Biochem. J. (1987) 245, 763–772 (Printed in Great Britain)

Extended and globular protein domains in cartilage proteoglycans

Mats PAULSSON,* Matthias MÖRGELIN,* Hanna WIEDEMANN,§ Matthew BEARDMORE-GRAY,† David DUNHAM,† Timothy HARDINGHAM,† Dick HEINEGÅRD,‡ Rupert TIMPL§ and Jürgen ENGEL*‖

*Abteilung Biophysikalische Chemie, Biozentrum, CH-4056 Basel, Switzerland, †Division of Biochemistry, Kennedy Institute, London W6 7DW, U.K., ‡Department of Physiological Chemistry, University of Lund, S-22100 Lund, Sweden, and §Abteilung Bindegewebsforschung, Max-Planck-Institut für Biochemie, D-8033 Martinsried, Federal Republic of Germany

Electron microscopy after rotary shadowing and negative staining of the large chondroitin sulphate proteoglycan from rat chondrosarcoma, bovine nasal cartilage and pig laryngeal cartilage demonstrated a unique multidomain structure for the protein core. A main characteristic is a pair of globular domains (diameter 6–8 nm), one of which forms the *N*-terminal hyaluronate-binding region. They are connected by a 25 nm-long rod-like domain of limited flexibility. This segment is continued by a 280 nm-long polypeptide strand containing most chondroitin sulphate chains (average length 40 nm) in a brush-like array and is terminated by a small *C*-terminal globular domain. The core protein showed a variable extent of degradation, including the loss of the *C*-terminal globular domain and sections of variable length of the chondroitin sulphate-bearing strand. The high abundance (30–50%) of the *C*-terminal domain in some extracted proteoglycan preparations indicated that this structure is present in the cartilage matrix rather than being a precursor-specific segment. It may contain the hepatolectin-like segment deduced from cDNA sequences corresponding to the 3′-end of protein core mRNA [Doege, Fernandez, Hassell, Sasaki & Yamada (1986) J. Biol. Chem. **261**, 8108–8111; Sai, Tanaka, Kosher & Tanzer (1986) Proc. Natl. Acad. Sci. **83**, 5081–5085; Oldberg, Antonsson & Heinegård (1987) Biochem. J. **243**, 255–259].

INTRODUCTION

Cartilage proteoglycans are complex macromolecules (M_r 1×10^6–4×10^6) in which many chondroitin sulphate and keratan sulphate chains and oligosaccharides are linked to a large extended protein core (see Hardingham, 1986; Heinegård & Paulsson, 1984; Hassell *et al.*, 1986). It has been established that the protein core of the proteoglycan contains distinct regions. These include an *N*-terminal globular domain that provides the sites for interaction with hyaluronate and with link protein. It is a region of low carbohydrate content and contains no chondroitin sulphate (Bonnet *et al.*, 1985; Heinegård & Hascall, 1974). A major segment of the protein core contains most of the chondroitin sulphate chains and half the keratan sulphate chains. This region has a distinctive amino acid composition of predominantly serine, glycine, glutamine (glutamate) and proline. A further carbohydrate-rich region has been identified by its resistance to trypsin and chymotrypsin digestion and contains about 50% of the keratan sulphate of the typical proteoglycan (Heinegård & Axelsson, 1977). It appears to be located between the hyaluronate-binding region and the major extended chondroitin sulphate-bearing region.

This basic model of protein core structure is generally supported by electron-microscopic observations of proteoglycans spread in monolayer by the Kleinschmidt technique (Rosenberg *et al.*, 1970; Thyberg *et al.*, 1975; Heinegård *et al.*, 1978; Buckwalter & Rosenberg, 1982). Our initial results from rotary shadowing (Wiedemann *et al.*, 1984) showed in addition the presence of a second globular protein domain which had not previously been detected by biochemical analysis. In the present study the structures revealed by rotary shadowing have been comprehensively analysed and correlated with protein core regions identified in biochemical studies.

MATERIALS AND METHODS

Materials for the preparation of proteoglycan components from pig laryngeal cartilage were as described by Bonnet *et al.* (1985) and from bovine nasal-septal cartilage and rat chondrosarcoma as described by Heinegård *et al.* (1985).

Preparation of proteoglycans from cartilage

Proteoglycan monomers from the Swarm rat chondrosarcoma were prepared by extraction with 4 M-guanidinium chloride/50 mM-sodium acetate buffer, pH 5.8, containing the proteinase inhibitors 5 mM-benzamidinium chloride, 0.1 M-6-aminohexanoic acid and 10 mM-EDTA (Oegema *et al.*, 1975), and purified by direct CsCl-density-gradient centrifugation as described by Heinegård & Paulsson (1984). Proteoglycan aggregates and monomers from bovine nasal cartilage were prepared by extraction as above, followed by sequential associative and dissociative CsCl-density-gradient centrifugation (for details and references, see Heinegård & Paulsson, 1984). Bovine nasal-cartilage proteoglycans were subfractioned into subpopulations of M_r 1.3×10^6 and 3.5×10^6 respectively as described previously (Heinegård *et al.*, 1985). Proteoglycans were extracted

‖ To whom requests for reprints should be addressed.

from pig laryngeal cartilage in 4 M-guanidinium chloride in the presence of proteinase inhibitors and were purified in CsCl density gradients as described by Hardingham (1979).

Chondroitinase digestion of proteoglycan monomer

Proteoglycan core protein was prepared from bovine nasal-cartilage proteoglycan monomers by digestion with chondroitinase ABC (Seikagaku Kogyo Ltd., Tokyo, Japan) (10 units/g for 5 h at 37 °C) in the presence of ovomucoid (10 mg/ml) as a proteinase inhibitor, followed by chromatography of the digest on Sepharose 6B eluted with 4 M-guanidinium chloride to remove oligosaccharides, enzymes and inhibitors (Heinegård & Axelsson, 1977). Proteoglycan from pig laryngeal cartilage dissolved in 0.1 M-Tris/HCl, pH 7.3, containing proteinase inhibitors (as above; Oegema *et al.*, 1975) was digested with chondroitinase ABC (3.5 units/g) for 3 h at 37 °C, dialysed against 0.1 M-Tris/HCl, pH 7.3, and chromatographed on a column (11 mm × 1200 mm) of Sepharose CL 6B to purify the core protein.

Preparation of proteoglycan fragments

Hyaluronate-binding region was obtained by digestion of bovine nasal-cartilage proteoglycan aggregates (10 mg/ml) with trypsin (40 μg/ml for 8 h at 37 °C), followed by purification of the resulting hyaluronate-binding-region–link-protein complex by associative CsCl-density-gradient centrifugation and Sepharose 2B gel filtration and, finally, dissociation of the complex in 4 M-guanidinium chloride and isolation of binding region by chromatography on Sephadex G-200 in 4 M-guanidinium chloride (Heinegård & Hascall, 1974; Heinegård & Axelsson, 1977). The hyaluronate-binding region and keratan sulphate-rich fragment were prepared from pig laryngeal cartilage as described by Bonnet *et al.* (1985). The fragment containing both globules G1 and G2 was prepared from pig laryngeal cartilage proteoglycan aggregates by mild trypsin digestion (200 μg of trypsin/g of aggregate for 10 min at 37 °C) and appeared as a product of intermediate stability during digestion. It was purified by size-exclusion chromatography on a column (7 mm × 300 mm) of TSK-4000 run in 4 M-guanidinium chloride/0.5 M-sodium acetate, pH 5.8. Iodination of the double globe preparation with ^{125}I was performed with chloramine-T as the oxidizing agent (Greenwood *et al.*, 1963) as described by Ratcliffe & Hardingham (1983). Bovine serum albumin (100 μg/ml) was added as carrier protein after iodination.

SDS/polyacrylamide-gel electrophoresis

Electrophoresis was carried out on slab gels [5% (w/v) polyacrylamide] essentially as described by Fairbanks *et al.* (1971). Protein bands were detected by silver staining (Morrissey, 1981) and gels were autoradiographed using Kodak X-Omat S film and intensifying screens (Cronex Lightning Plus; du Pont). The autoradiographs were scanned with a Shimadzu densitometer.

Electron microscopy

Samples were stored in aqueous solutions at concentrations of about 1 mg/ml in 50 μl aliquots at −20 °C and diluted into 0.2 M-NH_4HCO_3 shortly before use. The dilute proteoglycan or fragment solution (about 20 μg/ml) was mixed with an equal volume of glycerol immediately before spraying on to mica discs and shadowed on a rotating table with carbon/platinum at a 9° angle. For negative staining, particles were adsorbed from 0.2 M-NH_4HCO_3 solution on to carbon films, which were rendered hydrophilic by glow discharge. Staining was performed with a freshly prepared saturated uranyl formate solution of pH 4. Measurements were performed at a total magnification of 500000 on projections of the negatives on to a screen. Details of the electron-microscopic procedures and methods of data evaluation were described previously (Engel *et al.*, 1981).

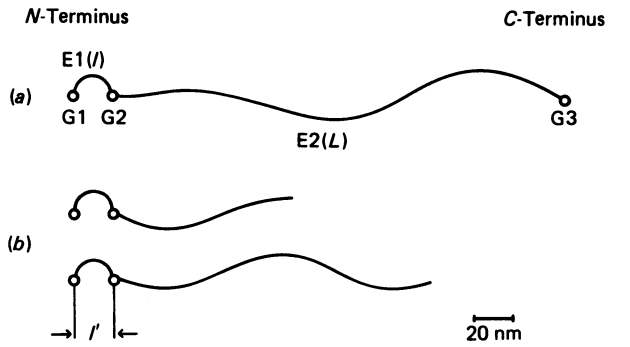

Fig. 1. Schematic representation of proteoglycan protein cores as visualized by rotary shadowing

(*a*) Intact particles (their fraction varies from 10 to 50%, depending on the preparation) exhibit two globular domains, G1 and G2, at the *N*-terminus, linked by an extended domain, E1, followed by a long strand, E2, which is terminated by a third globular domain, G3. The lengths $l = 30$ nm and $L = 280$ nm were measured between the centres of the terminating globes following the contours of E1 and E2 respectively. (*b*) Core particles which lack the *C*-terminal globular domain G3 exhibit lower values and a much broader distribution of L (average value 182 nm). The *N*-terminal part of these molecules appears to be unaltered. l' (= 21 nm) denotes the centre-to-centre distance between G1 and G2. Histograms of the length dimensions l, l' and L are shown in Figs. 3 and 5.

RESULTS

Proteoglycan monomers

Proteoglycans are observed by the use of rotary shadowing as extended particles with brush-like arrays of fine filaments in their central position and globular morphological units at both ends. These prominent features were most distinctly seen with monomers derived from Swarm rat chondrosarcoma (Figs. 1*a* and 1*b*), but very similar images were obtained for material from pig laryngeal or bovine nasal-septum cartilage (see also Wiedemann *et al.*, 1984).

The filaments in the particles correspond to the chondroitin sulphate chains, which with this technique usually remain extended (average length 40 nm) and do not form bundles or associate with the protein core as is frequently observed with the Kleinschmidt technique (Rosenberg *et al.*, 1975; Heinegård *et al.*, 1978). The glycosaminoglycan chains are revealed by their decoration with metal crystallites of 2–3 nm diameter. It is mainly the formation of ordered rows of crystallites which distinguishes them from the background of

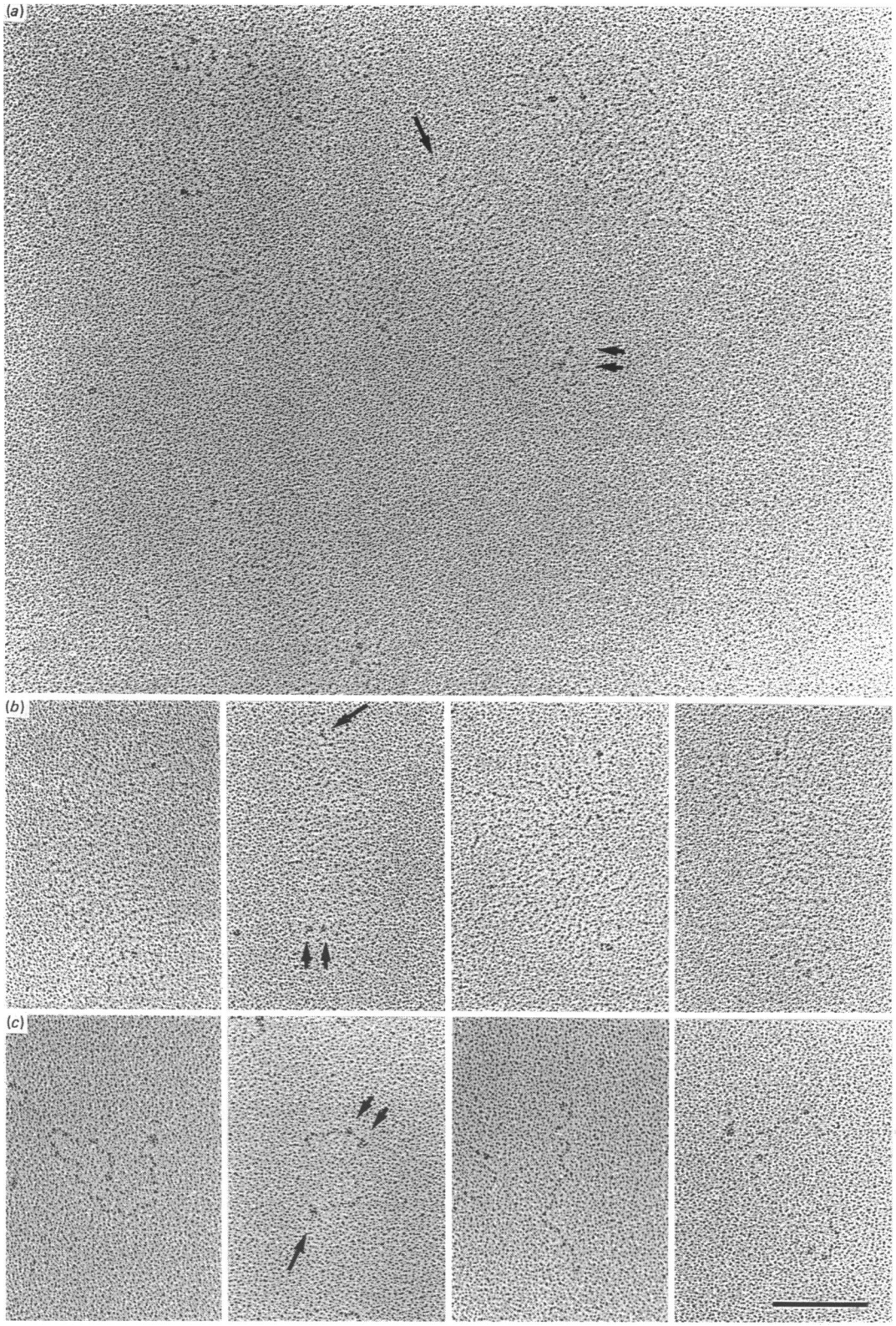

Fig. 2. Electron micrographs after rotary shadowing of (*a*) a field of proteoglycan monomers isolated from rat chondrosarcoma, (*b*) selected species of the same material, and (*c*) selected species of proteoglycan monomers from the same source in which the glycosaminoglycan side chains were collapsed to the protein core during sample preparation

Pairs of globules (distance about 21 nm) at one end of the molecules are indicated by short arrows and the globular domain at the opposite end (*C*-terminus) by a long arrow. The bar represents 100 nm. Originals of the photographs are available from J. E.

randomly orientated metal particles. Because of the weak decoration, however, images are of very low contrast, similar to those of heparan sulphate chains from other proteoglycans (Fujiwara *et al.*, 1984; Paulsson *et al.*, 1986) and of hyaluronate chains (M. Mörgelin, M. Paulsson, T. E. Heinegård & T. Engel, unpublished work). Because of the low contrast it is difficult to reproduce the images in print. It is easier to trace the chains by projecting the original negatives on a screen, but even then it is difficult to determine the number or length of the chondroitin sulphate side chains consistently. In the most favourable instances, up to 50 side chains were evident on each monomer of length 300 nm. This is greater than the number previously observed by electron microscopy using the Kleinschmidt technique (Thyberg *et al.*, 1975; Buckwalter & Rosenberg, 1982), but it is less than the estimated number of chains from composition and M_r analyses (Heinegård *et al.*, 1985; Hardingham, 1986). Occasionally a thin central strand can be observed which probably represents the polypeptide chain of the protein core.

The most striking new feature revealed by the rotary-shadowing technique is the presence of three more heavily decorated globules which, according to the experience gained by imaging many other multidomain proteins by the same technique (Engel & Furthmayr, 1987), may be interpreted as three distinct globular domains. Most of the particles showed, at one end, two adjacent globular domains G1 and G2 (Figs. 2*a* and 2*b*) spaced at 21 ± 4 nm distance (Fig. 3). A fraction of the molecules exhibits a third globular domain (G3) at the opposite end. The fraction of monomers possessing G3 varied among different proteoglycan preparations. It was highest (about 50%) for the proteoglycan from Swarm rat chondrosarcoma and amounted to 30% for proteoglycan from mature bovine nasal cartilage.

In intact proteoglycan monomers the extended and globular domains of the protein core are somewhat obscured by the many side-chain filaments. The occasional appearance of proteoglycan monomers with collapsed side chains was found to provide a clearer display of the protein core and the position of the globular domains (Fig. 2*c*). The collapsed chondroitin sulphate chains outlined the central filament E2 (for definition, see Fig. 1) with a much higher contrast than seen for a single polypeptide strand (Fig. 2*c*). The pair of globules at one end and the third globule at the opposite end are more clearly seen than in Figs. 2(*a*) and 2(*b*), and it is also more evident that globules G1 and G3 represent terminal domains. A schematic representation of these findings is given in Fig. 1(*a*). The collapsed structures of rat chondrosarcoma proteoglycans shown in Fig. 2(*c*) were obtained when freeze-dried samples were redissolved, whereas in most other experiments samples were stored frozen and diluted into the volatile buffer shortly before use. The collapse of side chains could, however, not be reproducibly obtained by freeze-drying, and it cannot be excluded that other factors in the preparation of samples for electron microscopy gave rise to this phenomenon. The lack of change in hexuronate content showed that the side chains were not lost from the preparation.

Protein core

Protein cores were prepared from bovine nasal-septum proteoglycan by chondroitinase treatment and examined by rotary shadowing (Fig. 4*a*) and negative staining (Fig. 4*b*). In its general appearance the protein core resembles the proteoglycan monomers with collapsed side chains (Fig. 2*c*), except that the long strand of the core between domains G2 and G3 was found to be decorated by metal to a lesser extent. The decoration is still, however, more pronounced than expected for a single polypeptide chain. This may be due to short oligosaccharide units of the chondroitin sulphate chains remaining after chondroitinase treatment, as well as to the presence of keratan sulphate and other oligosaccharides along the protein core (Heinegård & Paulsson, 1984).

The globular domains G1–G3 (see Fig. 1) are clearly visible in images after rotary shadowing (Fig. 4*a*). The fraction of particles which possessed domain G3 was about the same as for the original proteoglycan monomers, indicating that little or no proteolytic degradation had occurred during the chondroitinase

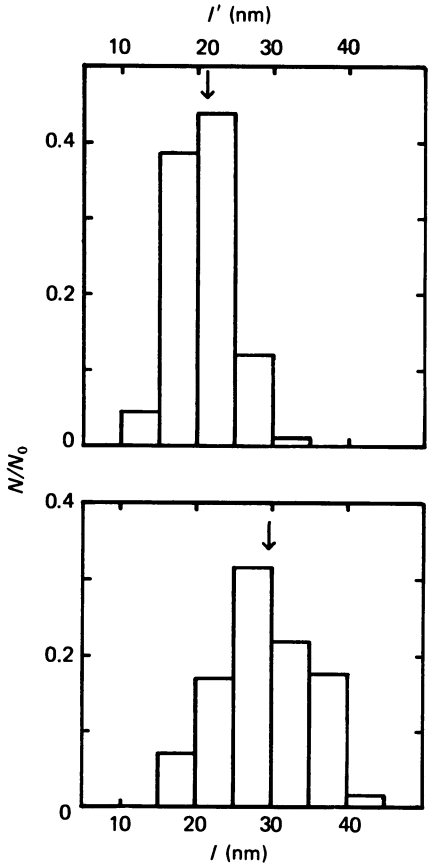

Fig. 3. Distribution of the centre-to-centre distance *l′* between globular domains G1 and G2 and of the length *l* measured between the centres of the globules, but tracing along the contour of the strand E1 connecting these domains

For definitions, see Fig. 1. N/N_0 denotes the fraction of molecules with a centre-to-centre distance (l') of ± 5 nm or a contour length (l) of ± 5 nm respectively. Average values (arrows) $\pm$S.D. are: l, 30 ± 5 nm; l', 21 ± 3 nm. The contour length of E1 (25 nm) is obtained by subtraction of the radii of G1 and G2 from the value of l. The distributions shown were measured on protein cores derived from bovine nasal-septum cartilage, but almost identical distributions were found for material from rat chondrosarcoma and from pig laryngeal cartilage.

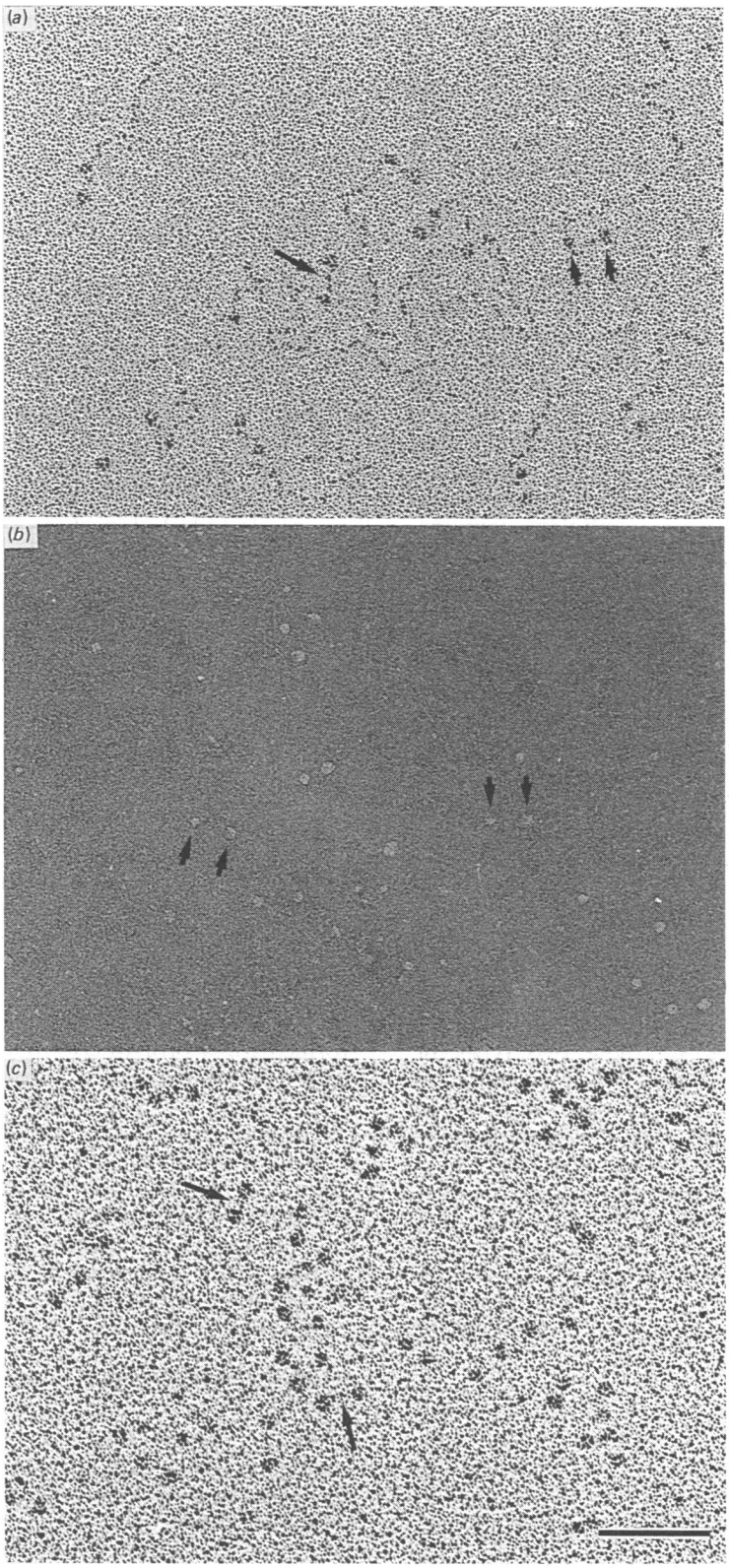

Fig. 4. Electron micrographs of proteoglycan protein core after rotary shadowing (*a*) and negative staining (*b*), and of double-globe fragment after rotary shadowing (*c*)

By the negative-staining technique only globular domains with a diameter of 6–8 nm are visible. Pairs of globules are indicated by short arrows. The connecting strand between the two globules is seen for some species after rotary shadowing (long arrows in *a* and *c*). The protein core shown was isolated from bovine nasal septum, but similar images were obtained for core protein prepared from pig laryngeal cartilage. Double-globe fragment was prepared from pig laryngeal cartilage as described in the Materials and methods section. The bar represents 100 nm.

digestion. Globular domains could also be revealed by negative staining of the protein core (Fig. 4*b*), but not the extended regions of the protein core. Pairs of globules apparently corresponding to domains G1 and G2 could be clearly identified. In these pairs one of the globes appeared with a somewhat larger diameter (8 ± 1.2 nm) than the other (6 ± 1.4 nm) in measurements of about 100 particles. Because of the lack of visibility of the extended domains E1 and E2, an unambiguous assignment to domains G1 and G2 was not possible. A tentative assignment of the larger globule to domain G1 can be made on the basis of rotary-shadowed material, where the terminal globule G1 appeared with a larger diameter than G2 (11 ± 2 nm as compared with 9 ± 2 nm). It is known from experience with other systems (Engel & Furthmayr, 1987) that diameters measured by the rotary-shadowing technique are overestimated by 2–3 nm due to the decoration of protein by metal crystallites, whereas the negative-staining technique provides approximately correct values for globular domains. Domain G3 could not be identified with certainty in negatively stained preparations. The double globule structures of Fig. 4(*b*), however, were always accompanied by a small globule (diameter 3.5 ± 1 nm) showing a random distribution. This globule may correspond to G3, as is also indicated by rotary-shadowing, which gives a diameter of 3.5 nm for this domain. Negative staining of intact proteoglycan monomers produced pictures (not shown) very similar to those shown for the core protein in Fig. 4(*b*). Again, only globular domains, but no other parts of the molecules, were visible.

For a preparation of nasal cartilage proteoglycan the length dimensions of the extended domains E1 and E2 as defined in Fig. 1 were measured for about 300 rotary-shadowed protein cores possessing the third globule G3 and for about the same number of particles which lacked this domain. With regard to the length of the strand E1 connecting domains G1 and G2, no difference was observed between these two types of particles. Tracing along the strand E1, the length (l) between the centres of G1 and G2 was found to be 30 ± 5 nm (Fig. 3). By subtracting the radii of the globules, the length of the linking segment E1 is approximately 25 nm. The direct centre-to-centre distance between G1 and G2 was also measured, and a value of 21 ± 3 nm was found. This value and its distribution (Fig. 3) corresponds well with the value of $l' = 21 \pm 4$ nm determined for intact proteoglycans. The contour length of the connecting strand E1 could not be measured in intact monomers, since it was obscured by the side chains.

Large differences were found for the length, L, of the long strand E2, depending on whether or not the terminal domain G3 was present (Fig. 5). When the strand was terminated by domain G3, a relatively sharp distribution with a maximum at $L = 280$ nm and an average value of $\langle L \rangle = 260 \pm 39$ nm were observed. For particles lacking the third globule the distribution was found to be much broader (average value of $L = 182 \pm 56$ nm), with not very well defined maxima at 200 and 260 nm. For about 50% of these particles the length of the long core strand E2 was even found to be smaller than 200 nm, whereas such short lengths of E2 were not observed when E2 was terminated by domain G3.

Similar measurements (not shown) were also performed for a protein core prepared from pig laryngeal cartilage. Almost identical results were obtained for l and l' of domain E1. This particular preparation, however, did not contain particles with a terminal domain G3 and showed a much wider distribution of L (140 ± 100 nm) than the preparation from bovine nasal septum. Preliminary findings with other preparations indicated that the length distribution of the extended domain in proteoglycans not only depend on the source of cartilage but also on other parameters, such as the extraction and purification procedure.

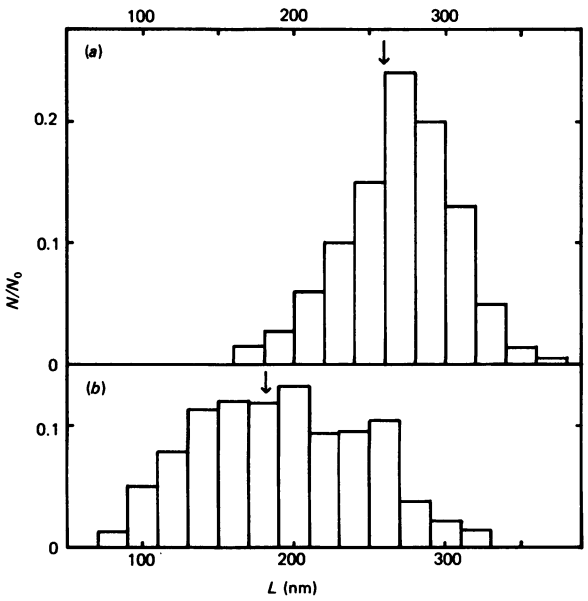

Fig. 5. Comparison of the length distributions of the long core segment (E2) following the globular domain (G2) (see Fig. 1) for protein cores which possess the *C*-terminal globular domain G3 (*a*) and for particles in which this domain is not present (*b*)

N/N_0 denotes the fraction of molecules in which the long core segment has a length (L) of ± 10 nm. Average values (arrows) $\pm$S.D. are: for (*a*), 260 ± 39 nm; for (*b*), 182 ± 56 nm. The value of l for (*a*) at the maximum 280 nm, and the distribution (*b*) does not show a clear maximum. About 300 particles each were measured in (*a*) and (*b*) respectively. The material was obtained from bovine nasal-septum cartilage and contained about 30% of particles with a *C*-terminal globule (G3).

Isolated binding region (G1) and double globe fragment (G1–E1–G2)

Binding region was prepared from proteoglycan aggregates from bovine nasal septum and from pig laryngeal cartilage by proteolysis with trypsin under standard conditions. When the binding-region preparations were examined by rotary shadowing (Fig. 6*a*) or by negative staining (Fig. 6*b*), large fields with globular particles of uniform size became apparent. The diameters measured by the two techniques were 11 ± 1 nm and 8 ± 1 nm respectively. No differences were found between preparations from bovine nasal and pig laryngeal cartilage.

A larger fragment containing binding region was generated by milder proteolysis with trypsin. This fragment migrated on SDS/polyacrylamide-gel electro-

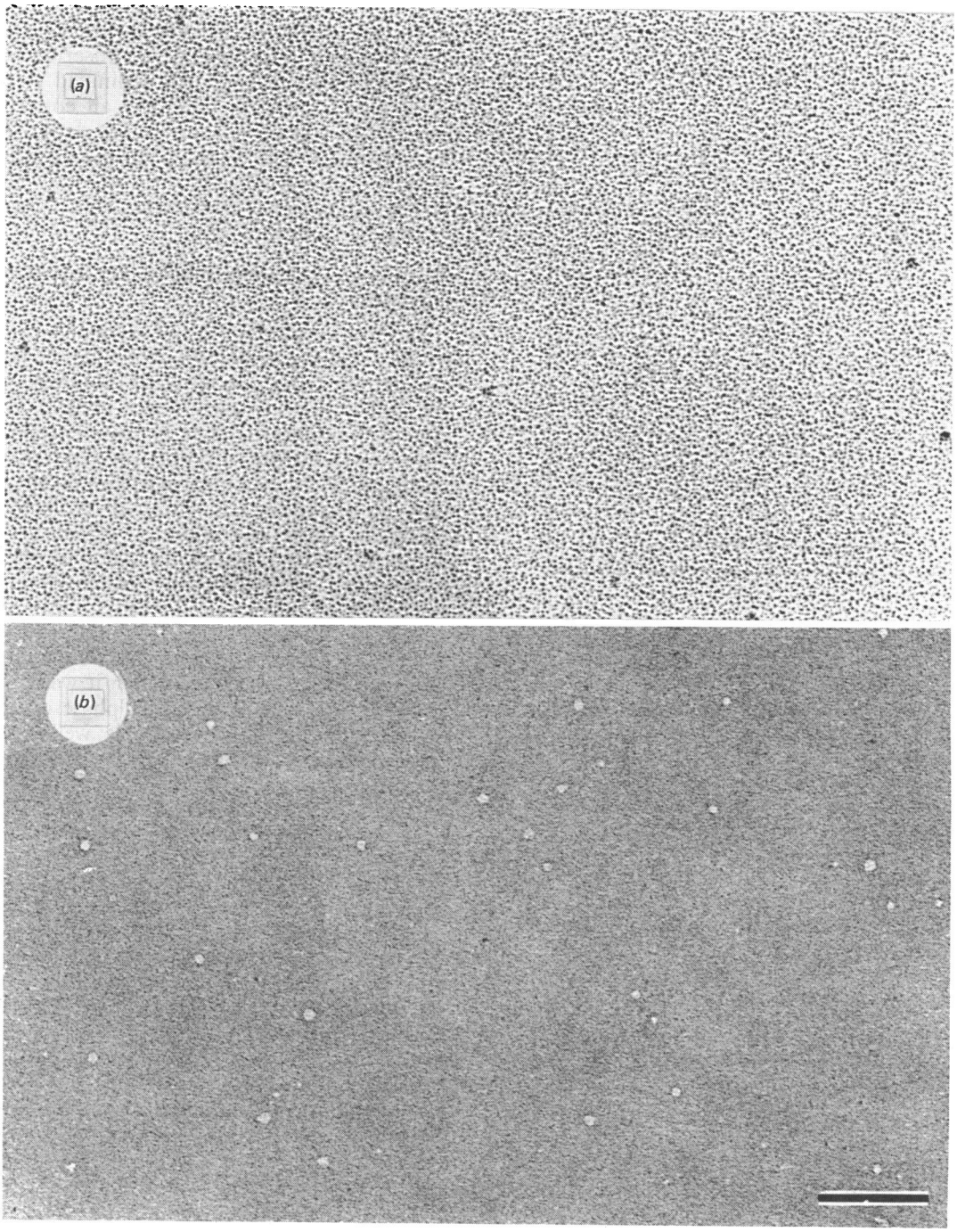

Fig. 6. Electron micrographs of isolated binding region after rotary shadowing (*a*) and negative staining (*b*)

The particles appear as monomeric globules without any indication of aggregate formation. The electron micrograph shows binding region derived from bovine nasal-septum cartilage. Identical pictures were obtained for binding region derived from pig laryngeal cartilage. The bar represents 100 nm.

phoresis as a single broad band with apparent M_r 150000. The band sharpened and was slightly decreased in size after keratanase digestion, but was unaffected by chondroitinase ABC (M. Beardmore-Gray, D. Dunham & T. Hardingham, unpublished work). The fragment thus contained some keratan sulphate but no chondroitin sulphate. Electron-microscopic inspection of the 150000-M_r fragment revealed that it comprised the double-globe region of the proteoglycan monomers (Fig. 4*c*). The contour length of the connecting strand, *l*, and the centre-to-centre distance between the two globules was found to be essentially identical in the double-globe fragment and in the core protein.

Evidence that the two globular domains, G1 and G2, are structurally dissimilar was obtained in experiments on fragments derived from the double-globe preparation. In initial attempts to iodinate the double-globe preparations, some partial fragmentation occurred. This

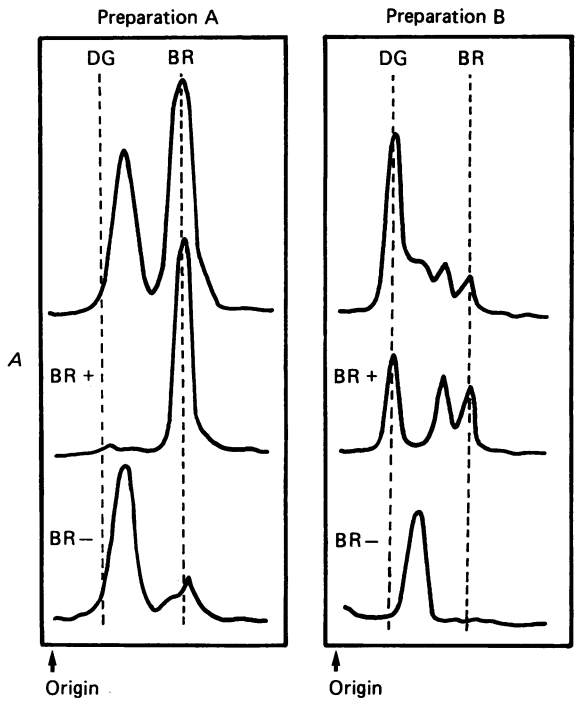

Fig. 7. SDS/polyacrylamide-gel-electrophoretic analysis of ^{125}I-labelled double-globe preparations A and B

Samples were iodinated and electrophoresed as described in the Materials and methods section. Immunoprecipitates with anti-(binding region) antiserum were isolated by using *Staphylococcus aureus* (Ratcliffe & Hardingham, 1983). Samples of the pellet and supernatant were electrophoresed to reveal binding-region-positive (BR+) and binding-region-negative (BR−) fragments respectively. Autoradiographs of the gels were scanned with a densitometer. The mobilities of unlabelled double-globe fragment (DG) and binding region (BR) are indicated (vertical broken line).

may have resulted from proteinase contamination of the bovine serum albumin added after iodination as carrier protein. On the first occasion (Preparation A), two major products were formed, and on repeating the procedure (Preparation B), less degradation occurred. Subsequent iodinations with a new batch of bovine serum albumin as carrier did not result in any detectible cleavage of the double-globe fragment. However, the fragments created in the first two iodinations proved useful in further investigations of the structure. Preparation A contained two major components on SDS/polyacrylamide-gel electrophoresis both smaller than intact double globes (Fig. 7). Immunoprecipitation with an antiserum specific for binding region (G1) (Ratcliffe & Hardingham, 1983) showed that only the smaller component (M_r 65000) was recognized, whereas the large fragment (M_r 110000) was not precipitated. With preparation B, where less fragmentation had occurred, binding region was present in three components. The largest of these was of similar mobility to intact double globe, and the smallest was similar to isolated binding region. An intermediate fragment (M_r 90000) containing binding region was also present. A large fragment (M_r 110000) was again not precipitated by the antibodies to binding region. These results showed that a large fragment derived from the double-globe structure did not contain any of the epitopes recognized by anti-(binding region) antibodies and was therefore of different structure. In further experiments with a polyclonal antiserum raised against intact proteoglycan, all the iodinated fragments were precipitated, showing that they all contained epitopes present on the protein core. The large fragment not recognized by anti-(binding region) was thus shown to contain its own distinct protein-core epitopes. The protein structures of G1 (binding region) and G2 are thus concluded to be different.

M_r estimates from electron micrographs of proteoglycan monomers

Electron micrographs of proteoglycan monomers usually reveal a distribution of sizes. This is, however, difficult to quantify in terms of molecular mass because the difficulties in tracing the protein core and the glycosaminoglycan chains in these complex molecules. An empirical relationship between the M_r and the area occupied by the monomer particles on the mica surface may be useful for estimation of molecular mass and particular mass distributions, which are difficult to obtain by other methods. In order to establish such a relationship, fractions of proteoglycan monomers isolated from bovine nasal-septum cartilage (Heinegård *et al.*, 1985) were studied by the rotary-shadowing technique. Particles exhibited a similar appearance to those shown in Fig. 2. Although it was not possible to trace individual chondroitin sulphate chains, the areas occupied by the molecules was well defined by the brush-like arrays of fine filaments (see Fig. 2). For two fractions, M_r 1.3×10^6 and 3.5×10^6 (Heinegård *et al.*, 1985), the average areas per molecule were determined to be 7500 ± 2800 nm^2 and 19200 ± 5300 nm^2 respectively. The large standard deviations may not only be due to errors in the determinations of the areas of about 30 particles, but also to a distribution of sizes within each fraction. Ratios of molecular mass to area were calculated to be 172 nm^{-2} and 182 nm^{-2} for the low- and high-molecular mass fractions respectively. The value for the molecular mass per unit area may also depend on the conditions of specimen preparation, and it is therefore advisable to use calibration standards of known molecular mass. These should be prepared for microscopy under conditions identical with those used for the samples of interest.

DISCUSSION

Important information on the principal structure of proteoglycan molecules has been obtained from earlier electron-microscopic work using the Kleinschmidt technique. This technique, by which the molecules are spread on a surface film of cytochrome *c* and imaged after negative staining, however, failed to provide sufficient details of the protein part of the molecules. When rotary shadowing after spraying from glycerol/buffer was applied to cartilage proteoglycans (Wiedemann *et al.*, 1984), a number of interesting features of the protein core became apparent, including two globular domains, which were desginated 'G1' and 'G2' in the present study. Applying the same technique it was now possible to detect a third globular domain ('G3') and to derive accurate dimensions of the two extended domains (designated 'E1' and 'E2') of the protein core (Fig. 1).

The detection of domain G3 is of particular relevance

in view of recent sequence data obtained by cDNA cloning (Doege *et al.*, 1986; Sai *et al.*, 1986; Oldberg *et al.*, 1987). A clone which corresponded to the 3'-end of RNA was isolated and found to encode for about 50 residues of a sequence in which, on average, every third residue is either serine or threonine, followed by 220 residues of a sequence with characteristic features for a globular domain. It was concluded that the serine-and-threonine-rich segment represents the *C*-terminal portion of the chondroitin sulphate-attachment region of the protein core, which is terminated by a globular domain of molecular mass about 20 kDa. Our electron-microscopic findings indicate that a terminal globular domain of this size is indeed expressed in the protein. Its presence in a significant proportion of proteoglycan monomers from mature cartilage suggest that this domain is an integral part of proteoglycans which is still present in the proteoglycan after secretion and matrix deposition. The relatively constant length of the extended domain E2 when terminated by domain G3 suggests that the protein core is initially synthesized with a defined length, in support of earlier studies of biosynthesis showing that polydispersity among newly synthesized proteoglycans is mainly due to variation in glycosylation (Fellini *et al.*, 1981). The observation of shorter core proteins with a much wider distribution of the length of E2 when G3 is missing supports the idea that slow continuous degradation of the core protein occurs while proteoglycans are in the extracellular matrix. It has been suggested that this degradation starts with the chondroitin sulphate region and accounts in part for the observed size heterogeneity of mature proteoglycans (for references, see Heinegård & Paulsson, 1984). The concept of a slow degradation starting from the *C*-terminal end is also consistent with the observation that the fraction of molecules with the terminal globule G3 is largest in the material from Swarm rat chondrosarcoma, which, because of rapid growth, may consist mainly of newly synthesized proteoglycan. More work is needed to exclude the less likely possibility that domain G3 is present only in a precursor form of the proteoglycan and is removed before incorporation into the cartilage matrix, or that it is present only on a genetically distinct subpopulation of proteoglycans. At present the function of the newly discovered domain is not known. Since it shows sequence homology with chick hepatic lectin (Doege *et al.*, 1986; Sai *et al.*, 1986), it was speculated that the domain interacts with some other matrix component by providing a binding site for *N*-acetylglucosamine.

The *N*-terminal domain G1 represents the hyaluronate-binding domain. This was most clearly shown from studies of the interaction of core protein, 150 kDa fragment and isolated binding region with hyaluronate (Hardingham *et al.*, 1986; M. Mörgelin, M. Paulsson, T. E. Hardingham, D. Heinegård & T. Engel, unpublished work). These studies also demonstrated that domain G2 is neither directly involved in the binding to hyaluronate nor in the stabilization of this interaction by link protein.

The immunochemical results also showed that G1 and G2 were structurally dissimilar. The size of the double-globe fragment also excluded the possibility that G2 represented the keratan sulphate-attachment region of the protein core previously identified (Heinegård & Axelsson, 1977). The isolated keratan sulphate region migrates on SDS/polyacrylamide-gel electrophoresis with an apparent M_r of 140000, as large as the whole double-globe structure; furthermore, it appears extended, with no globular protein domains, when examined by rotary shadowing (M. Mörgelin, M. Paulsson, J. Engel, D. Dunham & T. E. Hardingham, unpublished work).

Domain G2 is connected with the hyaluronate-binding domain G1 by another extended strand-like domain of the core E1. The length of this segment was found to be very constant (25 nm) in all monomers and fragments studied. It keeps G1 and G2 at an average centre-to-centre distance of 21 nm and exhibits limited flexibility. The functional significance of this arrangement and of domain G2 is not clear at present.

With the known M_r of the protein part of isolated binding region (domain G1) of about 50000 [the M_r was determined to be 65000, with about 15000 made up of glycoconjugate (Bonnet *et al.*, 1985)], an estimated M_r of domain G2 of 40000 and that of G3 about 20000 (from sequence data), the minimum M_r of the globular portions of the core protein is estimated to be 115000. The minimum value for the strands E1 (25 nm) and E2 (280 nm) may be estimated from their lengths, with an M_r/length ratio for the completely extended polypeptide chain (M_r/L) of 330 nm^{-1}. M_r values are 8000 for E1 and 100000 for E2. The resulting minimum M_r for the unglycosylated protein core is 225000. This value is significantly lower than the estimates of 340000 for core protein translated in cell-free systems (Upholt *et al.*, 1979; Treadwell *et al.*, 1980). A more satisfactory agreement would be obtained when an α-helical conformation ($M_r/L = 750$ nm^{-1}) is assumed for the extended domains. With this assumption the calculated M_r for the core protein is 345000. Further work is needed to elucidate details of the structure of the various protein domains in cartilage proteoglycan and to learn more about their specific functions.

We gratefully acknowledge the support of the Swedish Medical Research Council, Folksams yrkesskadors stiftelse, the Medical Research Council (U.K.), the Arthritis and Rheumatism Council (U.K.) and the Swiss National Science Foundation.

REFERENCES

Bonnet, F., Dunham, D. G. & Hardingham, T. E. (1985) Biochem. J. **228**, 77–85

Buckwalter, J. A. & Rosenberg, L. A. (1982) J. Biol. Chem. **257**, 9830–9839

Doege, K., Fernandez, P., Hassell, J. R., Sasaki, M. & Yamada, Y. (1986) J. Biol. Chem. **261**, 8108–8111

Engel, J. & Furthmayr, H. (1987) Methods Enzymol. **145**, 1–47

Engel, J., Odermatt, E., Engel, A., Madri, J. A., Furthmayr, H., Rohde, H. & Timpl, R. (1981) J. Mol. Biol. **150**, 97–120

Fairbanks, G., Steck, T. L. & Wallach, D. F. H. (1971) Biochemistry **10**, 2606–2617

Fellini, S. A., Kimura, J. H. & Hascall, V. C. (1981) J. Biol. Chem. **256**, 7883–7889

Fujiwara, S., Wiedemann, H., Timpl, R., Lustig, A. & Engel, J. (1984) Eur. J. Biochem. **143**, 145–157

Greenwood, E. C., Hunter, W. M. & Glover, J. S. (1963) Biochem. J. **89**, 114–123

Hardingham, T. E. (1979) Biochem. J. **177**, 237–247

Hardingham, T. E. (1986) in Rheumatology, Vol. 10, Connective Tissue: Biological and Clinical Aspects (Kühn, K. & Krieg, T., eds.), pp. 143–183, Karger, Basel

Hardingham, T. E., Beardmore-Gray, M., Dunham, D. G. & Ratcliffe, A. (1986) Ciba Found. Symp. **124**, 30–46
Hassell, J. R., Kimura, J. H. & Hascall, V. C. (1986) Annu. Rev. Biochem. **55**, 539–567
Heinegård, D. (1977) J. Biol. Chem. **252**, 1980–1989
Heinegård, D. & Axelsson, I. (1977) J. Biol. Chem. **252**, 1971–1979
Heinegård, D. & Hascall, V. C. (1974) J. Biol. Chem. **249**, 4250–4256
Heinegård, D. & Paulsson, M. (1984) in Extracellular Matrix Biochemistry (Piez, K. & Reddi, A., eds.), pp. 277–328, Elsevier, New York
Heinegård, D., Lohmander, S. & Thyberg, J. (1978) Biochem. J. **175**, 913–919
Heinegård, D., Wieslander, J., Sheehan, J., Paulsson, M. & Sommarin, Y. (1985) Biochem. J. **225**, 95–106
Morrissey, J. H. (1981) Anal. Biochem. **117**, 307–310
Oegema, T. R., Hascall, V. C. & Dziewiatkowski, D. D. (1975) J. Biol. Chem. **250**, 6151–6159
Paulsson, M., Fujiwara, S., Dziadek, M., Timpl, R., Pejler, G., Bäckström, G., Lindahl, U. & Engel, J. (1986) Ciba Found. Symp. **124**, 189–203
Périn, J. P., Bonnet, F., Jollès, J. & Jollès, P. (1984) FEBS Lett. **176**, 37–42
Ratcliffe, A. & Hardingham, T. (1983) Biochem. J. **213**, 371–378
Rosenberg, L., Hellmann, W. & Kleinschmidt, A. K. (1970) J. Biol. Chem. **245**, 4123–4130
Rosenberg, L., Hellmann, W. & Kleinschmidt, A. (1975) J. Biol. Chem. **250**, 1877–1883
Sai, S., Tanaka, T., Kosher, R. A. & Tanzer, M. L. (1986) Proc. Natl. Acad. Sci. **83**, 5081–5085
Thyberg, J., Lohmander, S. & Heinegård, D. (1975) Biochem. J. **151**, 157–166
Treadwell, B. V., Mankin, D. P., Ho, P. K. & Mankin, H. J. (1980) Biochemistry **19**, 2269–2275
Upholt, W. B., Vertel, B. M. & Dorfman, A. (1979) Proc. Natl. Acad. Sci. U.S.A. **76**, 4847–4851
Wiedemann, H., Paulsson, M., Timpl, R., Engel, J. & Heinegård, D. (1984) Biochem. J. **224**, 331–333

Received 27 November 1986/11 February 1987; accepted 23 April 1987