Biochem. J. (1992) 288, 565–569 (Printed in Great Britain)

Characterization of a human seminal plasma glycosaminoglycan-bearing polypeptide

François BONNET, Jean-Pierre PERIN, Philippe MAILLET, Pierre JOLLES and Patrick M. ALLIEL
Laboratoire des Protéines, URA C.N.R.S. No. 1188, Université de Paris V, 45 rue des Saints-Pères, F 75270 Paris, Cedex 06, France

A glycosaminoglycan-bearing polypeptide (S.GP), present in human seminal plasma, was purified to homogeneity by a combination of CsCl density-gradient centrifugation, f.p.l.c. ion-exchange chromatography on a Mono Q HR column and Superose 6 gel filtration. The observed polydispersity of S.GP was attributed to the heterogeneity of its glycosaminoglycan content. Enzymic deglycosylation experiments and *N*-terminal amino-acid sequence determination indicate that it consists of a polypeptide (apparent molecular mass approx. 18 kDa) bearing both chondroitin and heparan sulphate chains. Evidence is given that S.GP contains a glycosaminoglycan-linkage domain of a so far uncharacterized gene product, proteolytically processed in the genital tract.

INTRODUCTION

Proteoglycan polymorphism is a well-established concept. Distinct gene products have been characterized as proteoglycan cores. New developments concerning the potential roles of proteoglycan have emerged: they are now thought not only to be involved in hydrodynamic or structural roles in tissues, but also to be implicated in various regulatory effects (Ruoslahti, 1988, 1989). In the testis, peritubular cells and Sertoli cells have been shown to act co-operatively in co-culture to synthesize different components of an extracellular matrix (Tung & Fritz, 1980; Skinner *et al.*, 1985); the latter regulates testicular cord formation and germ cell development *in vitro* (Hadley *et al.*, 1985). It has been suggested that the synthesis of some of its components, namely proteoglycans, is an age-related process which probably plays an active role in spermatogenesis (Rodriguez & Minguell, 1989*a*). Somatic rat testicular cells secrete two types of proteoglycans: one containing both heparan/chondroitin, the other chondroitin/dermatan sulphate as glycosaminoglycan chains (Skinner & Fritz, 1985; Rodriguez & Minguell, 1989*b*). Proteoglycans from rat peritubular cells and Sertoli cells have been partially characterized, mainly as a function of the nature of the glycosaminoglycan side-chains which they bear (Skinner & Fritz, 1985). Molecular characterization of the proteoglycans present in the seminal plasma is lacking. The aim of this work was to characterize, at a molecular level, the proteoglycans in human seminal plasma.

EXPERIMENTAL

Materials

The chromatographic media, namely Mono Q (HR 5/5 column) and Superose 6 (HR 10/30 column), were obtained from Pharmacia (Uppsala, Sweden); they were fitted to the Pharmacia f.p.l.c. apparatus and elution-profile recordings were obtained with the aid of a Pharmacia U.V.-1 monitor at 280 nm; Fractogel TSK HW-55(F) and TSK HW-40(S) were from Merck (Darmstadt, Germany). Chondroitinase ABC (chondroitin ABC lyase; EC 4.2.2.4) was from I.C.N. Biomedicals Ltd. (High Wycombe, Bucks., U.K.) and heparinase III (EC 4.2.2.8) was obtained from Sigma Chemical Co. (St. Louis, MO, U.S.A.). All other reagents were from Prolabo (Paris, France), Merck (Darmstadt, Germany), Farmitalia/Carlo Erba (Milan, Italy), or Sigma Chemical Co. (St. Louis, MO, U.S.A.). The reagents used for protein sequencing experiments were from Applied Biosystems (Forster City, CA, U.S.A.).

Preparative procedures

Human seminal plasma was prepared from ejaculates of healthy donors according to Guess *et al.* (1986). Seminal plasma was mixed with 2 vol. of extracting buffer [6 M-guanidinium hydrochloride (Gdn-HCl)/50 mM-sodium acetate, pH 5.8, containing 5 mM-benzamidine, 10 mM-6-aminohexanoic acid, and 10 mM-EDTA as protease inhibitors]. Extraction was carried out for 2 h at 4 °C with constant stirring, the supernatant was recovered after centrifugation at 1500 *g* for 15 min and the pellet was extracted twice, for 15 min, with 10 ml of 4.5 M-Gdn-HCl/50 mM-sodium acetate, pH 5.8, containing the protease inhibitors. Solid CsCl was added to the pooled extracts up to a density of 1.45 g/ml. The mixture was submitted to density-gradient centrifugation for 40 h at 20 °C (35000 rev./min in a Beckman L8-55 ultracentrifuge, 60 Ti rotor). Eight fractions were collected from the gradient and exhaustively dialysed against distilled water in the presence of the protease inhibitors listed above. The hexuronate-containing fractions were mixed with an equal volume of 7.6 M-urea/50 mM-Tris/HCl buffer, pH 7.6, (buffer A) and applied to a Mono Q HR 5/5 column equilibrated in buffer A. The column was washed with 0.3 M-NaCl/buffer A; elution was performed with a linear 95 ml NaCl gradient (0.3–1.5 M) in buffer A (at a flow rate of 0.5 ml/min).

The hexuronate-containing fractions were pooled, diluted with 4 vol. of buffer A and applied to a Q-Sepharose column (2 cm × 0.4 cm) equilibrated in the same buffer. Washings (5 ml) were carried out with 0.05 M-Tris/HCl buffer, pH 7.6, hexuronates were eluted with 6 M-Gdn-HCl and precipitated by the addition of 4 vol. of cold ethanol (20 h, −20 °C). The pellet was briefly washed three times with 70% (v/v) cold ethanol and solubilized in water.

Reduction (3 h, 37 °C) was performed in 150 μl of a 4.5 M-Gdn-HCl/0.1 M-Tris/HCl buffer, pH 7.6 (buffer B), by the addition of 1 μmol of 2-mercaptoethanol to 120 μg of hexuronate. The reduced material was precipitated with 4 vol. of cold ethanol (1 h, −20 °C). The pellet was rapidly washed twice with 1 ml of 70% (v/v) cold ethanol, and solubilized in 50 μl of buffer B. Alkylation (1 h, 20 °C) was performed by the addition

Abbreviations used: Gdn-HCl, guanidinium hydrochloride; S.GP, seminal plasma glycosaminoglycan-bearing polypeptide; RA-S.GP, reduced and alkylated S.GP; PTH, phenylthiohydantoin.

of 1 μmol (50 μCi) of [^{14}C]iodoacetamide in 250 μl of the same buffer. After precipitation with 4 vol. of cold ethanol (20 h, −20 °C), followed by three washings with 70% (v/v) cold ethanol, the reduced and alkylated material was freeze-dried.

Analytical procedures

Uronate concentrations were determined by the carbazole method (Dische, 1947). Polyacrylamide/agarose gel electrophoresis was performed in accordance with the method of McDevitt & Muir (1971) on 2.4% (w/v) polyacrylamide/1.2% (w/v) agarose gels; proteoglycans were stained with 0.1% Toluidine Blue in 1% acetic acid. SDS/PAGE was performed by the procedure of Schägger & Von Jagow (1987), and scintillation measurements were carried out after solubilization of gel slices (10 mm^3) in 30% (w/v) hydrogen peroxide according to Tishler & Epstein (1968).

Superose 6 gel filtration was carried out in buffer B on an HR 10/30 column, at a flow rate of 0.5 ml/min; in this experiment, proteoglycans were detected by a slot-blot assay: adsorption was performed in 10% (w/v) acetic acid on Hybond-N$^+$ membranes (Amersham International, Aylesbury, Bucks., U.K.) followed by Toluidine Blue staining. TSK HW-55(F) gel filtrations were performed on a column (0.3 cm × 22 cm) using buffer B as eluent; fractions (120 μl) were collected and analysed by scintillation measurements; TSK HW-40(S) gel filtrations for the removal of Gdn-HCl after TSK HW-55(F) filtration were performed on a column of the same size in a 50 mM-ammonium acetate, pH 8, buffer. *N*-Terminal amino-acid sequences were determined with an Applied Biosystems 470 A sequencer equipped with a model 120 phenylthiohydantoin (PTH) analyser; ^{14}C-labelled *S*-carboxymethylcysteine PTH-derivatives were identified by scintillation measurement (Kontron Instruments, Montigny, France; β counter).

Chondroitinase ABC and heparinase III digestions were carried out at 37 °C over 4 h in 40 μl of a 1 mM-$CaCl_2$/50 mM-sodium acetate (pH 7.5) buffer containing 50 mM-6-aminohexanoic acid, 5 mM-benzamidine hydrochloride and 1 mM-phenylmethanesulphonyl fluoride as protease inhibitors, with 5×10^{-3} enzyme unit/μg of hexuronate.

RESULTS

Purification of S.GP

Upon dissociative CsCl-gradient centrifugation of human seminal plasma extract, 95% of the hexuronate was recovered in the bottom quarter of the gradient. By this procedure 20 ejaculates yielded 400 ± 25 μg of hexuronate in this fraction, which was then applied on the Mono Q column equilibrated in buffer A. After washing of the column with 20 ml of 0.3 M-NaCl/buffer A, hexuronates mainly eluted between 0.65–0.85 M-NaCl; this material was termed S.GP (Fig. 1). The first minor eluted peak (0.6 M-NaCl) was not further studied.

On composite polyacrylamide/agarose gels, S.GP behaved as a single broad band, the electrophoretic mobility of which was similar to that of human platelet proteoglycan (Périn *et al.*, 1988), but consistently different from rat chondrosarcoma proteoglycan (Oegema *et al.*, 1975) and free chondroitin sulphate chains (Fig. 2).

[^{14}C]Iodoacetamide was used to form reduced and alkylated S.GP (RA-S.GP). After this treatment hexuronates were detected in a broad peak and eluted from a Superose 6 column ($K_{av.} \approx 0.3$) (slot-dot assay detection); this purified fraction, termed RA-S.GP-II, contained 80% of the incorporated radioactivity (Fig. 3). On SDS/PAGE the radiolabelled, purified RA-S.GP-II co-migrated with the Toluidine-Blue-stained product (results not shown).

Identification of glycosaminoglycan side-chains in the S.GP

Chondroitinase ABC treatment of S.GP resulted in an extensive decrease of the material detected with Toluidine Blue on composite polyacrylamide/agarose gel electrophoresis, whereas heparinase III was significantly less active. Combined treatment

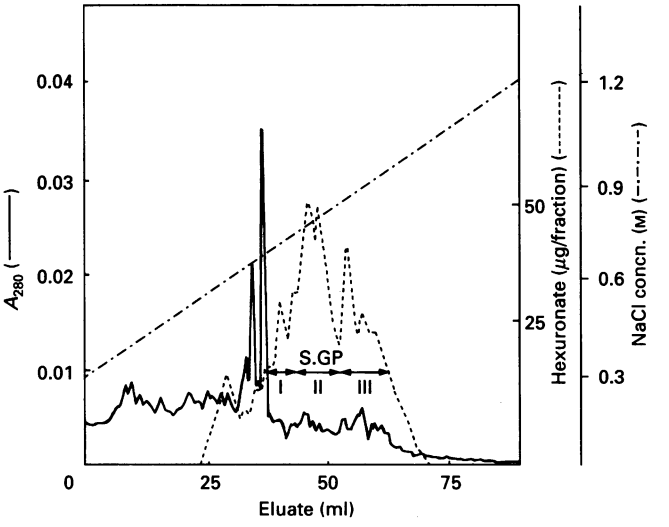

Fig. 1. Ion-exchange chromatography of S.GP on a Mono Q column

The material recovered in the bottom quarter of the CsCl density gradient was applied to the Mono Q HR 5/5 column equilibrated in buffer A and f.p.l.c. was run as indicated in the Experimental section. The hexuronate-containing fractions were alternatively pooled, either as S.GP (0.65–0.85 M-NaCl) or as subfractions (I, II and III).

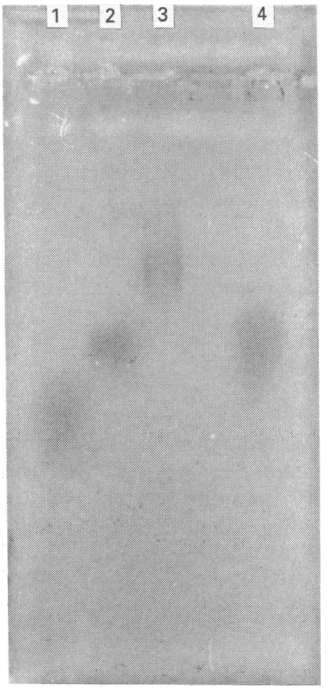

Fig. 2. Composite polyacrylamide/agarose gel electrophoresis of S.GP

Lane 1, free chondroitin sulphate chains (A and C isomers); lane 2, platelet proteoglycans (see Périn *et al.*, 1988); lane 3, rat chondrosarcoma proteoglycan (see Oegema *et al.*, 1975); lane 4, human seminal plasma proteoglycan. Proteoglycans were detected by Toluidine Blue staining.

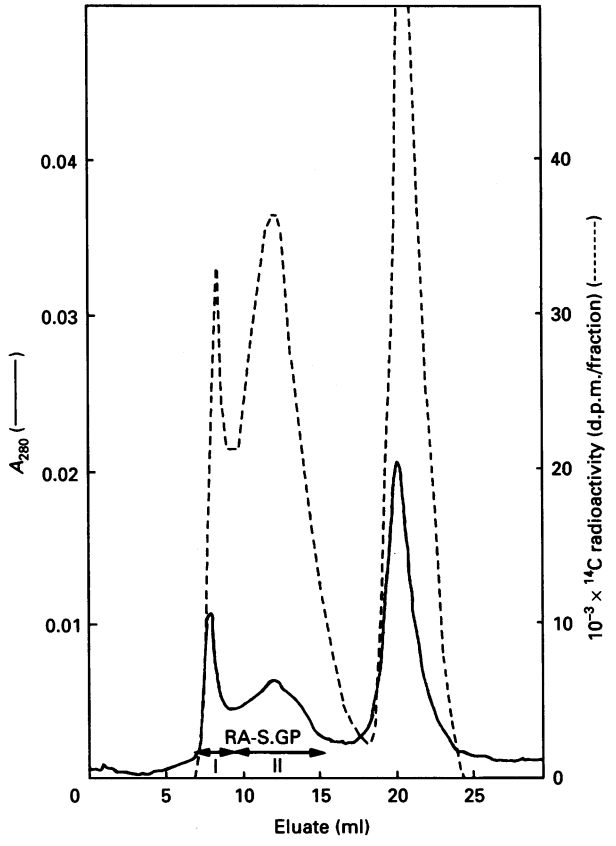

Fig. 3. Superose 6 gel filtration of RA-S.GP

S.GP recovered after Mono Q ion-exchange chromatography was reduced and alkylated using [^{14}C]iodoacetamide as the alkylating reagent, as indicated in the Experimental section. The sample was applied in 200 μl of buffer B on an HR 10/30 column. The hexuronate-containing fractions, detected by slot-blot assay, were pooled, yielding the RA-S.GP-II fraction.

with both enzymes resulted in an almost total disappearance of the Toluidine Blue staining (Fig. 4). This result indicates that both chondroitin and heparan sulphate chains are present in the S.GP fraction, and that chondroitin sulphate is the major glycosaminoglycan.

Identification of the RA-S.GP protein core

A set of two analytical experiments was carried out to characterize the RA-S.GP protein core.

In the first series of experiments, reduced and ^{14}C-labelled carboxymethylated S.GP was purified by Q-Sepharose ion-exchange chromatography. This material, containing both RA-S.GP-I and RA-S.GP-II, was submitted to various enzymic deglycosylation experiments and the resulting products were analysed by TSK HW-55(F) gel filtration (Fig. 5). When compared with the untreated sample ($K_{av.} \approx 0.3$), chondroitinase treatment resulted in an extensive broadening of the peak containing the radiolabelled material, while upon heparinase III treatment, the elution profile was less affected. Additionally, after combined treatment with both enzymes, a new peak ($K_{av.} \approx 0.6$–0.8) was characterized. However, a significant portion of the radiolabelled material still eluted at a position similar to that of the untreated sample (fraction I in Fig. 5); this material could not be shifted towards a lower molecular mass species even after an additional deglycosylation attempt. On SDS/PAGE, fraction I in Fig. 5 was shown to contain two different entities of $\approx$ 94 kDa and $\approx$ 45 kDa (Fig. 6*a*). The material eluted at a $K_{av.}$

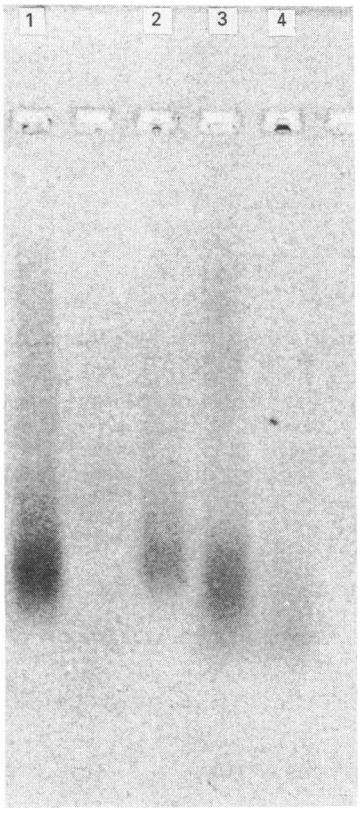

Fig. 4. Composite polyacrylamide/agarose gel electrophoresis of S.GP after enzymic deglycosylations

The sensitivity of S.GP (lane 1) to chondroitinase ABC (lane 2), heparinase III (lane 3) and chondroitinase ABC/heparinase III (lane 4) was estimated by Toluidine Blue staining of the resulting products.

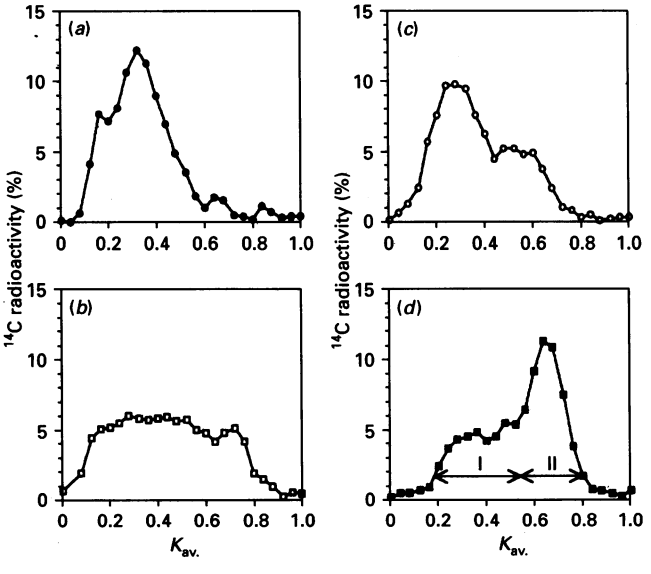

Fig. 5. TSK W-55(F) gel filtrations of RA-S.GP before and after enzymic deglycosylations

Gel filtrations were carried out on a column (0.3 cm × 22 cm) using 4.5 M-Gdn-HCl/0.1 M-Tris/HCl, pH 7.6, buffer as eluent (120 μl/fraction). Results are expressed as percentages of total radioactivity. Key to panels: (*a*), untreated RA-S.GP; (*b*), chondroitinase ABC-digested RA-S.GP; (*c*), heparinase III-digested RA-S.GP; and (*d*), chondroitinase ABC/heparinase III-digested RA-S.GP. Fractions I and II, characterized after chondroitinase ABC/heparinase III treatment, were pooled as indicated by the bars.

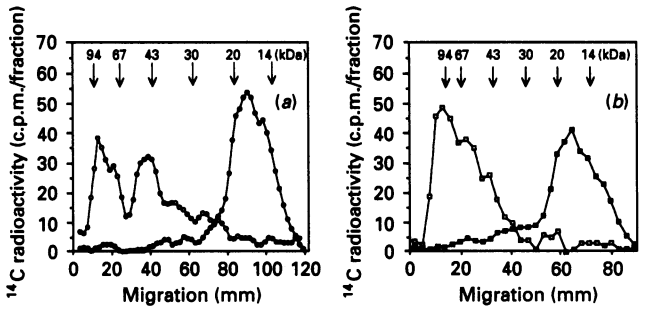

Fig. 6. SDS/PAGE analysis of chondroitinase ABC/heparinase III-digested RA-S.GP and RA-S.GP-II

(*a*) Fraction I (○) and fraction II (●) were isolated after TSK HW-55(F) gel filtration. (*b*) (□), undigested RA-S.GP-II; (■), digested RA-S.GP-II. Gel slices were solubilized in 30% (w/v) hydrogen peroxide before scintillation measurements. Standard molecular mass markers were: phosphorylase *b* (94 kDa), BSA (67 kDa), ovalbumin (43 kDa), carbonic anhydrase (30 kDa), soybean trypsin inhibitor (20 kDa) and α-lactalbumin (14.4 kDa).

Table 1. *N*-terminal amino-acid sequence of RA-S.GP-II

Values in parentheses indicate the amounts of identified PTH derivatives (pmoles); ? represents an undetermined amino acid.

Residue number	Amino acid
1	Lys (36)
2	Gln (30)
3	Gly (28)
4	Ala (25)
5	Val (20)
6	Ser (18)
7	Cys *
8	Glu (14)
9	Glu (14)
10	Glu (14)
11	Gln (13)
12	Glu (12)
13	Thr (10)
14	?
15	Gly (10)
16	Asp (6)
17	Phe (6)
18	Gly (8)
19	?
20	Gly (5)
21	Gly (5)
22	Ser (5)
23	Val (3)

* Cys was identified as an *S*-carboxymethyl derivative in position 7 by scintillation measurement.

≈ 0.6–0.8 (fraction II in Fig. 5) and behaved as a lower molecular mass compound in the 18 kDa range (Fig. 6*a*). From these results, the 18 kDa fragment, which was only detected after chondroitinase ABC/heparinase III treatment, was suspected to be derived from a chondroitin/heparan-bearing polypeptide.

Its identity was ascertained by a second experiment, using RA-S.GP-II purified using Superose 6. After combined treatment with both chondroitinase ABC and heparinase III, this fraction only gave rise to the 18 kDa species on SDS/PAGE (Fig. 6*b*).

Identification of the *N*-terminal amino-acid sequences of S.GP and RA-S.GP-II

A unique *N*-terminal amino-acid sequence was characterized for Superose-6-purified RA-S.GP-II (Table 1). The polydispersity of the material, indicated during ion-exchange chromatography, reflects a heterogeneity in the glycosaminoglycan/sulphate groups rather than a heterogeneity of the protein core. No protein sequence could be characterized for RA-S.GP-I.

When sequencing experiments were performed on S.GP, two sequences (*a* and *b*) could be characterized. One of them (*a*) corresponded to the sequence established for RA-S.GP-II.

$$(a)\ \overset{1}{K}\ Q\ G\ A\ V\ S\ ?\ E\ E\ \overset{10}{E}\ Q\ E\ T....$$

$$(b)\ K\ A\ T\ Q....$$

where ? represents an unidentified amino acid.

The Q/A ratio in step 2 was 52/48, but the level of PTH derivatives belonging to the (*b*) sequence dropped after step 4. Even when further purified by Superose 6 gel filtration, a non-reduced and non-alkylated sample still gave rise to these two *N*-terminal amino-acid sequences. Identical results were obtained when three different subfractions of S.GP (I, II and III in Fig. 1) were subjected to recurrent Edman degradation. Identical results were also obtained for a preparation carried out from a whole semen sample which had been immediately frozen after collection.

DISCUSSION

This work was initiated to characterize, at a molecular level, proteoglycans present in human seminal plasma; all the results reported in this study are consistent with the statement that this biological fluid contains a glycosaminoglycan-bearing polypeptide (S.GP) as the major proteoglycan entity.

As a function of deglycosylation experiments, the occurrence of both heparan and chondroitin sulphate has been shown in the purified S.GP, as well as in RA-S.GP-II.

From this set of experiments, it appeared that the 'KQGA .....' polypeptide, previously characterized in Superose-6-purified RA-S.GP-II, consists of a peptidic core bearing chondroitinase ABC/heparinase III-sensitive glycosaminoglycan side-chains. However, as chondroitinase treatment alone gave rise to a low amount of material in the peak II $K_{av.}$ range (Fig. 5*b*), the possibility that some RA-S.GP only contains chondroitin sulphate has to be considered.

Two glycosylation sites can be postulated in positions 14 and 19 in the RA-S.GP-II *N*-terminal sequence. The surrounding amino acids fit well with the general consensus sequence E/D-X-S-G for a glycosaminoglycan linkage, involving a serine residue; X represents one or perhaps two amino acids and either a hydrophobic or small neutral residue (Bourdon, 1990). Thus we propose to assign serine residues to positions 14 and 19 (see Table 1); as a consequence RA-S.GP-II could contain two clustered glycosaminoglycan-attachment sites. One of them, ETSGDF, (underlined serine residue corresponds to a glycosaminoglycan linkage) closely resembles the ETSGEN sequence, which was suggested to be a heparan sulphate-attachment site in murine syndecan (Saunders *et al.*, 1989). The other, DFSGG, does not share the constant feature of such a site (i.e. an acidic amino-acid residue on the *C*-terminus side of SG); it therefore could be relevant to a chondroitin-sulphate-bearing site. However, these assumptions remain to be confirmed as only a few glycosylation sites, corresponding to chondroitin/dermatan sulphate chains, have so far been clearly demonstrated, the others still being speculative (Bourdon, 1990). Furthermore distinct post-translational modifications on a single protein core have already been reported for syndecan (Sanderson & Bernfield,

1988) and serglycin (Tantravahi *et al.*, 1986), depending on the cellular type.

S.GP was isolated to a degree of purity allowing the determination of a single *N*-terminal amino-acid sequence of its reduced and alkylated form (RA-S.GP-II). However, two *N*-terminal sequences were characterized for a non-reduced sample. Contamination of the S.GP preparation by a foreign protein can be ruled out as a function of the preparative procedure used in this work: recovery in a fraction of density > 1.5 g/ml after density-gradient centrifugation, high acidic characteristic during anionic exchange chromatography. Several observations: (i) the equimolar ratio of Q/A at step 2 during protein sequencing of a non-reduced sample; (ii) the characterization of the KATQ..... sequence in three different pools of S.GP (I, II and III in Fig. 1); and (iii) its loss after reduction and alkylation, are indicative that the KATQ..... entity did not bear glycosaminoglycan side-chains but represents a short peptidic moiety, bound via cysteine residue(s) to the KQGA..... material. Preliminary studies, dealing with cDNA cloning, reveal the presence of these two amino-acid sequences on a single protein core (P. M. Alliel, J. P. Périn, P. Jollès and F. Bonnet, unpublished work). A search in the National Biomedical Research Foundation protein database revealed that the RA-S.GP-II sequence had no significant sequence similarity with any protein described so far.

Particular care has been taken to include protease inhibitors up to the Mono Q purification step in order to avoid proteolytic degradation of the material during the preparative procedure. In addition, both preparations carried out on seminal plasma and whole semen, immediately frozen after collection, gave rise to the same *N*-terminal amino-acid sequences. The possibility that S.GP consists of an ejaculated glycosaminoglycan-bearing region, arising from a larger proteoglycan proteolytically processed in the genital tract, should therefore be considered.

Received 24 April 1992/8 June 1992; accepted 11 June 1992

We thank Professor J. P. Dadoune for providing seminal plasma. The skilful technical assistance of Miss Michelle Rougeot (I.N.S.E.R.M.) and Mr. Ly Quan Le (C.N.R.S.) is gratefully acknowledged. This research was supported by the C.N.R.S. and by the I.N.S.E.R.M. (Contrat de Recherche Externe No. 910714).

REFERENCES

Bonner, E. M. & Laskey, R. A. (1974) Eur. J. Biochem. **46**, 83–88
Bourdon, M. A. (1990) Extracellular Matrix Genes (Sandell, L. J. & Boyd, C. D., eds.), pp. 157–174, Academic Press, San Diego, U.S.A.
Dische, Z. (1947) J. Biol. Chem. **167**, 189–198
Guess, M., Sautiere, E., Belaiche, D., Martinage, A., Roux, C., Dadoune, J. P. & Chevaillier, P. (1986) Biochim. Biophys. Acta **884**, 124–134
Hadley, M. A., Byers, S. W., Suraez-Quian, C. A., Kleinman, H. K. & Dym, M. (1985) J. Cell Biol. **101**, 1511–1522
McDevitt, C. A. & Muir, H. (1971) Anal. Biochem. **44**, 612–622
Oegema, T. R., Hascall, V. C. & Dziewiatkowski, D. D. (1975) J. Biol. Chem. **250**, 6151–6159
Périn, J.-P., Bonnet, F., Maillet, P. & Jollès, P. (1988) Biochem. J. **255**, 1007–1013
Rodriguez, J. P. & Minguell, J. J. (1989*a*) Int. J. Androl. **12**, 231–239
Rodriguez, J. P. & Minguell, J. J. (1989*b*) Cell Biochem. Funct. **7**, 293–300
Ruoslahti, E. (1988) Annu. Rev. Cell Biol. **4**, 229–255
Ruoslahti, E. (1989) J. Biol. Chem. **264**, 13369–13372
Sanderson, R. D. & Bernfield, M. (1988) Proc. Natl. Acad. Sci. U.S.A. **85**, 9562–9566
Saunders, S., Jalkanen, M., O'Farrell, S. & Bernfield, M. (1989) J. Cell Biol. **108**, 1547–1556
Schägger, H. & Von Jagow, G. (1987) Anal. Biochem. **166**, 368–379
Skinner, M. K. & Fritz, I. B. (1985) J. Biol. Chem. **260**, 11874–11883
Skinner, M. K., Tung, P. S. & Fritz, I. B. (1985) J. Cell Biol. **100**, 1941–1947
Tantravahi, R. M., Stevens, R. L., Austen, K. F. & Weis, J. H. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 9207–9210
Tishler, P. V. & Epstein, C. J. (1968) Anal. Biochem. **22**, 89–98
Tung, P. S. & Fritz, I. B. (1980) Biol. Reprod. **23**, 207–217