Determination of the Glycosaminoglycan-Protein Linkage Region
Oligosaccharide Structures of Proteoglycans from Drosophila melanogaster and Caenorhabditis elegans*

Shuhei Yamada†, Yukihiko Okada†, Momoyo Ueno†, Satomi Iwata†, S. S. Deepa‡, Shuji Nishimura‡, Masaki Fujita‡, Irma Van Die§, Yoshio Hirabayashi‡, and Kazuyuki Sugahara¶

From the †Department of Biochemistry, Kobe Pharmaceutical University, Higashinada-ku, Kobe 658-8558, Japan, the ¶Department of Molecular Cell Biology, Vrije Universiteit Medical Center, 1081 BT Amsterdam, The Netherlands, and the §Neuronal Circuit Mechanisms Research Group, RIKEN Brain Science Institute, Wako, Saitama 351-01, Japan

Caenorhabditis elegans and Drosophila melanogaster are relevant models for studying the roles of glycosaminoglycans (GAG) during the development of multicellular organisms. The genome projects of these organisms have revealed the existence of multiple genes related to GAG-synthesizing enzymes. Although the putative genes encoding the enzymes that synthesize the GAG-protein linkage region have also been identified, there is no direct evidence that the GAG chains bind covalently to core proteins. This study aimed to clarify whether GAG chains in these organisms are linked to core proteins through the conventional linkage region tetrasaccharide sequence found in vertebrates and whether modifications by phosphorylation and sulfation reported for vertebrates are present also in invertebrates. The linkage region oligosaccharides were isolated from C. elegans chondroitin in addition to D. melanogaster heparan and chondroitin sulfate after digestion with the respective bacterial eliminases and were then derivatized with a fluorophore 2-aminobenzamide. Their structures were characterized by gel filtration and anion-exchange high performance liquid chromatography in conjunction with enzymatic digestion and matrix-assisted laser desorption ionization time-of-flight spectrometry, which demonstrated a uniform linkage tetrasaccharide structure of -GlC(UA)-Gal-Gal-Xyl- or -GlC(UA)-Gal-Gal-Xyl(2-O-phosphate)- for C. elegans chondroitin and D. melanogaster CS, respectively. In contrast, the unmodified and phosphorylated counterparts were demonstrated in heparan sulfate of adult flies at a molar ratio of 7:93, which suggests that the linkage region in the fruit fly first becomes phosphorylated uniformly on the Xyl residue and then dephosphorylated. It has been established here that GAG chains in both C. elegans and D. melanogaster are synthesized on the core protein through the ubiquitous linkage region tetrasaccharide sequence, suggesting that indispensable functions of the linkage region in the GAG synthesis have been well conserved during evolution.

* The work at Kobe Pharmaceutical University was supported in part by a science research promotion fund from the Japan Private School Promotion Foundation and by Grants-in-aid for Encouragement of Young Scientists 11771474 (to S. Y.) and Scientific Research (B) Promotion Foundation and by Grants-in-aid for Encouragement of Science Research Promotion Fund from the Japan Private School Promotion Foundation. The nematode Caenorhabditis elegans and the fruit fly Drosophila melanogaster are ideal model organisms for studying a wide range of fundamental biological disciplines in development. The genetic studies have established that proteoglycans and their associated glycosaminoglycans (GAGs) are required for normal development of these organisms (Refs. 1 and 2; for a review, see Ref. 3). The unc-52 gene encodes the nematode homolog of mammalian perlecan, the major heparan sulfate (HS) proteoglycan of extracellular matrix. The unc-52 gene plays an essential role in the myofilalement assembly in the body wall muscle during the embryonic development (for a review, see Ref. 4). Glypican is a family of HS proteoglycans that are linked to the cell surface by a glycosylphosphatidylinositol anchor, and two family gene members, division abnormally delayed (daily) and daily-like (dly), have been identified in D. melanogaster. The D. melanogaster glypicans encoded by these genes have been implicated in Wingless (Wg)-mediated patterning of the embryo and play a critical role during the development (for a review, see Ref. 5).

Mutations affecting the genes encoding putative proteins related to GAG biosynthetic enzymes have also been described for these organisms. Mutations in the toll velu (ttv) gene of D. melanogaster cause defects in Hedgehog movement in mosaic wing discs (1, 6). The ttv gene is a putative ortholog of vertebrate EXT1, which encodes a heparan polymerase and is associated with the hereditary multiple exostoses syndrome in humans (7). The pipe gene, which affects dorsal-ventral patterning of D. melanogaster development, encodes a homolog of vertebrate HS 2-O-sulfotransferase (2, 8). The sugarless and sulfateless genes, both of which affect the fibroblast growth factor signaling during the D. melanogaster development, en-

1 The abbreviations used are: GAG, glycosaminoglycan; 2AB, 2-aminobenzamide; DE MALDI-TOF, delayed extraction matrix-assisted laser desorption ionization time-of-flight mass spectrometry; HPLC, high performance liquid chromatography; CS, chondroitin sulfate; DS, dermatan sulfate; HS, heparan sulfate; PG, proteoglycan; GlcUA, D-glucuronic acid; GlcN, D-glucosamine; GalNAc, N-acetyl-D-galactosamine; GalNAc-c, N-acetyl-D-galactosaminitol; Pen, pentose; Hex, hexose; HexUA, hexuronic acid; HexNAc, N-acetyl-D-glucosamine; HexUA, 4-deoxy-L-hex-4-enepyranosyluronic acid; 2P, 2-O-phosphate; 4S, 4-O-sulfate; 6S, 6-O-sulfate; Di-diSD, DiHS-diS 2, 4-O-sulfate; Di-4S, DiHS-4S; Di-3S, DiHS-3S; Di-2S, DiHS-2S; Di, DiHS; 6S, HexUA1-3GalNAc(6-O-sulfate); 6S, HexUA1-3GalNAc(4-O-sulfate); ADi-6S, ADiHS-6S; ADi, ADiHS; Di-3GalNAc(4-O-sulfate); Di, DiHS; HexUA1-3GalNAc(6-O-sulfate); HexUA1-3GalNAc(4-O-sulfate); ADiHs-II, ADiHS-II; 2P, 2-O-sulfate; 4S, 4-O-sulfate; 6S, 6-O-sulfate; 4S, 4-O-sulfate; 6S, 6-O-sulfate.
code homologs of the essential vertebrate enzymes for HS biosynthesis, UDP-glucose dehydrogenase, and HS N-deacetylase/N-sulfotransferase, respectively (9–12). The sqv-3 and -8 genes, identified on the basis of the common vulval invagination defect in \textit{C. elegans}, encode homologs of galactosyltransferase I and glucuronyltransferase I, respectively, both of which are required for the synthesis of the GAG-protein linkage region tetrasaccharide sequence -GlucUAβ1–3Galβ1–3Galβ1–4Xylβ1– on specific Ser residues of the core proteins (13–15).

Compared with the genetic analysis of proteoglycans and GAGs in \textit{C. elegans} and \textit{D. melanogaster}, their biochemical analysis has been less advanced. GAG chains in \textit{D. melanogaster} have been detected in the tissue extracts, based on the \textsuperscript{35}S)sulfate incorporation and the sensitivity of the materials to chondroitinas and nitrous acid treatments (16), whereas those in \textit{C. elegans} have been detected in cross-sections of all organs using an electron-dense dye in conjunction with GAG lysis digestion (17). It was also reported, based on the sensitivity to heparitinase and nitrous acid but not to chondroitinase ABC, that a single syndecan homolog is expressed as an HS proteoglycan in \textit{D. melanogaster} (18). The disaccharide compositions of HS and chondroitin sulfate (CS) chains in \textit{C. elegans} and \textit{D. melanogaster} have recently been determined (19, 20).

The genome projects of \textit{C. elegans} and \textit{D. melanogaster} have revealed the existence of multiple genes that are putative homologs of HS-synthesizing enzymes not only for the disaccharide repeating region but also for the linkage region (for a review, see Ref. 21). Although GAGs in vertebrates are covalently linked to Ser residues of their core proteins through the common tetrasaccharide linkage region (for reviews, see Refs. 22 and 23), there is no direct evidence that these GAG chains are covalently attached to core proteins through the conventional linker in invertebrates. Notably, it was suggested that unusual N-linked heparan sulfate and chondroitin sulfate chains are produced in addition to the classical SH and HS chains by some cultured cell lines (24) and by bovine lung tissues (25).

In the present study, we isolated and characterized the protein linkage region of \textit{C. elegans} chondroitin in addition to \textit{D. melanogaster} HS and CS to investigate whether these GAGs in invertebrates are attached to core proteins through the conventional linkage region tetrasaccharide and whether the tetrasaccharide is modified by phosphorylation and sulfation as in vertebrates.

### EXPERIMENTAL PROCEDURES

#### Materials—GAG lyases, sulfatases, and unsaturated CS disaccharides were obtained from Seikagaku Corp. (Tokyo, Japan). Calf intestine alkaline phosphatase (EC 3.1.3.1) of special quality for molecular biology was from Roche Molecular Biochemicals (Tokyo, Japan). Actinase \( E \) was purchased from Kaken Pharmaceutical Co. (Tokyo, Japan). Lysyl endopeptidase (\textit{Achromobacter lyticus}) was from Wako Pure Chemical Industries (Osaka, Japan). 2-Aminobenzamide (2AB) was purchased from Nacalai Tesque (Kyoto, Japan). Sephadex G-50 (fine) was purchased from Pharmacia Fine Chemicals (Osaka, Japan). 2-Aminobenzamide (2AB) was purchased from Sigma (St. Louis, MO). 2-AB-derivative of the linkage hexasaccharide, which was purified by gel filtration HPLC on an Asahipak GS320 column (29), was mixed in an equal volume of the aqueous solution (10 mg/ml) of a matrix, 2,5-dihydroxybenzoic acid. An aliquot (1 ml) of this sample-matrix mixture was placed on the sample plate well, dried under a stream of air, and analyzed.

#### Preparation of 2AB-derivatized GAGs from Adult Flies (\textit{D. melanogaster})—Adult flies (3.5 g) were extracted with chloroform/methanol (2:1), homogenized in acetone, and air-dried. The dried materials (790 mg) were treated in 36 ml of boiling water for 10 min, cooled, and exhaustively digested with lysyl endopeptidase using 4.3 mg (10 absorbance units) of enzyme. The reaction was terminated by heating at 100 °C for 1 min. The digest was mixed with 20 μl of water and 10 μl of 250 mM sodium acetate, pH 6.0, containing 5 mM of chondroitinase AC-II and incubated at 37 °C for 1 h. The reaction was terminated as above. Each digest was analyzed by anion-exchange HPLC on a PA-03 column (30) or by gel filtration HPLC on a column of Asahipak GS320 (7.6 × 500 mm) as described previously (32).

#### Digestion of the 2AB-derivatized Chondroitin from \textit{C. elegans} with \textit{Chondroitinases} ABC and AC-II—The 2AB-derivatives of the \textit{C. elegans} chondroitin (0.5 mg of the dry powder of the worm homogenate as the starting material) was digested with 12.5 μIU of chondroitinase ABC in a total volume of 20 μl of 50 mM Tris-HCl, pH 8.0, containing 60 mM sodium acetate at 57 °C for 1 h (29). The enzymatic reaction was terminated by heating at 100 °C for 1 min. The digest was mixed with 20 μl of water and 10 μl of 250 mM sodium acetate, pH 6.0, containing 5 mM of chondroitinase AC-II and incubated at 37 °C for 1 h. The reaction was terminated as above. Each digest was analyzed by anion-exchange HPLC on a PA-03 column (30) or by gel filtration HPLC on a column of Asahipak GS320 (7.6 × 500 mm) as described previously (32).

#### Separator Extraction Mass-assisted Laser desorption ionization Time-of-flight (DE MALDI-TOF) Mass Spectrometry (MS)—DE MALDI-TOF mass spectra in the positive or negative ion mode of the 2AB-derivatized linkage hexasaccharide from \textit{C. elegans} chondroitin were recorded on a Voyager DE-Pr/Pro (PerSeptive Biosystems, Framingham, MA) in the linear mode (33). An aqueous solution of the 2AB-derivatization of the linkage hexasaccharide, which was purified by gel filtration HPLC on an Asahipak GS320 column (29), was mixed in an equal volume of the aqueous solution (10 mg/ml) of a matrix, 2,5-dihydroxybenzoic acid. An aliquot (1 ml) of this sample-matrix mixture was placed on the sample plate well, dried under a stream of air, and analyzed.
above, followed by gel filtration on a PD-10 column using 50 mM pyridine acetate, pH 5.0, as the eluent. The flow-through fractions containing 2AB-labeled GAGs were dried and subjected to anion-exchange chromatography using a Sep-Pak cartridge (1.0 ml) of AccellTM Plus QMA, which had been equilibrated with 300 mM sodium phosphate, pH 6.0, containing 0.15 M NaCl. After washing with the buffer, the column was eluted stepwise with the same buffer containing 0.075, 0.15, or 0.5 M NaCl, respectively. The 0.5 M NaCl fraction, which contained essentially all the 2AB-labeled CS chains and 88% of the HS chains (see “Results”), was desalted by gel filtration on a PD-10 column using 50 mM pyridine acetate, pH 5.0, as the eluent. The flow-through fraction was dried and used for the analyses of the disaccharide composition, molecular sizes, and protein linkage regions of GAGs.

**Disaccharide Composition Analysis of D. melanogaster CS and HS—**
The 2AB-derived GAGs from adult flies were purified by ion-exchange chromatography using a Sep-Pak cartridge of AccellTM Plus QMA, and the purified 2AB-derived GAG preparation was digested either with chondroitinases ABC, AC-I, or AC-II or with a mixture of heparanase and heparitinase. In the case of the S2 cell-derived GAGs, the GAG-peptide preparation was subjected to the above lyase digestions. Each digest was treated with 2AB, and the 2AB-derived fraction was analyzed by anion-exchange HPLC on a PA-03 column (30). In the case of the adult fly sample, chondroitinase-produced unsaturated disaccharides were detectable also by ultraviolet absorption because of the large amounts.

**Gel Filtration Chromatography of CS and HS Chains from Adult Flies—**The 2AB-derived GAG fraction prepared above from adult flies was analyzed by gel filtration chromatography on a column (10 × 300 mm) of Superdex 200 eluted with 0.2 M CH₃COONH₄ at 80 °C, at a flow rate of 0.3 ml/min in an FPLC system (Amersham Biosciences). Fractions were collected at 3-min intervals, lyophilized, and digested with chondroitinase ABC or a mixture of heparanase and heparitinase (35, 36). The digests were derivatized with 2AB, then analyzed by anion-exchange HPLC on an amine-bound PA-03 column (30).

**HPLC Analysis of the 2AB-derivatives of the Linkage Region Oligosaccharides from Adult Flies—**The 2AB-derived GAG fraction prepared above was digested either with chondroitinases ABC and then AC-II or with a mixture of heparanase and heparitinase as described previously (34, 36). For the chondroitinase digestion, the 2AB-derivatized GAG fraction corresponding to 390 mg of the dried flies was digested successively with 75 μIU of chondroitinase ABC and 50 μIU of chondroitinase AC-II. Another aliquot of the isolated GAG fraction corresponding to 70 mg of the dried flies was used for the heparanase/heparitinase digestion. Each digest was analyzed by anion-exchange HPLC on a PA-03 column (30).

**Alkaline Phosphatase Digestion—**Alkaline phosphatase treatment was carried out using 4 IU of the enzyme in a total volume of 100 μl of 80 mM glycine/NaOH buffer, pH 9.9, containing 0.5 mM MgCl₂ at 37 °C for 30 min (32). The enzymatic reaction was terminated by heating at 100 °C for 5 min. Each enzyme digest was analyzed by anion-exchange HPLC on a PA-03 column (30, 32).

**Preparation of 2AB-derivatized GAGs from D. melanogaster S2 Cell Line—**S2 cells were grown in a spinner flask for suspension culture using the Schneider’s insect medium (Invitrogen) supplemented with 10% fetal calf serum and kanamycin sulfate (60 μg/ml) at 25 °C. Approximately 7 × 10⁶ cells were homogenized with ice-cold acetone and vacuum-dried. The dried materials were boiled in water for 10 min, cooled, and exhaustively digested overnight with actinase E at 60 °C as described above. The digest was treated with 5% trichloroacetic acid, and the acid-soluble fraction was extracted with ether. The aqueous phase was desalted by gel filtration chromatography on a PD-10 column. The purified GAG-peptide preparation was used for disaccharide composition analysis (see above). Another aliquot was treated with 0.5 M LiOH to liberate O-linked saccharides, and the liberated linkage oligosaccharides were derivatized with 2AB as described above.

**Simultaneous LiOH Treatment of Desialylated Bovine Submaxillary Mucin and Peptide CS from Whale Cartilage—**Highly purified bovine submaxillary mucin (donated by Prof. K. Kakehi, Kinki University, Osaka, Japan) was treated with 2 μg actinase E at 80 °C for 80 h to remove sialic acid and neutralized with alkali, and the desialylated mucin was recovered by ethanol precipitation. This preparation (48 mg), corresponding to 60 pmol of GalNAc (37), was mixed with a purified preparation of linkage region hexasaccharide peptides (30 pmol as hexasaccharides) prepared by chondroitinase ABC digestion of peptide CS-A from whale cartilage (27), dried, and treated with 50 μl of 0.5 M LiOH at 4 °C for 18 h. The mixture was neutralized with acetic acid, and the sample was treated with anion-exchange resin AG 50W-X2 (H⁺ form). The unbound fraction was neutralized with 1 × NH₄HCO₃ and was derivatized with 2AB as described above (30). One fifth of the sample was analyzed by gel filtration HPLC on a column of Asahipak GS20 (see above). Only the four expected hexasaccharide peaks (27, 30) were detected, suggesting that the GalNAc–O-Ser/Thr linkage is resistant to the LiOH treatment. The desialylated mucin showed a predominant peak of GalNAc–oH on HPLC as expected when treated with 1.0 M NaBH₄, 0.05 M NaOH in a control experiment.

**RESULTS**

**Structural Analysis of the GAG-Protein Linkage Region of C. elegans—**Previous studies have shown chondroitin and HS in C. elegans (19, 20). The acetone powder of the crude homogenate of the adult worms was digested with actinase E, and the resultant GAG-peptides were purified as described in “Experimental Procedures.” The GAG-peptides were treated with LiOH to liberate O-glycan chains including GAGs from the core proteins (27, 31). The liberated saccharides were labeled with a fluorophore 2AB, and the excess 2AB reagent was removed by paper chromatography. The sample extracted with water was then subjected to gel filtration chromatography and monitored by fluorescence intensity (Fig. 1). The fluorescence intensity of the sample prepared without the LiOH treatment was significantly weaker, compared with that of the LiOH-treated sample, suggesting the presence of C. elegans GAG attached to core proteins via O-glycosidic bonds. The flow-through fractions of both samples were pooled (see bars in Fig. 1) and analyzed further. The samples were digested with chondroitinase ABC and/or AC-II, and the digests were analyzed by anion-exchange or gel filtration HPLC. A major peak was observed at the position of the authentic 2AB-labeled nonsulfated hexasaccharide, ΔHexUA1–3GalNAcβ1–4GlcUAβ1–3Galβ1–4Xyl1–2AB (ΔHexUA represents 4-deoxy-a-L-threo-hex-4-enepyransoronic acid), when a chondroitinase ABC digest of the 2AB-derivative prepared after the LiOH treatment was analyzed by anion-exchange HPLC (Fig. 2A). No significant peaks were detected for the control sample prepared without the LiOH treatment (data not shown). Upon subsequent chondroitinase AC-II digestion, this peak was shifted to the position of the authentic 2AB-labeled nonsulfated tetrasaccharide, ΔHexUA1–3Galβ1–3Galβ1–4Xyl1–2AB (2B). The 2AB-labeled oligosaccharide generated by chondroitinase ABC digestion was isolated by gel filtration HPLC as shown in Fig. 3 and analyzed by DE MALDI-TOF/MS with 2,5-dihydroxybenzoic acid as a matrix (Fig. 4). Negative and positive ion mode DE MALDI-TOF/MS analyses showed molecular ion signals at m/z 1131 (Fig. 4) and 1133 (data not shown), corresponding to the molecular ions [M – H]⁻ and [M + H]⁺ of ΔHexUA, HexUA, HexNAc, Hex, and Pen, respectively. Together these results indicate that the C. elegans chondroitin is synthesized as proteoglycans, being covalently bound probably to a Ser residue of the core protein through the nonsulfated linkage region hexasaccharide, -GlcUAβ1–3GalNAcβ1–4GlcUAβ1–3Galβ1–3Galβ1–4Xyl1-. The structure of the GAG-protein linkage region of C. elegans HS could not be detected because of the limited amount of the sample, even when a sample corresponding to as much as 25 mg of the dried worms was used for a single injection for HPLC. The failure was consistent with the finding that the total amounts of HS disaccharides were 0.71 nmol (19) or 310 ng (20) per 25 mg of dried worms. Another aliquot of the sample prepared in the present study showed comparable amounts of 2AB-labeled oligosaccharides to those from the previous studies, confirming the reproducibility of the results. An amount of the sample larger by 1 order of magnitude may be required for the chemical detection,
Chondroitinase digestion analysis as shown in Fig. 2. QMA, which had been equilibrated with 300 mM sodium phosphate, pH 6.0, containing 0.015 M NaCl. The column was progressively with chondroitinases ABC and then AC-II, and were digested with chondroitinase ABC or a mixture of heparinase and heparitinase, respectively. The products were then derivatized with a fluorophore 2AB and analyzed by anion-exchange HPLC as described under “Experimental Procedures.” Compared with the calibration plot generated using the data obtained with size-defined commercial polysaccharides (Fig. 5, inset), the average molecular sizes of the D. melanogaster CS and HS were estimated to be 70 and 20 kDa, respectively. The C. elegans chondroitin gave a broad peak in the range of 40–50 kDa, suggesting that the heterogeneous populations were consistent with previous observations (19). It was not possible to analyze the size of the C. elegans HS chains because of the limited amount.

Structural Analysis of the GAG-Protein Linkage Region of the D. melanogaster CS and HS Chains—From the above mentioned 2AB-derivatized GAG preparation by gel filtration chromatography on a column of Superdex 200. To monitor small amounts of GAGs, aliquots of individual fractions were lyophilized and digested with chondroitinase ABC or a mixture of heparinase and heparitinase, respectively. The products were then derivatized with a fluorophore 2AB and analyzed by anion-exchange HPLC as described under “Experimental Procedures.” Compared with the calibration plot generated using the data obtained with size-defined commercial polysaccharides (Fig. 5, inset), the average molecular sizes of the D. melanogaster CS and HS were estimated to be 70 and 20 kDa, respectively. The C. elegans chondroitin gave a broad peak in the range of 40–50 kDa, suggesting that the heterogeneous populations were consistent with previous observations (19). It was not possible to analyze the size of the C. elegans HS chains because of the limited amount.

Analyses of the Disaccharide Compositions and Molecular Sizes of the D. melanogaster CS and HS—Prior to the analysis of the GAG-protein linkage regions of D. melanogaster CS and HS, the CS and HS chains were characterized for the disaccharide compositions and molecular sizes. For this purpose and for the subsequent analysis of the linkage region, fluorescently labeled free GAG chains were prepared from the acetone powder of adult flies by lysyl endopeptidase digestion followed by LiOH treatment and derivatization with a fluorophore 2AB as described under “Experimental Procedures.” The resultant 2AB-derivatized GAGs were subjected to ion-exchange chromatography using a Sep-Pak cartridge (1.0 ml) of Accell\textsuperscript{TM} Plus QMA, which had been equilibrated with 300 mM sodium phosphate, pH 6.0, containing 0.015 M NaCl. The column was washed with the equilibration buffer and then eluted stepwise with the same buffer containing 0.075, 0.15, 0.5, or 1.0 M NaCl. The 0.5 M NaCl fraction contained essentially all the 2AB-labeled CS chains and 88% of the HS chains as examined by disaccharide analysis using chondroitinase AC-II or a mixture of heparinase and heparitinase in conjunction with anion-exchange HPLC on a PA-03 column (data not shown). The 0.5 M NaCl fraction was desalted and used for disaccharide composition analysis, which was carried out by digestion with either chondroitinase ABC or a mixture of heparinase and heparitinase, respectively, followed by anion-exchange HPLC. The results, summarized in Table I, were comparable with those obtained for the CS and HS chains from the S2 cells (see below), and are basically in agreement with the previously reported results by Toyoda et al. (20). Chondroitinase AC-I or AC-II digestion gave a comparable disaccharide composition with that obtained by chondroitinase ABC, and chondroitinase B digestion gave no appreciable disaccharides, indicating that most, if not all, of the chondroitinase ABC-degraded materials are CS but not dermatan sulfate (DS), being consistent with the previous findings (20).

The molecular sizes of the GAGs were also analyzed using the 2AB-derivatized GAG preparation by gel filtration chromatography on a column of Superdex 200. To monitor small amounts of GAGs, aliquots of individual fractions were lyophilized and digested with chondroitinase ABC or a mixture of heparinase and heparitinase, respectively. The products were then derivatized with a fluorophore 2AB and analyzed by anion-exchange HPLC as described under “Experimental Procedures.” Compared with the calibration plot generated using the data obtained with size-defined commercial polysaccharides (Fig. 5, inset), the average molecular sizes of the D. melanogaster CS and HS were estimated to be 70 and 20 kDa, respectively. The C. elegans chondroitin gave a broad peak in the range of 40–50 kDa, suggesting that the heterogeneous populations were consistent with previous observations (19). It was not possible to analyze the size of the C. elegans HS chains because of the limited amount.

Structural Analysis of the GAG-Protein Linkage Region of the D. melanogaster CS and HS Chains—From the above mentioned 2AB-derivatized GAG preparation, the 2AB-labeled tetrasaccharides derived from the CS-protein linkage region were prepared by digesting the repeating disaccharide region successively with chondroitinases ABC and then AC-II, and were analyzed by anion-exchange HPLC on a PA-03 column. Depolymerization of CS chains synthesized on the conventional protein linkage region tetrasaccharide by chondroitinase ABC results in sulfated disaccharide units and core hexasaccharides derived from the linkage region (38). Chondroitinase AC-II will degrade a linkage region hexasaccharide into a disaccharide unit and a core tetrasaccharide (39). When the 2AB-derivatized linkage region prepared by chondroitinase AC-II digestion were analyzed by anion-exchange HPLC, only a single predominant peak was observed at the elution position of the authentic 2AB-tetrasaccharide \( \text{HexUA}_1-3\text{Gal}1-3\text{Gal}1-4\text{Xyl}2-O\)-phosphate)1–2AB (Fig. 6B). This sample was co-chromatographed (data not shown) with the corresponding standard linkage tetrasaccharide (Fig. 6A), confirming the identity of the peak. Upon subsequent alkaline phosphatase digestion, the peak was shifted by 13 min to the position of the nonsulfated tetrasaccharide (data not shown), suggesting that the compound in the peak contained a phosphate group most likely on the C-2 position of the Xyl residue in the linkage region tetrasaccharide structure (32, 40–42), and accounted for most, if not all, of the linkage regions of the CS-proteoglycans in Drosophila.

The 2AB-labeled tetrasaccharides derived from the HS-pro-
tein linkage region were also analyzed by HPLC after digesting the repeating disaccharide region using a mixture of bacterial heparinase and heparitinase. Because heparitinase cleaves the innermost glucosaminidic bond of HS chains (43, 44), the enzyme digestion results in various sulfated disaccharide units and linkage region core tetrasaccharides. As shown in Fig. 6C, two major peaks were observed at the elution positions of the authentic 2AB-tetrasaccharides, ΔHexUAα1→3Galβ1→3Galβ1→4Xyl1→2AB and ΔHexUAα1→3Galβ1→3Galβ1→4Xyl1→2AB, in a molar ratio of 73:27. The two peaks were co-eluted with the corresponding standards upon co-chromatography (data not shown). The peak, eluted at the position of ΔHexUA-Gal-Gal-Xyl(2-O-phosphate)1→2AB, was shifted by subsequent alkaline phosphatase digestion to the elution position of the nonsulfated tetrasaccharide (data not shown), suggesting that the compound in the peak contained a phosphate group most likely on the C-2 position of the Xyl residue in the linkage region tetrasaccharide sequence. The data obtained from the analyses of the chain sizes, disaccharide compositions, and linkage region structures of the adult flies are summarized in the diagrams shown in Fig. 7.

The GAG-protein linkage regions of CS and HS derived from the cultured S2 cells of D. melanogaster were also analyzed. Prior analysis of the disaccharide composition by HPLC on an amine-bound silica column after chondroitinases ABC, AC-I, or AC-II digestion of the 2AB-derivatized GAG fraction followed...
The disaccharide compositions of the D. melanogaster CS and HS chains

The 2AB-derivatized GAG fraction from adult flies was prepared as described under "Experimental Procedures" after protease digestion followed by LiOH treatment and derivatization with a fluorophore 2AB. The resultant 2AB-derivatized GAGs from adult flies or the GAG-peptide preparation from the D. melanogaster S2 cells were digested either with chondroitinase ABC, AC-I, or AC-II or with a mixture of heparinase and heparitinase. Each digest was labeled with 2AB, and the 2AB-derivatized fraction was analyzed by HPLC as described under "Experimental Procedures." The results from the chondroitinase digests were comparable and those obtained with the chondroitinase ABC digests are shown. In the case of the sample for adult flies, the large amount of the sample allowed us to detect the disaccharides by ultraviolet absorbance without 2AB labeling.

| Disaccharides                      | Adult flies | S2 cells |
|-----------------------------------|-------------|----------|
| CS disaccharides                  |             |          |
| ΔHexUA1-3GalNAc                   | 80          | 79       |
| ΔHexUA1-3GalNAc(4-O-sulfate)      | 20          | 21       |
| HS disaccharides                  |             |          |
| ΔHexUA1-4GlcNAC                   | 28          | 33       |
| ΔHexUA1-4GlcNAC(6-O-sulfate)      | 8           | 2        |
| ΔHexUA1-4GlcN(2-N-sulfate)        | 26          | 37       |
| ΔHexUA1-4GlcN(2-N, 6-O-sulfate)   | 21          | 17       |
| ΔHexUA1-4GlcN(2-N)-phosphate      | 20          | 21       |
| ΔHexUA1-4GlcN(2-N, 6-O-sulfate)   | 12          | 10       |
| ΔHexUA1-4GlcN(2-N, 6-O)-phosphate| 5           | 1        |

by 2AB-derivatization of the resultant disaccharides resulted in 2AB-labeled ΔDi-OS and ΔDi-4S in a molar ratio of 79:21, being in good agreement with the undersulfation observed for that of adult flies. Analysis after digestion with a mixture of heparinase and heparitinase gave 2AB-labeled disaccharides with sulfation profiles comparable with those obtained for the adult flies (Table I). The disaccharide compositions obtained by digests with chondroitinases ABC, AC-I, or AC-II were comparable, and chondroitinase B digestion gave no appreciable disaccharides, suggesting that the S2 cells produce CS but not DS. It is interesting from the molecular evolution point of view that Drosophila produces 4-O-sulfated CS but no DS which is also 4-O-sulfated on GalNAc residues, and C. elegans synthesizes only nonsulfated chondroitin. A CS/DS-specific epimerase does not appear to be expressed in these organisms. For the analysis of the CS- or HS-protein linkage region, the 2AB-labeled GAG chains were digested with chondroitinases ABC and then AC-II or a mixture of heparinase and heparitinase, respectively. The chondroitinases digest was analyzed by HPLC, and a single peak was observed at the elution position of ΔHexUA-Gal-Gal-Xyl(2-O-phosphate)-2AB (data not shown). This peak was shifted to the elution position of the nonsulfated tetrasaccharide upon alkaline phosphatase digestion (data not shown). The 2AB-labeled oligosaccharides derived from the HS-protein linkage region were also analyzed by HPLC, and two major peaks were observed at the elution positions of the authentic 2AB-tetrasaccharides, ΔHexUA-Gal-Gal-Xyl-2AB and ΔHexUA-Gal-Gal-Xyl(2-O-phosphate)-2AB (data not shown), in a molar ratio of 7:93. The results from the linkage region analyses of the CS and HS chains were in agreement with the findings obtained from the analysis of adult flies, although the phosphorylated component was the predominant component in the S2 cells, which suggests that the Xyl residue first becomes uniformly phosphorylated and then dephosphorylated.

**DISCUSSION**

The genes encoding the putative enzymes that are involved in the biosynthesis of the GAG-protein linkage region have been identified in the C. elegans and D. melanogaster genomes (for a review, see Ref. 21). The product of the C. elegans sqv-3 gene, which is involved in vulval invagination and oocyte development, has 38% sequence homology to human galactosyltransferase I (46). The catalytic activities of the expressed proteins of these cDNAs have been demonstrated in vitro (47). Xylosyltransferase, which transfers a xylose residue from UDP-Xyl to Ser residues at the GAG acceptor sites of core proteins, has also been cloned, and homologous genes are found in C. elegans and D. melanogaster genomes (48). Recently, Wilson (49) and Deepa et al. detected the xylosyltransferase activity for the expressed fly homolog. Bai et al. (50) reported the cloning of the Chinese hamster ovary galactosyltransferase II cDNA and the presence of its orthologs in C. elegans and D. melanogaster genomes.

The linkage region tetrasaccharide GlcUA-Gal-Gal-Xyl structure was presumed to exist but has now been isolated for

---

**TABLE I**

**GAG-Protein Linkage Region of D. melanogaster and C. elegans**

The 2AB-derivatized oligosaccharide fraction prepared from C. elegans chondroitin by chondroitinase ABC digestion was subfractionated on gel filtration HPLC as shown in Fig. 3, and its DE MALDI-TOF/MS was recorded in the negative ion mode with 2,5-dihydroxybenzoic acid as the matrix.

FIG. 4. DE MALDI-TOF/MS of the 2AB-derivatized oligosaccharide from the C. elegans chondroitin. The 2AB-derivatized oligosaccharide fraction prepared from C. elegans chondroitin by chondroitinase ABC digestion was subfractionated on gel filtration HPLC as shown in Fig. 3, and its DE MALDI-TOF/MS was recorded in the negative ion mode with 2,5-dihydroxybenzoic acid as the matrix.
the first time as discrete structures from the *C. elegans* chondroitin and *D. melanogaster* HS and CS. Although it was not detected for *C. elegans* HS, this was most likely the result of the limited amount of HS. This is the first study that demonstrates the existence of the linkage region tetrasaccharide in invertebrates. Despite a wide distribution of GAGs in different classes of invertebrates (for a review, see Ref. 51), it has not been clarified how these GAGs are attached to the presumable core proteins. Some bacterial strains, *Escherichia coli* K4 and K5 and *Streptococcus pyrogenes*, synthesize polysaccharides with the same backbone structures as CS, heparin/HS, and hyaluronic acid, respectively (52–54). These GAG-like polymers are produced as extracellular polysaccharide capsules that serve as virulence factors. These polysaccharides are not attached to core proteins and are synthesized on the inner surface of the cytoplasmic membrane (for a review, see Ref. 55). Studies on the GAG-protein linkage region of invertebrate GAGs are limited, and the reducing end structure of the invertebrate GAG chains has not been characterized. It has now been clarified that at least the conventional linkage region tetrasaccharide GlcUA-Gal-Gal-Xyl of proteoglycans has been conserved through evolution, although other linkage region structures have not been excluded. It has been suggested that unusual N-linked HS and CS chains, which are released by peptide-N-glycosidase digestion, are produced by some cultured cell lines such as bovine pulmonary arterial endothelial (CPAE) and human erythroleukemia (K562) cell lines (24) as well as by bovine lung tissues (25). Such possibilities were not investigated rigorously in this study and remain to be explored. The possibility that chondroitin chains are attached to core proteins via O-linked GalNAc residues (56) will be discussed below.

In this study, phosphorylation of the Xyl residue was demonstrated in the linkage region of both CS and HS from *D. melanogaster*. Although only the phosphorylated structure in the former, and both phosphorylated and nonphosphorylated forms in the latter, were detected, the existence of the phosphorylated Xyl in both HS and CS chains is consistent with the previous findings (23, 34). Recent simultaneous analysis of both CS and HS chains of the hybrid type proteoglycan syndecan-1 molecule has suggested the Xyl phosphorylation of both CS and HS chains (34). Structural analysis of the biosynthetic intermediate oligosaccharides formed in the cultured rat fibroblasts has suggested that dephosphorylation takes place soon after the transfer reaction of GlcUA to the trisaccharide (57–59). In *vitro* enzyme assays using crystallized human glucuronyltransferase I as enzyme (60) and Gal-Gal-Xyl(2-O-phosphate)-Ser as acceptor have demonstrated the higher enzyme activity with the phosphorylated acceptor substrate.3 Together, these results suggest that the Xyl phosphorylation may play an important role for the transfer reaction of the first GlcUA residue to the linkage region trisaccharide, and may be required for the efficient maturation of the linkage region tetrasaccharide.

3 Y. Tone, L. Pedersen, H. Kitagawa, T. Yamamoto, J. Nishihara, J. Tamura, T. A. Darden, M. Negishi, and K. Sugahara, manuscript in preparation.

![Figure 5. Molecular size analysis of the *D. melanogaster* CS and HS as well as the *C. elegans* chondroitin by gel filtration chromatography.](image-url)
rasaccharide as prerequisites for the assembly of GAG chains.

In strong contrast, the phosphorylated Xyl was not detected in the linkage region of C. elegans chondroitin in the present study. This may suggest that the glycosytransferases involved in the synthesis of the C. elegans chondroitin have different specificities from those in higher organisms and do not require the phosphate group on the Xyl. In fact, in the C. elegans genome there is no ortholog of the recently cloned human N-acetylglactosaminyltransferase that transfers the first GalNAc to the linkage region tetrasaccharide (61). Instead, the C. elegans ortholog of human chondroitin synthase (62) may possess this enzyme activity. Thus, although C. elegans chondroitin is formed on the conventional linkage region tetrasaccharide structure, the biosynthetic mechanism appears to be somewhat different from those of D. melanogaster or higher organisms including humans. In the present study no sulfation of the Gal residues was observed in the linkage region of either C. elegans chondroitin or D. melanogaster CS, although it has been reported for the linkage region of CS and DS chains from various vertebrate tissues (23, 34). No sulfation of the Gal residues in the linkage region of C. elegans chondroitin may be related to the finding that chondroitin is not sulfated in this organism. No homologs of human CS/DS 4-O-sulfotransferases (63, 64) are found in the C. elegans genome. In Drosophila, for the synthesis of chondroitin 4-sulfate, the 4-O-sulfation of the Gal residue in the linkage region may not be required as a prerequisite.

Although low sulfated CS has been found in various animal tissues such as human inter-α-trypsin inhibitor (65), CS from human placenta (66), mollusc (67), and human cornea (68), reports on completely nonsulfated chondroitin chains are limited to polysaccharides from squid skin (69–71). Although the corresponding acidic polysaccharides isolated from bovine cornea in the pioneering work by Davidson and Meyer (72) was named chondroitin, it contained reportedly a small proportion (2.1%) of sulfate groups. Despite the high chondroitin content in the nematode, it contained reportedly a small proportion of the structures of the C. elegans chondroitin in the linkage region tetrasaccharide standards are shown in panel A: 1, ΔHexUAα1–3Galβ1–3Galβ1–4Xyl1–2AB (12.2 pmol); 2, ΔHexUAα1–3Galαβ3Galβ1–3Galβ1–4Xyl1–2AB (9.3 pmol); 3, ΔHexUAα1–3Galβ1–3Galβ1–4Xyl1–2AB (5.7 pmol). The fluorescence intensity of the chromatograms in panels A and B has been attenuated by a factor of 4 compared with that in panel C.

**FIG. 6.** HPLC analysis of the linkage oligosaccharide fraction prepared by enzymatic digestion of the D. melanogaster CS and HS. The 2AB-derivatized GAG preparation from adult flies was digested with chondroitinases ABC and then AC-II (panel B) or a mixture of heparinase and heparitinase (panel C). Each digest was analyzed by HPLC on an amine-bound silica PA-03 column using a linear gradient of NaH2PO4 as indicated by the dashed lines. The elution positions of authentic 2AB-derivatized linkage tetrasaccharide standards are shown in panel A: 1, ΔHexUAα1–3Galβ1–3Galβ1–4Xyl1–2AB (12.2 pmol); 2, ΔHexUAα1–3Galαβ3Galβ1–3Galβ1–4Xyl1–2AB (9.3 pmol); 3, ΔHexUAα1–3Galβ1–3Galβ1–4Xyl1–2AB (5.7 pmol). The fluorescence intensity of the chromatograms in panels A and B has been attenuated by a factor of 4 compared with that in panel C.
were standard mild conditions (1 M NaBH₄, 100 mM NaOH), and hence decomposition of chondroitin chains by peeling reactions or random cleavage was unlikely. In addition, the large amounts of the isolated oligosaccharides appear to indicate that they were released directly from the core peptides. In the present study, such small truncated oligosaccharides, if any, would have been resistant to the LiOH treatment as confirmed by simultaneously treating desialylated bovine submaxillary mucin and a standard linkage region hexasaccharide peptide preparation followed by HPLC analysis. Only the linkage region hexasaccharides but no mucin-derived GalNAc (37) were detected after 2AB-derivationization (data not shown), suggesting that the GaINAcO1-0-Ser/Thr linkage is resistant to the mild alkaline treatment with LiOH (see “Experimental Procedures”). The unsaturated hexasaccharide, \( \text{HexA} \text{GaINAc-GlcUA} \text{Gal-Xyl-2AB} \), was the only linkage region oligosaccharide obtained in this study after exhaustive chondroitinase ABC digestion of the *C. elegans* chondroitin polysaccharide (Figs. 2 and 3), and neither \( \text{HexUA} \text{GaINAc-GlcUA-Gal-Xyl-2AB} \) nor \( \text{HexUA} \text{GaINAc-Gal-Xyl-2AB} \) was discerned. The failure to detect even GalNAc-GlcA-Gal-Gal-Xyl-2AB (56) appears to suggest that such small truncated oligosaccharides might have escaped at the step of the gel filtration column chromatography on Sephadex G-50 after Actinase E digestion (see “Experimental Procedures”). The present findings may suggest that polymer chondroitin chains are polymerized only on the conventional common linkage region tetrasaccharide structure by putative chondroitin synthsase but not on the short oligosaccharides, with a reducing terminal GaINAc residue but without the conventional linkage region, which are likely to build on the core proteins but may not be elongated to a large extent. Identification of presumable glycosyltransferases responsible for the synthesis of these oligosaccharides and isolation of glycopeptides derived from the putative chondroitin oligosaccharides with a reducing terminal GaINAc residue would give insights into the mechanism of evolution of polymer chondroitin in *C. elegans*. Dally and Dly, the orthologous proteins of vertebrate glypicans, in addition to *Drosophila* sydnean (18), have been found as the major HS-PGs in *D. melanogaster* or HS- and chondroitin-PGs in *C. elegans* have been identified and the *Drosophila* sydnean core protein does not bear CS chains (18), unlike mammalian sydnean-1 or -4 (75, 76). Although *C. elegans* unc-2, the ortholog of the perlecan HS-PG gene, has been shown to affect muscle attachment and sarcomere organization (45), little of the biochemical nature is known for the GAG chains of this PG. Recently, Bulik et al. (47) reported that there are numerous chondroitin-modified PGs in *C. elegans* by Western blotting analysis using anti-\( \text{HexUA} \text{GaINAc} \) antibody after chondroitinase ABC digestion. However, because short chondroitin-like oligosaccharides, GlcUA\( \beta_1-3 \)GaINAc\( \beta_1-4 \)GlcUA\( \beta_1-3 \)GaINAc\( \beta_1-4 \)GlcUA\( \beta_1-3 \)GaINAc, identified in *C. elegans* would also form the \( \text{HexUA} \text{GaINAc} \)-structure on chondroitinase ABC digestion, the heterogeneity of the core protein of chondroitin-PG may be the result of the heterogeneous O-glycosylation by such short chondroitin-like oligosaccharides. The progress of the *C. elegans* genome project has revealed the genes homologous to those of the core proteins of vertebrate CS-PGs such as bamacan, aggrecan, decorin, and phosphacan. Further studies are needed for identification of the core protein(s) of the *C. elegans* chondroitin oligo- and polysaccharides to gain insights into the biological functions of *C. elegans* chondroitin oligo- and polysaccharides.
52. Kendall, F. E., Heidelberger, M., and Dawson, M. H. (1937) *J. Biol. Chem.* 118, 61–69
53. Vann, W. F., Schmidt, M. A., Jann, B., and Jann, K. (1981) *Eur. J. Biochem.* 116, 359–364
54. Rodriguez, M. L., Jann, B., and Jann, K. (1988) *Eur. J. Biochem.* 177, 117–124
55. DeAngelis, P. L. (2002) *Glycobiology* 12, 9R–16R
56. Guejar, Y., Balanzino, L., Maes, E., Leroy, Y., Coddeville, B., Oriol, R., and Strecker, G. (2001) *Biochem. J.* 357, 167–182
57. Moses, J., Oldberg, Å., Cheng, F., and Fransson, L.-Å. (1997) *Eur. J. Biochem.* 248, 521–526
58. Moses, J., Oldberg, Å., Eklund, E., and Fransson, L.-Å. (1999) *Eur. J. Biochem.* 260, 879–884
59. Pedersen, L. C., Tsuchida, K., Kitagawa, H., Sugahara, K., Darden, T. A., and Negishi, M. (2000) *J. Biol. Chem.* 275, 34580–34585
60. Kitagawa, H., Uyama, T., and Sugahara, K. (2001) *J. Biol. Chem.* 276, 38721–38726
61. Yamauchi, S., Mita, S., Matsubara, T., Fukuta, M., Habuchi, H., Kimata, K., and Habuchi, O. (2000) *J. Biol. Chem.* 275, 8975–8981
62. Hiraoa, N., Nakagawa, H., Ong, E., Akama, T. O., Fukuda, M. N., and Fukuda, M. (2000) *J. Biol. Chem.* 275, 20188–20196
63. Yamada, S., Oyama, M., Kinugasa, H., Nakagawa, T., Kawasaki, T., Nagasawa, S., Khoo, K.-H., Morris, H. R., Dell, A., and Sugahara, K. (1995) *Glycobiology* 5, 335–341
64. Achur, R. N., Valiyaveettil, M., Alkhalil, A., Ockenhouse, C. F., and Gowda, D. C. (2000) *J. Biol. Chem.* 275, 40344–40356
65. Nader, H. B., Ferreira, T. M., Paiva, J. F., Medeiros, M. G., Jeronimo, S. M., Paiva, V. M., and Dietrich, C. P. (1984) *J. Biol. Chem.* 259, 1431–1435
66. Plaa, A. H., West, L. A., Thonar, E. J., Karig, Z. A., Smith, C. J., Klintworth, G. K., and Hascall, V. C. (2001) *J. Biol. Chem.* 276, 39788–39796
67. Anno, K., Kawai, Y., and Seno, N. (1964) *Biochim. Biophys. Acta* 83, 348–359
68. Karamanos, K. K., Aletras, A. J., Antonopoulos, C. A., Tsigas, T., Tsiganos, C. P., and Vynios, D. H. (1988) *Biochim. Biophys. Acta* 966, 36–43
69. Karamanos, K. K., Aletras, A. J., Antonopoulos, C. A., Hjerpe, A., and Tsiganos, C. (1990) *Eur. J. Biochem.* 192, 33–38
70. Davidson, E. A., and Meyer, K. (1954) *J. Biol. Chem.* 211, 605–611
71. Nakato, H., Futch, T. A., and Selleck, S. B. (1995) *Development* 121, 3687–3702
72. Baeg, G. H., Lin, X., Khare, N., Baumgartner, S., and Perrimon, N. (2001) *Development* 128, 87–94
73. Rapraeger, A., Jalkanen, M., Endo, E., Koda, J., and Bernfield, M. (1985) *J. Biol. Chem.* 260, 11046–11052
74. Shworak, N. W., Shirakawa, M., Mulligan, R. C., and Rosenberg, R. D. (1994) *J. Biol. Chem.* 269, 21204–21214