The major proteoglycan of adult rabbit skeletal muscle

Relationship to small proteoglycans of other tissues

Narayanan PARTHASARATHY,*‡ Lakshmi CHANDRASEKARAN† and Marvin L. TANZER†

*Department of Orthopaedic Surgery and †Department of BioStructure and Function, University of Connecticut Health Center, Farmington, CT 06032, U.S.A.

We have been interested in examining the putative biological role(s) of the major proteoglycan of adult skeletal muscle. The small proteoglycans of adult rabbit skeletal muscle and tendon were extracted and purified by sequential density-gradient ultracentrifugation, ion-exchange chromatography and gel filtration. They appeared to be homogeneous by the criterion of gel electrophoresis in SDS and to yield one major product, the core protein, after digestion with chondroitin ABC lyase, also observed after gel electrophoresis. Two major products were obtained when the intact proteoglycans were cleaved by CNBr, and those peptides were separated by SDS/PAGE and by ion-exchange chromatography. Sequencing of the *N*-terminal amino acids of either the intact proteoglycans or the CNBr-cleaved products allowed for comparison of the muscle and tendon proteoglycan with derived amino acid sequences previously reported for bovine bone proteoglycan. The bone and tendon proteoglycan sequences were remarkably similar, whereas those of the muscle proteoglycan differed from the other two molecules. The major site of glycosaminoglycan substitution was on a peptide fragment distant from the *N*-terminus, and a presumptive serine residue at position 4 from the *N*-terminus also appeared to be substituted, perhaps with a small glycosaminoglycan chain. These results provide some insight into the diversity of small proteoglycans of the PG-II class and provide a basis for exploring their mode of genetic expression.

INTRODUCTION

Proteoglycans are ubiquitous components of the extracellular matrix of all mammalian tissues [1]. Early studies [2–4] of the tissue proteoglycans provided analyses of the structure and function of the large chondroitin sulphate (CS-PG) from cartilage. However, it has become increasingly clear that proteoglycans with a much smaller core-protein size compared with cartilage CS-PG occur in a number of tissues, including skin [5,6], bone [7], tendon [8,9], cartilage [10], aorta [11,12], placenta [13], cornea [14], sclera [15] and periodontal ligament [16]. These proteoglycans (PG-II) have a core protein of apparent M_r 48 000 [17]. Although the functional significance of PG-II molecules has not been completely understood, these proteoglycans seem to affect collagen fibre diameter [18] and also have been implicated in binding to specific loci of collagen fibrils in many connective tissues [19].

Characterization of the proteoglycans of skeletal muscle has been slow, even though this tissue is among the most abundant in the body. Probably the low relative abundance of proteoglycans in adult tissue has made their isolation difficult. Some investigators have chosen to use embryonic tissues as their source, or have cultured cells from such tissues [20,21]. In both instances, metabolic radiolabelling of newly synthesized molecules has been the primary experimental approach, followed by their isolation and partial characterization. We have previously reported [22] the isolation and characterization, from adult rabbit skeletal muscle, of a small proteoglycan containing chondroitin sulphate A and C. It is the major proteoglycan of adult rabbit skeletal muscle, comprising about 44% of the total hexuronic acid of that tissue [22]. This CS-PG-II resembles other PG-II proteoglycans in that the residue remaining after chondroitin ABC lyase digestion had an average M_r of 50 000, which is near the range of M_r values (45 000–48 000) reported [17] for the core proteins of the PG-II class. Although the amino acid

composition of muscle CS-PG-II was generally comparable with that of the other PG-II molecules it was notable for its relatively high content of glycine and lower contents of aspartic acid, glutamic acid and leucine, the amino acids characteristic of the other small proteoglycans [17]. These differences imply that the core protein of muscle CS-PG-II may have some unique features in its amino acid sequence. In order to explore this possibility, we have determined the *N*-terminal sequences and some internal sequences of amino acids of muscle CS-PG-II. We compare these sequences with those of rabbit tendon dermatan sulphate DS-PG-II and with the deduced amino acid sequences of bovine bone CS-PG-II [23]. We also provide data that imply that a serine residue, the fourth amino acid from the *N*-terminus of the muscle CS-PG-II, is substituted with a moiety other than a conventional glycosaminoglycan.

MATERIALS AND METHODS

Materials

DEAE-Sephacel and Sepharose CL-4B were from Pharmacia; DEAE-cellulose (DE-52) was from Whatman; chondroitin ABC lyase was from Sigma Chemical Co.; CNBr was from Pierce Chemical Co.; frozen rabbit muscle and achilles tendon were obtained from Pel Freeze, Rogers, AR, U.S.A.

Purification of skeletal-muscle and tendon proteoglycans

Proteoglycans were extracted at 4 °C from rabbit skeletal muscle and tendon with 4 M-guanidinium chloride in the presence of proteinase inhibitors and then were separated by direct dissociative CsCl centrifugation [22]. The PG-II was purified from bottom fractions (D1 + D2, density > 1.38 g/ml) by DEAE-Sephacel adsorption in a column of 20 ml bed volume equilibrated with 4 M-urea/0.1 M-NaCl/0.05 M-sodium acetate buffer, pH 6.0. After washing the column with the same buffer,

Abbreviations used: PG-II, proteoglycans of small mass with very similar chemical features; CS-PG-II, chondroitin sulphate proteoglycans within the general class; DS-PG-II, dermatan sulphate proteoglycans within the general class.

‡ To whom correspondence should be addressed.

proteoglycans were eluted with 2 M-NaCl in the same urea buffer. Fractionation of the proteoglycans was achieved by another DEAE-Sephacel column (4 ml bed volume), with a linear concentration gradient of 0.1–1.0 M-NaCl in the urea buffer. Fractions from the column were assayed for glycosaminoglycans by the dye-binding method of Farndale *et al.* [24]. The proteoglycan-containing fractions were pooled, and a sample of the peak was labelled with [^{14}C]formaldehyde by our procedure [22] and mixed with the unlabelled material. The mixture was further purified by Sepharose CL-4B chromatography on a 1 cm $\times$ 50 cm column equilibrated with and eluted with 4 M-guanidinium chloride/0.05 M-sodium acetate buffer, pH 5.0.

Purity of proteoglycans

The isolated PG-II fractions from both muscle and tendon were analysed by SDS/PAGE [25] on vertical mini slab gels (1.5 mm thick) consisting of a linear 3–12 % -acrylamide gradient. Staining was accomplished with Coomassie Blue followed by Alcian Blue [7]. After chondroitin ABC lyase treatment [26], the core protein of the proteoglycan was also analysed by SDS/PAGE.

N-Terminal amino acid sequencing of the core proteins

After chondroitin ABC lyase digestion and electrophoresis by SDS/PAGE, the core proteins were transferred on to a poly(vinylidene difluoride) (Millipore) membrane [27] and stained with Coomassie Blue. Such membranes, containing about 100 pmol of protein, were sequenced with an Applied Biosystems 475A protein sequencer by Dr. Audree Fowler at the U.C.L.A. sequencing facility.

CNBr-cleavage peptide sequences

The proteoglycans were treated with CNBr [28] in a final volume of 100 μl containing 5 mg of CNBr in 70 % (v/v) formic acid, with incubation at 25 $^{\circ}\text{C}$ for 3–4 h in an atmosphere of N_2 . The peptide reaction products were initially dialysed through a Spectrapor membrane, which retarded substances larger than M_r 12000. The non-diffusible material was separated by SDS/PAGE on 15 %-acrylamide gels, transferred on to poly(vinylidene difluoride) membranes and sequenced. Glycosaminoglycan peptides were also isolated from the CNBr digest by chromatography on a DEAE-cellulose column (1 cm $\times$ 4 cm) in pyridine/formate buffer, pH 5.0 [22]. Portions of the fractions were assayed for glycosaminoglycans by a dye-binding method [24].

RESULTS

Isolation of proteoglycans from rabbit muscle and tendon

The isolation of CS-PG-II from rabbit skeletal muscle was carried out by minor modifications of our procedure [22]. The proteoglycan was bound to a DEAE-Sephacel column and eluted as a peak at 0.5 M-NaCl concentration. Upon gel filtration on Sepharose CL-4B the proteoglycan was eluted (Fig. 1) as a major peak. The DS-PG-II molecule from rabbit tendon was isolated by a similar procedure, except that the gel-filtration step was omitted.

Purity of proteoglycans

SDS/PAGE of CS-PG-II from skeletal muscle revealed one band, stained with Coomassie Blue and Alcian Blue, which migrated (Fig. 2, lane A) to a position consistent with an average M_r of 95000 (range 93000–100000). After digestion with chondroitin ABC lyase, the proteoglycan changed mobility to a band migrating (Fig. 2, lane B) at an apparent M_r of 48000. In the case of the DS-PG-II obtained from tendon, it initially

Fig. 1. Gel filtration of muscle proteoglycan on Sepharose CL-4B

Unlabelled purified proteoglycan from DEAE-Sephacel chromatography was mixed with previously purified proteoglycan which was radiolabelled by reductive alkylation [22]. The mixture, containing 2.0×10^6 c.p.m., was applied to a 1 cm $\times$ 50 cm column, which was equilibrated and eluted with 4 M-guanidinium chloride/0.05 M-sodium acetate buffer, pH 5.0. A sample from each 1 ml fraction was analysed for radioactivity. Recovery of the radioactivity applied to the column was complete.

Fig. 2. SDS/PAGE on a 3–12% linear gradient gel of the muscle proteoglycan pooled from the gel-filtration purification step (Fig. 1)

Lane A shows the intact proteoglycan, and lane B shows its migration after digestion with chondroitin ABC lyase. Lane C shows enzyme alone, which contains BSA as a carrier. Lane D shows M_r standards.

migrated (Fig. 3, lane A) at an M_r of 105000, and then became a more rapidly migrating band (Fig. 3, lane B) at M_r 48000 after digestion with chondroitin ABC lyase.

N-Terminal sequence of muscle CS-PG-II

Analysis of the amino acid sequence of muscle CS-PG-II was performed after its digestion by chondroitin ABC lyase, and is shown in Fig. 4. Since this sequence shows an indeterminate residue at position 4 from the N-terminus, it is most likely that this amino acid is substituted by a carbohydrate or other moiety.

Amino acid sequences of CNBr-cleavage peptides

After CNBr digestion and dialysis, the proteoglycan from

Fig. 3. SDS/PAGE of rabbit tendon proteoglycan (lane A), its digestion product after incubation with chondroitin ABC lyase (lane B) and enzyme alone (lane C)
Positions of M_r standards are indicated.

muscle gave a broad diffuse band (Fig. 5, lane A) with an apparent M_r of 29000 (range 27000–31000). A sharp band of M_r 14000 was also seen. The tendon proteoglycan peptides showed a similar electrophoretic pattern (results not shown). When the muscle proteoglycan CNBr-cleavage products were fractionated by DEAE-cellulose, the peptides containing glycosaminoglycans remained bound to the column under the

Fig. 5. SDS/PAGE on a 15% gel of the CNBr digest of muscle proteoglycan (lane A) and the glycosaminoglycan-containing peptide obtained from the DEAE-cellulose chromatography step (lane B)
Positions of M_r standards are indicated.

conditions employed. Elution with 2.5 M-pyridine/formate buffer, pH 5.0, yielded a major peptide which appeared as a broad band of M_r 29000 (Fig. 5, lane B). This peptide resembles that seen in the unfractionated mixture at the same location (Fig. 5, lane A). No other component was eluted from the column with 2.5 M-pyridine/formate buffer. Furthermore, only the material eluted with 2.5 M-pyridine/formate buffer produced the char-

N-Terminal sequences		Sequence identity with bone PG (%)
Skeletal muscle (rabbit)	D E A X G I G P D E L A P E L L D D L M	62
Tendon (bovine)	D E A S G I G P E E H F P E	100
Bone (bovine) (1–21)	D E A S G I G P E E H F P E V P E I E P M	
Internal sequences		
(i) 14000- M_r CNBr-cleavage peptide		
Skeletal muscle (rabbit)	G P V T P F F K Q S S L P V I V	50
Tendon (rabbit)	G P V C P F R C Q C H L R V V Q C S D L G L D K V P K D L P P	91
Bone (bovine) (22–52)	G P V C P F R C Q C H L R V V Q C S D L G L E L V P L D L P P	
(ii) 29000- M_r CNBr-cleavage peptide		
Skeletal muscle (rabbit)	M K K L S Y I C I A D T I K T K I K Q G L I P S S T I L Y L I	70
Tendon (rabbit)	M K K L S Y I R I A D T R I T T I P Q G L P P S L T E L	100
Bone (bovine) (170–200)	M K K L S Y I R I A D T N I T T I P Q G L P P S L T E L H L D	

Fig. 4. Comparison of N-terminal and internal amino acid sequences of small proteoglycans (PG-II) from skeletal muscle, tendon and bone

The N-terminal sequence of tendon proteoglycan was determined directly [30], whereas that of bone proteoglycan was deduced from the corresponding cDNA sequence [23]. The presence of alternative amino acids at a specific locus in the sequences indicates uncertainty of identification. Cysteine residues from direct sequencing are deduced from their degradation products.

acteristic chromophore in the dye-binding assay; this observation, coupled with the lack of another major broad band in Fig. 5, suggests that only one peptide containing glycosaminoglycan was produced by the CNBr reaction.

The amino acid sequence of these peptides is shown in Fig. 4, which also shows that the muscle peptide is very similar to the equivalent peptide from the tendon proteoglycan. The sequences differed at residues 8, 18 and 22, with the muscle proteoglycan showing cysteine, lysine and isoleucine respectively, whereas the corresponding tendon amino acid residues were arginine, proline and leucine. These sequences (Fig. 4) do not show any indeterminate sites, suggesting that the substituted glycosaminoglycan is at a serine residue further along the peptide chain.

The 14000- M_r peptide from both proteoglycans was also partially sequenced, as shown in Fig. 4. The narrow bands in the electrophoretic gel (Fig. 5) imply that these peptides lack glycosaminoglycan chains. This suggestion is consistent with the finding that they are eluted at much lower concentrations of pyridine/formate buffer than the 29000- M_r peptides. Comparison of the sequences of these peptides with the deduced amino acid sequence of bovine bone CS-PG-II [23] suggests that the 14000- M_r peptides are contiguous with the methionine residue at position 21. Fig. 4 also shows that the corresponding sequences of rabbit tendon DS-PG-II and bovine bone CS-PG-II are virtually identical (91–100% similarity), whereas the rabbit muscle sequences show less similarity (50–70%).

DISCUSSION

The predominant class of loose connective-tissue proteoglycans [29] is characterized by having a small core protein (M_r 48000) containing one or two glycosaminoglycan chains. Even though these proteoglycans differ from one another in tissue source as well as in iduronate content, their core proteins seem to have quite similar structures [29]. In immunological assays the PG-II from bone, cornea, sclera and tendon have been demonstrated to show considerable identity [17]. Furthermore, immunoblots have shown that antibodies raised against tendon DS-PG-II cross-react with cartilage and bone PG-II [30]. Identical amino acid sequences at the *N*-terminus have been found for the small proteoglycans of skin [31], tendon [30], bone [23] and cartilage [32]. In the present study we have shown that such sequences are virtually identical for muscle, tendon and bone proteoglycans (Fig. 4). Furthermore, the internal amino acid sequences indicate that tendon and bone have the greatest similarity (Fig. 4). Although internal sequences have previously been deduced by recombinant-DNA methods, our data provide direct sequence information.

One consensus sequence for potential initiation of a glycosaminoglycan chain on a core protein is Ser-Gly-Xaa-Gly [33]. Although this consensus sequence occurs at the *N*-terminus of the muscle CS-PG-II, assuming that the indeterminate residue is serine, our data for CNBr-cleavage peptides did not detect a glycosaminoglycan substituent on the cleaved peptide. Presumably another type of substituent is present at residue 4, although we cannot exclude a short glycosaminoglycan moiety. It has been postulated that such short chains, containing five to 15 disaccharide units, may be present in the DS-PG-II synthesized in chick-embryo skeletal muscle [34]. It is likely that these peptides were lost during the dialysis procedure after cleavage by CNBr.

Cell-free translation studies by Termine and co-workers [35] have shown that the recombinant protein expressed by cDNA clones encoding CS-PG-II from bone cross-reacted with polyclonal and monoclonal antisera to PG-II molecules from skin, cartilage, tendon, smooth muscle and cornea. Sandell *et al.* [36] have shown that smooth-muscle PG-II precursor core protein

is immunologically related to both cartilage and fibroblast core proteins. Although the structural similarities of these various core proteins indicate their close relatedness, it is not known whether, in fact, they are encoded by a single gene. Determination of this point will require more detailed recombinant-DNA studies.

The PG-II molecules can also be distinguished by the observation that those of skeletal muscle [22] and bone [7] contain chondroitin sulphate chains, whereas those of skin [5], tendon [8] and cornea [14] are substituted with dermatan sulphate chains. However, cartilage PG-II molecules contain either chondroitin sulphate [32] or dermatan sulphate [37] chains. Thus, although the same consensus sequences for glycosylation are present in the core proteins of both groups, other information must be provided to determine what the final glycosaminoglycan will be.

Binding of these small proteoglycans to other components of the extracellular matrix is a prominent phenomenon. Thus they have been shown to bind to collagen [18] and to fibronectin fibrils [38]. The tendon DS-PG-II inhibits collagen fibril formation *in vitro* [9], suggesting a regulatory role in matrix assembly. Which domains of the proteoglycans are important for binding is not known. Conceivably, both the carbohydrate and protein regions are potential participants. Both enzymic and chemical methods, as well as site-directed mutagenesis, could be used to find which portions of the molecule are essential for interactions with other components of the extracellular matrix.

This work was supported in part by research grants from the University of Connecticut Research Foundation to N.P. and by National Institutes of Health Grant AR-12683 to M.L.T.

REFERENCES

- Heinegård, D. & Paulsson, M. (1984) in *Extracellular Matrix Biochemistry* (Piez, K. A. & Reddi, A. H., eds.), pp. 278–328, Elsevier, New York.
- Hascall, V. C. & Kimura, J. H. (1982) *Methods Enzymol.* **82**, 769–800.
- Carney, S. L. & Muir, H. (1988) *Physiol. Rev.* **68**, 858–910.
- Fosang, A. J. & Hardingham, T. E. (1989) *Biochem. J.* **261**, 801–809.
- Damle, S. P., Cöster, L. & Gregory, J. D. (1982) *J. Biol. Chem.* **257**, 5523–5527.
- Swann, D. A., Garg, H. G., Jung, W. & Hermann, H. (1985) *J. Invest. Dermatol.* **84**, 527–531.
- Fischer, L. W., Termine, J. D., Dejter, S. W., Jr., Whitson, S. W., Yanagishita, M., Kimura, J. H., Hascall, V. C., Kleinman, H. K., Hassell, J. R. & Nilsson, B. (1983) *J. Biol. Chem.* **258**, 6588–6594.
- Vogel, K. G. & Heinegård, D. (1985) *J. Biol. Chem.* **260**, 9298–9306.
- Vogel, K. G., Paulsson, M. & Heinegård, D. (1984) *Biochem. J.* **223**, 587–597.
- Rosenberg, L. C., Choi, H. U., Tang, L.-H., Johnson, T. L., Pal, S., Webber, C., Reiner, A. & Poole, A. R. (1985) *J. Biol. Chem.* **260**, 6304–6313.
- Kapoor, R., Phelps, C. F., Cöster, L. & Franzén, A. (1981) *Biochem. J.* **197**, 259–268.
- Rowe, H. A. & Wagner, W. D. (1985) *Arteriosclerosis* **5**, 101–109.
- Brennan, M. J., Oldberg, A., Pierschbacher, M. D. & Ruoslahti, E. (1984) *J. Biol. Chem.* **259**, 13742–13750.
- Gregory, J. D., Cöster, L. & Damle, S. P. (1982) *J. Biol. Chem.* **257**, 6965–6970.
- Cöster, L. & Franzén, A. (1981) *Biochem. J.* **193**, 143–153.
- Pearson, C. H. & Gibson, G. J. (1982) *Biochem. J.* **201**, 27–37.
- Heinegård, D., Björne-Persson, A., Cöster, L., Franzén, A., Gardell, S., Malmström, A., Paulsson, M., Sandfalk, R. & Vogel, K. (1985) *Biochem. J.* **230**, 181–194.
- Scott, J. E. & Orford, C. R. (1981) *Biochem. J.* **197**, 213–216.
- Scott, J. E. & Haigh, M. (1985) *Biosci. Rep.* **5**, 765–774.
- Carrino, D. A. & Caplan, A. I. (1982) *J. Biol. Chem.* **257**, 14145–14154.
- Carrino, D. A., Pechak, D. G. & Caplan, A. I. (1984) *Exp. Biol. Med.* **9**, 80–86.

22. Parthasarathy, N. & Tanzer, M. L. (1987) *Biochemistry* **26**, 3149–3156
23. Day, A. A., McQuillan, C. I., Termine, J. D. & Young, M. R. (1987) *Biochem. J.* **248**, 801–805
24. Farndale, R. W., Sayers, C. A. & Barrett, A. J. (1982) *Connect. Tissue Res.* **9**, 247–248
25. Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685
26. Vogel, K. G. & Evanko, S. P. (1987) *J. Biol. Chem.* **262**, 13607–13613
27. Matsudaira, P. (1987) *J. Biol. Chem.* **262**, 10035–10038
28. Cheung, D. T., Benya, P. D., Hofer, D. P. & Nimni, M. E. (1981) *Collagen Relat. Res.* **1**, 247–256
29. Mörgelin, M., Paulsson, M., Malmström, A. & Heinegård, D. (1989) *J. Biol. Chem.* **264**, 12080–12090
30. Vogel, K. G. & Fisher, L. W. (1986) *J. Biol. Chem.* **261**, 11334–11340
31. Chopra, R. K., Pearson, C. H., Pringle, G. A., Fackre, D. S. & Scott, P. G. (1985) *Biochem. J.* **232**, 277–279
32. Choi, H. U., Johnson, T. L., Pal, S., Tang, L.-H., Rosenberg, L. & Neame, P. J. (1989) *J. Biol. Chem.* **264**, 2876–2884
33. Bourdon, M. A., Krusius, T., Campbell, S., Schwartz, N. B. & Ruoslahti, E. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 3194–3198
34. Young, H. E., Carrino, D. A. & Caplan, A. I. (1988) *Connect. Tissue Res.* **17**, 99–118
35. Day, A. A., Ramis, C. I., Fischer, L. W., Gehron-Robey, P., Termine, J. D. & Young, M. F. (1986) *Nucleic Acids Res.* **14**, 9861–9876
36. Sandell, L. J., Sawhney, R. S., Yeo, T.-K., Poole, A. R., Rosenberg, L. C., Kresse, H. & Wight, T. N. (1988) *Eur. J. Cell Biol.* **46**, 253–258
37. Sampaio, L. O., Bayliss, M. T., Hardingham, T. E. & Muir, H. (1988) *Biochem. J.* **254**, 757–764
38. Schmidt, G., Robenek, H., Harrach, B., Glössl, J., Nolte, V., Hörmann, H., Richter, H. & Kresse, H. (1987) *J. Cell Biol.* **104**, 1683–1691

Received 10 May 1990/23 July 1990; accepted 1 August 1990