Biochem. J. (1985) 232, 277–279 (Printed in Great Britain)

Dermatan sulphate is located on serine-4 of bovine skin proteodermatan sulphate

Demonstration that most molecules possess only one glycosaminoglycan chain and comparison of amino acid sequences around glycosylation sites in different proteoglycans

Ravi K. CHOPRA, C. Harold PEARSON,* Gordon A. PRINGLE, Daniel S. FACKRE and Paul G. SCOTT
Department of Oral Biology, University of Alberta, Edmonton, Alberta T6G 2N8, Canada

Digestions of bovine skin proteodermatan sulphate with cathepsin C proved that the dermatan sulphate was located on Ser-4 in most of the molecules. A Ser-Gly sequence is essential for xylosylation of the serine residue and sulphated galactosaminoglycan synthesis in different proteoglycans. Variations in adjoining sequences may be significant in relation to the glycosylation process in different tissues.

INTRODUCTION

Proteodermatan sulphate (PDS) is a small proteoglycan containing L-iduronate-rich dermatan sulphate (DS), which is widely distributed in soft connective tissues (Pearson *et al.*, 1983*a*). Its functions are probably related to specific interactions with collagen (Vogel *et al.*, 1984; Scott, 1984; Pringle, 1985). Amino acid sequencing of the *N*-terminal region of the protein core of bovine skin PDS showed that Ser-4 was *O*-substituted, but the substituent could not be identified after Edman degradation (Pearson *et al.*, 1983*a*). We now show that Ser-4 carries the DS chain. This has enabled us to compare amino acid sequences in the vicinity of attached galactosaminoglycan chains in different types of proteoglycans.

MATERIALS AND METHODS

The bovine skin PDS and its protein core derived by deglycosylation with anhydrous HF [PDS(HF)core] were prepared as previously described (Pearson *et al.*, 1983*a*). One portion of the PDS was reduced and alkylated (Hardingham *et al.*, 1976). A peptide-DS was purified (Pringle, 1985) from a CNBr digest of PDS (Pringle *et al.*, 1985). Bovine spleen cathepsin C (peptidylaminopeptidase I, EC 3.4.14.1; McDonald & Schwabe, 1977) was purchased from Sigma Chemical Co. (St. Louis, MO, U.S.A.) and was stated to contain 13.5 units/mg of protein. Digestions were usually performed at 37 °C, with 0.5–1.3 units of cathepsin C/mg of substrate protein, in the presence of a buffer/activator, which was prepared as described by Callahan *et al.* (1972), but the pH was 4.5. Controls contained no enzyme or no substrate. The digests were examined by sodium dodecyl sulphate/polyacrylamide-gel electrophoresis followed by scanning (Pearson & Gibson, 1982) and by cellulose acetate electrophoresis (Pearson *et al.*, 1986). After electrophoresis components of cathepsin C digests were separately eluted with water at 4 °C from unstained cellulose acetate strips, extensively dialysed at 4 °C and freeze-dried. The release of dipeptides from PDS was monitored by adding 12 μl of 2 M-HCl to 100 μl of the digestion mixture and analysing directly on a Beckman 121MB amino acid analyser with the standard procedures for determination of amino acids. Dipeptide standards Gly-L-Ile, Gly-L-Pro and L-Ala-L-Ser were obtained from Sigma Chemical Co. L-Asp-L-Glu was purchased as the *N*-benzyloxycarbonyl derivative (Vega Biotechnologies, Tucson, AZ, U.S.A.), from which it was regenerated by catalytic hydrogenation (Katsoyannis & Schwartz, 1977). Amino acids and hexosamines were determined as described (Pearson & Gibson, 1982).

RESULTS AND DISCUSSION

Release of the *N*-terminal dipeptide Asp-Glu from PDS by cathepsin C was detected after 1–2 min at 37 °C. Rapid digestion occurred whether or not the proteoglycan was reduced and alkylated, and the solution quickly became turbid. When the whole digest was examined by sodium dodecyl sulphate/polyacrylamide-gel electrophoresis it was clear that most of the DS (detected with Toluidine Blue) had also been released from the PDS and was located in a broad band with a much higher mobility than PDS (Fig. 1). At the same time a major protein band appeared that was not associated with DS and that was similar in size to the complete protein core prepared by deglycosylation with chondroitinase ABC (Pearson *et al.*, 1983*a*). These changes did not occur in control incubations. Thus a limited proteolysis of the *N*-terminal zone of the protein core had removed the DS, strongly suggesting that it was located on Ser-4, as this is the only serine residue in the first 24 residues (Pearson *et al.*, 1983*a*).

Confirmation of the attachment of DS to Ser-4 was obtained from the compositions of two peptidoglycans, (i) the dipeptide-DS liberated by cathepsin C (in 30 min at room temperature), which was isolated by preparative cellulose acetate electrophoresis, and (ii) the peptide-DS

Abbreviations used: PDS, proteodermatan sulphate; DS, dermatan sulphate; PDS(HF) core, protein core derived by deglycosylation of PDS with anhydrous HF.

* To whom correspondence should be addressed.

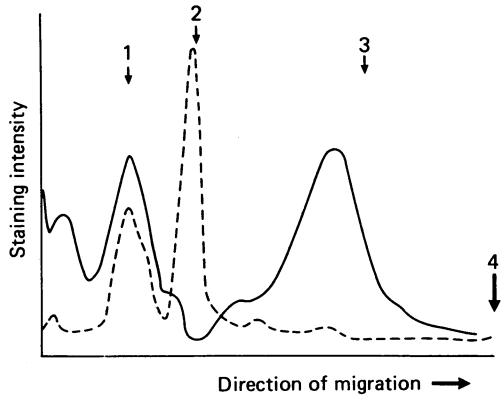

Fig. 1. Sodium dodecyl sulphate/polyacrylamide-gel electrophoresis of cathepsin C digest of PDS

Reduced and alkylated PDS (1.15 mg of protein) was incubated for 25 min at 37 °C with 1.5 units of cathepsin C in 1.3 ml of buffer/activator, pH 4.5. Portions (50 μl) of the digest were pretreated as described (Pearson & Gibson, 1982) before electrophoresis on 10% gels. Gels were stained with Toluidine Blue and scanned at 550 nm (——). After extensive destaining the same gels were stained with Coomassie Blue R250 and scanned at 560 nm (– – – –). Arrows indicate the mobilities of: 1, PDS control; 2, protein core of PDS derived by a treatment with chondroitinase ABC (Pearson *et al.*, 1983*a*); 3, DS (Miles Laboratories, Rexdale, Ontario, Canada); 4, Bromophenol Blue.

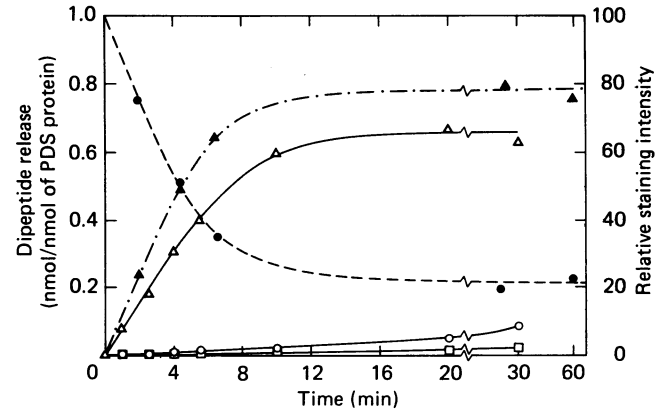

Fig. 2. Rates of release of dipeptides and dipeptide-DS from reduced and alkylated PDS

Digestion with cathepsin C was performed at 37 °C as described in Fig. 1 legend, but for various times. Left-hand scale: △, Asp-Glu; ○, Gly-Ile; □, Gly-Pro; all determined as described in the text. Right-hand scale: ▲, dipeptide-DS; ●, apparently unchanged PDS; both estimated from the relative areas of the major Toluidine Blue-stained peaks in scans of sodium dodecyl sulphate/polyacrylamide gels after electrophoresis.

isolated from a CNBr digest of PDS (see below). Acid hydrolysates of the dipeptide-DS contained, in addition to a large amount of galactosamine (but only traces of glucosamine), equimolar proportions of alanine and serine, as expected from the sequence and the action of the dipeptidyl peptidase (McDonald & Schwabe, 1977). Smaller proportions of amino acids that occupy positions close to Ala-Ser in the *N*-terminal zone of the protein cor were also present, namely Asp, 0.21, Glu, 0.36, and Gly, 0.29 (Ala = 1.00), whereas other amino acids were less than 0.1, except for Leu (0.11). The retention of aspartate, glutamate and glycine after cellulose acetate electrophoresis and extensive dialysis suggested that the cathepsin C (or possibly contaminating exopeptidases) to a small extent catalysed cleavages in addition to those expected. However, endopeptidase activity appeared to be minimal, because, apart from the almost intact protein core and residual PDS, other Coomassie Blue-stained bands were very weak, especially if the digestion was limited to 20–30 min at 37 °C (Fig. 1).

Fig. 2 shows the rates of release of dipeptides and the dipeptide-DS. Release of the dipeptide-DS reached a maximum after 15–20 min at 37 °C, closely paralleling the results for the *N*-terminal dipeptide Asp-Glu. In contrast, only a small proportion of the third dipeptide (Gly-Ile) in the sequence (Table 1) was liberated, even less of the fourth (Gly-Pro), and no other dipeptide was detected. It was rather surprising that the attachment of DS did not significantly inhibit the cleavage of the Ser–Gly linkage by the enzyme. We also concluded that the presence of the dipeptide-DS in the digests was not responsible for the poor release of Gly-Ile, because the results were similar when the PDS(HF) core (free of DS) was the substrate (results not shown). The precipitation of the protein core that occurred when Ala-Ser-DS was

Table 1. Amino acid sequences in different proteoglycans

Sequences in regions of protein cores containing attached sulphated galactosaminoglycans are compared. (*a*) Bovine skins PDS; (*b*) small proteoglycan isolated from human foetal membranes (Brennan *et al.*, 1984); (*c*)–(*e*) peptidogalactosaminoglycans derived from cartilage proteoglycans (Isemura *et al.*, 1981): (*c*) bovine and pig; (*d*) human; (*e*) shark. A repeating sequence corresponding to the first seven residues in (*c*) may be prominent in the protein core of bovine nasal proteoglycan (Bonnet *et al.*, 1983). The serine residue marked with an asterisk (*) is the location of the DS (*a*) or chondroitin sulphate [(*c*)–(*e*)] chain. For explanation of ? see the text.

(*a*)	Asp-Glu-Ala-Ser*-Gly-Ile-Gly-Pro-Glu-Glu-
(*b*)	Asp-Glu-Ala- ? -Gly-Ile-Gly-Pro-Glu-Val-
(*c*)	-Leu-Pro-Ser*-Gly-Gly-Pro-Glu-
(*d*)	-Leu-Pro-Ser*-Gly-Glu-
(*e*)	-Leu-Pro-Ser*-Gly-Asp-

removed may have slowed the attack on the next susceptible linkage, but does not seem to explain adequately the resistance of the Ile–Gly linkage to the enzyme. Knowledge of the sequence specificity of cathepsin C is incomplete, but the very low release of the fourth dipeptide, which included proline, was predictable (McDonald & Schwabe, 1977).

The peptide-DS liberated by CNBr cleavage contained galactosamine as a major component, but no glucosamine, and it had the following amino acid composition (residues/21 amino acid residues), with the values predicted for the *N*-terminal CNBr-cleavage peptide from the known sequence (Pearson *et al.*, 1983*a*) in parentheses: Asp, 1.3 (1); Ser, 1.3 (1); Glu, 6.0 (6); Pro, 3.9 (4); Gly, 2.1 (2); Ala, 1.0 (1); Val, 0.9 (1); Ile, 1.5 (2); Phe, 0.9 (1); His, 0.8 (1); Hse, 0.7 (1); other amino acids, 0.2 or less.

Estimates of the number of DS chains in PDS have varied from one to four (for references see Pearson *et al.*, 1983*a*), but the above investigations proved directly that Ser-4 was the only attachment site of DS in the majority of the bovine skin PDS molecules. The protein core liberated in good yield by cathepsin C was not metachromatic (Fig. 1), nor did it stain with Alcian Blue after cellulose acetate electrophoresis (Pearson *et al.*, 1986). However, part of the PDS, which we estimate to be about a quarter of the total, retained DS after repeated incubations with cathepsin C. Further work is required to investigate whether this fraction has a different structure from the rest.

Similar results were obtained when human PDS (isolated from gingiva; C. H. Pearson, S. Lehocky & E. H. Moase, unpublished work) was incubated with cathepsin C. It seems likely that the structure of bovine skin PDS is representative of small proteoglycans containing L-iduronate-rich DS in different species and tissues, although M_r values vary (80000–130000 estimated by sodium dodecyl sulphate/polyacrylamide-gel electrophoresis), probably mainly because of differences in the length of the DS chain (Pearson & Gibson, 1982; Pringle, 1985). Results of immunoblotting with monoclonal antibodies (Pringle *et al.*, 1985) suggest that the protein cores of human and bovine PDS are similar but not identical (Pringle, 1985).

Other small proteoglycans, differing from PDS in some respects, may nevertheless have sulphated glycosaminoglycan located near the *N*-terminus of the protein core. The residue at position 4 in the protein core of a small proteoglycan associated with human foetal membranes was not identified after Edman degradation (Brennan *et al.*, 1984). We assume that this residue was serine carrying a glycosaminoglycan chain, by analogy with our own results and because the rest of the *N*-terminal amino acid sequence up to residue 9 was identical with that of PDS (Table 1). Remarkably, there was almost no correspondence for the next 14 residues (Brennan *et al.*, 1984). In addition, the sulphated galactosaminoglycan component differed, as it contained a much lower proportion of L-iduronate than did DS of PDS and it had an unusually high M_r value (Brennan *et al.*, 1984).

The conservation of the first nine amino acid residues surrounding the attached glycosaminoglycan in small proteoglycans of different species suggests that this sequence is functionally important. However, only the glycine residue following the serine seems to be essential for xylosylation of the serine residue and subsequent glycosaminoglycan synthesis, confirming suggestions from earlier work (reviewed by Rodén *et al.*, 1985). In cartilage peptidogalactosaminoglycans (derived from internal regions of the proteoglycans; Isemura *et al.*, 1981), the two amino acid residues preceding and the one following the Ser-Gly have very different side chains from those in the small proteoglycans of soft connective tissues (Table 1). Strikingly, the next three residues, Gly-Pro-Glu, are identical in different proteoglycans, but they may not be critical for the glycosylation process. Chondroitin sulphate chains were synthesized on a repeating Ser-Gly sequence (deduced from nucleotide sequencing) in a tumour proteoglycan (Bourdon *et al.*, 1985). The key feature is probably an ability of the region of the protein core containing the potential xylosylation site to form a stable β-bend (Pearson *et al.*, 1983*a*,*b*). This may explain why protein cores of proteoglycans are much more efficient acceptors in the xylosyltransferase-catalysed reaction than are oligopeptides (Rodén *et al.*, 1985). The significance of the differences in the sequences shown in Table 1 may become clearer when more is known about the enzymes involved in the glycosylation process; for instance it is not certain whether more than one xylosyltransferase exists or whether xylosylation occurs in the rough endoplasmic reticulum of fibroblasts as it does in chondrocytes (Rodén *et al.*, 1985).

We thank Mr. N. Winterbottom for providing the reduced and alkylated proteoglycan and other help and Mrs. S. Lehocky for technical assistance. The work was financed by the Canadian Medical Research Council, the University of Alberta Central Research Fund and the Alberta Heritage Foundation for Medical Research.

REFERENCES

Bonnet, F., Le Glédic, S., Perin, J.-P., Jollès, J. & Jollès, P. (1983) Biochim. Biophys. Acta **743**, 82–90
Bourdon, M. A., Oldberg, Å., Pierschbacher, M. & Ruoslahti, E. (1985) Proc. Natl. Acad. Sci. U.S.A. **82**, 1321–1325
Brennan, M. J., Oldberg, Å., Pierschbacher, M. & Ruoslahti, E. (1984) J. Biol. Chem. **259**, 13742–13750
Callahan, P. X., McDonald, J. K. & Ellis, S. (1972) Methods Enzymol. **25**, 282–298
Hardingham, T. E., Ewins, R. J. F. & Muir, H. (1976) Biochem. J. **157**, 127–143
Isemura, M., Hanyu, T., Kosaka, H., Ono, T. & Ikenaka, T. (1981) J. Biochem. (Tokyo) **89**, 1113–1119
Katsoyannis, P. G. & Schwartz, G. P. (1977) Methods Enzymol. **47**, 501–578
McDonald, J. K. & Schwabe, C. (1977) in Proteinases in Mammalian Cells and Tissues (Barrett, A. J., ed.), pp. 311–322, North-Holland Publishing Co., Amsterdam, New York and Oxford
Pearson, C. H. & Gibson, G. J. (1982) Biochem. J. **201**, 27–37
Pearson, C. H., Winterbottom, N., Fackre, D. S., Scott, P. G. & Carpenter, M. R. (1983*a*) J. Biol. Chem. **258**, 15101–15104
Pearson, C. H., Winterbottom, N., Scott, P. G. & Fackre, D. S. (1983*b*) Biochem. Soc. Trans. **11**, 747–749
Pearson, C. H., Chopra, R. K., Lehocky, S. & Moase, E. H. (1986) Biochem. Soc. Trans. **14**, in the press
Pringle, G. A. (1985) Ph.D. Thesis, University of Alberta
Pringle, G. A., Dodd, C. M., Osborn, J. W., Pearson, C. H. & Mosmann, T. R. (1985) Collagen Relat. Res. Clin. Exp. **5**, 23–39
Rodén, L., Koerna, T., Olson, C. & Schwartz, N. B. (1985) Fed. Proc. Fed. Am. Soc. Exp. Biol. **44**, 373–380
Scott, J. E. (1984) Biochem. J. **218**, 229–233
Vogel, K. G., Paulsson, M. & Heinegård, D. (1984) Biochem. J. **223**, 587–597

Received 18 July 1985/3 September 1985; accepted 9 September 1985