

Structural Analysis of Cartilage Proteoglycans and Glycoproteins Using Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry

Joseph Zaia, Beishan Liu, Raymond Boynton, and Frank Barry

Osiris Therapeutics, Inc., Baltimore, Maryland 21231

Received July 29, 1999

Cartilage extracellular matrix molecules synthesized and maintained by chondrocytes form a strong, elastic tissue functioning to cushion and protect the subchondral bone. Osteoarthritis is characterized by degradation of cartilage extracellular matrix molecules resulting in fibrillation, irreversible erosion, and eventual failure of the tissue. With recent interest in the degradation of cartilage extracellular matrix molecules, a need for more detailed structural information exists. Posttranslational modifications are believed to play a role in determining the susceptibility of these molecules to proteolytic degradation during the development of osteoarthritis. The purpose of this paper is to show how the application of matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry to extracellular matrix protein and proteoglycan structure will help elucidate problems in extracellular matrix biochemistry. Methodological issues relating to the high molecular weight, polydispersity, and high degree of posttranslational modification of these molecules are discussed. MALDI-TOF mass spectrometry provides an improved level of detail for extracellular matrix protein and proteoglycan structure and is useful in addressing issues surrounding the causes of degradation during osteoarthritis. © 2000 Academic Press

Key Words: aggrecan; biglycan; cartilage; COMP; decorin; glycosaminoglycan; MALDI-TOF; mass spectrometry; osteoarthritis; proteoglycan.

The normal functioning of articular cartilage depends on the synthesis and assembly of an integrated extracellular matrix (ECM).¹ Since the tissue is sub-

jected to large compressive and shear forces, the maintenance of integrity appears to depend on a variety of specific and organized interactions between components of the ECM and between chondrocytes and the ECM. A network, consisting primarily of collagen type II fibrils, whose assembly and structure are modulated by collagen types IX and XI (1), provides tensile strength. Aggrecan, in association with link protein and hyaluronan, provides swelling capacity which, when countered by the inelastic collagen network, results in a strong, elastic tissue (2, 3). Although aggrecan, with a molecular mass of 2.5 MDa (4) is the major proteoglycan by weight in cartilage, other proteoglycans are present in equimolar amounts (5). Decorin, biglycan, and fibromodulin are members of the family of small leucine-rich proteoglycans (6) which bind to collagen and regulate fibril diameter (7–10) and have been shown to bind other connective tissue macromolecules (11) and growth factors (12, 13). Type VI collagen is found specifically in the chondrocyte pericellular matrix and appears to play a role in the microenvironment surrounding the cells (14).

Homeostasis in articular cartilage depends on signals received by the chondrocytes in contact with the ECM. For instance, hyaluronan is anchored to the chondrocytes by the CD44 receptor (15) via a 100-amino-acid N-terminal domain with homology to link protein and to the aggrecan G1 domain (16–18). Each hyaluronan chain binds several aggrecan-link protein complexes and is known to associate with the collagen network (3). Collagens are also bound by anchorin CII (19) and chondronectin (20) associated with the chondrocyte membrane. While the mechanisms of signal

hexosamine; MALDI-TOF, matrix-assisted laser desorption/ionization time-of-flight; MS, mass spectrometry; *m/z* mass-to-charge ratio; NeuAc, *N*-acetylneuraminic acid; SEC, sedimentation equilibrium centrifugation; SLRP, small leucine-rich proteoglycan.

¹ Abbreviations used: CHCA, α -cyano-4-hydroxycinnamic acid; ECM, extracellular matrix; COMP, cartilage oligomeric matrix protein; GlcNAc, *N*-acetylglucosamine; Hex, hexose; HexNAc, *N*-acetyl-

transduction remain unclear, the interactions between ECM molecules and chondrocytes orchestrate tissue homeostasis. Maintenance of the ECM requires a relatively high rate of aggrecan turnover (21), and this involves the metabolically active chondrocytes. Link protein, which stabilizes the interaction between aggrecan and hyaluronan (22), is turned over at a lower rate in normal cartilage (23) and there is little remodeling of the collagen network (24, 25).

Osteoarthritis is a chronic degradative joint disease with a very high prevalence in aging populations. The disease is characterized by an initial fibrillation of the articular cartilage followed by gradual erosion in an irreversible manner, leading eventually to exposure of the subchondral bone. Cartilage tissue taken during joint replacement shows an enhanced rate of synthesis of collagen and proteoglycans in comparison to normal cartilage (26, 27), but with a concomitant net loss of proteoglycans from the matrix (28). These losses are observed earliest in the superficial zones of the cartilage, and contribute in a major way to the destruction of the tissue. Osteoarthritic cartilage resembles fetal (29) and osteophytic (30) cartilage in that type II collagen is actively synthesized and it appears that the chondrocytes receive similar signals from the ECM, at least with respect to synthesis of matrix components. Unlike fetal cartilage, however, the newly formed ECM components do not assemble into functional tissue in osteoarthritis.

Cleavage of the collagen triple helix and subsequent disruption of the fibril structure results in loss of aggrecan complexes from the tissue (31, 32). Cleavage of the aggrecan backbone itself leads to loss of part or all of the chondroitin sulfate domain, with consequent loss of swelling pressure and increase in load on the collagen network (33).

The purpose of this paper is to describe the application of matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry (34) to the analysis of ECM components. The methods described here allow an increased level of structural detail of fragments arising from matrix turnover. MALDI-TOF complements techniques based on neopeptide-specific antibodies, protein sequencing, amino acid analysis, and monosaccharide analysis and allows a higher level of sensitivity and accuracy.

The MALDI-TOF spectrum of a protein consists usually of singly and doubly charged molecular ions and may contain ions corresponding in mass-to-charge ratio (m/z) to protein multimers. The singly charged molecular ion is often the most abundant ion in the spectrum (an exception is discussed under Results). The molecular weight of a protein can be measured as a function of enzyme digestion or other perturbation. The technique is useful in profiling ECM proteins (35). The complexity of ECM proteins is a result of their inherent polydispersity, large size, and high degree of posttranslational modification, and as such great care

is required in their analysis. Furthermore, some ECM molecules form stable aggregates and special sample handling techniques are required.

For the purposes of this study we have selected a panel of ECM molecules that are characteristic of or specific to articular cartilage, and we have studied these using MALDI-TOF mass spectrometry. Data are presented from the analysis of cartilage oligomeric matrix protein (COMP), a pentameric member of the thrombospondin family present in articular cartilage. Intact COMP is a rather large molecule with a mass of 420,000 Da. Despite its large size MALDI-TOF spectra can be obtained using intact COMP and valuable information regarding structure can be elucidated after reduction or after treatment with glycosidases. The proteoglycans aggrecan, decorin, and biglycan have been analyzed and the specific experimental approach and the information obtained differ in each case. In the case of aggrecan we have studied the G1 domain tryptic fragment. This fragment has a tendency to associate to form multimers. This behavior impacts analysis by MALDI-TOF MS and we discuss treatment with formic acid as a means of reducing the degree of aggregation. Sequential specific deglycosylation with the generation of intermediate spectra is a very powerful approach for determining the nature and extent of glycosylations. When used in combination with glycopeptide analysis, a great deal of information may be obtained. Both decorin and biglycan are substituted with chondroitin/dermatan sulfate, a modification that represents a special problem with respect to ionization and mass analysis. These are also discussed.

EXPERIMENTAL

Keratanase and keratanase II were purchased from Seikagaku America, Inc. (Ijamsville, MD). *N*-Glycanase, neuraminidase and *O*-glycanase were purchased from Genzyme Diagnostics (Cambridge, MA). Sequencing grade trypsin was purchased from Boehringer-Mannheim, Inc. (Indianapolis, IN); sinapinic acid solution and α -cyano-4-hydroxycinnamic acid (CHCA) solution were from Hewlett-Packard Co. (Palo Alto, CA); and formic acid was from Sigma Chemicals, Inc. (St. Louis, MO).

MALDI-TOF mass spectra were acquired using a Hewlett-Packard G2025A mass spectrometer. Protein samples were prepared for MALDI-TOF mass spectrometry by mixing the sample solution with sinapinic acid solution to achieve a protein concentration of 1–10 pmol/ μ l. An aliquot (0.8 μ l) of this solution was dried on a metal target, forming cocrystals of matrix and protein. Salts were washed away, by dipping the target into a solution of 0.1% TFA. Details in sample preparation are given below. Spectra of proteins were calibrated externally using a mixture containing bovine serum albumin (monomer and dimer), equine myoglo-

bin, and equine cytochrome *c* (Hewlett–Packard). Peptides were prepared in a similar manner substituting CHCA as the matrix and using a standard peptide mix (Hewlett–Packard) for external calibration.

COMP was purified as previously described from mature bovine articular cartilage obtained from metacarpal phalangeal joints (36). The tryptic G1 domain of aggrecan was purified from pig laryngeal cartilage as previously described (37).

Aggrecan tryptic G1 (100 μ g) was reduced with dithiothreitol and alkylated with iodoacetic acid as described previously (37). An aliquot (0.5 μ l) was removed after each enzyme digestion step and mixed with sinapinic acid solution (2.5 μ l). An aliquot (0.8 μ l) of this mixture was dried on the MALDI target, washed by immersing in 0.1% TFA, and dried again. An aliquot (0.8 μ l) of a formic acid solution (100:100:35 water:acetonitrile:formic acid) was placed on top of the crystals and allowed to dry. Reduced and alkylated aggrecan G1 (10 μ g in 10 μ l 0.1 M Tris–HCl, pH 7.4) was digested with keratanase (2 mU) and keratanase II (2 mU) for 15 h at 37°C. The protein was then digested with *N*-glycanase (300 mU) for 15 h at 37°C followed by neuraminidase (12.4 mU) for 2 h at 37°C and *O*-glycanase (300 μ U) for 15 h at 37°C.

Aggrecan G1 (90 μ g) was dissolved in 0.1 M Tris–HCl, pH 8.5, and digested with trypsin (0.9 μ g) for 6 h at 37°C. The solution was acidified with TFA and the tryptic peptides were fractionated using a reversed-phase HPLC column (Vydac, Hesperia, CA, C18, 2.1 mm $\times$ 250 mm protein column) with a gradient of 0–55% B in 50 min (A = H₂O, 0.1% TFA, B = 70% acetonitrile, 0.8% TFA, 350 μ l/min) using a Hewlett–Packard 1090 liquid chromatograph. Fractions were dried using a centrifugal vacuum concentrator and dissolved in 0.1% TFA. MALDI-TOF spectra were acquired using 1% of the peptide volume diluted in CHCA solution. Aliquots of fractions of each fraction were digested with *N*-glycanase and MALDI-TOF mass spectra were acquired.

Decorin and biglycan from bovine skin (51, 53) were gifts of L. Rosenberg (Orthopaedic Research Laboratory, Montefiore Medical Center, Bronx, NY).

Protein sequencing was carried out using a Hewlett–Packard G1000A protein sequencer using routine 3.0 methods.

RESULTS

COMP

Figure 1a shows the MALDI-TOF mass spectrum of pentameric bovine COMP with a mass of $413,200 \pm 1100$ Da. The peak at m/z $208,500 \pm 400$ corresponds to the pentamer with two positive charges, at half the m/z of the singly charged ion. The other peaks correspond to COMP pentamer with three to five positive charges, as noted in the figure. Reducing the disulfide bonds

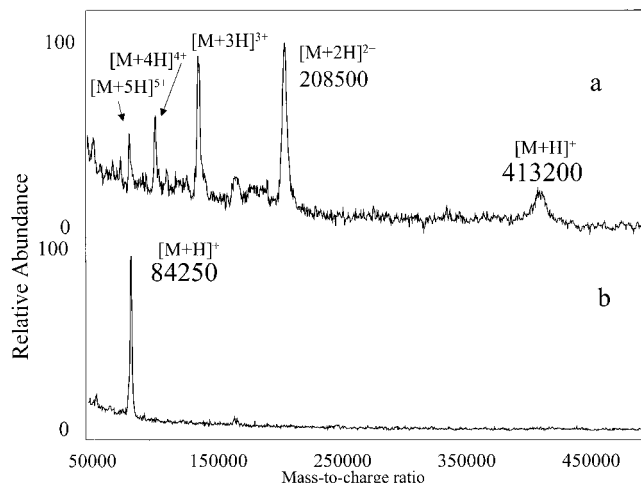


FIG. 1. MALDI-TOF mass spectra of (a) bovine COMP and (b) bovine COMP reduced with triethylphosphine.

with triethylphosphine produces an abundant ion at m/z $84,250 \pm 30$ (Fig. 1b), consistent with monomeric COMP, and a very low abundance peak at m/z $168,000 \pm 400$ corresponding to the protein dimer. Observation of a protein dimer, presumed to be a non-covalent complex of two protein molecules with one proton charge, is typical of MALDI-TOF spectra of intact proteins. Both spectra were calibrated externally using a mixture of bovine serum albumin (monomer and dimer), equine myoglobin, and equine cytochrome *c*. The m/z values and standard errors reported were taken from three replicate spectra. The observed m/z value divided by five gives a calculated monomer m/z of $82,640 \pm 230$, a 2% deviation of the m/z value observed in Fig. 1b. The difference in mass is due to the low signal-to-noise ratio observed for the pentameric molecule in combination with the necessity of extrapolating the calibration curve in Fig. 1a. The data show that useful mass measurement can be obtained for ECM proteins above 400 kDa in mass.

Aggrecan

Aggrecan monomer may be treated with trypsin to produce a fragment containing amino acids 1–367 or 1–375 (37) that includes G1 and part of the interglobular domain. This G1 fragment was purified from pig laryngeal cartilage and analyzed by MALDI-TOF using SA as the matrix (37). The sample produces a series of peaks (Fig. 2a), the masses of which correspond to G1 monomer through pentamer as indicated. This spectrum is unusual compared to a typical protein in this mass range due to the high abundance of peaks corresponding to G1 multimers relative to that of the monomer. The abundance of the multimer ions may be caused by association between G1 molecules that are stable in the MALDI matrix solution (100 mM solution

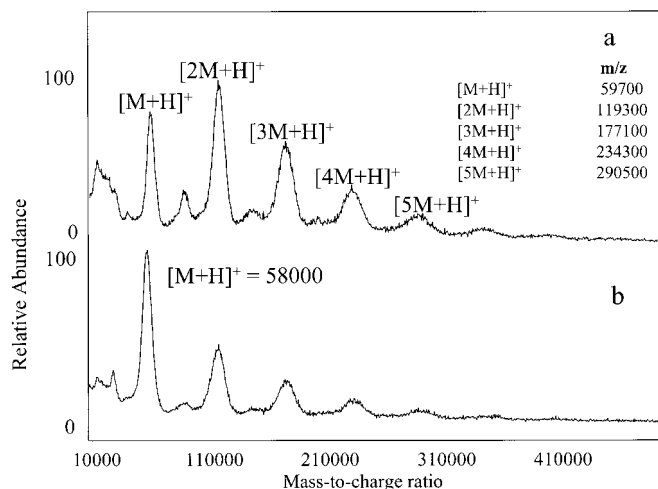


FIG. 2. MALDI-TOF mass spectra of (a) the G1 domain fragment of porcine aggrecan, generated by digestion with trypsin and (b) the same sample after treatment with formic acid.

of SA in acetonitrile/methanol/water, pH 3.70). Treatment of the dried sample–matrix crystals with formic acid (Fig. 2b) changes the pattern such that the monomer is the most abundant peak, as is observed for most proteins. This suggests that lowering the pH breaks the association between G1 monomers. Since the acidified sample produces one major ion, corresponding to the singly charged protein monomer, the signal-to-noise ratio is improved and interpretation is more straightforward. Accordingly, this treatment was used for determination of mass as a function of enzyme treatment. Figure 3 shows spectra obtained on G1 after sequential treatment with keratanase/keratanase II, *N*-glycanase, neuraminidase, and *O*-glycanase. Masses obtained are given in Table 1. The data show that treatment with all four enzymes (lower trace) resulted in a mass of $46,866 \pm 55$, with a total change relative to the untreated sample of $11,138 \pm 326$ Da, corresponding to the total carbohydrate. The mass removed by keratanase treatment accounts for $76 \pm 6\%$ (8425 ± 368 Da) of the carbohydrate and subsequent treatment with *N*-glycanase removed glycans accounting for $20 \pm 2\%$ (2225 ± 145 Da) of the glycan mass. This analysis provides a useful estimation of the glycosylation of the core protein.

To analyze individual glycans it is necessary to digest G1 with a proteolytic enzyme and to fractionate using chromatography. MALDI-TOF mass spectrometry can be used to identify sites of glycosylation and to determine the mass and heterogeneity of glycans. This approach was taken with the pig aggrecan G1 sample toward the end of characterizing the N-linked glycan heterogeneity, particularly in the vicinity of the stromelysin and aggrecanase cleavage sites (Asn³⁴¹-Phe³⁴² and Glu³⁷³-Ala³⁷⁴, respectively). G1 was reduced, alkylated, digested with trypsin, and fractionated by

HPLC and MALDI-TOF spectra were acquired on all fractions. Fraction 9, eluting at 17.2 min, displayed a cluster of peaks between m/z 3100 and 5000, in addition to an ion at m/z 2257.6 ± 1.1 (Fig. 4a). Treatment with *N*-glycanase (Fig. 4b) caused the spectrum to collapse to a single peak at m/z 1907.5 ± 0.5 , consistent with protonated pig G1 peptide 311–327 (TVYLHAN-QTGYPDPSSR, calculated m/z 1907.05). Automated Edman degradation determined a sequence of TVYL-HAXQTG, confirming the assignment and indicating a glycosylated Asn residue. The m/z values of the labeled peaks in Fig. 4a can be accounted for by adding the mass of the N-linked core structure (GlcNAc)₂(Man)₃ (calculated residue mass 892.82) to the mass of the deglycosylated peptide and then by adding masses corresponding to HexNAc, Hex, or NeuAc. The compositions calculated for each ion are shown in Table 2. The calculated protonated masses of the glycopeptides listed fall within 0.1% of the corresponding observed m/z value.

The ion at m/z 2257.6 ± 1.1 is consistent with modification of the peptide with a fucose and an *N*-acetylhexosamine residue. Given that the N-linked core structure is often fucosylated at the innermost *N*-acetylglucosamine residue, this observation indicates that the glycan has been truncated by an enzyme whose activity is similar to that of endoglycosidase H. The ion at m/z 3149.0 ± 1.2 corresponds to the addition of (HexNAc)₂(Hex)₃ (calculated mass 892.82 Da) to the ion at m/z 2257.6 ± 1.1 , indicating that the glycopeptide producing this ion contains the fucosylated N-linked core structure with the addition of an *N*-acetylhexosamine residue to one of the mannose branches. This structure matches one commonly found during

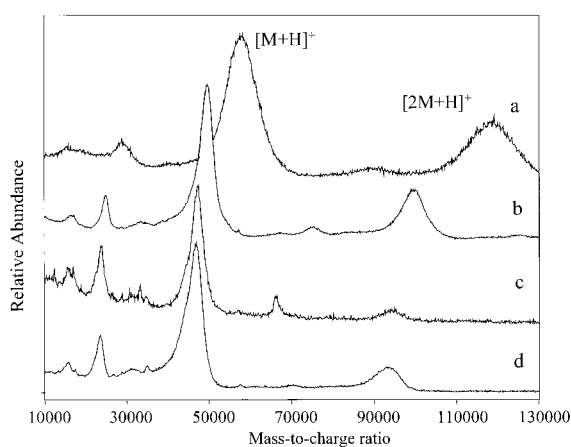


FIG. 3. Measurement of glycan masses of the tryptic G1 fragment of porcine aggrecan determined by reduction in mass from treatment with specific glycosidases. Spectra corresponding to (a) the untreated G1 fragment and (b) after digestion with keratanase/keratanase II, (c) keratanase/keratanase II and *N*-glycanase, and (d) keratanase/keratanase II, *N*-glycanase, neuraminidase, and *O*-glycanase. The masses for each spectrum are shown in Table 1.

TABLE 1

Measured Masses of Glycan Substituents on the Tryptic G1 Fragment Derived from Pig Aggrecan Obtained by Sequential Digestion with Specific Glycosidases (See Fig. 3)

Glycosidase treatment	Observed m/z	Glycan type	Glycan mass
Untreated	$58,004 \pm 271$		
Keratanase/keratanase II	$49,579 \pm 97$	Keratan sulfate	8425 ± 368
Keratanase/keratanase II/ <i>N</i> -glycanase	$47,354 \pm 48$	N-linked oligosaccharide	2225 ± 145
Keratanase/keratanase II/ <i>N</i> -glycanase/neuraminidase (spectrum not shown)	$47,169 \pm 26$	O-linked sialic acid	185 ± 74
Keratanase/keratanase II/ <i>N</i> -glycanase/neuraminidase/ <i>O</i> -glycanase	$46,866 \pm 55$	O-linked oligosaccharide	303 ± 81

the processing of asparagine-linked oligosaccharides in the medial Golgi cisternae after removal of mannose residues by Golgi α -mannosidase II (38). Since the glycan compositions in Table 2 indicate that the structure is of the complex type, no mannose residues beyond the core structure are likely to be present. The ion at m/z 3311.0 ± 0.9 therefore corresponds to the addition of a hexose residue (calculated mass 162.14 Da) and it is reasonable to assume that it is added to the outer chain *N*-acetylglucosamine residue to form a lactosamine disaccharide (Gal β 1-4GlcNAc β 1-3, calculated mass 365.34 Da) as is common in complex N-linked structures. The addition of a second *N*-acetylglucosamine (calculated mass 203.20 Da) and a second hexose are indicated by ions at m/z 3515.1 ± 1.2 and 3676.4 ± 0.6 , respectively, possibly forming a second branch to the complex structure. The ion at m/z 4042.1 ± 0.8 is consistent with the addition of a third lactosamine disaccharide and that at m/z 4332.0 ± 0.6 with the addition of an *N*-acetylneuraminic acid residue. A fourth lactosamine disaccharide addition is indicated by the ion at m/z 4697.4 ± 0.5 and a second *N*-acetylneuraminic acid is indicated by the ion at m/z 4988.2.

Decorin and Biglycan

Figure 5a shows the MALDI-TOF spectrum of bovine decorin collected using sinapinic acid as the matrix. The peak centered at m/z 58,000 is broad and a poor signal-to-noise ratio is evident. The mass is somewhat lower than the 70 kDa expected for decorin (51, 52). Figure 5b shows the positive ionization mode MALDI-TOF spectrum of bovine biglycan. The broad peak is centered at m/z 70,000, significantly higher than that observed for decorin in Fig. 5a and below the sedimentation equilibrium centrifugation (SEC) measurement for biglycan (93–99 kDa) (49). Figure 5c shows the spectrum for biglycan after treatment with chondroitinase ABC with a mass of 42,556. The signal-to-noise ratio is greatly enhanced by removal of the chondroitin/dermatan sulfate chains.

DISCUSSION

COMP is a disulfide-linked pentameric glycoprotein with 84-kDa subunits (cysteines not alkylated) (35). Found in tracheal, nasal, and articular cartilage (39), this molecule is believed to play a role in the assembly of the ECM. Bovine fetal and adult forms are known to differ in the extent of glycosylation (35). Mutation in COMP gives rise to cartilage abnormalities including pseudoachondroplasia and multiple epiphyseal dysplasia (40).

MALDI-TOF MS can be used to produce useful mass information on ECM molecules such as COMP in excess of 400 kDa as shown in Fig. 1. The data may be produced on a small amount of material (1–10 pmol) and may be of value in studying alternative splicing of chondrocyte genes or the formation of covalent complexes. A better signal-to-noise ratio is produced on the reduced COMP monomer and improved accuracy and precision result.

Aggrecan consists of a 200–250 kDa core protein with three globular domains, G1, G2, and G3. G1, the N-terminal domain, and link protein are homologous and bind to hyaluronan. A 140-amino-acid interglobular domain separates G1 and G2 and a large extended region separates G2 from the C-terminal G3 domain.

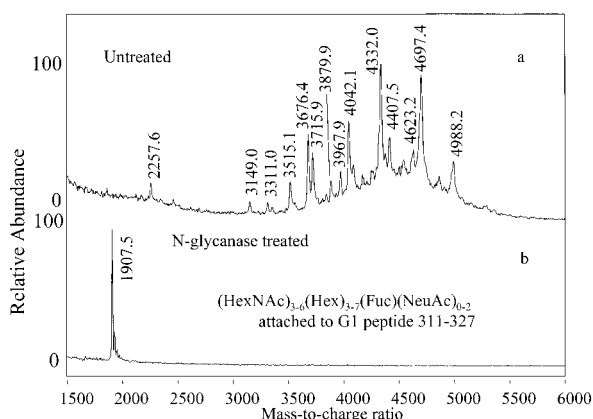
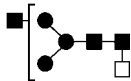
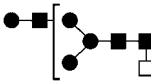
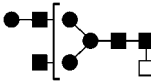
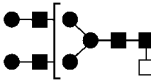
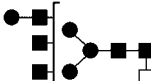
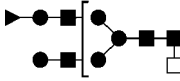
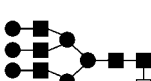
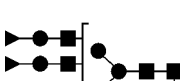


FIG. 4. MALDI-TOF spectra of a tryptic peptide generated from reduced and carboxymethylated G1 fragment from porcine aggrecan before (a) and after (b) treatment with *N*-glycanase. The masses labeled indicate that the peptide is variably modified with the oligosaccharides whose composition is shown in Table 2.

TABLE 2
Observed and Calculated Masses and Calculated Glycan Composition for Porcine Aggrecan Tryptic Glycopeptide 311–327, Spectra Shown in Fig. 4

Observed m/z	Calculated m/z	Glycan composition	Structure ^a
1907.5 $\pm$ 0.5 2257.6 $\pm$ 1.1	1907.05 2256.39	De-N-glycosylated peptide (HexNAc)(Fuc)	
3149.0 $\pm$ 1.2	3149.21	(HexNAc) ₃ (Hex) ₃ (Fuc)	
3311.0 $\pm$ 0.9	3311.35	(HexNAc) ₃ (Hex) ₄ (Fuc)	
3515.1 $\pm$ 1.2	3514.55	(HexNAc) ₄ (Hex) ₄ (Fuc)	
3676.4 $\pm$ 0.6	3676.69	(HexNAc) ₄ (Hex) ₅ (Fuc)	
3715.9 $\pm$ 0.8	3717.74	(HexNAc) ₅ (Hex) ₄ (Fuc)	
3879.9 $\pm$ 1.9	3879.88	(HexNAc) ₅ (Hex) ₅ (Fuc)	
3967.9 $\pm$ 2.0	3967.94	(HexNAc) ₄ (Hex) ₅ (Fuc)(NeuAc)	
4042.1 $\pm$ 0.8	4042.03	(HexNAc) ₅ (Hex) ₆ (Fuc)	
4332.0 $\pm$ 0.6	4333.29	(HexNAc) ₅ (Hex) ₆ (Fuc)(NeuAc)	
4407.5 $\pm$ 0.8	4407.36	(HexNAc) ₆ (Hex) ₇ (Fuc)	
4623.2 $\pm$ 0.8	4624.54	(HexNAc) ₅ (Hex) ₆ (Fuc)(NeuAc) ₂	
4697.4 $\pm$ 0.5	4698.63	(HexNAc) ₆ (Hex) ₇ (Fuc)NeuAc	
4988.2 $\pm$ 0.3	4989.89	(HexNAc) ₆ (Hex) ₇ (Fuc)(NeuAc) ₂	

^a The symbols used: ■, HexNAc; ●, Hex; ▲, NeuAc; □, Fuc.

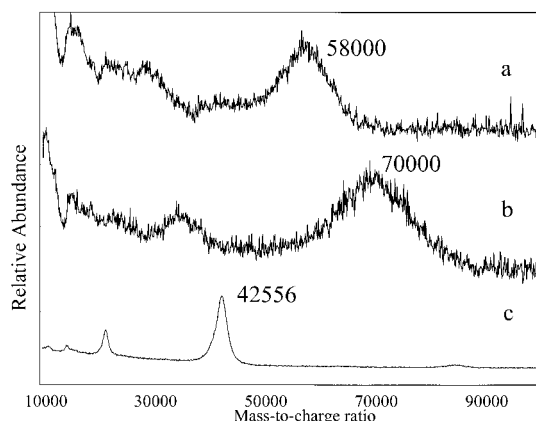


FIG. 5. MALDI-TOF spectra of (a) bovine decorin, (b) bovine biglycan and (c) biglycan after treatment with chondroitinase ABC.

The extended region carries approximately 100 chondroitin sulfate chains and is also modified with keratan sulfate. Each chondroitin sulfate chain has approximately 1 sulfate per disaccharide unit or 100 per chain, conferring approximately 10,000 fixed negative charges per aggrecan molecule. To date it does not appear to be possible to analyze a proteoglycan as large as intact aggrecan (2.5 MDa in size) using MALDI-TOF MS and it not clear if such a measurement would be of value considering the polydispersity of the molecule. The G1 fragment of aggrecan is amenable to MALDI-TOF MS, however, and a carbohydrate composition can be generated with glycosidase enzymes (Fig. 3).

Endogenous proteolytic cleavage of aggrecan in the interglobular domain generates GAG-containing fragments that do not bind hyaluronan and are lost from the cartilage matrix. Two proteolytic sites in the interglobular domain have been identified, the metalloproteinase site between Asn³⁴¹ and Phe³⁴² and the aggrecanase site between Glu³⁷³ and Ala³⁷⁴. While cleavage of both sites occurs *in vivo* and *in vitro*, fragments generated by aggrecanase have been observed in the synovial fluids of patients with osteoarthritis and other joint disorders, indicating that the enzyme may have an important role in aggrecan catabolism. ADMP-1 (aggrecanase), an enzyme cleaving the Glu³⁷³-Ala³⁷⁴ bond, has recently been purified (41) and cloned (42).

The extent of glycosylation of aggrecan G1-interglobular domain region is known to vary with age (43) and this may change the susceptibility of the molecule to degradation. Supporting this hypothesis, deglycosylated aggrecan has been observed to be degraded at a rate different than the untreated monomer *in vitro* (44, 45). The application of mass spectrometry to aggrecan G1 structure allows detailed information to be generated to correlate the susceptibility toward proteolytic cleavage with a detailed structure with respect to glycosylation. That G1 monomers remain associated after dilution in sinapinic acid solution and are detected as

multimers (Fig. 2a) is interesting in that most proteins do not display this degree of multimer formation. Treatment of the crystals with formic acid breaks the association and produces a spectrum dominated by monomers that has a higher signal-to-noise ratio and is thus preferred for the purpose of measuring changes in mass caused by glycosidase treatment.

Keratan sulfate, like other sulfated glycosaminoglycans, is not amenable to MALDI-TOF analysis. Measurement of the mass of KS modifying aggrecan in the interglobular domain is important because these chains may play a role in aggrecanase binding to aggrecan. Measurement of G1 mass as a function of glycosidase treatment provides a means of measuring keratan sulfate masses and provides a rapid means of comparing aggrecan G1 isolated from different tissue sources, i.e., normal vs osteoarthritic cartilage.

Mass spectrometric data of the type described in Fig. 4 and Table 2 complement data generated by monosaccharide analysis, nuclear magnetic resonance, and enzyme-based structural analysis. The spectra were produced on a few picomoles of material and considerable information was gained from the masses of individual glycoforms when analyzed in conjunction with a database of known carbohydrate structures (46). The glycan structure has an overall composition of (HexNAc)₃₋₆(Hex)₃₋₇(Fuc)(NeuAc)₀₋₂ and is attached at Asn³¹⁴, the only consensus site for N-glycosylation in the tryptic peptide. Because this peptide is proximal to the stromelysin and aggrecanase cleavage sites, it is of interest to compare the extent of glycosylation in G1 purified from normal and diseased tissue. The goal of continuing work is to determine if certain glycoforms can be correlated with aggrecanase cleavage of the aggrecan core protein. Knowing the HPLC retention time, this strategy allows the focus of analysis to be specific HPLC fractions, or specific regions of the chromatogram.

The small leucine-rich repeat proteoglycans (SLRPs) are a family of molecules believed to function in the regulation of collagen fibril formation and in the interactions between collagens and other molecules in the ECM. These molecules are found in connective tissues including articular and nasal cartilage, tendon, skin, smooth muscle, sclera, gingiva, and intervertebral disk. The core proteins are characterized by leucine-rich repeats consisting of 20–29 residues with leucines at conserved positions (47, 48). In addition, the SLRPs have a cluster of four cysteines in the N-terminal region of the core protein organized in a metallothionein-like pattern (49). The number and composition of GAG chains on these molecules vary from one (decorin) or two (biglycan) to four (fibromodulin). The SLRPs have a horseshoe shape, the concave inner surface of which binds to the collagen triple helix (7, 10). Degradation of these molecules has potential effects on the size of collagen fibrils and thus the tensile strength and elas-

ticity of cartilage. In addition, degradation may alter the way SLRPs interact with growth factors such as TGF- β s and adhesion proteins such as fibronectin.

Decorin isolated from a variety of connective tissues migrates on a SDS-PAGE gel at the same position as an unbranched 100-kDa standard protein (50–54). Because the molecule is branched and carries many sulfate groups, decorin would not be expected to bind SDS molecules in the same ratio as observed for standard proteins and would travel at a different rate through the polyacrylamide matrix toward the anode. The relative mobility of biglycan, with two dermatan sulfate chains, falls in the 90- to 220-kDa range, reflecting the tendency of biglycan to form multimers under certain SDS-PAGE conditions (52, 53, 55). The molecular weight of monomeric bovine articular cartilage biglycan has been measured to be 93–99 kDa by SEC (49) and is generally regarded as the best measurement of biglycan molecular weight. With one fewer dermatan sulfate chain, decorin is therefore expected to have a molecular weight of 70 kDa and analytical ultracentrifugation measurements give a range of 70–100 kDa (56).

The MALDI-TOF masses observed for decorin (58,000, Fig. 5a) and biglycan (Fig. 5b) are significantly lower than the values measured by SEC for these molecules. It may be that the high mass of the attached chondroitin sulfate chains and large number of sulfate groups cause fragmentation and a decrease in the detected mass, even at low laser energy. On the other hand, the nature of the chondroitin chains may prevent an accurate measurement of the partial specific volume of decorin and biglycan and therefore skew the mass measured by sedimentation equilibrium centrifugation. Efforts are underway to characterize the mass spectrometric behavior of SLRPs using other MALDI matrices, such as dihydroxybenzoic acid, known to produce less fragmentation than sinapinic acid. In addition, a mass spectrometer equipped with an ion reflector and infrared laser may shed light on the mass accuracy problem. Quantitation of chondroitin sulfate disaccharides produced by chondroitinase ABC treatment of decorin and biglycan relative to the core protein will provide a benchmark for mass measurement.

The high negative charge on the dermatan sulfate chains on decorin and biglycan combined with the high degree of polydispersity of these molecules make it difficult to determine the causes for the discrepancy between the molecular weights measured by MALDI-TOF and SEC. It may be that the populations of biglycan and decorin that contain the largest dermatan sulfate chains do not ionize by MALDI. Sulfated glycosaminoglycans do not ionize under the conditions used to acquire these spectra, and it may be that core proteins modified above a certain percentage of total mass by sulfated glycosaminoglycan do not ionize by MALDI. If this were true, then decorin and biglycan,

with core proteins of similar mass and composition, would be observed at the same MALDI-TOF m/z . The observation of an ion at m/z 70,000 for biglycan (Fig. 5b) and m/z 58,000 for decorin spectrum (Fig. 5a) indicates that no phenomenon of selective ionization is observed. To resolve the discrepancy between MALDI-TOF and SEC for sulfated proteoglycan mass measurement, it will be necessary to quantitate the dermatan sulfate/protein core ratio by combination of monosaccharide and amino acid analyses.

ACKNOWLEDGMENT

Decorin and biglycan were generously given by Lawrence Rosenberg, Orthopaedic Research Laboratory, Montefiore Medical Center, Bronx, New York.

REFERENCES

1. Mendler, M., Eich-Bender, S. G., Vaughan, L., Winterhalter, K. H., and Bruckner, P. (1989) Cartilage contains mixed fibrils of collagen types II, IX, and XI. *J. Cell Biol.* **108**, 191–197.
2. Roughley, P. J., and Lee, E. R. (1994) Cartilage proteoglycans: Structure and potential functions. *Microsc. Res. Technol.* **28**, 385–397.
3. Poole, A. R., Pidoux, I., Reiner, A., and Rosenberg, L. C. (1982) An immunoelectron microscope study of the organization of proteoglycan monomer, link protein, and collagen in the matrix of articular cartilage. *J. Cell Biol.* **93**, 921–937.
4. Hascall, V. C., and Sajdera, S. W. (1970) Physical properties and polydispersity of proteoglycan from bovine nasal cartilage. *J. Biol. Chem.* **245**, 4920–4930.
5. Poole, A. R., Rosenberg, L. C., Reiner, A., Ionescu, M., Bogoch, E., and Roughley, P. J. (1996) Contents and distributions of the proteoglycans decorin and biglycan in normal and osteoarthritic human articular cartilage. *J. Orthop. Res.* **14**, 681–689.
6. Neame, P. J., Choi, H. U., and Rosenberg, L. C. (1989) The primary structure of the core protein of the small, leucine-rich proteoglycan (PG I) from bovine articular cartilage. *J. Biol. Chem.* **264**, 8653–8661.
7. Scott, J. E., Orford, C. R., and Hughes, E. W. (1981) Proteoglycan–collagen arrangements in developing rat tail tendon. An electron microscopical and biochemical investigation. *Biochem. J.* **195**, 573–581.
8. Scott, J. E. (1984) The periphery of the developing collagen fibril. Quantitative relationships with dermatan sulphate and other surface-associated species. *Biochem. J.* **218**, 229–233.
9. Scott, J. E. (1996) Proteodermatan and proteokeratan sulfate (decorin, lumican/fibromodulin) proteins are horseshoe shaped. Implications for their interactions with collagen. *Biochemistry* **35**, 8795–8799.
10. Weber, I. T., Harrison, R. W., and Iozzo, R. V. (1996) Model structure of decorin and implications for collagen fibrillogenesis. *J. Biol. Chem.* **271**, 31767–31770.
11. Lewandowska, K., Choi, H. U., Rosenberg, L. C., Zardi, L., and Culp, L. A. (1987) Fibronectin-mediated adhesion of fibroblasts: Inhibition by dermatan sulfate proteoglycan and evidence for a cryptic glycosaminoglycan-binding domain. *J. Cell Biol.* **105**, 1443–1454.
12. Yamaguchi, Y., Mann, D. M., and Ruoslahti, E. (1990) Negative regulation of transforming growth factor-beta by the proteoglycan decorin. *Nature* **346**, 281–284.
13. Border, W. A., Noble, N. A., Yamamoto, T., Harper, J. R., Yamaguchi, Y., Pierschbacher, M. D., and Ruoslahti, E. (1992)

- Natural inhibitor of transforming growth factor-beta protects against scarring in experimental kidney disease. *Nature* **360**, 361–364.
14. Poole, C. A., Ayad, S., and Schofield, J. R. (1988) Chondrons from articular cartilage. I. Immunolocalization of type VI collagen in the pericellular capsule of isolated canine tibial chondrons. *J. Cell Sci.* **90**, 635–643.
 15. Salter, D. M., Godolphin, J. L., Gourlay, M. S., Lawson, M. F., Hughes, D. E., and Dunne, E. (1996) Analysis of human articular chondrocyte CD44 isoform expression and function in health and disease. *J. Pathol.* **179**, 396–402.
 16. Stamenkovic, I., Amiot, M., Pesando, J. M., and Seed, B. (1989) A lymphocyte molecule implicated in lymph node homing is a member of the cartilage link protein family. *Cell* **56**, 1057–1062.
 17. Goldstein, L. A., Zhou, D. F., Picker, L. J., Minty, C. N., Bargatze, R. F., Ding, J. F., and Butcher, E. C. (1989) A human lymphocyte homing receptor, the hermes antigen, is related to cartilage proteoglycan core and link proteins. *Cell* **56**, 1063–1072.
 18. Yang, B., Yang, B. L., Savani, R. C., and Turley, E. A. (1994) Identification of a common hyaluronan binding motif in the hyaluronan binding proteins RHAMM, CD44 and link protein. *EMBO J.* **13**, 286–296.
 19. Mollenhauer, J., Bee, J. A., Lizarbe, M. A., and von der Mark, K. (1984) Role of anchorin CII, a 31,000-mol-wt membrane protein, in the interaction of chondrocytes with type II collagen. *J. Cell Biol.* **98**, 1572–1579.
 20. Hewitt, A. T., Varner, H. H., Silver, M. H., Dessau, W., Wilkes, C. M., and Martin, G. R. (1982) The isolation and partial characterization of chondronectin, an attachment factor for chondrocytes. *J. Biol. Chem.* **257**, 2330–2334.
 21. Maroudas, A., Bayliss, M. T., Uchitel-Kaushansky, N., Schneiderman, R., and Gilav, E. (1998) Aggrecan turnover in human articular cartilage: Use of aspartic acid racemization as a marker of molecular age. *Arch. Biochem. Biophys.* **350**, 61–71.
 22. Hardingham, T. E. (1979) The role of link-protein in the structure of cartilage proteoglycan aggregates. *Biochem. J.* **177** 237–247.
 23. Nguyen, Q., Liu, J., Roughley, P. J and Mort, J. S. (1991) Link protein as a monitor *in situ* of endogenous proteolysis in adult human articular cartilage. *Biochem. J.* **278**, 143–147.
 24. Repo, R. U., and Mitchell, N. S. (1971) Collagen synthesis in mature articular cartilage of the rabbit. *J. Bone Joint Surg. [Br].* **53**, 541–548.
 25. Aigner, T., Stoss, H., Weseloh, G., Zeiler, G., and von der Mark, K. (1992) Activation of collagen type II expression in osteoarthritic and rheumatoid cartilage. *Virchow's Arch. B Cell Pathol. Incl. Mol. Pathol.* **62**, 337–345.
 26. Rizkalla, G., Reiner, A., Bogoch, E., and Poole, A. R. (1992) Studies of the articular cartilage proteoglycan aggrecan in health and osteoarthritis. Evidence for molecular heterogeneity and extensive molecular changes in disease. *J. Clin. Invest.* **90**, 2268–2277.
 27. Cs-Szabo, G., Roughley, P. J., Plaas, A. H., and Glant, T. T. (1995) Large and small proteoglycans of osteoarthritic and rheumatoid articular cartilage. *Arthritis Rheum.* **38**, 660–668.
 28. Mankin, H. J., Dorfman, H., Lippiello, L., and Zarins, A. (1971) Biochemical and metabolic abnormalities in articular cartilage from osteoarthritic human hips. II. Correlation of morphology with biochemical and metabolic data. *J. Bone Joint Surg. [Am.]* **53**, 523–537.
 29. Vornehm, S. I., Dudhia, J., Von der Mark, K., and Aigner, T. (1996) Expression of collagen types IX and XI and other major cartilage matrix components by human fetal chondrocytes *in vivo*. *Matrix Biol.* **15**, 91–98.
 30. Aigner, T., Dietz, U., Stoss, H., and von der Mark, K. (1995) Differential expression of collagen types I, II, III, and X in human osteophytes. *Lab Invest* **73**, 236–243.
 31. McDevitt, C. A., Muir, H., and Gilbertson, E. M. (1975) Impairment of proteoglycan-collagen interaction in early experimental osteoarthritis. *Ann. Rheum. Dis.* **34**, 541–542.
 32. Poole, A. R. (1986) Changes in the collagen and proteoglycan of articular cartilage in arthritis. *Rheumatology* **10**, 316–371.
 33. Maroudas, A. (1976) Balance between swelling pressure and collagen tension in normal and degenerate cartilage. *Nature* **260**, 808–809.
 34. Karas, M., and Hillenkamp, F. (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal. Chem.* **60**, 2209–2301.
 35. Zaia, J., Boynton, R. E., McIntosh, A., Marshak, D. R., Olsson, H., Heinegård, D., and Barry, F. P. (1997) Post-translational modifications in cartilage oligomeric matrix protein. Characterization of the N-linked oligosaccharides by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *J. Biol. Chem.* **272**, 14120–14126.
 36. Hedbom, E., Antonsson, P., Hjerpe, A., Aeschlimann, D., Paulsson, M., Rosa-Pimental, E., Sommarin, E., Wendel, M., Oldberg, A., and Heinegård, D. (1992) Cartilage matrix proteins. An acidic oligomeric protein (COMP) detected only in cartilage. *J. Biol. Chem.* **267**, 6132–6136.
 37. Barry, F. P., Gaw, J. U., Young, C. N., and Neame, P. J. (1992) Hyaluronan-binding region of aggrecan from pig laryngeal cartilage. Amino acid sequence, analysis of N-linked oligosaccharides and location of the keratan sulphate. *Biochem. J.* **286**, 761–769.
 38. Kornfeld, R., and Kornfeld, S. (1985) Assembly of asparagine-linked oligosaccharides. *Annu. Rev. Biochem.* **54**, 631–634.
 39. Morgelin, M., Heinegård, D., Engel, J., and Paulsson, M. (1992) Electron microscopy of native cartilage oligomeric matrix protein purified from the Swarm rat chondrosarcoma reveals a five-armed structure. *J. Biol. Chem.* **267**, 6137–6141.
 40. Hecht, J. T., Nelson, L. D., Crowder, E., Wang, Y., Elder, F. F. B., Harrison, W. R., Francomano, C. A., Prange, C. K., Lennon, G. G., Deere, M., and Lawler, J. (1995) Mutations in exon 17B of cartilage oligomeric matrix protein (COMP) cause pseudoachondroplasia. *Nat. Genet.* **10**, 325–329.
 41. Arner, E. C., Pratta, M. A., Trzaskos, J. M., Decicco, C. P., and Tortorella, M. D. (1999) Generation and characterization of aggrecanase. A soluble, cartilage-derived aggrecan-degrading activity. *J. Biol. Chem.* **274**, 6594–6601.
 42. Tortorella, M. D., Burn, T. C., Pratta, M. A., *et al.* (1999) Purification and cloning of aggrecanase-1: A member of the ADAMTS family of proteins. *Science* **284**, 1664–1666.
 43. Barry, F. P., Rosenberg, L. C., Gaw, J. U., Gaw, J. U., Koob, T. J., and Neame, P. J. (1995) N- and O-linked keratan sulfate on the hyaluronan binding region of aggrecan from mature and immature bovine cartilage. *J. Biol. Chem.* **270**, 20516–20524.
 44. Pratta, M. A., Tortorella, M., and Arner, E. C. (1999) Aggrecan glycosylation affects cleavage by aggrecanase. *Trans. 45th Ann. Meet. Orthop. Res. Soc.*, 407.
 45. Hörber, C., Büttner, F. H., Hughes, C. E., Caterson, B., and Bartnik, E. (1999) Glycosylation of the recombinant aggrecan RAGG1MUT affects susceptibility to aggrecanase. *Trans. 45th Ann. Meet. Orthop. Res. Soc.*, 676.
 46. Doubet, S., and Albersheim, P. (1992) CarbBank. *Glycobiology* **2**, 505.
 47. Day, A. A., Ramis, C. I., Fisher, L. W., Gehron-Robey P., Termine, J. D., and Young, M. F. (1986) Characterization of bone PG II cDNA and its relationship to PG II mRNA from other connective tissues. *Nucleic Acids Res.* **14**, 9861–9876.

48. Krusius, T., and Ruoslahti, E. (1986) Primary structure of an extracellular matrix proteoglycan core protein deduced from cloned cDNA. *Proc. Natl. Acad. Sci. USA* **83**, 7683–7687.
49. Liu, J., Laue, T. M., Choi, H. U., Tang, L-H., and Rosenberg, L. (1994) The self-association of biglycan from bovine articular cartilage. *J. Biol. Chem.* **269**, 28366–28373.
50. Ramamurthy, P., Hocking, A. M., and McQuillan, D. J. (1996) Recombinant decorin glycoforms. Purification and structure. *J. Biol. Chem.* **271**, 19578–19584.
51. Rosenberg, L., Choi, H. U., Tang, L-H., Johnson, T. L., Pal, S., Webber, C., Reiner, A., and Poole, A. R. (1985) Isolation of dermatan sulfate proteoglycans from mature bovine articular cartilages. *J. Biol. Chem.* **260**, 6304–6313.
52. Roughley, P. J., and White, R. J. (1989) Dermatan sulphate proteoglycans of human articular cartilage. The properties of dermatan sulphate proteoglycans I and II. *Biochem. J.* **262**, 823–827.
53. Choi, H. U., Johnson, T. L., Pal, S., Tang, L-H., Rosenberg, L., and Neame, P. J. (1989) Characterization of the dermatan sulfate proteoglycans, DS-PGI and DS-PGII, from bovine articular cartilage and skin isolated by octyl-Sepharose chromatography. *J. Biol. Chem.* **264**, 2876–2884.
54. Little, C. B., Ghosh, P., and Bellenger, C. R. (1996) Topographic variation in biglycan and decorin synthesis by articular cartilage in the early stages of osteoarthritis: An experimental study in sheep. *J. Orthop. Res.* **14**, 433–444.
55. Hocking, A. M., Strugnell, R. A., Ramamurthy, P., and McQuillan, D. J. (1996) Eukaryotic expression of recombinant biglycan. Post-translational processing and the importance of secondary structure for biological activity. *J. Biol. Chem.* **271**, 19571–19577.
56. Cöster, L., Fransson, L. A., Sheehan, J., Nieduszynski, I. A., and Phelps, C. F. (1981) Self-association of dermatan sulphate proteoglycans from bovine sclera. *Biochem. J.* **197**, 483–490.