein, such as viral glycoprotein VSV-G, to Sar1 promotes its efficient ER export (32). Therefore, the weaker interaction with Sar1 was considered to lead to the partial ER distribution of lGlcAT-P (Fig. 3*A*). On the other hand, as shown in Fig. 7, *D* and *E*, the attachment of another 13 amino acids (from human B4GalT-I) to sGlcAT-P caused neither ER entrapment nor an activity decrease. Further, P-N13, but not G-N13, had similar effects, such as a localization change and reduced activity, on another glycosyltransferase, PST (Figs. 6 and 7), which has a similar cytoplasmic tail with a dibasic motif to sGlcAT-P. These results refuted the idea that any peptide near the dibasic motif had an influence on glycosyltransferase localization and activity, rather suggesting a sequence-specific role of P-N13. However, the dibasic motif is also present in G-N13 near the N terminus (RLR sequence). Although this sequence might also be bound with Sar1 and serve to effectively convey G-N13-GlcAT-P and G-N13-PST from ER to the Golgi apparatus, Maccioni's group (16) demonstrated that the dibasic motif is most effective when located proximal to the transmembrane border, and they showed that the Golgi localization of SialT2, which has two dibasic motifs near the N terminus and the transmembrane region, was not affected by mutation of the first motif or by the deletion of 15 amino acids, including the first motif. Moreover, we demonstrated that the mutant G-N13-GlcAT-P lacking the motif near the N terminus (G-N13AA-GlcAT-P) showed similar distribution to G-N13-GlcAT-P and significantly higher intracellular activity than lGlcAT-P (supplemental Fig. 1). These findings suggested that the motif in G-N13 was ineffective in the case of G-N13-GlcAT-P and G-N13-PST. Based on these findings, we assumed that an unidentified protein specifically binds with P-N13 to cause these effects, but further experiments are required to explore this possibility.

We also found that the production of cleaved secreted sGlcAT-P was higher than that of lGlcAT-P (Fig. 3*B*) and that the cleavage occurred at the stem domain of this enzyme (Fig. 3*D*), as in the case of other glycosyltransferases. This is the first report that GlcAT-P is also cleavable and secreted, but the cleavage mechanism, including the responsible protease and cleavage site, is largely unknown. Although it is well known that some glycosyltransferases are cleaved and secreted into body fluids (34), the cleavage mechanisms and roles of the soluble enzymes remain to be clarified except for in the cases of ST6Gal-I and GnT-V (35, 36). As for GlcAT-P cleavage, we revealed that the entry of GlcAT-P into the Golgi apparatus was at least required because the disruption of GlcAT-P export from ER led to a significant decrease in GlcAT-P secretion (Fig. 4*C*). We also found that a soluble fraction as well as a membrane

fraction prepared from a mouse brain homogenate contained glucuronyltransferase activity *in vitro*, whereas those from GlcAT-P-deficient mice did not (data not shown), indicating that GlcAT-P was also cleaved and existed as a soluble enzyme *in vivo*. We are now trying to identify the cleaving protease(s) and to elucidate the function of extracellular soluble GlcAT-P.

As shown in Fig. 2*B*, sGlcAT-P exhibited significantly higher biosynthetic activity than lGlcAT-P in cells. Regarding the cause of this phenomenon, we consider that there are two possibilities. One is that a larger production of soluble enzymes from sGlcAT-P caused a higher expression of product glycans. In the case of ST6Gal-I, the enhanced production of soluble ST6Gal-I induced by the overexpression of BACE1, which is the protease responsible for its cleavage, led to greater sialylation of soluble acceptor glycoproteins without alteration of the sialylation of membrane-anchored acceptors (35). Another possibility is that the strict Golgi accumulation of sGlcAT-P is the main cause of the higher intracellular activity, since ER-distributed lGlcAT-P cannot serve as a functional enzyme because of the lack of an acceptor substrate with a complex type *N*-glycan. We think that the latter possibility is more plausible for the following reasons. Although soluble GlcAT-P may also be freely accessible to soluble acceptor glycoproteins in the Golgi lumen like soluble ST6Gal-I, our results showed that sGlcAT-P biosynthesized a significantly larger amount of glycan epitopes for both membrane-anchored (cellular) and soluble (extracellular) glycoproteins (Fig. 2*B* and data not shown). In addition, in our experiments, we could not detect cleaved soluble PST secreted into the culture medium, suggesting that the reduced activity of PST with the attachment of P-N13 was irrelevant concerning its cleavage and secretion. Rather, it was likely that the lower polysialylation by P-N13-PST was caused by its significant loss of Golgi distribution (Fig. 6). Overall, we considered that the increased secretion of sGlcAT-P was not the main cause of the enhanced activity but a consequence of the distribution change and that P-N13 influenced enzyme export from ER by hampering the interaction with Sar1, leading to activity regulation.

Various roles of the cytoplasmic tail of glycosyltransferases have gradually been uncovered, as described (see Introduction). However, we considered that the cytoplasmic tail of each glycosyltransferase has a distinct role, since each tail consists of amino acids markedly different in length and sequence. The dibasic motif in the cytoplasmic tail does not exist in all glycosyltransferases, and even GlcAT-S, the glycosyltransferase most highly homologous to GlcAT-P, does not have this motif. The mechanisms underlying the selective ER export of glycosyltransferases without the dibasic motif remain to be elucidated. Glycosyltransferase isoforms differing in the cytoplasmic tail have hardly been identified except for B4GalT-I. The B4GalT-I isoforms and GlcAT-Ps examined here showed the common feature that the longer form has an additional N-terminal 13 amino acids, and functional roles of G-N13 have been reported (*i.e.* in cell surface expression of the enzyme and cycling between *trans*-Golgi and *trans*-Golgi network) (14, 15). On the other hand, we revealed that P-N13 contained a different signal from that of B4GalT-I (*i.e.* one involved in the exit of the enzyme from the ER which eventually led to the change of the product glycan expression). Now we are searching for proteins binding to P-N13 and investigating whether and how the balance of the expression levels of the two GlcAT-P mRNAs is regulated. If we are successful, a novel molecular mechanism to regulate the specific glycan expression in the nervous system will be uncovered.

Acknowledgment—We thank Dr. K. Nakayama (Kyoto University) for valuable discussion.

REFERENCES

1. Ohtsubo, K., and Marth, J. D. (2006) *Cell* **126,** 855–867
2. Lowe, J. B., and Marth, J. D. (2003) *Annu. Rev. Biochem.* **72,** 643–691
3. Nieke, J., and Schachner, M. (1985) *Differentiation* **30,** 141–151
4. Terayama, K., Oka, S., Seiki, T., Miki, Y., Nakamura, A., Kozutsumi, Y., Takio, K., and Kawasaki, T. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 6093–6098
5. Seiki, T., Oka, S., Terayama, K., Imiya, K., and Kawasaki, T. (1999) *Biochem. Biophys. Res. Commun.* **255,** 182–187
6. Bakker, H., Friedmann, I., Oka, S., Kawasaki, T., Nifant'ev, N., Schachner, M., and Mantei, N. (1997) *J. Biol. Chem.* **272,** 29942–29946
7. Yamamoto, S., Oka, S., Inoue, M., Shimuta, M., Manabe, T., Takahashi, H., Miyamoto, M., Asano, M., Sakagami, J., Sudo, K., Iwakura, Y., Ono, K., and Kawasaki, T. (2002) *J. Biol. Chem.* **277,** 27227–27231
8. Kizuka, Y., Matsui, T., Takematsu, H., Kozutsumi, Y., Kawasaki, T., and Oka, S. (2006) *J. Biol. Chem.* **281,** 13644–13651
9. de Graffenried, C. L., and Bertozzi, C. R. (2004) *Curr. Opin. Cell Biol.* **16,** 356–363
10. Breton, C., Mucha, J., and Jeanneau, C. (2001) *Biochimie (Paris)* **83,** 713–718
11. Opat, A. S., van Vliet, C., and Gleeson, P. A. (2001) *Biochimie (Paris)* **83,** 763–773
12. Uliana, A. S., Giraudo, C. G., and Maccioni, H. J. (2006) *Traffic* **7,** 604–612
13. Kakuda, S., Sato, Y., Tonoyama, Y., Oka, S., and Kawasaki, T. (2005) *Glycobiology* **15,** 203–210
14. Hathaway, H. J., Evans, S. C., Dubois, D. H., Foote, C. L., Elder, B. H., and Shur, B. D. (2003) *J. Cell Sci.* **116,** 4319–4330
15. Schaub, B. E., Berger, B., Berger, E. G., and Rohrer, J. (2006) *Mol. Biol. Cell* **17,** 5153–5162
16. Giraudo, C. G., and Maccioni, H. J. (2003) *Mol. Biol. Cell* **14,** 3753–3766
17. Schmitz, K. R., Liu, J., Li, S., Setty, T. G., Wood, C. S., Burd, C. G., and Ferguson, K. M. (2008) *Dev. Cell* **14,** 523–534
18. Tu, L., Tai, W. C., Chen, L., and Banfield, D. K. (2008) *Science* **321,** 404–407
19. Mitsumoto, Y., Oka, S., Sakuma, H., Inazawa, J., and Kawasaki, T. (2000) *Genomics* **65,** 166–173
20. Yamamoto, S., Oka, S., Saito-Ohara, F., Inazawa, J., and Kawasaki, T. (2002) *J. Biochem.* **131,** 337–347
21. Tagawa, H., Kizuka, Y., Ikeda, T., Itoh, S., Kawasaki, N., Kurihara, H., Onozato, M. L., Tojo, A., Sakai, T., Kawasaki, T., and Oka, S. (2005) *J. Biol. Chem.* **280,** 23876–23883
22. Ma, J., Qian, R., Rausa, F. M., and Colley, K. J. (1997) *J. Biol. Chem.* **272,** 672–679
23. Fenteany, F. H., and Colley, K. J. (2005) *J. Biol. Chem.* **280,** 5423–5429
24. Kitazume, S., Tachida, Y., Oka, R., Kotani, N., Ogawa, K., Suzuki, M., Dohmae, N., Takio, K., Saido, T. C., and Hashimoto, Y. (2003) *J. Biol. Chem.* **278,** 14865–14871
25. Paulson, J. C., and Colley, K. J. (1989) *J. Biol. Chem.* **264,** 17615–17618
26. Eckhardt, M., Muhlenhoff, M., Bethe, A., Koopman, J., Frosch, M., and Gerardy-Schahn, R. (1995) *Nature* **373,** 715–718
27. Kojima, N., Yoshida, Y., and Tsuji, S. (1995) *FEBS Lett.* **373,** 119–122
28. Weinhold, B., Seidenfaden, R., Rockle, I., Muhlenhoff, M., Schertzinger, F., Conzelmann, S., Marth, J. D., Gerardy-Schahn, R., and Hildebrandt, H. (2005) *J. Biol. Chem.* **280,** 42971–42977
29. Kojima, N., Tachida, Y., and Tsuji, S. (1997) *J. Biochem. (Tokyo)* **122,** 1265–1273

30. Virtanen, I., Ekblom, P., and Laurila, P. (1980) *J. Cell Biol.* **85,** 429–434
31. Barlowe, C., d'Enfert, C., and Schekman, R. (1993) *J. Biol. Chem.* **268,** 873–879
32. Antonny, B., and Schekman, R. (2001) *Curr. Opin. Cell Biol.* **13,** 438–443
33. Stroud, W. J., Jiang, S., Jack, G., and Storrie, B. (2003) *Traffic* **4,** 631–641
34. Gerber, A. C., Kozdrowski, L., Wyss, S. R., and Berger, E. G. (1979) *Eur. J. Biochem.* **93,** 453–460
35. Sugimoto, I., Futakawa, S., Oka, R., Ogawa, K., Marth, J. D., Miyoshi, E., Taniguchi, N., Hashimoto, Y., and Kitazume, S. (2007) *J. Biol. Chem.* **282,** 34896–34903
36. Nakahara, S., Saito, T., Kondo, N., Moriwaki, K., Noda, K., Ihara, S., Takahashi, M., Ide, Y., Gu, J., Inohara, H., Katayama, T., Tohyama, M., Kubo, T., Taniguchi, N., and Miyoshi, E. (2006) *FASEB J.* **20,** 2451–2459

BJ www.biochemj.org

BJ Structure

Biochem. J. (2009) 421, 157–162 (Printed in Great Britain) doi:10.1042/BJ20090474

online data

ACCELERATED PUBLICATION

FAM20B is a kinase that phosphorylates xylose in the glycosaminoglycan–protein linkage region

Toshiyasu KOIKE*[1], Tomomi IZUMIKAWA*[1], Jun-Ichi TAMURA† and Hiroshi KITAGAWA*[2]

*Department of Biochemistry, Kobe Pharmaceutical University, Higashinada-ku, Kobe 658-8558, Japan, and †Department of Regional Environment, Faculty of Regional Sciences, Tottori University, Tottori 680-8551, Japan

2-O-phosphorylation of xylose has been detected in the glycosaminoglycan–protein linkage region, GlcAβ1-3Galβ1-3Galβ1-4Xylβ1-*O*-Ser, of proteoglycans. Recent mutant analyses in zebrafish suggest that xylosyltransferase I and FAM20B, a protein of unknown function that shows weak similarity to a Golgi kinase encoded by *four-jointed*, operate in a linear pathway for proteoglycan production. In the present study, we identified FAM20B as a kinase that phosphorylates the xylose residue in the linkage region. Overexpression of *FAM20B* increased the amount of both chondroitin sulfate and heparan sulfate in HeLa cells, whereas the RNA interference of *FAM20B* resulted in a reduction of their amount in the cells. Gel-filtration analysis of the glycosaminoglycan chains synthesized in the overexpressing cells revealed that the glycosaminoglycan chains had a similar length to those in mock-transfected cells. These results suggest that FAM20B regulates the number of glycosaminoglycan chains by phosphorylating the xylose residue in the glycosaminoglycan–protein linkage region of proteoglycans.

Key words: chondroitin sulfate, FAM20B, glycosaminoglycan–protein linkage region, heparan sulfate, phosphorylation, xylose kinase.

INTRODUCTION

Sulfated GAGs (glycosaminoglycans), including HS (heparan sulfate) and CS (chondroitin sulfate), are linear polysaccharides consisting of a repetition of [(-4GlcAβ1-4GlcNAcα1-)$_n$] and [(-4GlcAβ1-3GalNAcβ1-)$_n$] disaccharide units respectively. The assembly of GAG chains is initiated by the synthesis of the so-called common GAG–protein linkage region (GlcAβ1-3Galβ1-3Galβ1-4Xylβ1-*O*-Ser), which is attached to specific serine residues of different core proteins. The linkage region tetrasaccharide is formed by sequential stepwise addition of monosaccharide residues by the respective specific glycosyltransferases, XylT (xylosyltransferase), GalT-I (galactosyltransferase I), galactosyltransferase II and GlcAT-I (β1,3-glucuronyltransferase I) [1]. The repeating disaccharide region of HS is synthesized on the linkage region by the HS co-polymerase complex of EXT1 and EXT2 [1]. In contrast, the repeating disaccharide region of CS is synthesized on the linkage region by enzyme complexes of bifunctional chondroitin synthases-1 [2], -2 [3] and -3 [4], and chondroitin polymerizing factor [5]. After synthesis of the GAG sugar backbone on this tetrasaccharide, numerous modifications, including sulfation, epimerization and desulfation, are performed in a spatiotemporal manner, producing mature and functional GAG chains that exert biological functions dependent on their specific structure [6].

To date, structural studies have shown the presence of various modifications of the GAG–protein linkage region. One of the modifications is phosphorylation of the xylose residue at position 2 [1,6]. This modification has been detected in both HS and CS derived from cell-rich tissues [7,8], and appears to affect the transfer of galactose and D-glucuronic acid residues by GalT-I and GlcAT-I respectively [9,10]. *In vitro* experiments using authentic substrates showed that phosphorylation of the xylose residue prevents the transfer of a galactose residue by GalT-I [9]. In contrast, GlcAT-I could efficiently transfer a D-glucuronic acid residue to the phosphorylated trisaccharide *in vitro* [10]. These results suggested that phosphorylation of the xylose residue takes place after transfer of the first galactose residue by GalT-I and before transfer of the D-glucuronic acid residue by GlcAT-I. In fact, phosphorylation of the xylose residue is most prominent after the addition of two galactose residues [11,12], although the biological role of this modification and the enzyme responsible for the phosphorylation remain unclear.

Recently, Eames et al. [13] reported that FAM20B and XylT-1 drive cartilage matrix production and inhibit perichondral bone formation during endochondral ossification. In addition, double-mutant analyses indicated that FAM20B and XylT-1 operate in a linear pathway for proteoglycan production. FAM20B is a member of a family of related proteins that has been named a family with sequence similarity 20 (FAM20) with three members (FAM20A, FAM20B and FAM20C) in mammals [14]. Although the function of FAM20 proteins is as yet unknown, they are reported to have weak similarity to a protein named Four-jointed, which was recently identified as a Golgi kinase that phosphorylates serine or threonine residues within the extracellular cadherin domain of Fat and its transmembrane ligand, Dachsous [15]. Hence, we hypothesized that FAM20B might be a kinase that phosphorylates the xylose residue of the GAG–protein linkage region. In the present paper, we describe the identification and characterization of FAM20B as a xylose kinase.

Abbreviations used: CS, chondroitin sulfate; EGFP, enhanced green fluorescent protein; ER, endoplasmic reticulum; FAM20, family with sequence similarity 20; GAG, glycosaminoglycan; GalT-I, galactosyltransferase I; GFP, green fluorescent protein; GlcAT-I, β1,3-glucuronyltransferase I; HS, heparan sulfate; RT, reverse transcription; shRNA, small hairpin RNA; α-TM, α-thrombomodulin; XylT, xylosyltransferase.

[1] These authors contributed equally to this work.

[2] To whom correspondence should be addressed (email kitagawa@kobepharma-u.ac.jp).

The nucleotide sequence data reported in this paper for FAM20B have been submitted to the DDBJ, EMBL, GenBank® and GSDB Nucleotide Sequence Databases under accession number AB480690.

EXPERIMENTAL

Materials

[γ-^{32}P]ATP (3000 Ci/mmol) was purchased from PerkinElmer. Unlabelled ATP and bovine liver β-glucuronidase (EC 3.2.1.31) were obtained from Sigma. Shrimp alkaline phosphatase was purchased from Roche Molecular Biochemicals. α-TM (α-thrombomodulin) with a truncated linkage region tetrasaccharide, GlcAβ1-3Galβ1-3Galβ1-4Xyl, was purified and structurally characterized as described previously [16,17]. Galβ1-3Galβ1-4Xyl(2-*O*-phosphate)β1-*O*-Ser was chemically synthesized [18]. Galβ1-3Galβ1-4Xylβ1-*O*-Ser was prepared by digestion of Galβ1-3Galβ1-4Xyl(2-*O*-phosphate)β1-*O*-Ser with alkaline phosphatase.

Construction of a soluble form of FAM20B

The cDNA fragment of a truncated form of FAM20B, lacking the first 30 N-terminal amino acids, was amplified by PCR with KIAA 0475 cDNA obtained from the Kazusa DNA Research Institute (Chiba, Japan) as a template using a forward primer (5′-CGGGATCCTCAGCTGCCAACCGGGAGGAC-3′) containing an in-frame BamHI site and a reverse primer (5′-CGGGATCCACTTCCAATCCATCTCATACC-3′) containing a BamHI site located 67 bp downstream of the stop codon. PCR was carried out with *KOD*-Plus DNA polymerase (Toyobo) for 30 cycles at 94 °C for 30 s, 58 °C for 30 s and 68 °C for 150 s in 5 % (v/v) DMSO. The PCR fragment was subcloned into the BamHI site of pGIR201protA [19], resulting in the fusion of the insulin signal sequence and the Protein A sequence present in the vector, as described in [4].

Expression of a soluble form of FAM20B and enzyme assays

The expression plasmid (6.0 μg) was transfected into COS-1 cells on 100-mm-diameter plates using FuGENE™ 6 (Roche Molecular Biochemicals) according to the manufacturer's instructions. At 2 days after transfection, 1 ml of the culture medium was collected and incubated with 10 μl of IgG–Sepharose (GE Healthcare) for 1 h at 4 °C. The beads recovered by centrifugation at 150 ***g*** for 2 min were washed with and then resuspended in the assay buffer, and tested for kinase activity, as described below. Assays for kinase activity were carried out using α-TM, Galβ1-3Galβ1-4Xyl(2-*O*-phosphate)β1-*O*-Ser or Galβ1-3Galβ1-4Xyl-*O*-Ser as an acceptor and ATP as a phosphate donor. Kinase reactions were incubated in reaction mixtures containing the following constituents in a total volume of 20 μl: 1 nmol of α-TM, Galβ1-3Galβ1-4Xyl(2-*O*-phosphate)β1-*O*-Ser or Galβ1-3Galβ1-4Xyl-*O*-Ser, 10 μM [γ-^{32}P]ATP (1.11×10^5 d.p.m.), 50 mM Tris/HCl (pH 7.0), 10 mM $MnCl_2$, 10 mM $CaCl_2$, 0.1 % BSA and 10 μl of the resuspended beads. The mixtures were incubated for 4 h at 37 °C.

Characterization of the reaction products

Products of kinase reactions on α-TM were isolated by gel filtration on a Superdex peptide column with 0.2 M ammonium bicarbonate as the eluent. The ^{32}P-labelled tetrasaccharide chains were released from α-TM by mild alkaline treatment using 0.5 M LiOH, derivatized with 2-aminobenzamide and then isolated by gel-filtration HPLC as described in [17]. The isolated tetrasaccharides were digested with β-glucuronidase in 50 mM sodium acetate (pH 5.2) or alkaline phosphatase in 50 mM Tris/HCl (pH 8.0) and 1 mM $MgCl_2$ at 37 °C. Periodate oxidation was performed in 0.02 M $NaIO_4$ and 0.05 M sodium formate (pH 3.0) as described in [8]. An aliquot of the compound was subjected to gel-filtration chromatography on a column (7.6 mm × 500 mm) of Asahipak GS-320 (Asahi Chemical Industry) using 50 mM ammonium acetate as the eluent at a flow rate of 1 ml/min as described in [17].

Subcellular localization

The cDNA fragment encoding FAM20B was amplified using a forward primer (5′-CGGAATTCCAGGGGAGAAGGAAAAGAG-3′) containing an EcoRI site and a reverse primer (5′-CGGGATCCAAGTGTGAGAGAGGCATCCT-3′) containing a BamHI site. PCR was carried out with *KOD*-Plus DNA polymerase for 30 cycles at 94 °C for 30 s, 53 °C for 42 s and 68 °C for 180 s in 5 % (v/v) DMSO. The PCR fragment was subcloned into the pEGFP-N1 expression vector (Clontech). The Golgi marker vector (pDsRed-Golgi) was constructed using the pECFP-Golgi vector (Clontech) that harbours a sequence encoding the N-terminal 81 amino acids of human β1-4-galactosyltransferase [20]. The region from pECFP-Golgi was digested with NheI and BamHI, and subcloned into pDsRed-N1. Combinations of GFP (green fluorescent protein)-tagged and DsRed-Monomer-tagged expression vectors (3.0 μg each) were transfected into HeLa cells on glass-bottom dishes (Matsunami Glass) using FuGENE™ 6 according to the manufacturer's instructions. Fluorescent images were obtained using a laser-scanning confocal microscope, FluoView (Olympus).

Creation of stably transfected cell lines

The cDNA fragment encoding FAM20B was amplified from KIAA 0475 cDNA as a template using a forward primer (5′-CGGAATTCGTGTCGAGAATTACAAGTGTG-3′) containing an EcoRI site and a reverse primer (5′-CGGGATCCACTTCCAATCCATCTCATACC-3′) containing a BamHI site. PCR was carried out with *KOD*-Plus DNA polymerase for 30 cycles at 94 °C for 30 s, 53 °C for 42 s and 68 °C for 180 s in 5 % (v/v) DMSO. The PCR fragments were subcloned into the EcoRI/BamHI site of the pCMV expression vector (Invitrogen). The expression plasmid was transfected into HeLa cells, and colonies were selected as described in [3].

FAM20B silencing in HeLa cells was performed using MISSION shRNA (short hairpin RNA) (Sigma). A hairpin construct identified by The RNAi Consortium clone number TRCN0000138872 was used. The shRNA plasmid (6.7 μg) was transfected into HeLa cells on 100-mm-diameter plates using FuGENE™ 6 according to the manufacturer's instructions. Transfectants were cultured in the presence of 0.4 μg/ml puromycin. Resultant colonies were then picked up and propagated for experiments.

Quantitative real-time RT (reverse transcription)–PCR

The cDNA was synthesized from total RNA extracted from HeLa cells as described in [4]. Primer sequences were as follows: *FAM20B*, forward primer 5′-AGAGATCAAACCTGTCGCC-3′ and reverse primer 5′-CCAAAGTGTGACAGATCCCT-3′; and glyceraldehyde-3-phosphate dehydrogenase, forward primer 5′-ATGGGTGTGAACCATGAGAAGTA-3′ and reverse primer 5′-GGCAGTGATGGCATGGAC-3′. Quantitative real-time RT–PCR was performed as described in [4].

GAG analysis

GAGs from HeLa cells were prepared as described in [3]. The purified GAG fraction was digested with chondroitinase ABC or a mixture of heparinase and heparitinase, and then the digests

Table 1 Kinase activities of fusion proteins secreted into the culture medium by the transfected COS-1 cells

α-TM contains a tetrasaccharide linkage GlcAβ1-3Galβ1-3Galβ1-4Xyl [17]. Activity values are the means ± S.E.M. of three determinations. ND, not detected (<0.1 pmol/h per ml of medium).

Acceptor	Kinase activity (pmol/h per ml of medium)
α-TM	102.7 ± 5.4
Galβ1-3Galβ1-4Xylβ1-*O*-Ser	128.2 ± 24.2
Galβ1-3Galβ1-4Xyl(2-*O*-phosphate)β1-*O*-Ser	ND

were derivatized with 2-aminobenzamide and analysed by HPLC as described in [3].

Gel-filtration chromatography of GAGs

To determine the chain length of GAGs, the purified GAG fraction was subjected to reductive β-elimination using $NaBH_4$/NaOH, and then analysed by gel-filtration chromatography on a column (10 mm × 300 mm) of Superdex 200 eluted with 0.2 M ammonium bicarbonate at a flow rate of 0.4 ml/min. Fractions were collected at 3.0 min intervals, freeze-dried and digested with chondroitinase ABC or a mixture of heparinase and heparitinase. The digests were derivatized with 2-aminobenzamide, and then analysed by HPLC on an amine-bound PA-03 column [3].

RESULTS AND DISCUSSION

FAM20B consists of 409 amino acids with type II transmembrane protein topology and shows weak similarity (21.4%) to a protein named Four-jointed, which was recently identified as a Golgi kinase that phosphorylates a subset of cadherin domains (see Supplementary Figure S1 at http://www.BiochemJ.org/bj/421/bj4210157add.htm) [15]. To examine whether FAM20B could phosphorylate xylose on the GAG–protein linkage region, a soluble form of FAM20B was generated by replacing the first 30 amino acids of FAM20B with a cleavable insulin signal sequence and a Protein A IgG-binding domain, as described in the Experimental section, and then soluble FAM20B was expressed in COS-1 cells as a recombinant protein fused with the Protein A IgG-binding domain. When the expression plasmid containing the putative kinase–Protein A fusion was expressed in COS-1 cells, an approx. 76 kDa protein was secreted, as shown by Western blotting (see Supplementary Figure S2 at http://www.BiochemJ.org/bj/421/bj4210157add.htm). The fused putative kinase expressed in the medium was adsorbed on IgG–Sepharose beads to eliminate endogenous kinases, and then the protein-bound beads were used as an enzyme source. The bound fusion protein was assayed for xylose kinase activity using various linkage region compounds as acceptor substrates. As shown in Table 1, marked kinase activity was detected with α-TM containing a linkage region tetrasaccharide, GlcAβ1-3Galβ1-3Galβ1-4Xylβ1, on native core protein [17] and non-phosphorylated trisaccharide–serine Galβ1-3Galβ1-4Xylβ1-*O*-Ser, but not with the phosphorylated counterpart Galβ1-3Galβ1-4Xyl(2-*O*-phosphate)β1-*O*-Ser as the acceptor substrate. These findings indicate that the expressed protein is a kinase that phosphorylated the GAG–protein linkage region.

To identify kinase reaction products, the representative acceptor substrate, α-TM, was labelled by the kinase reaction using [γ-^{32}P]ATP, and the linkage tetrasaccharide was liberated from α-TM by mild alkaline treatment, as described in the Experimental section. The isolated linkage tetrasaccharide fraction was derivatized with a fluorophore, 2-aminobenzamide, and fractionated by gel-filtration HPLC, resulting in two fluorescent components: ^{32}P-labelled (peak I) and unlabelled tetrasaccharides (peak II) (Figure 1A). Peak I was shifted to a position corresponding to 2-aminobenzamide-derivatized GlcAβ1-3Galβ1-3Galβ1-4Xyl by alkaline phosphatase digestion. Peak II was co-eluted with authentic 2-aminobenzamide-derivatized GlcAβ1-3Galβ1-3Galβ1-4Xyl when they were co-injected, indicating that it was derived from the unused acceptor substrate. Thus peak I was analysed below. The ^{32}P-labelled tetrasaccharide (peak I) was converted into ^{32}P-labelled trisaccharide after treatment with β-glucuronidase (Figure 1B). This ^{32}P-labelled trisaccharide was resistant to periodate oxidation (Figure 1C). In contrast, after phosphate removal by alkaline phosphatase (Figure 1D), the trisaccharide became sensitive to periodate oxidation (Figure 1E). As expected, the 2-aminobenzamide-derivatized peaks in Figures 1(D) and 1(E) were eluted at the positions of the 2-aminobenzamide-derivatized Galβ1-3Galβ1-4Xyl derived from peak II (Figure 1A) treated with β-glucuronidase (Figure 1F) and the products obtained after periodate oxidation of the 2-aminobenzamide-derivatized Galβ1-3Galβ1-4Xyl (Figure 1G) respectively. As the 2-aminobenzamide was attached to C-1 of xylose, the phosphate must have been attached to C-2, making the C-1/C-2 and the C-2/C-3 glycols resistant to periodate oxidation. These findings clearly showed that FAM20B is a xylose kinase that phosphorylates C-2 of the xylose residue in the GAG–protein linkage region of the tetrasaccharide sequence of α-TM.

To examine the intracellular localization of xylose kinase, a full-length form of FAM20B fused with EGFP (enhanced GFP) at the C-terminus (FAM20B–EGFP) was generated as described in the Experimental section. FAM20B–EGFP was then co-expressed with a Golgi marker (Golgi–DsRed) or an ER (endoplasmic reticulum) marker (ER–DsRed) in HeLa cells and analysed by confocal microscopy. FAM20B–EGFP (Figure 2A) was co-localized with the Golgi–DsRed marker (Figure 2C), whereas FAM20B–EGFP (Figure 2D) was not completely co-localized with the ER–DsRed marker (Figure 2F). These results suggested that FAM20B acts as a xylose kinase in the Golgi apparatus.

To examine further the physiological relevance of the xylose kinase, we investigated whether overexpression or knockdown of *FAM20B* changes the amount of GAG in HeLa cells. As shown in Table 2, the disaccharide composition and the amount of GAG isolated from each of the representative stable clones were analysed by HPLC, as described in the Experimental section. The results showed a correlation between the expression levels of *FAM20B* and the total amounts of both CS and HS in these stable clones. In addition, while no significant change was detected in the disaccharide composition of HS among these stable clones, differences in that of CS were observed corresponding to the expression level of *FAM20B*. Notably, the increase in the expression level of *FAM20B* was concomitant with the increase in the proportion of CS 6-sulfate to 4-sulfate. Together, results these indicated that the xylose kinase regulates the sulfation profile of CS chains as well as the total amount of GAG synthesized in cells.

Prompted by these observations, we next examined whether the increase in the amount of GAG in *FAM20B*-overexpressing cells is caused by an increased number of GAG chains or by greater elongation of individual GAG chains. For this analysis, the length of CS and HS chains obtained from *FAM20B*-overexpressing and mock-transfected cells was compared. Gel-filtration analysis using a Superdex 200 column revealed that the length of CS and HS chains in *FAM20B*-overexpressing cells was similar to that in mock-transfected cells (Figure 3), although the short CS chains in the overexpressing cells were particularly augmented. These findings indicated that the increase in the amount of CS and HS in *FAM20B*-overexpressing cells was mainly caused by

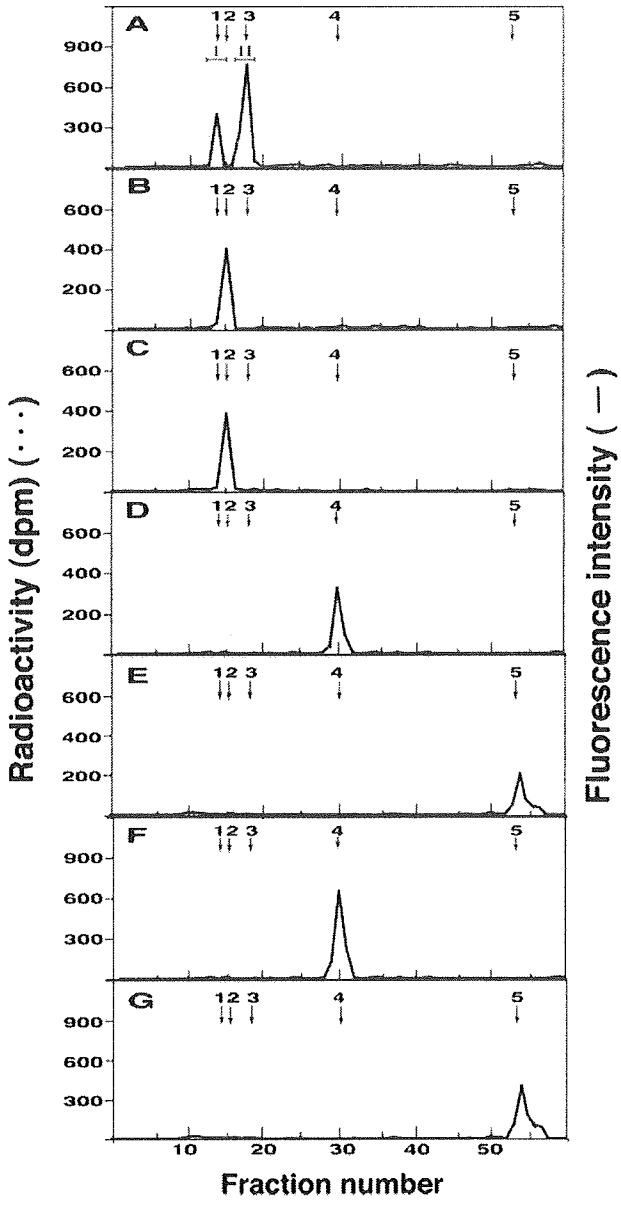

Figure 1 Identification of FAM20B reaction products as 2-O-phosphorylated xylose

^{32}P-labelled kinase reaction products were isolated by gel filtration, and ^{32}P-labelled tetrasaccharide chains were released by mild alkaline treatment, and then derivatized with 2-aminobenzamide (2-AB). The resultant ^{32}P-labelled tetrasaccharide fractions were subjected to gel-filtration HPLC on a column of GS-320 (7.6 mm × 500 mm) using 50 mM ammonium acetate as the eluent at a flow rate of 1 ml/min, as described in the Experimental section. (**A**) 2-AB-labelled tetrasaccharide fractions derived from FAM20B reaction products, which were released from α-TM by mild alkaline treatment. Peak I, ^{32}P-labelled tetrasaccharides; peak II, unlabelled tetrasaccharides. (**B**) ^{32}P-labelled compound (peak I) treated with β-glucuronidase. (**C**) Products obtained in (**B**) after periodate oxidation. (**D**) The same product after treatment with alkaline phosphatase. (**E**) The same product after digestion with alkaline phosphatase followed by periodate oxidation. (**F**) The unlabelled tetrasaccharide (peak II) treated with β-glucuronidase. (**G**) The products obtained in (**F**) after periodate oxidation. The eluates were monitored by fluorescence (solid line) at an excitation wavelength of 330 nm and an emission wavelength of 420 nm, and by ^{32}P radioactivity (broken line). Note that the ^{32}P-labelled tetrasaccharide (peak I) and the ^{32}P-labelled compound treated with β-glucuronidase were eluted at fractions 14 and 15 respectively. The arrows denote the elution positions of the following compounds: 1, GlcAβ1-3Galβ1-3Galβ1-4Xyl(2-*O*-phosphate)-2AB or free [^{32}P]phosphate; 2, Galβ1-3Galβ1-4Xyl(2-*O*-phosphate)-2AB; 3, GlcAβ1-3Galβ1-3Galβ1-4Xyl-2AB; 4, Galβ1-3Galβ1-4Xyl-2AB; 5, 2-AB-C_1-fragments.

Table 2 Disaccharide composition of GAGs in control and transfected HeLa cells

HeLa cells were transfected with vector alone (mock), plasmids containing *FAM20B* or *FAM20B* shRNA. For details, see the Experimental section. The values are expressed as pmol of disaccharide per mg of dried homogenate of these cells and are the means ± S.E.M. of three determinations. ΔHexA, GalNAc, GlcN and GlcNAc are unsaturated hexuronic acid, *N*-acetylgalactosamine, glucosamine and *N*-acetylglucosamine respectively, and 6S, 4S, NS and 2S represent 6-*O*-sulfate, 4-*O*-sulfate, 2-*N*-sulfate and 2-*O*-sulfate respectively. For relative expression, relative amounts of the *FAM20B* transcript were quantified by quantitative real-time RT–PCR. Data were normalized using glyceraldehyde-3-phosphate dehydrogenase mRNA levels. ND, not detected (<0.01 pmol/mg).

	Composition [pmol/mg (mol%)]		
Disaccharide	FAM20B	Mock	*FAM20B* shRNA
ΔHexAα1-3GalNAc	2.4 ± 0.3 (3)	1.7 ± 0.2 (3)	2.4 ± 0.3 (6)
ΔHexAα1-3GalNAc(6S)	16.0 ± 1.5 (21)	8.7 ± 1.3 (17)	3.7 ± 1.0 (10)
ΔHexAα1-3GalNAc(4S)	53.4 ± 6.7 (69)	37.9 ± 3.3 (74)	29.7 ± 1.1 (79)
ΔHexA(2S)α1-3GalNAc(6S)	2.3 ± 0.3 (3)	1.2 ± 0.3 (2)	0.5 ± 0.2 (1)
ΔHexAα1-3GalNAc(4S,6S)	2.8 ± 0.2 (4)	1.9 ± 0.5 (4)	1.6 ± 0.1 (4)
Total CS	76.9 ± 7.9	51.4 ± 3.7	37.9 ± 2.4
ΔHexAα1-4GlcNAc	105.8 ± 12.9 (45)	49.4 ± 4.6 (44)	43.1 ± 2.5 (50)
ΔHexAα1-4GlcNAc(6S)	8.1 ± 1.2 (3)	1.8 ± 0.4 (1)	0.2 ± 0.1 (1)
ΔHexAα1-4GlcN(NS)	76.8 ± 10.0 (33)	36.2 ± 4.9 (32)	29.8 ± 3.0 (34)
ΔHexAα1-4GlcN(NS,6S)	9.5 ± 1.4 (4)	2.5 ± 0.7 (2)	ND
ΔHexA(2S)α1-4GlcN(NS)	18.6 ± 1.7 (8)	10.9 ± 1.1 (10)	8.2 ± 0.7 (9)
ΔHexA(2S)α1-4GlcN(NS,6S)	17.3 ± 1.8 (7)	12.2 ± 1.5 (11)	5.6 ± 0.3 (6)
Total HS	236.1 ± 16.7	113.0 ± 10.9	86.9 ± 5.3
Relative expression	2.4	1.0	0.3

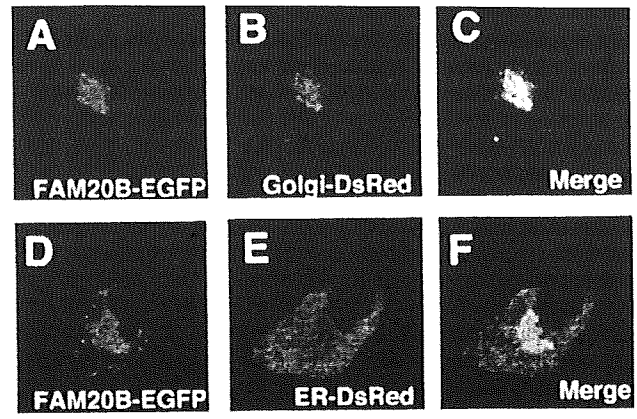

Figure 2 Subcellular localization of FAM20B

HeLa cells co-expressing FAM20B–EGFP (**A**) and Golgi–DsRed (Golgi marker) (**B**), co-expressing FAM20B–EGFP (**D**) and ER–DsRed (ER marker) (**E**) were analysed by confocal microscopy, as described in the Experimental section. The merged images show the precise co-localization of FAM20B with a Golgi–DsRed marker (**C**), whereas FAM20B–EGFP was not completely co-localized with an ER–DsRed marker (**F**).

the increased number of CS and HS chains. Although we cannot rule out the possibility that FAM20B may act on other substrates that could contribute to the enhanced effect on GAG biosynthesis, since a direct effect at the cellular level on xylose phosphorylation remains to be formally established, these results suggest that FAM20B regulates the amount of GAG chains by controlling the number of GAG chains and plays an important role in the biosynthesis of GAG.

We demonstrated recently that GlcAT-I could efficiently transfer a D-glucuronic acid residue to the phosphorylated

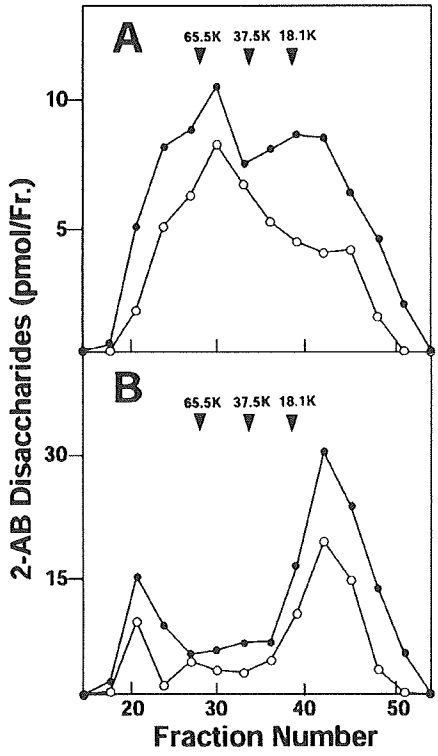

Figure 3 Analysis of the length of CS (A) and HS (B) chains

The purified GAG fraction was subjected to reductive β-elimination using $NaBH_4$/NaOH, and then analysed by gel-filtration chromatography on a column (10 mm × 300 mm) of Superdex 200. The digests of individual fractions obtained with chondroitinase ABC (**A**) or a mixture of heparinase and heparitinase (**B**) were derivatized with 2-aminobenzamide (2-AB), and then analysed by HPLC. The amounts of the 2-aminobenzamide derivatives of unsaturated disaccharides were calculated on the basis of fluorescence intensity. The samples from *FAM20B*-overexpressing cells (●) and the mock-transfected cells (○) are shown. Arrowheads indicate the size of molecular-mass standards (sizes are given in kDa). Results are representative of two series of independent experiments, where the two series of experiments gave essentially identical results.

trisaccharide–serine Galβ1-3Galβ1-4Xyl(2-*O*-phosphate)β1-*O*-Ser rather than to the non-phosphorylated counterpart Galβ1-3Galβ1-4Xylβ1-*O*-Ser [10]. In addition, in rat articular cartilage explants, the introduction of *GlcAT-I* enhanced GAG synthesis was attributable to an increase in the abundance rather than the length of GAG chains, whereas antisense inhibition of *GlcAT-I* expression impaired GAG synthesis [21]. Moreover, Bai et al. [22] showed that the transfection of Chinese-hamster ovary cell mutants defective in *GlcAT-I* with *GlcAT-I* cDNA augmented GAG synthesis to levels approximately double that in wild-type cells, suggesting that GlcAT-I regulates the expression of GAGs. Hence, phosphorylation of the xylose residue may be required for biosynthetic maturation of the linkage region tetrasaccharide sequence, which may be a prerequisite for the initiation and efficient elongation of the repeating disaccharide region of GAG chains. Xyl-2-*O*-phosphate has been found in both HS and CS from *Drosophila* to mammals [1,6,23]. In fact, a homologue of human *FAM20B* is present in *Drosophila*, suggesting that the possible involvement of phosphorylation of the xylose residue by FAM20B in the processing and maturation of the growing linkage region might be conserved during evolution.

To date, in addition to FAM20B, two FAM members (FAM20A and FAM20C) have been reported in mammals [14]. Database searches suggested that the amino acid sequence of FAM20B displays 35.7 and 39.1 % identity with FAM20A and FAM20C respectively. It was also reported that the *FAM20A* gene displays the most restricted expression pattern, whereas *FAM20B* and *FAM20C* are expressed in a wider variety of tissues and their expression patterns are very similar [14]. It is therefore possible that FAM20A and FAM20C are also involved in phosphorylation of the xylose residue in the linkage region, but exhibit distinct or overlapping acceptor substrate specificities. Characterization of FAM20A and FAM20C is now in progress.

AUTHOR CONTRIBUTION

Toshiyasu Koike and Tomomi Izumikawa performed the research, Hiroshi Kitagawa designed the research, Toshiyasu Koike, Tomomi Izumikawa and Hiroshi Kitagawa analysed the data, Jun-ichi Tamura contributed new reagents, and Tomomi Izumikawa and Hiroshi Kitagawa wrote the paper.

FUNDING

This work was supported in part by a Scientific Research Promotion Fund from the Japan Private School Promotion Foundation, Core Research for Evolutional Science and Technology (CREST) of the Japan Science and Technology (JST) Corporation (to H. K.), and Grants-in-aid for Scientific Research-B [grant numbers 19390025 and 21390025] (to H. K.).

REFERENCES

1 Sugahara, K. and Kitagawa, H. (2000) Recent advances in the study of the biosynthesis and functions of sulfated glycosaminoglycans. Curr. Opin. Struct. Biol. **10**, 518–527
2 Kitagawa, H., Uyama, T. and Sugahara, K. (2001) Molecular cloning and expression of a human chondroitin synthase. J. Biol. Chem. **276**, 38721–38726
3 Izumikawa, T., Uyama, T., Okuura, Y., Sugahara, K. and Kitagawa, H. (2007) Involvement of chondroitin sulfate synthase-3 (chondroitin synthase-2) in chondroitin polymerization through its interaction with chondroitin synthase-1 or chondroitin polymerizing factor. Biochem. J. **403**, 545–552
4 Izumikawa, T., Koike, T., Shiozawa, S., Sugahara, K., Tamura, J. and Kitagawa, H. (2008) Identification of chondroitin sulfate glucuronyltransferase as chondroitin synthase-3 involved in chondroitin polymerization: chondroitin polymerization is achieved by multiple enzyme complexes consisting of chondroitin synthase family members. J. Biol. Chem. **283**, 11396–11406
5 Kitagawa, H., Izumikawa, T., Uyama, T. and Sugahara, K. (2003) Molecular cloning of a chondroitin polymerizing factor that cooperates with chondroitin synthase for chondroitin polymerization. J. Biol. Chem. **278**, 23666–23671
6 Mizumoto, S., Uyama, T., Mikami, T, Kitagawa, H. and Sugahara, K. (2005) Biosynthetic pathways for differential expression of functional chondroitin sulfate and heparan sulfate. In Handbook of Carbohydrate Engineering (Yarema, K. J., ed.), pp. 289–324, CRC Press, Boca Raton
7 Oegema, Jr, T. R., Kraft, E. L., Jourdian, G. W. and Van Valen, T. R. (1984) Phosphorylation of chondroitin sulfate in proteoglycans from the swarm rat chondrosarcoma. J. Biol. Chem. **259**, 1720–1726
8 Fransson, L.-Å, Silverberg, I. and Carlstedt, I. (1985) Structure of the heparan sulfate-protein linkage region: demonstration of the sequence galactosyl-galactosyl-xylose-2-phosphate. J. Biol. Chem. **260**, 14722–14726
9 Gulberti, S., Lattard, V., Fondeur, M., Jacquinet, J. C., Mulliert, G., Netter, P., Magdalou, J., Ouzzine, M. and Fournel-Gigleux, S. (2005) Phosphorylation and sulfation of oligosaccharide substrates critically influence the activity of human β1,4-galactosyltransferase 7 (GalT-I) and β1,3-glucuronosyltransferase I (GlcAT-I) involved in the biosynthesis of the glycosaminoglycan-protein linkage region of proteoglycans. J. Biol. Chem. **280**, 1417–1425
10 Tone, Y., Pedersen, L. C., Yamamoto, T., Izumikawa, T., Kitagawa, H., Nishihara, J., Tamura, J., Negishi, M. and Sugahara, K. (2008) 2-O-phosphorylation of xylose and 6-O-sulfation of galactose in the protein linkage region of glycosaminoglycans influence the glucuronyltransferase-I activity involved in the linkage regions synthesis. J. Biol. Chem. **283**, 16801–16807
11 Moses, J., Oldberg, Å., Cheng, F. and Fransson, L.-Å. (1997) Biosynthesis of the proteoglycan decorin-transient 2-phosphorylation of xylose during formation of the trisaccharide linkage region. Eur. J. Biochem. **248**, 521–526

12 Moses, J., Oldberg, Å. and Fransson, L.-Å. (1999) Initiation of galactosaminoglycan biosynthesis: separate galactosylation and dephosphorylation pathways for phosphoxylosylated decorin protein and exogenous xyloside. Eur. J. Biochem. **260**, 879–884
13 Eames, F. B., Swartz, E. M. and Kimmel, C. B. (2008) Fam20b and xylosyltransferase1 (Xylt1) drive cartilage matrix production and inhibit perichondral bone during endochondral ossification. Dev. Biol. **319**, 480
14 Nalbant, D., Youn, H., Nalbant, S. I., Sharma, S., Cobos, E., Beale, E. G., Du, Y. and Williams, S. C. (2005) FAM20: an evolutionarily conserved family of secreted proteins expressed in hematopoietic cells. BMC Genomics **6**, 11
15 Ishikawa, H. O., Takeuchi, H., Haltiwanger, R. S. and Irvine, K. D. (2008) Four-jointed is a Golgi kinase that phosphorylates a subset of cadherin domains. Science **321**, 401–404
16 Nawa, K., Sakano, K., Fujiwara, H., Sato, Y., Sugiyama, N., Teruuchi, T., Iwamoto, M. and Marumoto, Y. (1990) Presence and function of chondroitin-4-sulfate on recombinant human soluble thrombomodulin. Biochem. Biophys. Res. Commun. **171**, 729–737
17 Nadanaka, S., Kitagawa, H. and Sugahara, K. (1998) Demonstration of the immature glycosaminoglycan tetrasaccharide sequence GlcAβ1-3Galβ1-3Galβ1-4Xyl on recombinant soluble human α-thrombomodulin: a possible mechanism generating "part-time" proteoglycans. J. Biol. Chem. **273**, 33728–33734
18 Tamura, J. and Nishihara, J. (2001) Synthesis of phosphorylated and sulfated glycosyl serines in the linkage region of the glycosaminoglycans. J. Org. Chem. **66**, 3074–3083
19 Kitagawa, H. and Paulson, J. C. (1994) Cloning of a novel α2,3-sialyltransferase that sialylates glycoprotein and glycolipid carbohydrate groups. J. Biol. Chem. **269**, 1394–1401
20 Watzele, G. and Berger, E. G. (1990) Near identity of HeLa cell galactosyltransferase with the human placental enzyme. Nucleic Acids. Res. **18**, 7174
21 Venkatesan, N., Barré, L., Benani, A., Netter, P., Magdalou, J., Fournel-Gigleux, S. and Ouzzine, M. (2004) Stimulation of proteoglycan synthesis by glucuronosyltransferase-I gene delivery: a strategy to promote cartilage repair. Proc. Natl. Acad. Sci. U.S.A. **101**, 18087–18092
22 Bai, X., Wei, G., Sinha, A. and Esko, J. D. (1999) Chinese hamster ovary cell mutants defective in glycosaminoglycan assembly and glucuronosyltransferase I. J. Biol. Chem. **274**, 13017–13024
23 Yamada, S., Okada, Y., Ueno, M., Iwata, S., Deepa, S. S., Nishimura, S., Fujita, M., Van Die, I., Hirabayashi, Y. and Sugahara, K. (2002) Determination of the glycosaminoglycan–protein linkage region oligosaccharide structures of proteoglycans from *Drosophila melanogaster* and *Caenorhabditis elegans*. J. Biol. Chem. **277**, 31877–31886

Received 23 March 2008/12 May 2009; accepted 28 May 2009
Published as BJ Immediate Publication 28 May 2009, doi:10.1042/BJ20090474

Biochem. J. (2009) **421**, 157–162 (Printed in Great Britain) doi:10.1042/BJ20090474

online data

SUPPLEMENTARY ONLINE DATA

FAM20B is a kinase that phosphorylates xylose in the glycosaminoglycan–protein linkage region

Toshiyasu KOIKE*[1], Tomomi IZUMIKAWA*[1], Jun-Ichi TAMURA† and Hiroshi KITAGAWA*[2]

*Department of Biochemistry, Kobe Pharmaceutical University, Higashinada-ku, Kobe 658-8558, Japan, and †Department of Regional Environment, Faculty of Regional Sciences, Tottori University, Tottori 680-8551, Japan

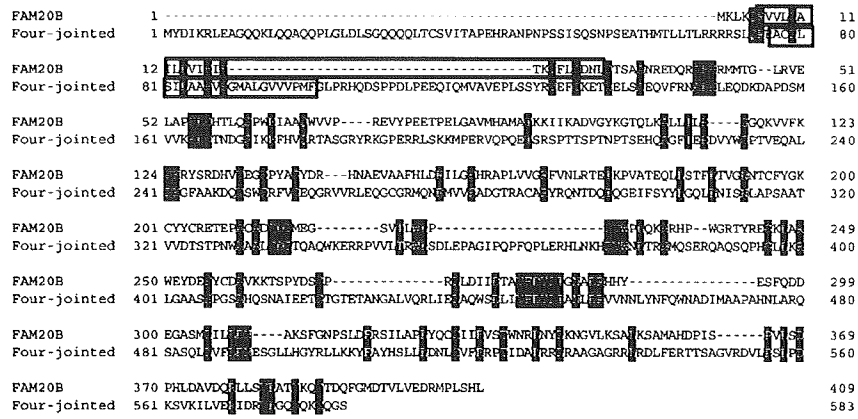

Figure S1 Comparison of the predicted amino acid sequence between human FAM20B and *Drosophila* Four-jointed

The predicted amino acid sequences were aligned using the program ClustalW (version 1.83). Closed boxes indicate that the predicted amino acid in the alignment is identical between the two sequences. Gaps introduced for maximal alignment are indicated by dashes. The putative membrane-spanning domains are boxed.

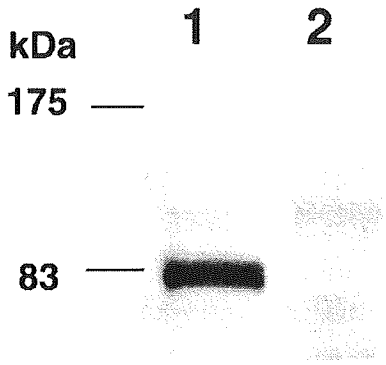

Figure S2 Western blot analysis of FAM20B

Culture medium from cells transfected with *FAM20B* or vector alone was purified with IgG–Sepharose and subjected to SDS/PAGE, and the expression of each Protein A-tagged protein was examined using anti-(mouse IgG) antibody. Lane 1, FAM20B–Protein A; lane 2, vector alone. Molecular masses are indicated in kDa.

Received 23 March 2008/12 May 2009; accepted 28 May 2009
Published as BJ Immediate Publication 28 May 2009, doi:10.1042/BJ20090474

[1] These authors contributed equally to this work.
[2] To whom correspondence should be addressed (email kitagawa@kobepharma-u.ac.jp).
The nucleotide sequence data reported in this paper have been submitted to the DDBJ, EMBL, GenBank® and GSDB Nucleotide Sequence Databases under accession number AB480690.

BJ www.biochemj.org

BJ Structure

Biochem. J. (2009) 419, 387–399 (Printed in Great Britain) doi:10.1042/BJ20081639

Chondroitin 4-*O*-sulfotransferase-1 is required for somitic muscle development and motor axon guidance in zebrafish

Shuji MIZUMOTO*†[1], Tadahisa MIKAMI*[1,2], Daiki YASUNAGA*, Naoki KOBAYASHI*, Hajime YAMAUCHI‡, Ayumi MIYAKE‡, Nobuyuki ITOH‡, Hiroshi KITAGAWA* and Kazuyuki SUGAHARA*†[2]

*Department of Biochemistry, Kobe Pharmaceutical University, Higashinada-ku, Kobe 658-8558, Japan, †Laboratory of Proteoglycan Signaling and Therapeutics, Graduate School of Life Science, Hokkaido University, Frontier Research Center for Post-Genomic Science and Technology, Nishi 11-choume, Kita 21-jo, Kita-ku, Sapporo, Hokkaido 001-0021, Japan, and ‡Department of Genetic Biochemistry, Kyoto University of Graduate School of Pharmaceutical Sciences, Yoshida-Shimoadachi, Sakyo-ku, Kyoto 606-8501, Japan

CS (chondroitin sulfate) has been implicated in a variety of biological processes during development. Its biological functions are closely associated with characteristic sulfated structures. Here, we report the characterization of a zebrafish counterpart of *C4ST-1* (chondroitin 4-*O*-sulfotransferase-1) and its functional importance in embryogenesis. Recombinant C4ST-1 showed a substrate preference for chondroitin and catalysed the 4-O-sulfation of GalNAc residues, a highly frequent modification of CS in the embryos of zebrafish as well as other vertebrates. Whole-mount *in situ* hybridization revealed that *C4ST-1* showed a distinct spatiotemporal expression pattern in the developing zebrafish embryo. During the segmentation stages, strong expression was observed along the body axis including the notochord and somites. Functional knockdown of *C4ST-1* with specific antisense morpholino-oligonucleotides led to a marked decrease in the 4-O-sulfation and amount of CS in the embryos. Consistent with the preferential expression in the rostrocaudal axis, *C4ST-1* morphants displayed morphological defects exemplified by a ventrally bent trunk and a curled and/or kinky tail, largely due to misregulated myotomal *myod* expression, implying perturbation of axial muscle differentiation in somites. Furthermore, the aberrant projection of spinal motor axons, which extended ventrally at the interface between the notochord and individual somites, was also observed in *C4ST-1* morphants. These results suggest that 4-O-sulfated CS formed by C4ST-1 is essential for somitic muscle differentiation and motor axon guidance in zebrafish development.

Key words: chondroitin sulfate, embryogenesis, glycosaminoglycan, sulfation, sulfotransferase, zebrafish.

INTRODUCTION

CS (chondroitin sulfate), one of the major sulfated GAGs (glycosaminoglycans), is ubiquitously distributed at the cell surface and in extracellular matrices as PGs (proteoglycans), in which CS chains are covalently attached to a panel of core proteins [1]. Mounting evidence suggests a substantial contribution of CS moieties to various physiological functions of CS-PGs such as cytokinesis, neuronal network formation, morphogenesis and infections with viruses and bacteria [2–5]. The chondroitin backbone is composed of repetitive disaccharide units [GlcA-GalNAc]$_n$. The building blocks can be replaced with sulfate groups at various positions, mainly at the C-2 position of GlcA and at the C-4 and/or C-6 positions of GalNAc residues, in various combinations, thereby producing characteristic sulfation patterns and structural heterogeneity [3,5]. The sulfation profiles are spatiotemporally tuned during development [6–8]. *In vitro* studies over the past decade have demonstrated that CS polymers and oligosaccharides possess neuroregulatory functions such as neuronal cell adhesion and neurite outgrowth, and interact with a wide range of signalling molecules including Hep (heparin)-binding growth factors in a sulfation-dependent manner [2,3,9,10], suggesting that CS chains differing in sulfation position and degree perform distinct functions *in vivo*.

In biosynthesis, the structural variability of CS is generated under the control of multiple sulfotransferases, and GlcA C-5-epimerase(s) that catalyse the epimerization of GlcA to IdoA (iduronic acid), converting CS into its stereoisomer DS (dermatan sulfate) [5,11,12]. Since each biosynthetic enzyme also exhibits a cell-type-specific and tissue-specific pattern of expression, substantial heterogeneity exists even in CS isolated from a single species. 4-O-sulfation of GalNAc residues is a typical modification found in CS/DS at higher frequency in vertebrates. To date, four phylogenetically related sulfotransferases, C4ST-1 (chondroitin 4-*O*-sulfotransferase-1), C4ST-2, C4ST-3 and D4ST-1 (dermatan 4-*O*-sulfotransferase-1), responsible for the 4-O-sulfation of CS/DS have been identified and characterized in mammals [13–17]. All these enzymes except C4ST-3 show broad, overlapping mRNA tissue distributions [13–16], and therefore are assumed to be, at least in part, functionally redundant, providing a plausible explanation for a predominance of the 4-O-sulfation over the other sulfations, each of which is catalysed by a single or two specific sulfotransferases [5,11]. However, a single deficiency in *C4ST-1* in sog9 cells, a mouse L cell mutant, has been demonstrated to

Abbreviations used: 2-AB, 2-aminobenzamide; BMP, bone morphogenetic protein; C4ST, chondroitin 4-*O*-sulfotransferase; CS, chondroitin sulfate; DAB, diaminobenzidine; D4ST, dermatan 4-*O*-sulfotransferase; DS, dermatan sulfate; EST, expressed sequence tag; GAG, glycosaminoglycan; Hep, heparin; hpf, hours post fertilization; HS, heparan sulfate; IdoA, iduronic acid; MO, morpholino-oligonucleotide; PAPS, 3′-phosphoadenosine 5′-phosphosulfate; PFA, paraformaldehyde; PG, proteoglycan; sema5A, semaphorin 5A; ΔHexA, 4,5-unsaturated hexuronic acid or 4-deoxy-α-L-*threo*-hex-4-enepyranosyluronic acid; ΔDi-0S, $\Delta^{4,5}$HexAα1–3GalNAc; ΔDi-4S, $\Delta^{4,5}$HexAα1–3GalNAc(4-*O*-sulfate); ΔDi-6S, $\Delta^{4,5}$HexAα1–3GalNAc(6-*O*-sulfate); ΔDi-diS$_D$, $\Delta^{4,5}$HexA(2-*O*-sulfate)α1–3GalNAc(6-*O*-sulfate); ΔDi-diS$_E$, $\Delta^{4,5}$HexAα1–3GalNAc(4,6-*O*-disulfate); ΔDi-triS, $\Delta^{4,5}$HexA(2-*O*-sulfate)α1–3GalNAc(4,6-*O*-disulfate).

[1] These authors have contributed equally to this work.

[2] Correspondence may be addressed to either of these authors (email tmikami@kobepharma-u.ac.jp or k-sugar@sci.hokudai.ac.jp).

The nucleotide sequence reported in this paper will appear in GenBank®, EMBL, DDBJ and GSDB Nucleotide Sequence Databases under the accession numbers AB097217, AB122017 and AB089140 for zebrafish C4ST-1, C4ST-2 and D4ST-1 respectively.

lead to a drastic decrease in the 4-O-sulfated structures and the amount of CS [18]. In addition, Klüppel et al. [19] have previously reported that a gene trap mutation of *C4ST-1* in mice causes severe chondrodysplasia characterized by a disorganized growth plate as well as specific alterations in the orientation of chondrocyte columns, where a strong reduction in 4-O-sulfated CS occurs. These findings suggest that C4ST-1 contributes substantially to the *in vivo* construction of 4-O-sulfated CS, and that the resultant CS plays crucial roles in embryonic development and physiological phenomena. However, despite the widespread expression of *C4ST-1*, additional phenotypic features resulting from loss of function of *C4ST-1* were not found, largely due to neonatal lethality with respiratory distress in the *C4ST-1*-deficient mice [19].

The zebrafish is an emerging model system for the study of vertebrate development and diseases [20], because of its well-characterized embryonic morphogenesis and suitability for forward and reverse genetics [21]. In zebrafish embryos, intense CS immunoreactivity was originally detected in the spinal cord and at the interface between the notochord and individual somites, where spinal motor axons extend ventrally to establish the midsegmental ventral motor nerves innervating the somatic musculature [22]. The labelled regions include at least two important embryonic structures, somites and peripheral spinal motor axons, both of which are formed as repetitive, morphologically similar units along the body axis through the segmentation processes that commonly occur during vertebrate development. Enzymatic removal of CS by injection of bacterial chondroitinase ABC in the trunk induces abnormal projections of the ventral motor axons, suggesting the involvement of CS in zebrafish motor axon growth [23]. In vertebrates, the motor axon growth is tightly associated with somite patterning [24]. Generally, each somite has two functionally different components, myotome and sclerotome, giving rise to skeletal muscle of trunk (and tail) and vertebral column respectively. Several lines of evidence suggest the importance of the sclerotomal component in the motor axon guidance in amniotes, whereas that in zebrafish embryos is linked with myotome formation [24]. The reason for such disparity can be explained by the fact that, in contrast with amniote embryos, the zebrafish somite consists predominantly of the myotome, and hence that the sclerotome is a relatively minor component [25]. In addition, the sclerotomal cells in teleosts including zebrafish are believed to differentiate into osteoblasts that produce bone matrix, but not into chondrocytes [26], implying that less cartilaginous anlagen forms in the trunk skeleton during embryogenesis, although cartilaginous skeletons are formed mainly in skull regions during early larval period. These characteristics of the zebrafish embryogenesis would facilitate the exploration of cryptic functions of CS and its biosynthetic enzymes in embryonic developmental processes besides endochondral skeletogenesis, which is severely affected in *C4ST-1*-deficient mice [19]. In the present study, we identified zebrafish *C4ST-1*, and examined its expression and *in vivo* functions during embryogenesis by functional knockdown of the enzyme using specific antisense MOs (morpholino-oligonucleotides).

EXPERIMENTAL

Materials

^{35}S-labelled PAPS (3′-phosphoadenosine 5′-phosphosulfate; 1.35 Ci/mmol) was purchased from PerkinElmer Life Sciences. Unlabelled PAPS was obtained from Sigma. The following sugars and enzymes were purchased from Seikagaku Corporation: chondroitin (a chemically desulfated derivative of whale cartilage CS-A); CS-B (porcine skin DS); six unsaturated standard disaccharides derived from CS {i.e. ΔDi-0S ($\Delta^{4,5}$HexAα1–3GalNAc, where ΔHexA is 4,5-unsaturated hexuronic acid or 4-deoxy-α-L-*threo*-hex-4-enepyranosyluronic acid), ΔDi-6S [$\Delta^{4,5}$HexAα1–3GalNAc(6-*O*-sulfate)], ΔDi-4S [$\Delta^{4,5}$HexAα1–3GalNAc(4-*O*-sulfate)], ΔDi-diS$_D$ [$\Delta^{4,5}$HexA(2-*O*-sulfate)α1–3GalNAc(6-*O*-sulfate), ΔDi-diS$_E$ [$\Delta^{4,5}$HexAα1–3GalNAc(4,6-*O*-disulfate)] and ΔDi-triS [$\Delta^{4,5}$HexA(2-*O*-sulfate)α1–3GalNAc(4,6-*O*-disulfate)]}; conventional chondroitinase ABC (EC 4.2.2.4) from *Proteus vulgaris*; chondroitinase AC-I (EC 4.2.2.5) from *Flavobacterium heparinum*; chondroitinase AC-II (EC 4.2.2.5) from *Arthrobacter aurescens*; and chondroitinase B (EC 4.2.2) from *F. heparinum*. Partially desulfated DS preparations were prepared by solvolysis using porcine skin DS with a high content (approx. 90%) of 4-O-sulfated disaccharide units [17]. For enzymatic assays of zebrafish sulfotransferases, a partially desulfated DS preparation that contained 4-O-sulfated disaccharide units at 31% was used, since the preparation has been found to be beneficial for comparative analysis of the substrate preferences of human enzymes, C4ST-1, C4ST-2 and D4ST-1 [17].

Fish maintenance

Zebrafish (*Danio rerio*) embryos were obtained by natural spawning and cultured at 28.5 °C in a 1/3 Ringer solution (39 mM NaCl, 0.97 mM KCl, 0.6 mM $CaCl_2$ and 1.67 mM Hepes, pH 7.2) [27].

Cloning of zebrafish cDNA encoding C4ST-1 and its isoenzymes

A search of the zebrafish EST (expressed sequence tag) database at 'Human BLAST against Zebrafish Shotgun Sequences' using the query sequences of human C4ST-1, C4ST-2, C4ST-3 and D4ST-1 [14–17] retrieved the respective zebrafish 5′- and 3′-ESTs, except for that of C4ST-3, which carried predicted start and stop codons respectively. The cDNAs covering the entire coding regions of the putative zebrafish C4ST-1, C4ST-2 and D4ST-1 were amplified from zebrafish embryonic cDNA by PCR using specific primer sets corresponding to the 5′- and 3′-non-coding regions, based on their respective ESTs. Each amplified cDNA fragment was subcloned into a pGEM®-T Easy vector (Promega) and sequenced in an ABI PRISM™ 377 DNA sequencer (Applied Biosciences). The cDNA sequences have been submitted to the GenBank® Nucleotide Sequence Database with accession numbers AB097217, AB122017 and AB089140 for zebrafish *C4ST-1* (*chst11*), *C4ST-2* (*zC4ST-2*) and *D4ST-1* (*d4st1*) respectively.

Construction of expression vectors encoding soluble forms of zebrafish C4ST-1

The cDNA encoding a truncated form of zebrafish C4ST-1 lacking the first N-terminal 59 amino acid residues was amplified by PCR with the pGEM®-T Easy vector containing the full coding sequence of the protein using a 5′-primer containing an in-frame BamHI site (5′-GCGGATCCCTACAGGAGCTGTAC-3′) and a 3′-primer containing a BamHI site located 19 bp downstream from the stop codon (5′-GCGGATCCCTCAGGAAGCGTGT-3′). PCR was carried out with Pfu DNA polymerase (Promega) by 30 cycles at 95 °C for 42 s, 59 °C for 42 s and 72 °C for 3 min. The PCR products were digested with BamHI, cloned into the BamHI site of an expression vector, pEF-BOS/IP [28], and sequenced. The resultant vector contained cDNA encoding a fusion protein that had an N-terminal cleavable insulin signal peptide and a Protein A IgG-binding domain followed by a truncated form of zebrafish C4ST-1.

Expression and sulfotransferase activity of the soluble forms of recombinant C4ST-1

The expression plasmid (6.7 μg) was transfected into COS-1 cells in 100 mm plates using FuGENE™ 6 transfection reagent (Roche Applied Science) according to the manufacturer's instructions. Two days after transfection, a 1 ml aliquot of the culture medium was incubated with 10 μl of IgG–Sepharose beads (GE Healthcare Bio-Sciences) for 2 h at 4 °C. The enzyme-bound beads were washed with TBS (Tris-buffered saline) containing 0.05 % (w/v) Tween 20, resuspended in the assay buffer described below and subsequently used as enzyme sources for sulfotransferase assays.

Sulfotransferase activities towards chondroitin and partially desulfated DS were assayed as described previously [17]. Briefly, the standard mixture (60 μl) contained 10 μl of the resuspended beads, 50 mM imidazole/HCl (pH 6.8), 2 mM dithiothreitol, 10 μM [^{35}S]PAPS (1 or 3 × 10^5 d.p.m.) and chondroitin or partially desulfated DS (10 nmol as GlcA) as an acceptor. The reaction mixtures were incubated at 28 °C for 30 or 60 min and subjected to gel filtration using a syringe column packed with Sephadex G-25 (superfine). The incorporation of [^{35}S]sulfate into chondroitin or partially desulfated DS was quantified by determination of the radioactivity in the flow-through fractions by liquid-scintillation counting. For identification of the transferase reaction products, the radioactive flow-through fractions were individually subjected to exhaustive digestion with chondroitinase ABC, AC-I or B. Each digest was analysed by anion-exchange HPLC on an amine-bound silica PA-03 column (YMC).

Whole-mount *in situ* hybridization

Digoxigenin-labelled RNA probes were transcribed *in vitro* using T7 or SP6 RNA polymerase, with the linearized pGEM®-T EASY vector containing the full-length cDNA of zebrafish *C4ST-1* (approx. 1.2 kb). Other probes for *shha* (*sonic hedgehog*; UniGene Cluster Dr.36074) [29] and *myod* (*myogenic differentiation*; UniGene Cluster Dr.36017) [30] were also synthesized. These cDNA fragments were amplified by PCR using embryonic zebrafish cDNA. For a cDNA fragment (approx. 700 bp) of zebrafish *shha*, the first PCR was performed with 5′-TTCGGCTCTGGTCTCGCTCCA-3′ as a forward primer, 5′-ATATCCCAACTCGCTCGAGCC-3′ as a reverse primer, and followed by nested PCR with 5′-GGAGGACAGAAGGCCGTGAAG-3′ and 5′-CTCTCACTCTCGCTCTCTCTC-3′. A cDNA fragment (approx. 640 bp) of zebrafish *myod* was amplified with 5′-CGAGCATCACCACATCGAGGA-3′ and 5′-TCGTTTAAGGTCGGATTCGCC-3′, followed by nested PCR with 5′-AGCACGTGAGGGCGCCCAGTG-3′ and 5′-CCTGAGCCTGCTGTTGAGGGC-3′. Each fragment was subcloned into a pGEM®-T EASY vector and the resultant plasmids were used as a template for the synthesis of the riboprobes.

Whole-mount *in situ* hybridization was carried out as described previously [31] with slight modifications. Zebrafish embryos were permeabilized by treatment with proteinase K (10 μg/ml), and then fixed with 4 % (w/v) PFA (paraformaldehyde)/PBS. The embryos were then treated with a hybridization buffer [50 % (v/v) formamide, 5× SSC (1× SSC is 0.15 M NaCl/0.015 M sodium citrate), 0.1 % (w/v) Tween 20, 50 μg/ml tRNA, 0.1 mg/ml single-stranded DNA and 50 μg/ml Hep] at 55 °C for 4 h, followed by hybridization with the appropriate digoxigenin-labelled RNA probe at 55 °C for 16 h. Thereafter, the embryos were treated with RNase A at 37 °C for 1 h in order to remove the excess probes; immersed in a blocking solution [150 mM maleate, 100 mM NaCl, 0.1 % (w/v) Tween 20 and 2 % (w/v) blocking reagent (Roche Applied Science), pH 7.5] at 4 °C for 4 h; incubated with anti-digoxigenin antibody conjugated with alkaline phosphatase (Roche) diluted 1:5000 with the blocking solution at room temperature (25 °C) for 4 h; and stained with an alkaline phosphatase substrate, BM Purple (Roche). The stained embryos were postfixed with 4 % PFA/PBS at room temperature for 30 min and then soaked in 50 % (w/v) glycerol/PBS overnight at 4 °C. Images were obtained using stereoscopic (SZX12) and confocal (BX51) microscopes (Olympus) equipped with a digital camera (CX380; Olympus).

Gene knockdown by morpholinos

MOs directed against *C4ST-1* were synthesized by Gene-Tools (Figure 4). The sequences of the MOs used are as follows: MO1, the sequence (25 bases) of which corresponds to that around the start codon, 5′-GGTCCAGTATGGTTTGTTTCATGGC-3′; control MO with five-base-mismatched nucleotides compared with MO1 (5-mis MO1), 5′-GGTaCAcTATGcTTTcTTTCATcGC-3′; control MO with the inverted sequence of MO1 (inverted MO1), 5′-CGGTACTTTGTTTGGTATGACCTGG-3′; MO2, the sequence (25 bases) of which corresponds to the 5′-non-coding region (–56 to –32), 5′-CTGCCGAGCCGAGCCCCGTTCAGCG-3′. Each MO was dissolved in distilled water at a final concentration of 8 mM. The stock solution was diluted to a working concentration of 0.8 mM in 1× Danieau solution (58 mM NaCl, 0.7 mM KCl, 0.4 mM $MgSO_4$, 0.6 mM $CaCl_2$ and 5 mM Hepes, pH 7.6), optionally with a non-toxic tracer, rhodamine–dextran (Molecular Probes) [32].

Each MO (approx. 3–5 ng) was individually injected into two- or four-cell embryos. MO-injected embryos (morphants) were incubated in a 1/3 Ringer solution containing 1 % penicillin/streptomycin at 28.5 °C for 16–50 h.

Inhibitory effects of the MOs on *in vitro* translation of zebrafish C4ST-1

For the preparation of mRNA, Myc-tagged sulfotransferase constructs were generated by in-frame insertion of the cDNA fragment of zebrafish *C4ST-1*, *C4ST-2* or *D4ST-1* into a pcDNA3.1/myc-His (–) vector (Invitrogen), and subsequently subcloned into a pEU3-NII vector (Toyobo). mRNAs encoding the fusion proteins with a C-terminal Myc tag were transcribed *in vitro* using ScriptMAX™ (Toyobo) with each pEU3-NII vector construct as a template. *In vitro* translation in the presence or absence of MOs was performed by using a cell-free protein synthesis kit, PROTEIOS™ (Toyobo), with the following modifications: in a 25 μl reaction, the transcribed mRNA (7 μg) and each MO at a final concentration of 100 μM were mixed with a reaction mixture containing components essential for *in vitro* translation and incubated at 26 °C for 24 h. An aliquot of the reaction mixture was heated at 95 °C for 5 min and then subjected to SDS/PAGE and proteins were transferred to a PVDF membrane. After pretreatment with a blocking solution [PBS containing 0.05 % (w/v) Tween 20 and 5 % (w/v) ECL® advance blocking reagent (GE Healthcare Bio-Sciences)], the membrane was treated with an anti-Myc antibody (Invitrogen), and subsequently with a horseradish peroxidase-conjugated anti-mouse IgG (GE Healthcare Bio-Sciences). Proteins bound to the antibody were visualized with an ECL® advance kit (GE Healthcare Bio-Sciences).

Rescue of morphological phenotypes of the MO1-injected embryos by injection with zebrafish *C4ST-1* RNA

A capped zebrafish *C4ST-1* sense RNA was synthesized using an mMASSAGE mMACHINE kit (Ambion) according to the manufacturer's protocol from a linearized pCS2+ vector carrying

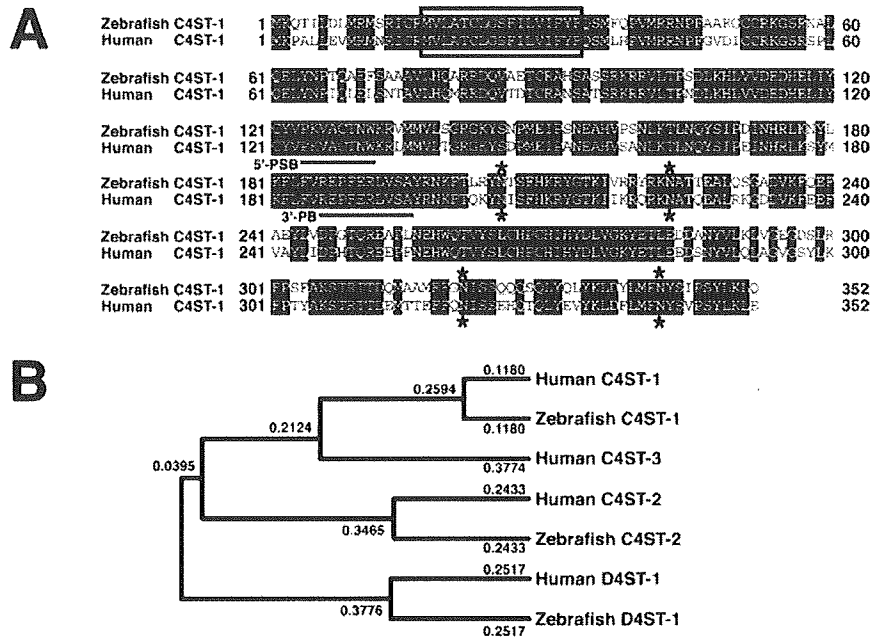

Figure 1 Comparison of the amino acid sequences of zebrafish and human C4ST-1 (A) and phylogenetic relations of subfamily members (B)

(**A**) Alignment of zebrafish and human C4ST-1 was performed using GENETYX-MAC (version 10.1) software. The aligned amino acids are shaded black if identical. The single transmembrane domain and potential N-linked glycosylation sites are indicated by a square box and asterisks respectively. Putative binding sites for the 5′-phosphosulfate group (5′-PSB) and 3′-phosphate group (3′-PB) of PAPS are underlined. (**B**) Apparent evolutional relationships of zebrafish and human C4ST/D4ST subfamily members were examined using GENETYX-MAC software. The length of each horizontal line is proportional to the degree of the divergence of the amino acid sequence.

the entire coding region of zebrafish *C4ST-1*. Zebrafish *C4ST-1* sense RNA (4–6 pg) was injected separately, immediately after the injection of MO1 into two-cell embryos.

Disaccharide composition analysis

GAG peptides derived from zebrafish embryos were prepared as described previously [33] but with minor modifications: dechorionated wild-type and MO1-treated embryos at 50 hpf (hours post fertilization) (approx. 160 and 300 embryos respectively) were homogenized in acetone and air-dried. The dried materials (9 and 16 mg respectively) were treated with boiling water for 10 min, cooled and exhaustively digested with heat-pretreated actinase E in 1 ml of 0.1 M borate buffer (pH 8.0), containing 10 mM calcium chloride at 60 °C for 16 h. The digest was treated with 5 % (v/v) trichloroacetic acid, and the acid-soluble fraction was extracted with diethyl ether. The aqueous phase was neutralized with 1 M sodium carbonate and adjusted to contain 80 % ethanol. The resultant precipitate was dissolved in water and subjected to gel filtration on a PD-10 column (GE Healthcare Bio-Sciences) with 50 mM pyridine acetate (pH 5.0) as eluent. The flow-through fractions were collected, evaporated dry and dissolved in water. An aliquot of the sample was digested with a mixture of chondroitinases ABC and AC-II in 50 mM Tris/HCl and 60 mM sodium acetate (pH 8.0) at 37 °C for 2 h. The digests were derivatized with a fluorophore, 2-AB (2-aminobenzamide), and then analysed by anion-exchange HPLC on a PA-03 column [34].

Immunohistochemistry

Wild-type and MO1-treated embryos (48–51 hpf) were fixed in 4 % PFA/PBS, dehydrated through graded concentrations of methanol and stored in methanol at −20 °C until used. The embryos were rehydrated in a stepwise fashion through a PBS-T (PBS containing 0.1 % Tween 20)/methanol series. After rinsing in PBS-T, the embryos were treated sequentially with the following solutions: (i) 3 % (w/v) H_2O_2 in methanol for 5 min; (ii) proteinase K (10 μg/ml) in PBS at 37 °C for 60 min; (iii) 4 % PFA/PBS for 20 min for postfixation; (iv) pre-chilled acetone at −20 °C for 8 min; (v) the above-mentioned blocking solution at 4 °C for 2 h; (vi) anti-acetylated α-tubulin (1:100, clone 6-11B-1; Sigma) in the blocking solution overnight at 4 °C; (vii) horseradish peroxidase-conjugated anti-mouse IgG (1:200) in the blocking solution overnight at 4 °C; (viii) 0.05 % DAB (diaminobenzidine) in PBS for 20 min; and (ix) 0.01 % DAB/3 % H_2O_2 in PBS for 5–15 min. Finally, the embryos were fixed with 4 % PFA/PBS and then soaked in 50 % glycerol/PBS.

RESULTS

Identification of zebrafish C4ST-1 and its isoenzymes

Our preliminary analysis of the disaccharide composition of CS present in the zebrafish embryos revealed that the 4-O-sulfation of GalNAc was a prominent modification of CS in the embryos as well as in the adult fish as recently reported [35]. In support of the prior analyses, a search of the zebrafish EST database showed the existence of at least three distinct zebrafish GalNAc 4-*O*-sulfotransferases potentially responsible for the 4-O-sulfation of CS/DS. Predicted open reading frames of the candidate enzymes were amplified by PCR using primers designed from the retrieved ESTs and the cDNA library of the 24-hpf zebrafish embryo as a template. One of the cDNA fragments (GenBank® accession number AB097217) obtained encoded a protein of 352 amino acid residues with a transmembrane domain, the putative PAPS-binding motifs, and 79 % identity to human C4ST-1 [14] at the amino acid level (Figure 1A). The deduced amino acid

Table 1 Comparison of the acceptor specificity of zebrafish and human C4ST-1

The recombinant zebrafish and human C4ST-1 were assayed using chondroitin and partially desulfated DS as acceptors (10 nmol as GlcA) as described in the Experimental section. The reaction products were separated from [^{35}S]PAPS by gel filtration on a syringe column packed with Sephadex G-25 (superfine). The radioactivity was measured by using liquid-scintillation counting.

	C4ST-1*† (pmol per ml of medium/h)	
Acceptor	Zebrafish	Human‡
Chondroitin	489	609
Partially desulfated DS	330	206

*The values represent the averages for two independent experiments.

†No detectable sulfotransferase activity was detected towards chondroitin or partially desulfated DS when the medium from cells transfected with the empty vector was used as an enzyme source in the control experiment.

‡The sulfotransferase activity of human C4ST-1 in [17].

sequences of the other two cDNAs (GenBank® accession numbers AB122017 and AB089140) had 61 and 65% identity to human C4ST-2 [14] and D4ST-1 [16,17] respectively. Phylogenetic analysis indicated that each zebrafish protein could be assigned to one of the human C4ST/D4ST subfamily members, except for C4ST-3 [15,17] (Figure 1B). Consequently, the zebrafish gene products of AB097217, AB122017 and AB089140 were named zebrafish C4ST-1, C4ST-2 and D4ST-1 respectively.

Characterization of zebrafish C4ST-1

To assess the sulfotransferase activity of zebrafish C4ST-1, a soluble form of the protein was generated by replacing the first 59 amino acids with a cleavable insulin signal sequence and a Protein A IgG-binding domain as described in the Experimental section, and then the chimaeric proteins were transiently expressed in COS-1 cells. The recombinant enzyme secreted into the culture medium was absorbed to IgG–Sepharose beads to eliminate endogenous sulfotransferase activities, and then the enzyme-bound beads were used for the enzyme assays. As shown in Table 1, the recombinant zebrafish C4ST-1 transferred [^{35}S]sulfate from [^{35}S]PAPS to both chondroitin and partially desulfated DS (containing 4-O-sulfated units at 31%) with approx. 1.5-fold greater incorporation into the former than into the latter. It should be noted that the partially desulfated DS used has been previously shown to be an excellent acceptor comparable with exhaustively desulfated DS under our assay conditions [17]. This substrate preference of the zebrafish enzyme was quite similar to that of human C4ST-1.

In order to identify the reaction products, the ^{35}S-labelled products were digested exhaustively with chondroitinase ABC, which cleaves both the GalNAc–GlcA and GalNAc–IdoA linkages in CS/DS, and then the digests were analysed by anion-exchange HPLC. In both digests of the reaction products, which were obtained by using chondroitin (Figure 2A) or partially desulfated DS (results not shown) as an acceptor, almost all of the total radioactivity was recovered at the elution position corresponding to the unsaturated disaccharide ΔDi-4S [ΔHexA-GalNAc(4S)], demonstrating that zebrafish C4ST-1 indeed catalysed the 4-O-sulfation of GalNAc residues in these polymers.

To further characterize the acceptor specificity of zebrafish C4ST-1, the radiolabelled 4-O-sulfation sites in the reaction products, which were obtained by incubation of C4ST-1 with

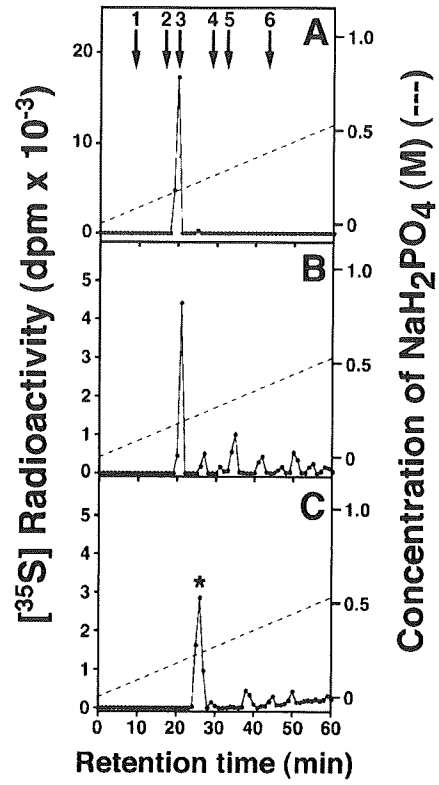

Figure 2 Identification of the zebrafish C4ST-1 reaction products prepared using chondroitin or partially desulfated DS as the acceptor substrate

^{35}S-labelled zebrafish C4ST-1 reaction products, which were obtained by using chondroitin (**A**) or partially desulfated DS (**B**, **C**) as the sulfate acceptor, were digested with chondroitinase ABC (**A**), chondroitinase AC-I (**B**) or chondroitinase B (**C**). Each digest was analysed by anion-exchange HPLC on an amine-bound silica PA-03 column using a linear gradient of NaH_2PO_4 from 16 to 530 mM over a 60 min period. The major radioactive peak marked by an asterisk in (**C**) is the putative monosulfated tetrasaccharide(s) identified previously [17]. The arrows indicate the elution positions of authentic unsaturated disaccharides: 1, ΔDi-0S; 2, ΔDi-6S; 3, ΔDi-4S; 4, ΔDi-diS_D; 5, ΔDi-diS_E; 6, ΔDi-triS.

partially desulfated DS, were structurally analysed, because porcine DS contains not only a large proportion of IdoA but also a very small proportion of GlcA. To this end, the reaction products were exhaustively digested with chondroitinase AC-I or chondroitinase B, which cleaves the GalNAc–GlcA linkages or GalNAc–IdoA linkages in CS/DS respectively. As reported previously [17], the ^{35}S-labelled DS preparations treated with the individual chondroitinases were decomposed into radioactive di-, tetra-, hexa-, octa- and higher oligosaccharides, depending on the distribution of GalNAc–GlcA/IdoA linkages in the polymer (results not shown). Particularly, quantification of the radio-labelled, 4-O-sulfated disaccharides derived from each digest provided important information regarding the preferable sequences (i.e. GalNAc residues flanked by either GlcA or IdoA) recognized by zebrafish C4ST-1. In the analysis of the radiolabelled reaction product of zebrafish C4ST-1, as much as 23% of total radioactivity was identified as ΔDi-4S in the chondroitinase AC-I digest (Figure 2B), whereas negligible radioactivity was detected at the position of ΔDi-4S in the chondroitinase B digest (Figure 2C), suggesting the strong preference of zebrafish C4ST-1 for the sequence -GlcA-GalNAc(4S)-GlcA- over -IdoA-GalNAc(4S)-IdoA-, this being consistent with the marked preference of human C4ST-1

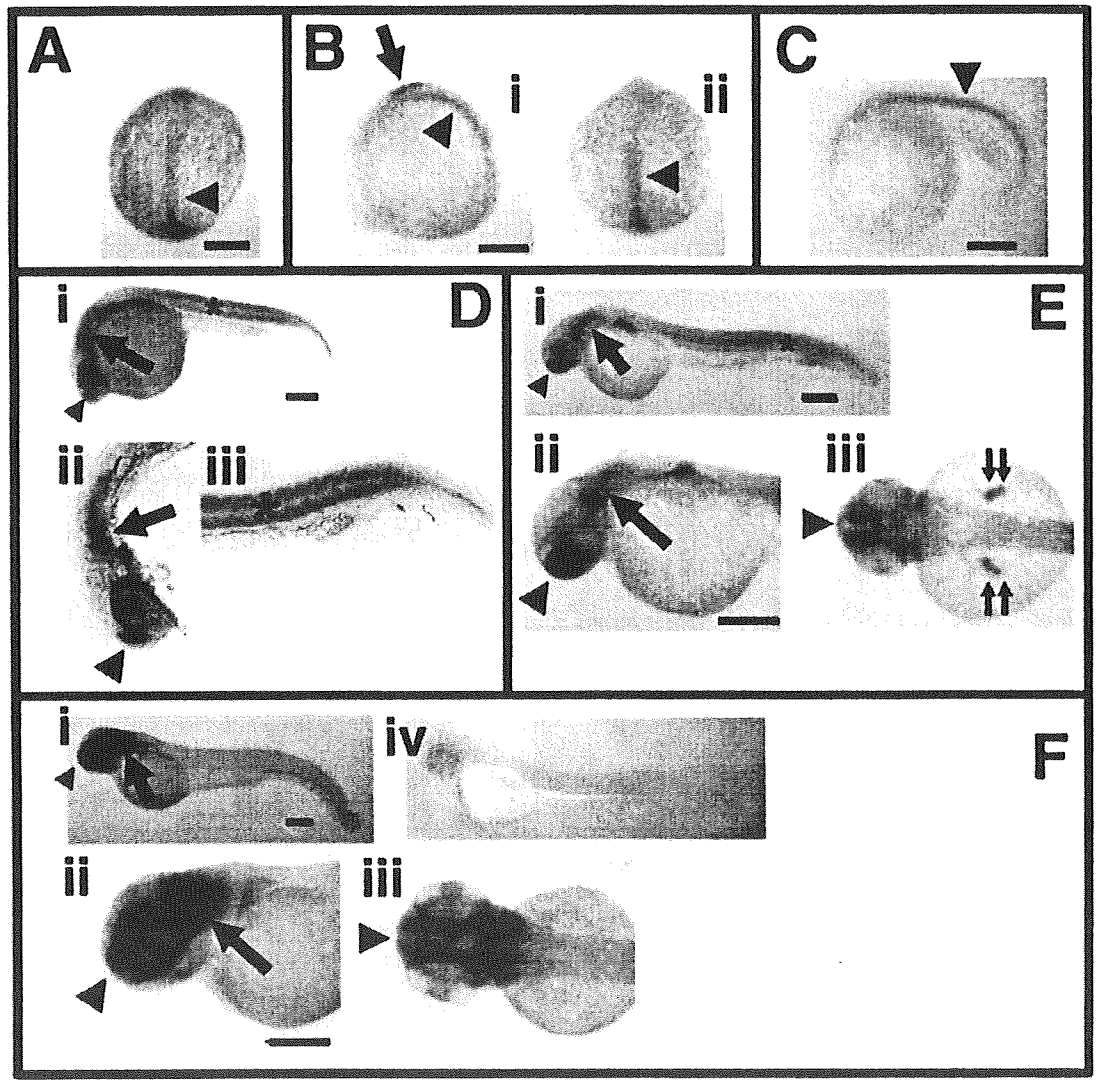

Figure 3 Expression of *C4ST-1* mRNA in zebrafish embryos

Expression of *C4ST-1* in zebrafish embryos at 10.5 (**A**), 13 (**B**), 18 (**C**), 24 (**D**), 36 (**E**) and 48 hpf (**F**). (**A–C**) *C4ST-1* was expressed in the notochord (arrowhead) and the spinal cord (arrow). (**D**, **E**) At 24 and 36 hpf, the expression was also observed in the somite (asterisk), the telencephalon (arrowhead), the rhombencephalon (arrow) and the pectoral fin buds (double arrows). Panels (**ii**) and (**iii**) of (**B–F**) are magnified images of the respective panels (**i**). (**F**) *C4ST-1* transcript was expressed in the telencephalon (arrowhead) and the rhombencephalon (arrow) at 48 hpf. No signal was detected by using a sense probe for *C4ST-1* [**F(iv)**]. [**A**, **B(ii)**] Dorsal view, anterior to the top. [**E(iii)**, **F(iii)**] Dorsal view, anterior to the left. [**B(i)**, **C**, **D**, **E(i)**, **E(ii)**, **F(i)**, **F(ii)**, **F(iv)**] Lateral view, anterior to the left. Scale bars, 200 μm.

for GalNAc residues in the GlcA-rich regions embedded in CS/DS [17]. We also detected GalNAc 4-*O*-sulfotransferase activities of zebrafish C4ST-2 and D4ST-1 towards chondroitin and partially desulfated DS. In terms of the preference for the isomeric uronic acids flanking the target GalNAc residues, their characteristics also closely resembled those of the respective human counterparts [17]: zebrafish D4ST-1 showed the opposite preference to zebrafish C4ST-1, and zebrafish C4ST-2 exhibited a less stringent structural requirement (results not shown). Taken together, the enzymatic specificities of these zebrafish enzymes strongly support orthologous relationships with their respective human counterparts (Figure 1B).

Because of an apparent predominance of C4ST-1 in CS biosynthesis and embryogenesis in mammals [18,19], compared with the other subfamily members, we focused on the biological functions of zebrafish C4ST-1 in the present study.

Expression pattern of C4ST-1 in zebrafish embryos

The developmental expression pattern of zebrafish *C4ST-1* was examined by whole-mount *in situ* hybridization using the specific antisense probe labelled with a digoxigenin. *C4ST-1* was detected in the notochord at 10.5–18 hpf within the segmentation period (Figures 3A–3C). A transient expression was also observed in the spinal cord at 13 hpf (Figure 3B). At 24 hpf, *C4ST-1* was strongly expressed in the somite, telencephalon and rhombencephalon, and the expression in the somite and brain persisted through 36 and 48 hpf respectively [Figures 3D, 3E and 3F(i–iii)]. In addition, the *C4ST-1* transcript was also transiently expressed in the fin buds at 36 hpf (Figure 3E). The sense probe for *C4ST-1* gave no signal throughout the entire embryo during all stages analysed [Figure 3F(iv) and results not shown], confirming the specific hybridization of the antisense probe.

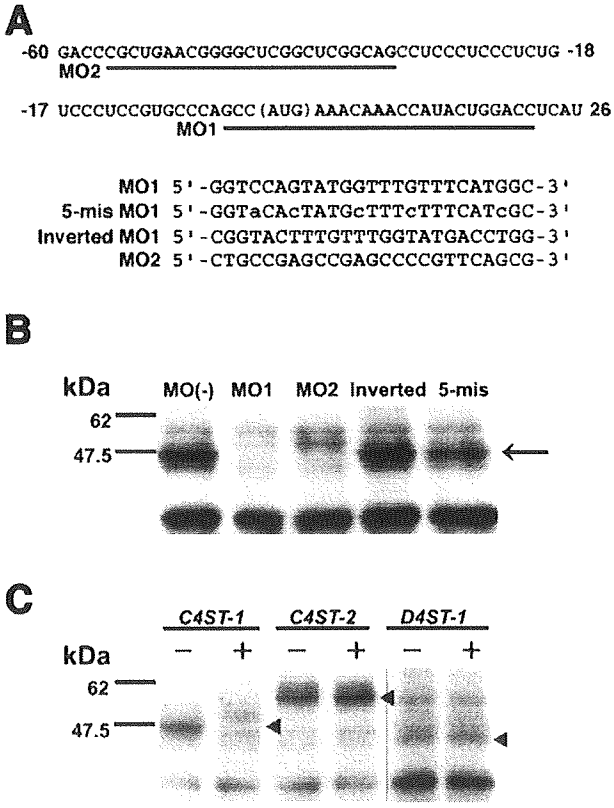

Figure 4 MOs directed against *C4ST-1* and their inhibitory effects on *in vitro* translation of *C4ST-1*

(**A**) *C4ST-1*-specific MOs, MO1 and MO2, were synthesized based on the nucleotide sequence of *C4ST-1* mRNA in the vicinity of the start codon (given in parentheses). 5-Mis MO1 and inverted MO1 were used as negative controls. The former has five-base-mismatched nucleotides compared with MO1 (indicated by lower-case letters), and the latter has the inverted sequence of MO1. (**B**) mRNA encoding C4ST-1/Myc was translated *in vitro* in the absence (−) or presence of each MO. (**C**) mRNAs encoding C4ST-1/Myc, C4ST-2/Myc and D4ST-1/Myc were individually translated *in vitro* in the absence (−) or presence (+) of MO1. The translated Myc-tagged proteins were analysed by Western blotting using an anti-Myc antibody (**B**, **C**). An arrow in (**B**) and arrowheads in (**C**) indicate the positions of C4ST-1/Myc and each Myc-tagged protein respectively. Bars on the left in (**B**) and (**C**) indicate the positions of standard protein markers.

Specificity of antisense MOs

The injection of antisense MOs has been demonstrated to inhibit the target gene functions in zebrafish embryos [32]. For functional knockdown of *C4ST-1*, we designed *C4ST-1*-specific MOs (MO1 and MO2) and control MOs (5-mis MO1 and inverted MO1) based on the nucleotide sequence of zebrafish *C4ST-1* mRNA in the vicinity of the start codon (Figure 4A and see the Experimental section). The specificity of the MOs was verified by an *in vitro* translational inhibition assay using a cell-free protein synthesis system. Incubation with MO1 or MO2 efficiently inhibited the translation of C4ST-1–Myc fusion protein, whereas no inhibitory effects were observed in the reaction mixture containing the individual control MOs (Figure 4B). In addition, MO1 did not interfere with the translation of the fusion proteins of C4ST-2 and D4ST-1 (Figure 4C).

Inhibition of C4ST-1 function causes morphological abnormalities in the trunk and tail

In order to assess the biological roles of 4-O-sulfated CS during zebrafish embryogenesis, *C4ST-1*-specific MOs and control MOs were individually injected into one- to four-cell stage embryos. The 48 hpf embryos injected with MO1 (approx. 3–5 ng per embryo) displayed a ventrally bent trunk and a curled and/or kinky tail (Figures 5B and 5B′). Furthermore, the injection of MO1 at higher dosages (> 5 ng per embryo) resulted in embryonic lethality by 48 hpf, and most of the embryos exhibited a severely shrunken body (Figure 5D). Thus, to avoid any non-specific toxicity of MOs, subsequent injections were conducted at dosages of 3–5 ng of MO per embryo. The morphological phenotypes of MO2-injected embryos were also similar to those of MO1-injected embryos (Figures 5C and 5C′), although the potential for functional knockdown of *C4ST-1* appeared to be less than that of MO1 (Figure 5H). In contrast, most of the embryos injected with control MOs, 5-mis or inverted MO1 developed normally (Figures 5E, 5F and 5H) as did wild-type embryos (Figures 5A and 5A′). In addition, co-injection of a zebrafish *C4ST-1* RNA (4–6 pg per embryo; $n = 20/31$, 64.5%; $P < 0.001$, χ^2 test) definitely rescued the morphological phenotypes of MO1-injected embryos (Figure 5G). These findings suggest that the morphological abnormalities in the body axis are caused by functional knockdown of *C4ST-1*, not by non-specific toxicity of MOs.

Analysis of CS of C4ST-1 morphants

To further verify the specificity of *C4ST-1*-specific MOs *in vivo*, GAG fractions prepared from wild-type and MO1-injected embryos at 50 hpf were chemically analysed. For exhaustive digestion of CS moieties, the purified GAG-peptide was treated with a mixture of chondroitinases ABC and AC-II. The resultant unsaturated CS disaccharides were derivatized with a fluorophore, 2-AB, followed by anion-exchange HPLC on a PA-03 column. As shown in Figure 6(A), in wild-type embryos, ΔDi-4S was the predominant sulfated disaccharide, accounting for 36.6% of all the disaccharides of CS (Table 2). As expected, the amount of ΔDi-4S in MO1-injected embryos was reduced by approximately one-third (Figure 6B and Table 2), suggesting a selective inhibitory effect of MO1 on the C4ST-1-mediated sulfation *in vivo*. Notably, the amount of CS (total disaccharides) in MO1-injected embryos was also approximately half that in wild-type embryos, due to significant reductions in non-sulfated disaccharide (ΔDi-0S) in addition to ΔDi-4S (Table 2). Together with previous biochemical analyses of CS in sog9 cells and in the growth plates of *C4ST-1*-deficient mice [18,19], these results indicate that C4ST-1 may also regulate the amount of CS, in addition to the 4-O-sulfation.

C4ST-1 is required for normal somite formation

Strong expression of *C4ST-1* in the notochord and somites (Figures 3A–3E), which are major components of the body axis during embryogenesis, indicated the aberrant morphology of the body axis observed in *C4ST-1* morphants to be caused by defects in the notochord and/or somites. To address this issue, we examined expression of the marker genes, *shha* [29] and *myod* [30], involved in the development of the notochord and skeletal muscle respectively, because zebrafish somites are almost exclusively occupied by myotomes, which give rise to the axial musculature through regulation of myogenic transcription factors including myoD [25,30]. In *C4ST-1* morphants, the signal of *shha* was observed along a rostrocaudal axis with a pattern of expression that was virtually the same as in wild-type embryos (Figures 7A and 7B) [29,36]. The profile of *myod* expression in the morphants at 16 hpf was also essentially indistinguishable from that in wild-type embryos, except that the expression in the morphants was expanded laterally to some extent (Figures 7C and 7D). Intriguingly, at 30 hpf, when development progressed

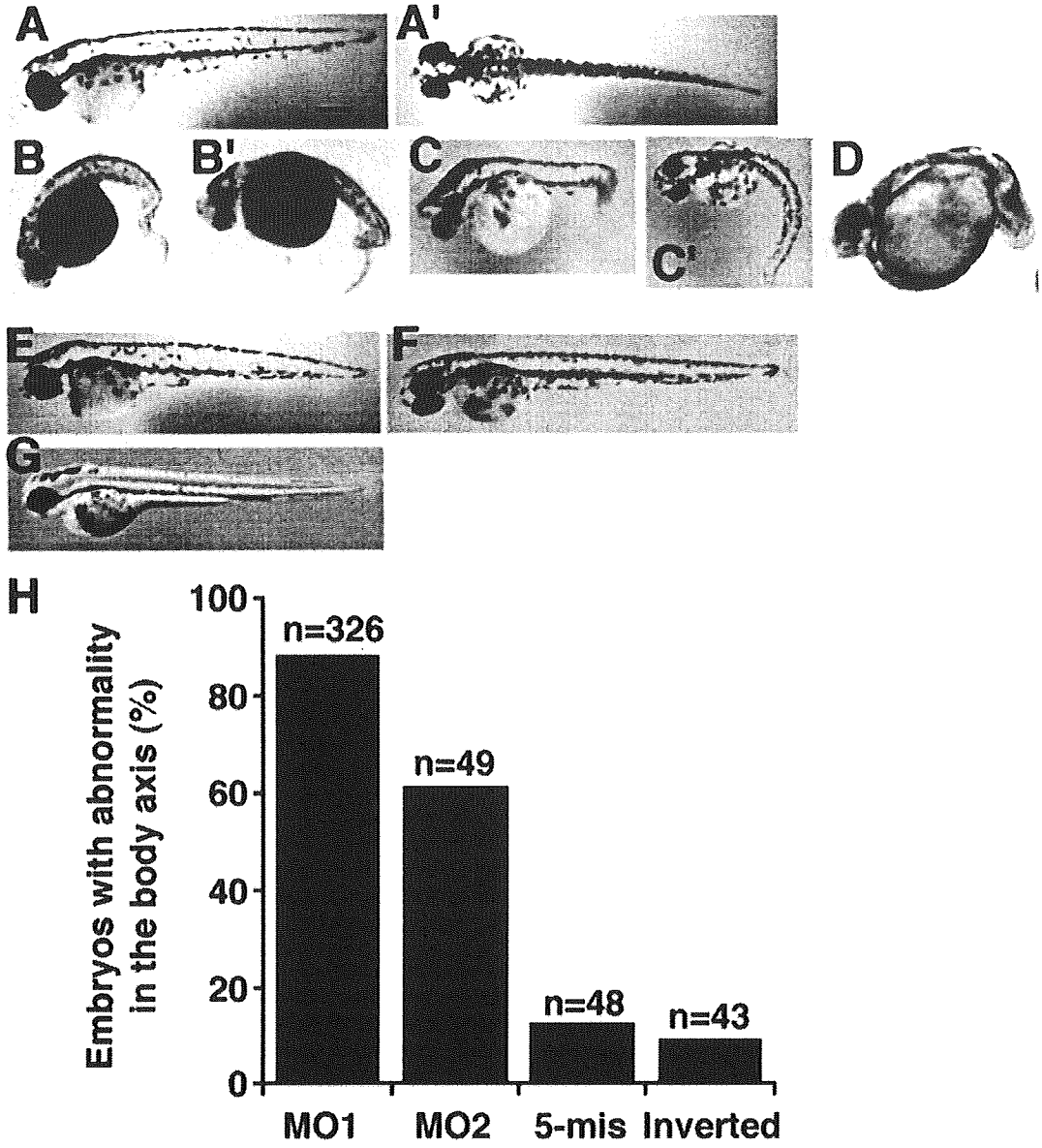

Figure 5 Morphological abnormality in *C4ST-1* morphants

(**A**, **A'**) A 48 hpf wild-type embryo. (**B**, **B'**, **C**, **C'**) Representative morphologies of 48 hpf embryos injected with 3–5 ng of MO1 (**B**, **B'**) or MO2 (**C**, **C'**). They had a ventrally bent trunk and a curled and/or kinky tail. (**D**) Injection of > 5 ng of MO1 was embryonic lethal by 48 hpf, and the morphology of the embryos was characterized by a severely shrunken body. (**E**, **F**) Embryos injected with 3–5 ng of 5-mis MO1 (**E**) or inverted MO1 (**F**) showed no obvious abnormality. (**G**) The morphological phenotypes of MO1-injected embryos were definitely rescued by injection with zebrafish *C4ST-1* RNA (4–6 pg per embryo). (**H**) Incidence of 48 hpf embryos with morphological abnormalities in the body axis after injection of each MO (3–5 ng). *n* represents the number of embryos injected with each MO. (**A**–**G**) Lateral view, anterior to the left. (**A'**, **B'**, **C'**) Dorsal view, anterior to the left.

further, *myod* in the somites disappeared in wild-type embryos (Figure 7E), whereas the high level of expression persisted in the *C4ST-1* morphants (Figure 7F). These observations suggest that the aberrant morphology in *C4ST-1* morphants is largely due to perturbation of somitic muscle development and not to notochordal defects.

Involvement of 4-O-sulfated CS in the motor axon guidance

A previous report by Bernhardt and Schachner [23] has suggested that CS has regulatory roles in axon guidance for ventral motor nerves in zebrafish embryos. This prompted us to investigate the overall pattern of the axonal projections from motor neurons in the *C4ST-1* morphants. Therefore, to visualize the ventral motor axons, 48–51 hpf MO1-treated and wild-type embryos were immunostained with anti-acetylated α-tubulin. As shown in Figures 7(G) and 7(I), the wild-type embryos exhibited regularly spaced, stereotyped ventral trajectories of motor axons along the midsegmental region of each somite. Since the segmentation process appeared morphologically normal in the *C4ST-1* morphants, the spatial arrangement of individual axons was almost identical with that in the wild-type embryos

Table 2 Disaccharide composition of CS from 50 hpf wild-type embryos and *C4ST-1* morphants

n.d., not detected.

	Wild-type embryos [pmol/mg of acetone powder (mol%)]	*C4ST-1* morphants [pmol/mg of acetone powder (mol%)]
ΔDi-0S	113.3 (53.6)	66.5 (58.4)
ΔDi-6S	17.9 (8.5)	16.7 (14.7)
ΔDi-4S	77.4 (36.6)	28.8 (25.3)
ΔDi-diS$_D$	2.9 (1.4)	1.9 (1.7)
ΔDi-diS$_E$	n.d.	n.d.
ΔDi-triS	n.d.	n.d.
Total CS/DS	211.5	113.9

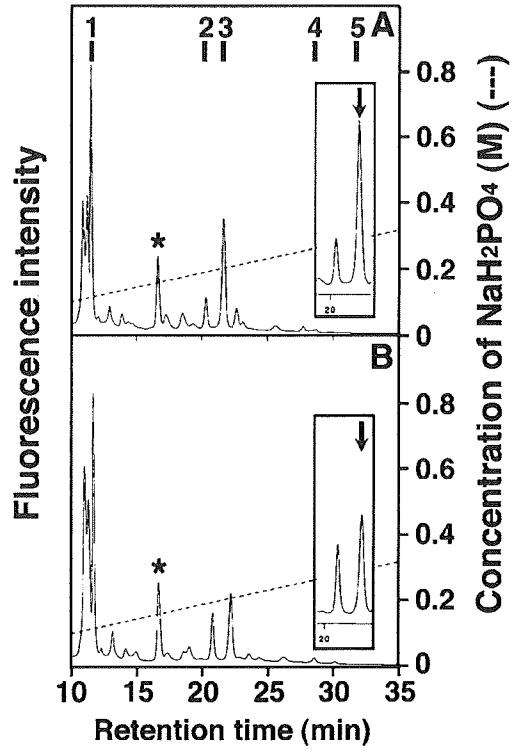

Figure 6 Anion-exchange HPLC analysis of CS derived from 50 hpf wild-type embryos and *C4ST-1* morphants

The GAG preparations purified from pooled 50 hpf wild-type embryos (**A**) and the *C4ST-1* morphants (**B**) were digested with a mixture of chondroitinases ABC and AC-II. The individual digests were labelled with the fluorophore 2-AB and analysed by anion-exchange HPLC as described in the legend of Figure 2. Insets of (**A**) and (**B**) are magnified chromatograms around the elution position corresponding to ΔDi-4S-2-AB (indicated by an arrow). Notably, the fluorescence intensity corresponding to the 4-O-sulfated disaccharide was significantly reduced in the *C4ST-1* morphants compared with the wild-type embryos. Vertical lines indicate the elution positions of 2-AB labelled, authentic unsaturated disaccharides: 1, ΔDi-0S-2-AB; 2, ΔDi-6S-2-AB; 3, ΔDi-4S-2-AB; 4, ΔDi-diS$_D$-2-AB; 5, ΔDi-diS$_E$-2-AB. The peaks marked by asterisks in (**A**) and (**B**) are unidentified materials.

(Figures 7H and 7J). However, many of the *C4ST-1* morphants had truncated, abnormally branched and/or misrouted axons (Figures 7H and 7J). Clearly recognizable ventral motor axons in the trunk and tail regions of these embryos were categorized into three groups by their morphological criteria (Table 3). Truncation and abnormal branching occurred significantly more often in the *C4ST-1* morphants than in the wild-type embryos. It should be noted that the axons observed in the morphants with a sharply bent trunk and/or a highly twisted tail were not scored for this classification to exclude the possibility that aberrant projections were caused by mechanical distortion of their profoundly altered body shapes. These results provide the first direct genetic evidence for the involvement of 4-O-sulfated CS in the axon guidance of ventral motor nerves.

Table 3 Incidence of abnormal motor axons in 48–51 hpf wild-type embryos and *C4ST-1* morphants

Ventral motor axons on only one body side per embryo were scored. Note that the aberrant axons in *C4ST-1* morphants with a sharply bent trunk and/or a highly twisted tail were not scored to exclude undesirable effects on axonal growth by mechanical distortion of their profoundly altered body shapes. The significance of wild-type embryos compared with *C4ST-1* morphants for each criterion was determined by a χ^2 test.

		n (% of total)		
	Nerves scored (total)	Truncation	Branching	Aberrant directions
Wild-type	162	3 (1.9)	0	0
C4ST-1 morphants	254	97 (38.2)*	11 (4.3)†	5 (2.0)‡

*$P < 0.001$.
†$P < 0.01$.
‡Not significant.

DISCUSSION

In the present study, a morpholino-based knockdown of zebrafish *C4ST-1*, whose product catalysed preferentially 4-O-sulfation of GalNAc residues in CS, was conducted to clarify the biological roles of CS in vertebrate embryogenesis. Biochemical analysis of CS derived from the *C4ST-1* morphants revealed that the enzyme is essential not only for 4-O-sulfation of CS but also for regulation of the amount of CS *in vivo*. In addition, the *C4ST-1* morphants displayed abnormal morphologies in the trunk and tail with perturbed expression of a myogenic marker *myod*, and in the projections of the ventral motor axons, providing the first genetic evidence for novel functions of 4-O-sulfated CS formed by C4ST-1 in skeletal muscle development and axon guidance during zebrafish embryogenesis.

As in mammals, the 4-O-sulfated structure was the predominant modification of CS in zebrafish embryos (Table 2). In addition, the existence of the sulfation machinery constituted by C4ST-1, C4ST-2 and D4ST-1 in zebrafish (Figure 1), and the strict substrate preferences towards chondroitin and partially desulfated DS (Figure 2 and Table 1), which were almost equivalent to those of the respective human counterparts [17], permitted the investigation of functions of these enzymes both in CS biosynthesis and in embryonic development using this representative model animal. It was recently reported that a single deficiency of *C4ST-1* in sog9 cells and a gene-trapped mutant mouse led to extensively reduced productivity in terms of the 4-O-sulfation and amount of CS [18,19]. Consistent with these observations, functional knockdown of *C4ST-1* in zebrafish embryos also resulted in a reduction in the 4-O-sulfation and amount of CS (Table 2), although the rate of decrease was milder with a MO-based knockdown than in sog9 cells and *C4ST-1*-deficient mice. This is probably attributable to the less severe penetrance in the phenotypes elicited by treatment with the *C4ST-1*-specific MO rather than to functional compensation by C4ST-2 and/or D4ST-1, because we used pooled MO1-treated embryos showing a mild phenotype in the body axis for the biochemical analysis of CS to avoid non-specific side effects of MO. Furthermore, an analysis using an *in vitro*

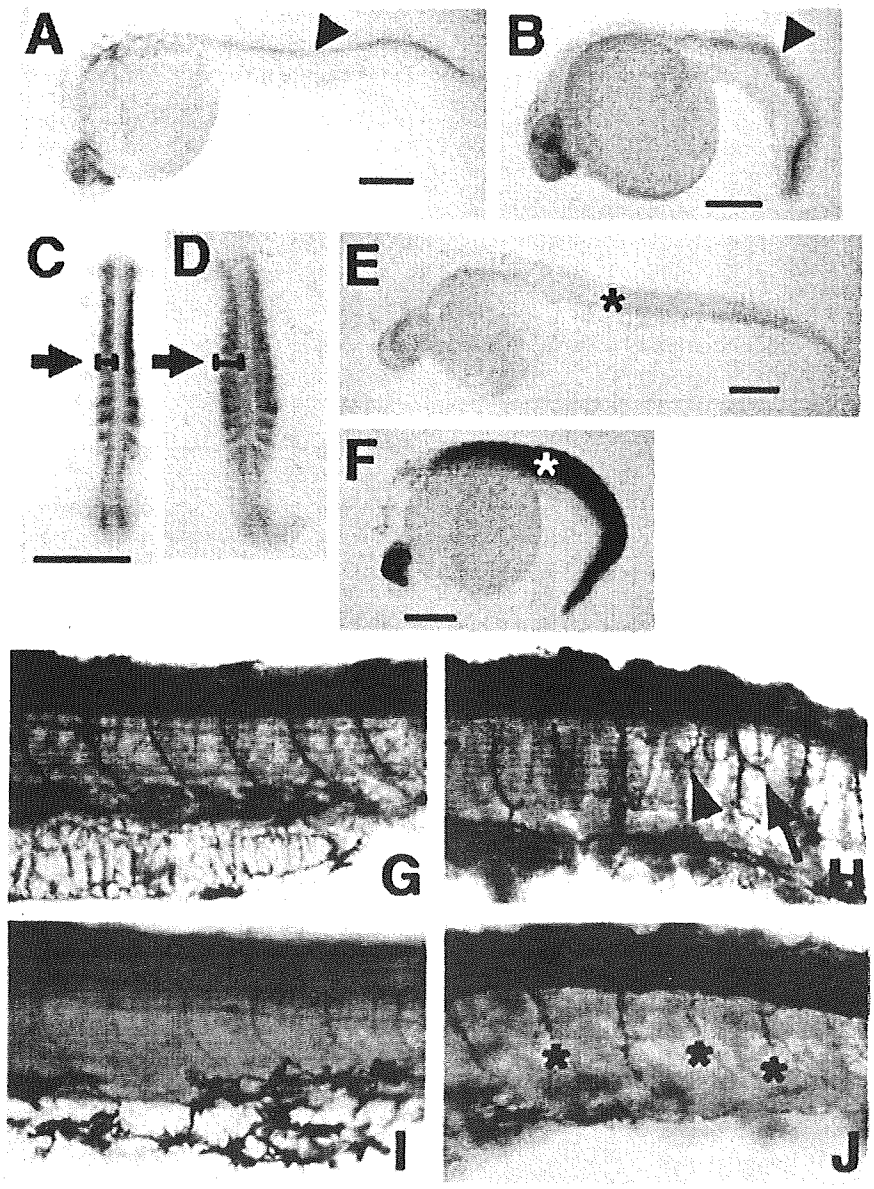

Figure 7 Expression patterns of *shha* and *myod* (A–F) and projection of ventral motor axons (G–J) in wild-type embryos and *C4ST-1* morphants

(**A–F**) Expression of *shha* (**A**, **B**) and *myod* (**C–F**) in wild-type embryos (**A**, **C**, **E**) and *C4ST-1* morphants (**B**, **D**, **F**) was detected by whole-mount *in situ* hybridization. There was no difference in notochordal *shha* expression (arrowhead) between wild-type embryos (**A**) and *C4ST-1* morphants (**B**) at 24 hpf. At 16 hpf, the expression pattern of *myod* in the morphants (**D**) had an intrinsically wild-type appearance (**C**), except for a slight lateral expansion (arrow). Of particular note, *myod* expression (asterisk) in the somites had largely disappeared in 30 hpf wild-type embryos (**E**), whereas it remained strong in 30 hpf *C4ST-1* morphants (**F**). (**A**, **B**, **E**, **F**) Lateral view, anterior to the left. (**C**, **D**) Dorsal view, anterior to the top. Scale bars, 200 μm. (**G–J**) Lateral views (anterior to the left, dorsal to the top) of trunk (**G**, **H**) and tail (**I**, **J**) regions in 48–51 hpf wild-type embryos (**G**, **I**) and the *C4ST-1* morphants (**H**, **J**) immunostained with anti-acetylated α-tubulin. In the *C4ST-1* morphants, truncated (asterisks), misrouted (arrowhead) and side-branched axons (arrow) were observed.

translation system revealed that MO1 inhibited the translation of C4ST-1 but not that of C4ST-2 or D4ST-1 (Figure 4), suggesting that *in vivo* expression of the two isoenzymes of C4ST-1 was not interfered with by the treatment with MO1. It should be noted that the overexpression of either C4ST-2 or D4ST-1, unlike C4ST-1, in sog9 cells cannot restore the proportion of 4-O-sulfated CS and amount of CS [18]. Together, these results suggest that C4ST-1 is the predominant sulfotransferase regulating the 4-O-sulfation and amount of CS during embryogenesis in zebrafish, as in mammals.

A previous immunohistochemical study has documented that CS is abundantly distributed at the interface between the notochord and individual somites [22]. In the present study, we also found the preferential expression of *C4ST-1* in the rostrocaudal axis including notochord and somites (Figure 3), suggesting important roles for CS in developmental processes in the body axis. Indeed, the morphological traits of 48 hpf *C4ST-1* morphants were characterized by a ventrally bent trunk and a twisted tail (Figure 5). Since the expression of the notochordal marker *shha* was unaffected in *C4ST-1* morphants (Figure 7), the notochordal formation appears to proceed normally. On the other hand, an unusually sustained expression of *myod* was observed in the *C4ST-1* morphants even at 30 hpf when the amount of its mRNA was