Biochemistry

 Volume 35, Number 27 July 9, 1996

New Concepts in Biochemistry

Proteodermatan and Proteokeratan Sulfate (Decorin, Lumican/Fibromodulin) Proteins Are Horseshoe Shaped. Implications for Their Interactions with Collagen†

John. E. Scott*

Department of Chemical Morphology, School of Biological Sciences, Chemistry Building, Manchester University, Oxford Road, Manchester M13 9PL, U.K.

Received April 1, 1996; Revised Manuscript Received May 14, 1996⊗

ABSTRACT: The small proteoglycans proteodermatan and proteokeratan sulfates organize collagen fibrils in extracellular matrix [Scott, J. E. (1992) *FASEB J. 6*, 2639−2645], thus helping to maintain tissue shape. Their interaction with fibrils is probably via the protein. They have been examined by rotary shadowing-electron microscopy, which showed that these leucine-rich-repeat proteins are horseshoe shaped. Morphometry and comparison with polypeptide sequences suggest ways in which decorin could interact with tissue collagen fibrils. It is proposed that decorin is a bidentate ligand attached to two parallel neighboring collagen molecules in the fibril, helping to stabilize fibrils and orient fibrillogenesis.

Two types of small proteoglycans (PGs),[1] proteodermatan and proteokeratan sulfates (PDSs and PKSs), containing leucine-rich-repeat proteins (Scott, P. G., 1993) are found in connective tissue extracellular matrices. They include decorin, biglycan (PDSs), lumican, and fibromodulin (PKSs). In this paper these terms refer to the protein cores, i.e., the gene product (Scott, 1994). One or more anionic glycosaminoglycan chains based on the polylactose backbone (Scott, 1992) are attached covalently to the protein. Electron micrographs of rotary shadowed decorin-containing PDSs showed that they were tadpole-shaped, with a dermochondan (formerly dermatan) sulfate (DS) glycan tail attached to a decorin protein head that appeared to be roughly globular, at the magnifications employed (Ward et al., 1987).

PDSs and PKSs are associated with collagen fibrils at specific binding sites, probably via the protein (Scott, 1988). Ultrastructural evidence indicates that they play a role in orienting and organizing the collagen fibrils, thus helping to establish and maintain tissue shape (Scott, 1992). The glycan tails form antiparallel aggregates between their protein carriers, which are attached to collagen fibrils at binding sites occurring regularly along the fibril. These aggregate bridges thus hold the fibrils in register, at a distance apart determined by the length of the glycan chain. The ability to aggregate is determined by the secondary structure of the glycans (Scott, et al., 1995).

The amino acid sequences of decorin, biglycan, fibromodulin, and lumican are known [see Scott, P. G. (1993) for a review], but evidence for the location of their collagen-binding regions is limited and sometimes conflicting. Knowledge of their secondary structures would greatly facilitate understanding of their interactions, but this is available only in spectroscopic and modelling terms, which suggested the presence in decorin of α-helix and β-turns (Scott P. G., 1993).

Recently, the first high-resolution X-ray study of one of the LRRPs (ribonuclease inhibitor, RI) reported a remarkably regular molecule, containing 15 tandem repeats of ∼28 residue sequences of amino acids, each comprising a β-turn and an α-helix. The β-turns formed a sheet, to one side of

† Supported by the Medical Research Council, London, U.K.
* Telephone: (44) 161 273 5330. Fax: (44) 161 275 4598.
⊗ Abstract published in *Advance ACS Abstracts*, June 15, 1996.
[1] Abbreviations: dermochondan (dermatan) sulfate, DS; leucine-rich-repeat protein, LRRP; proteoglycan, PG; proteodermochondan sulfate, PDS; proteokeratan sulfate, PKS; ribonuclease inhibitor, RI.

S0006-2960(96)00773-8 CCC: $12.00

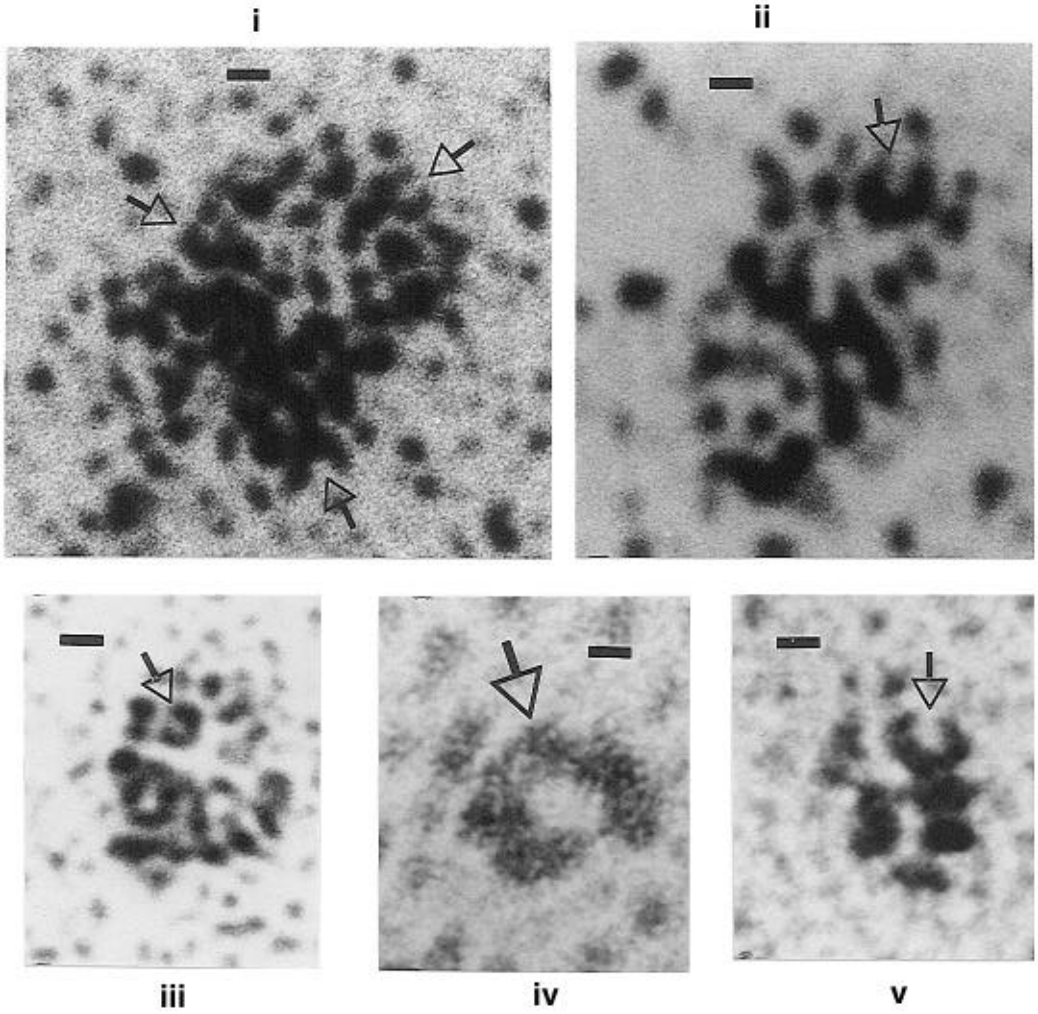

FIGURE 1: (i−iii) Small PDSs from bovine sclera isolated and prepared for rotary shadowing and electron microscopy (Ward et al., 1987). i and ii were Figure 3 (vi and xi) from Ward et al., further enlarged 12×; iv and v are bovine corneal PKSs. Horseshoes are arrowed. Bars (top of each picture); (i) 3.8, (ii) 4.5, (iii) 3.8, (iv) 1.6, and (v) 2.1 nm, respectively.

which the α-helices were attached. The line of close packed α-helices was longer than the β-sheet, thus forcing both elements into an incomplete circle, or horseshoe, to accommodate the differences in lengths (Kobe & Deisenhofer, 1994). The authors suggested that other LRRPs would resemble RI. Decorin, fibromodulin, etc. contain 10 or more leucine-rich repeating sequences, compared with 15 in RI.

The size of RI (7.0 × 6.2 × 3.2 nm) is such that similar structures would not have been readily seen in our original electron micrographs of rotary shadowed preparations. All fine detail was lost in the pixel-based publication−reproduction process (Ward et al., 1987; Scott et al., 1990). However, the photographic quality of the originals is high enough to withstand considerable magnification, and they were therefore reexamined. This showed that decorin and lumican and/or fibromodulin are indeed horseshoe shaped (this paper).

These results were presented at the 655th Meeting of the British Biochemical Society, June 1995.

EXPERIMENTAL PROCEDURES

Full details of how these pictures were obtained, together with the properties of the relevant PGs, were published (Ward et al., 1987; Scott et al., 1990). Briefly, a drop of PG solution was sandwiched between two mica leaves, dried in vacuo, shadowed with Pt/W, and viewed by electron microscopy.

RESULTS

Figure 1 i and ii were parts of published pictures, now considerably enlarged. The PG aggregates with the glycan chains radiating outward were described as "sunbursts" (Ward et al., 1987; Scott et al., 1990), which contain decorin in all orientations. At this higher magnification it is clear that many apparently single molecules of PDSs pictured in, e.g., Figure 3 of Ward et al. (1987) are aggregates, with the glycan tails aligned together and indistinguishable from each other.

Figure 1 iv and v show images of corneal PKSs. They were not antigenically characterized but probably included both lumican and fibromodulin.

The images reflect different dispositions of the horseshoes in the path of the shadowing metal. More than 30 examples of complete (i.e., in plan) horseshoes were seen in other pictures, plus many more with recognizable attributes ("partial horseshoes"). They were present in reduced and alkylated decorin [pictured at lower magnification in Figure 5 of Ward et al. (1987)]. Similar motifs were not seen in preparations of non-PGs (hyaluronan, alginate, chondroitin sulfates) and in pictures of "blank" preparations without polymer.

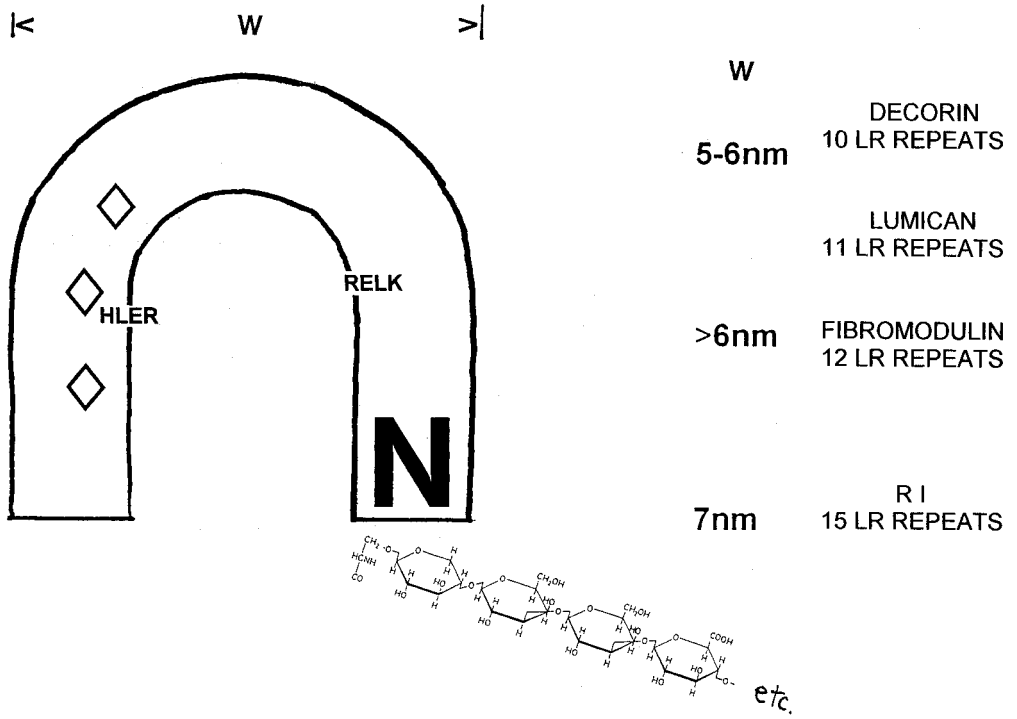

FIGURE 2: Proposed "map" of PDS with attached glycan chain (dermatan sulfate) linkage region (to scale). N = N-terminal. Approximate positions of the three N-linked oligosaccharides [represented by open diamonds (◇)] and putative markers (RELK, RELH) for binding sites to collagen are shown, assuming that the length along the horseshoe approximately parallels that along the primary sequence (see text). Residue numbering according to Scott, P. G. (1993). Overall widths of the horseshoes (W) in nm measured on electron micrographs are given for PDS and PKS(s). RI widths are from Kobe and Deisenhofer (1994).

The PG proteins appear slightly smaller than RI. Sizes across the horseshoes are given in Figure 2. Decorin was somewhat shorter (~3.8 nm) from the front of the jaws to the apex of the horshoe than from side to side. Seen in side views, they were about 2 nm thick. The distance between the jaws of the decoron horsehoe was 1.5−2.0 nm. The PKS proteins were slightly larger, compatible with having more leucine-rich repeats than PDSs. The contribution of the metal decoration is quantitatively uncertain. RI dimensions in Figure 2 are from the X-ray study (Kobe & Deisenhofer, 1994).

DISCUSSION

The pictures exclude that decorin, etc., are predominantly globular. Rather, they are approximately uniform in cross-section, quasicylindrical throughout most of their length. Some of the shapes seen in the pictures could be due to flexible cylinders of length 10−11 nm bending into curved forms. Since no long straight chains of this size were seen it follows that curved forms were more readily taken up, with a strong tendency to produce a symmetrical horseshoe.

Our results support the suggestion by Kobe and Deisenhofer (1994) that other LRRPs have something in common with RI. This is probably because the large number (10−12) of tandem leucine-rich-repeat modules force the two parallel lines of α-helices and β-turns into a curved shape, as in RI. Each α-β module is wedge-shaped, wider at the outside, narrowing toward the center. The shape is therefore intrinsic and not just one of many possibilities due to flexion of a cylinder. Thus, the three or four LRRPs so far examined are all horseshoe shaped. Although we have not studied biglycan it may be similar, given the very extensive homologies with decorin (Scott, P. G., 1993). Over 40 LRRPs are known.

RI, and by analogy PG proteins, are regular, comparable with collagen and other fibrillar proteins in that the number of amino acid residues per unit length is fairly constant. On this assumption particular residues in the primary sequence can be approximately located a priori within the horseshoe (Figure 2).

Interactions with Collagen. The horseshoe shape and the sizes of the various parts must be significant in the organisation of extracellular matrix, in particular the combination of PDS and PKS with collagen and collagen fibrils. Thus, the diameter of the space inside the horseshoes (~2 nm) is slightly larger than the diameter of a collagen molecule (1.5 nm). The arms of the horseshoes are about as thick as a collagen molecule. The horseshoe widths span slightly less than three collagen molecules aligned in parallel (Figure 3).

It was suggested that there were two collagen binding sites on decorin (Schonherr et al., 1993), one somewhere in the 158 residues at the N-terminal end and the second in the C-terminal half. In that case there would be a binding site on each arm of the horseshoe, separated by 2−5 nm across the horshoe. This implies that the binding sites on the collagen fibril(s) would also be paired and separated by 2−5 nm. They would be 7−17 amino acid residues apart, if both sites were on the same collagen molecule, with the horseshoe placed across the fibril. Alternatively, if the horseshoe axis were aligned along the fibril axis the binding sites would be separated by at least the full width of a collagen molecule (~1.5 nm), e.g., on two parallel collagen molecules.

The PDS and PKS binding sites in the d, e, a, and c bands, respectively, colocated with highly conserved 11-residue amino acid sequences (e.g., at residue 886, GDRGEPG-PAGP) which are found only in fibrillar collagens types I, II, and III in the α1 chains (Scott & Glanville, 1993). These were suggested to be markers of PG binding sites. Blotting

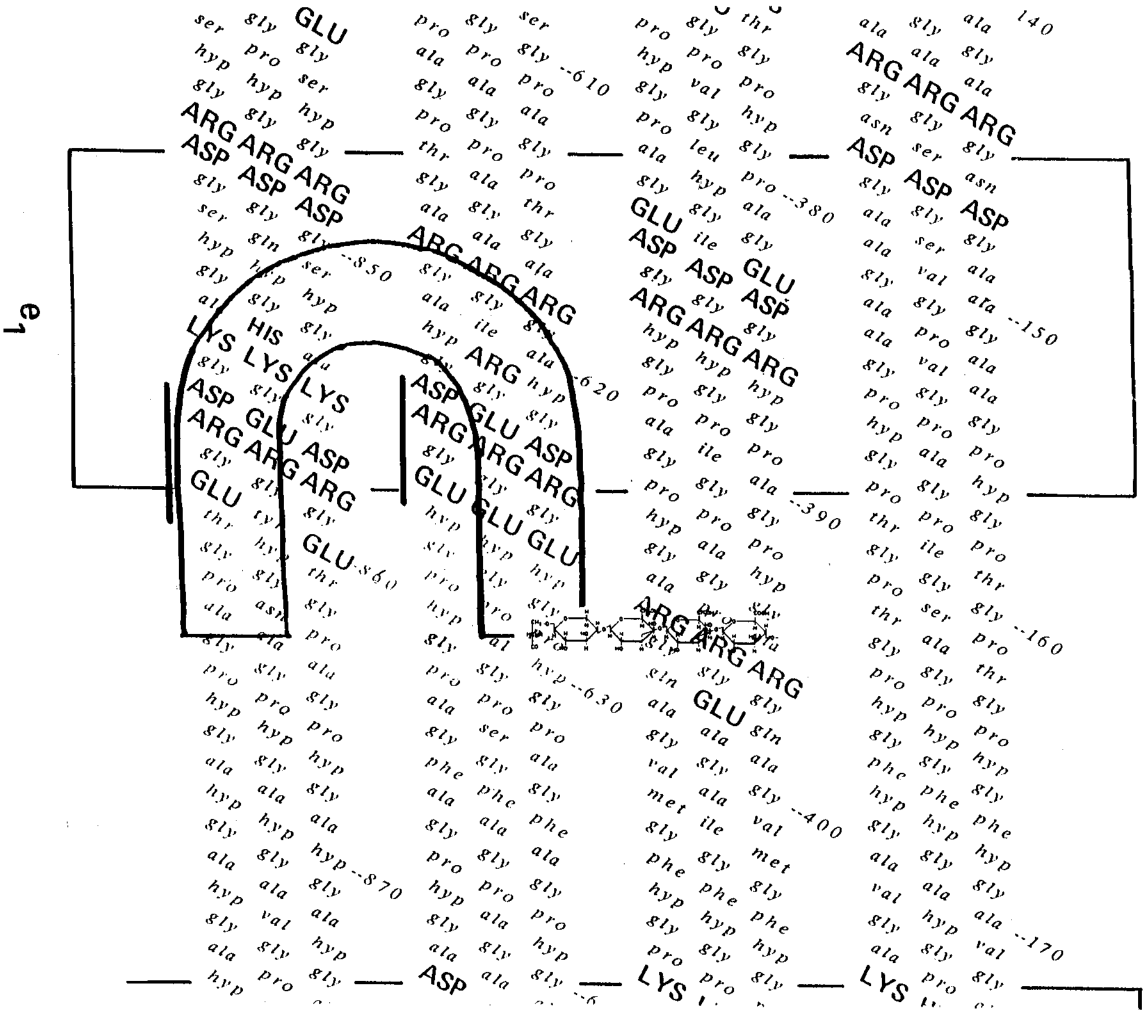

FIGURE 3: Proposed scheme of binding of decorin PG to collagen type I fibril. Four triple-helical collagen molecules at the gap zone align in parallel in a quarter-staggered array according to Chapman and Hulmes (1984). Two α1 chains flank an α2 chain in each collagen molecule. Amino acid residues are numbered. Charged residues are in bold caps. The box delineates the e_1 band. The decorin molecule is accommodated inside the e band, as suggested by ultrastructural studies (Scott, 1992). Assuming a rise of 0.3 nm for each residue in an α chain, the decorin molecule is drawn to scale and positioned to line up RELH and RELK sequences (see Figure 2) with the two GDRGE homologies in the collagen α1 chains, which are underlined. The highly anionic DS chain then approaches closely to the cationic nest of three arginine residues in the neighboring collagen molecule. This orients the glycan orthogonally to the collagen fibril, as is observed in tissues (Scott, 1992).

experiments showed that indeed the α1 chain of type I collagen, but not the α2 chain, binds to decorin (Sini, Denti, Tira, and Balduini, personal communication). Then, either a decorin molecule (a) combines with two α1 chains in the same collagen molecule, or (b) with two α1 chains from different collagen molecules. Alternative b would require the putative binding sequences to be present on immediately neighboring collagen molecules. This is the case in the e_1 band, where the 11-residue marker sequences (at residues 856 and 622, respectively) are exactly aligned side by side, even though the two neighboring collagen molecules involved are quarter-staggered by 234 amino acid residues with respect to each other (Figure 3).

All six 11-residue homologies from the collagen PG-binding regions contain a sequence, e.g., DRGE, at positions 2−5 that carries all the charge in the 11 residues (Scott & Glanville, 1993). This charge motif [−, +, 0, −] might be complemented by a similar motif with charges reversed [+, −, 0, +] in the PG, as a binding partner. Decorin indeed has two such, RELK (100−103) and RELH (242−245). The Swiss protein data base (release 31) found the LRELHL sequence only in the PGs decorin, biglycan, fibromodulin, and RI but in no other protein (Glanville and Scott, unpublished).

The two sequences in decorin are approximately equidistant (Figure 2) from the two termini, on each arm of the horseshoe. Interestingly, they run in opposite senses, RELK toward the N terminus and RELH toward the C-terminus. Since the decorin molecule turns through ∼180° from one arm to the other, both sequences might then run in parallel, as do all the putative binding sequences on the collagen fibril (Figure 3).

If two collagen molecules bind to a PG protein molecule, this may have a stabilizing and orienting effect during fibrillogenesis. This role would be additional to that of orientation and organization of collagen *fibrils* as seen in the tissue and exerted by the whole PG (Scott, 1992).

ACKNOWLEDGMENT

I thank Drs. L. Coster, S. P. Damle, J. D. Gregory, H. Greiling, and H. W. Stuhlsatz for PGs and Mrs. C. Cummings and Dr. N. Ward for electron microscopy.

REFERENCES

Chapman, J. A., & Hulmes, D. J. S. (1984) in *Ultrastructure of the Connective Tissue Matrix* (Ruggeri, A., & Motta, P. M., Eds.) pp 1−33, Martinus Nijhoff, Dordrecht, The Netherlands.

Kobe, B., & Deisenhofer (1994) *Trends Biochem. Sci. 19,* 415−421.

Schonherr, E., Winnemoller, M., Harrach, B., Robenek, H., & Kresse, H. (1993) in *Dermatan Sulphate Proteoglycans. Chemistry, Biology, Chemical Pathology* (Scott, J. E., Ed.) pp 241−247, Portland Press, London.

Scott J. E. (1988) *Biochem. J. 252* 313−323.

Scott, J. E. (1992) *FASEB J. 6,* 2639−2645.

Scott, J. E. (1993) *Glycoconj. J. 10,* 319−421.

Scott, J. E., & Glanville, R. W. (1993) *Biochem. Soc. Trans. 21,* 123S.

Scott, J. E., Cummings, C., Greiling, H., Stuhlsatz, H. W., Gregory, J. D., & Damle, S. P. (1990) *Int. J. Biol. Macromol. 12,* 180−184.

Scott J. E., Heatley F., & Wood, B. (1995) *Biochemistry 34,* 15467−15474.

Scott, P. G. (1993) in *Dermatan Sulphate Proteoglycans. Chemistry, Biology, Chemical Pathology* (Scott, J. E., Ed.) pp 81−102, Portland Press, London.

Ward, N. P., Scott, J. E., & Coster, L. (1987) *Biochem. J. 242,* 761−766.

BI960773T