The EMBO Journal vol.8 no.10 pp.2975-2981, 1989

Multiple domains of the large fibroblast proteoglycan, versican

Dieter R.Zimmermann and Erkki Ruoslahti

La Jolla Cancer Research Foundation, 10901 North Torrey Pines Road, La Jolla, CA 92037, USA

Communicated by C.G.Gahmberg

The primary structure of a large chondroitin sulfate proteoglycan expressed by human fibroblasts has been determined. Overlapping cDNA clones code for the entire 2389 amino acid long core protein and the 20-residue signal peptide. The sequence predicts a potential hyaluronic acid-binding domain in the amino-terminal portion. This domain contains sequences virtually identical to partial peptide sequences from a glial hyaluronate-binding protein. Putative glycosaminoglycan attachment sites are located in the middle of the protein. The carboxy-terminal portion includes two epidermal growth factor (EGF)-like repeats, a lectin-like sequence and a complement regulatory protein-like domain. The same set of binding elements has also been identified in a new class of cell adhesion molecules. Amino- and carboxy-terminal portions of the fibroblast core protein are closely related to the core protein of a large chondroitin sulfate proteoglycan of chondrosarcoma cells. However, the glycosaminoglycan attachment regions in the middle of the core proteins are different and only the fibroblast core protein contains EGF-like repeats. Based on the similarities of its domains with various binding elements of other proteins, we suggest that the large fibroblast proteoglycan, herein referred to as versican, may function in cell recognition, possibly by connecting extracellular matrix components and cell surface glycoproteins.

Key words: chondroitin-dermatan sulfate proteoglycan/complement regulatory protein-like/epidermal growth factor-like/hyaluronic acid/lectin

Introduction

Proteoglycans are a heterogenous group of proteins that are characterized by their content of glycosaminoglycan (Ruoslahti, 1988). Large aggregating proteoglycans that carry primarily chondroitin sulfate side chains are present in the extracellular matrix of a wide variety of tissues including cartilage, aorta, skin, tendon, placenta and brain (for reviews see Poole, 1986; Heinegård and Sommarin, 1987). The sizes of their protein cores range from 200 to 400 kd and, depending on the tissue origin, they may contain up to 100 chondroitin sulfate (or dermatan sulfate) side chains covalently attached to the core protein. These large proteoglycans are likely to play a role in cell-extracellular matrix interactions, as their core proteins bind to hyaluronic acid and contain potential binding sites for other molecules.

The complete amino acid sequence of the large aggregating proteoglycan from rat chondrosarcoma (Doege *et al.*, 1987) and partial sequences for its homologs from bovine (Oldberg *et al.*, 1987) and chicken cartilage (Sai *et al.*, 1986) have been deduced from cloned cDNAs. The core protein contains at its amino-terminus a hyaluronic acid-binding motif and close to its carboxy-terminal end is a lectin-like domain that has been shown to interact with monosaccharides such as fucose and galactose (Halberg *et al.*, 1988).

A partial amino acid sequence is known for a large fibroblast proteoglycan that also contains a lectin-like domain (Krusius *et al.*, 1987). Unlike the cartilage proteoglycan, the fibroblast proteoglycan includes two epidermal growth factor (EGF)-like repeats immediately amino-terminal of the lectin-like domain. These intriguing structural features encouraged us to complete the sequencing of the core protein.

In this paper we describe the cDNA sequence coding for the entire core protein of the large fibroblast proteoglycan. The sequence reveals a mosaic structure that is more complex than that of the large cartilage proteoglycan and shares structural motifs with two groups of adhesion molecules that participate in lymphocyte homing. We propose the name versican from 'versatile' and 'proteoglycan' for this proteoglycan.

Results

Versican cDNA clones and predicted amino acid sequence

Screening of a placental library resulted in the isolation of six overlapping cDNA clones (Figure 1) that together with the previously described clones 1A, 4A and 2B (Krusius *et al.*, 1987) cover the entire coding sequence of versican (7227 bp) as well as 266 bp of 5′ and 731 bp of 3′ flanking sequence (Figure 2). The deduced 2409-residue polypeptide, including signal sequence, has a molecular weight of 265 048. The amino-terminal 20 amino acids form a typical secretory signal sequence with a putative cleavage site between Ala-20 and Leu-21 (von Heijne, 1986). This site

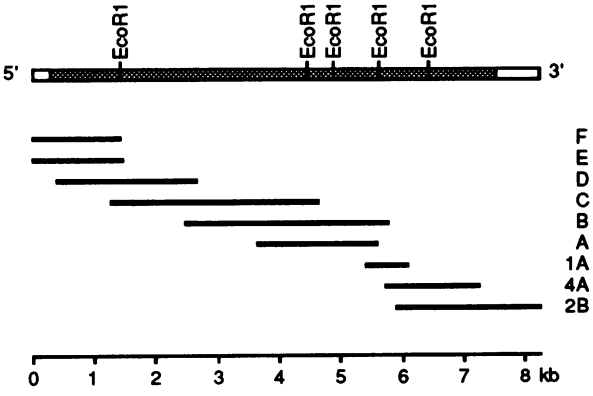

Fig. 1. Localization and nomenclature of the overlapping cDNA clones coding for versican.

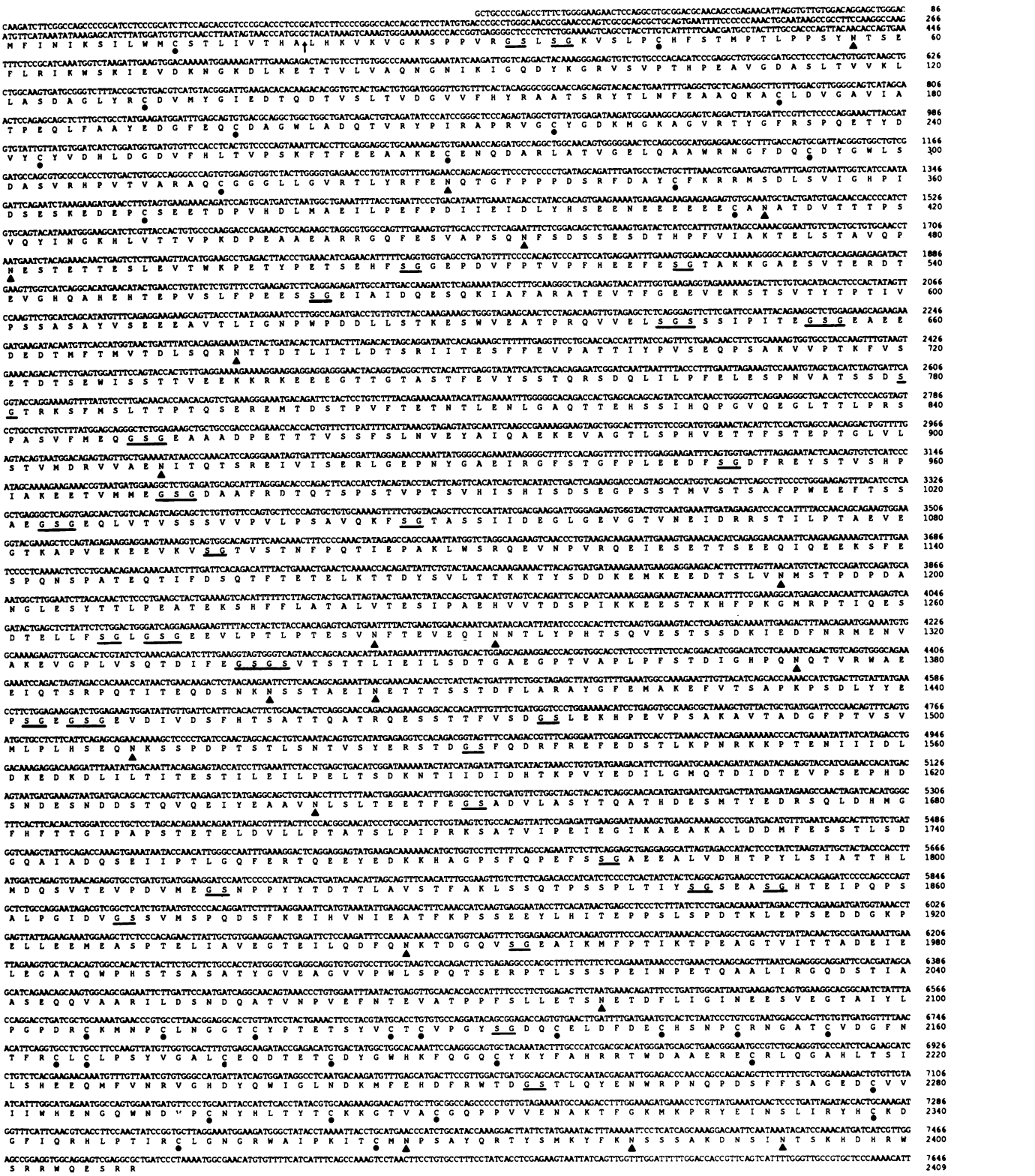

Fig. 2. cDNA and deduced protein sequences of versican. The arrow indicates the putative signal cleavage site, dots mark cysteine residues and triangles show potential *N*-glycosylation sites. Serine−glycine and glycine−serine sequences are underlined. Serine in position 371 (codon AGT) is substituted in clone E by threonine (codon ATT).

is in agreement with the (−3, −1) rule (von Heijne, 1986). The sequence immediately upstream of the ATG start codon conforms with the consensus sequence for translation initiation in eukaryotes (Kozak, 1984).

The core protein of versican is rich in glutamic acid, serine and threonine residues. It is highly negatively charged with a calculated pI of 4.2. There are a total of 34 cysteine residues in the mature protein. They are exclusively located

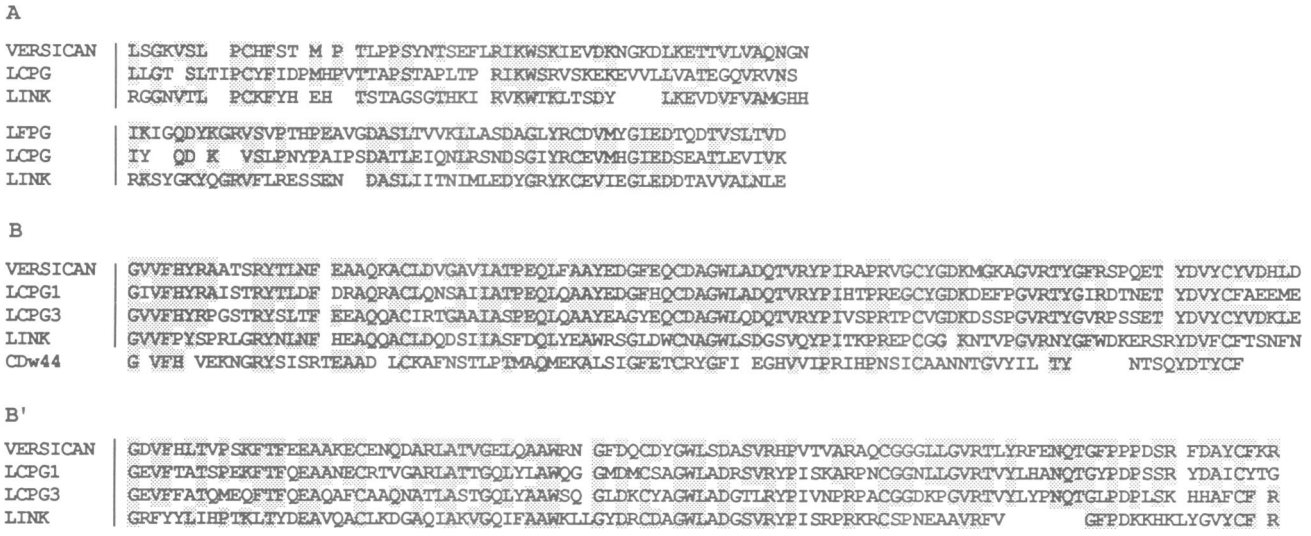

Fig. 3. Comparison of the link protein-like sequence of versican with corresponding domains in other proteins. Amino acid sequences of versican (residues 36–348) are aligned with the large cartilage proteoglycan domain 1 (LCPG1, residues 42–351) and domain 3 (LCPG3, residues 486–684) (Doege *et al.*, 1987), chicken link protein (LINK, residues 53–352) (Deák *et al.*, 1986) and CDw44/gp90^{Hermes} (CDw44, residues 32–119) (Stamenkovic *et al.*, 1989). Link protein loops A, B, B′ are compared separately. Residues matching with the versican sequence are shaded in the background.

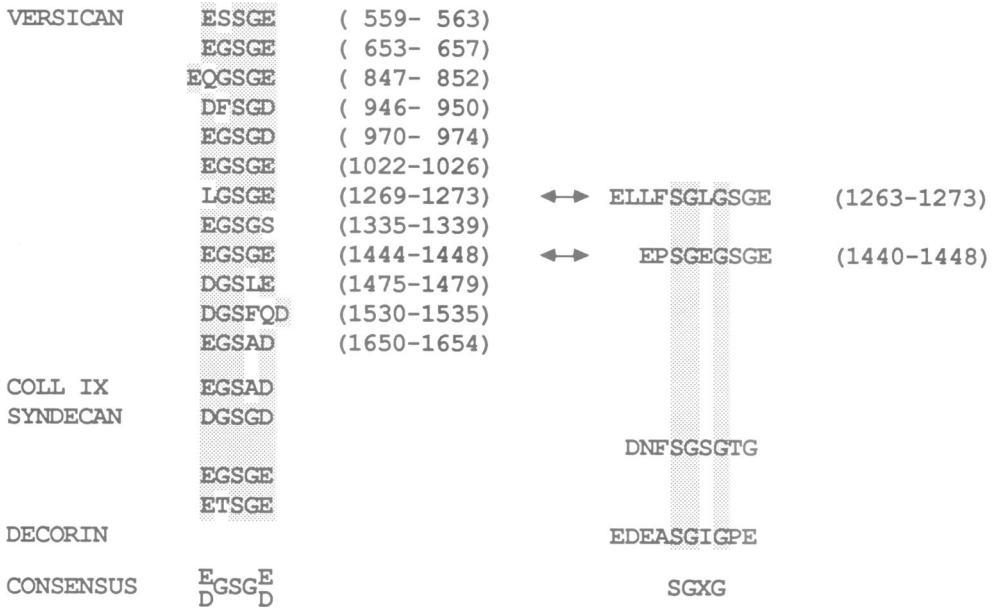

Fig. 4. Putative glycosaminoglycan attachment sites of versican. Comparison with sites in collagen type IX (Huber *et al.*, 1988), syndecan (Saunders *et al.*, 1989) and decorin (PG-40, Krusius and Ruoslahti, 1986). Shaded boxes mark residues conforming with the corresponding consensus sequences.

in the amino- and carboxy-terminal domains. Cysteine is absent over a stretch of almost 1700 amino acid residues in the middle of the core protein.

The versican core protein contains 20 potential *N*-glycosylation sites, and a series of serine–threonine clusters (22 sites with three and more consecutive serine or threonine residues) which may serve as acceptors for *O*-linked carbohydrates (Tomita *et al.*, 1978).

The amino-terminal domain of versican is similar to hyaluronic acid-binding proteins

The amino-terminal domain (residues 36–348) of the versican core protein is similar to the three loop structure of link protein (Deák *et al.*, 1986; Doege *et al.*, 1986; Kiss *et al.*, 1987) and to the link protein-like sequences in the large cartilage proteoglycan (Doege *et al.*, 1987) (Figure 3). The amino acid sequences of the B and B′ loops, which are known to mediate the hyaluronic acid binding of link protein (Goetinck *et al.*, 1987), are especially well conserved. Moreover, CDw44/gp90^{Hermes}, which is thought to be involved in lymphocyte homing, contains a link protein B-loop domain with 30% amino acid sequence identity to versican (Goldstein *et al.*, 1989; Stamenkovic *et al.*, 1989).

On the carboxy-terminal side of the link protein-like sequence is an ~200-amino acid domain containing two cysteines and an unusual cluster of glutamic acid residues (EENEEEEEE). This region is followed by the glycosaminoglycan attachment domain.

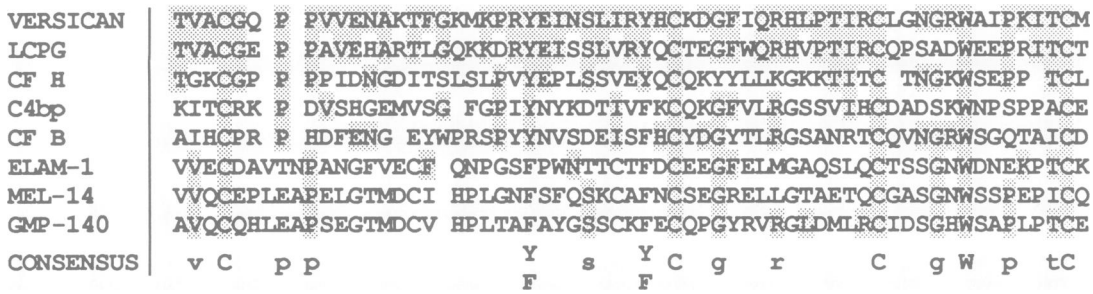

Fig. 5. Amino acid sequence comparison of complement regulatory protein-like domain of versican and other proteins. Residues 2306–2366 from versican were aligned with sequences from the large cartilage proteoglycan (LCPG, 2038–2098) (Doege *et al.*, 1987), complement factor H (CF H, 1111–1169) (Kristensen and Tack, 1986), complement C4 binding protein (C4bp, 188–247) (Chung *et al.*, 1985), complement factor B (CF B, 75–134) (Mole *et al.*, 1984), ELAM-1 (218–279) (Bevilacqua *et al.*, 1989), MEL-14 (156–217) (Siegelman *et al.*, 1989) and GMP-140 (280–341) (Johnston *et al.*, 1989). Shaded boxes mark residues matching with the versican sequence. The consensus motif includes residues identical to versican in all (upper case letters) or in more than half (lower case letters) of the compared sequences.

Glycosaminoglycan attachment domain

The region between amino acid residues 559 and 1654 contains a series of serine–glycine or glycine–serine sequences (Figure 4) that are similar to the chondroitin sulfate attachment site in collagen type IX (Huber *et al.*, 1988). These sequences are likely to carry one, or in the case of the sequence between residues 1335 and 1339 two, glycosaminoglycan chains attached to serine residues. Furthermore, two sites that are similar to the consensus glycosaminoglycan attachment sequence, SGXG, in various small proteoglycans (Bourdon *et al.*, 1987) are located in this domain. Versican contains 16 additional serine–glycine or glycine–serine sequences. However, they show little similarity with known chondroitin sulfate attachment sites. Finally, 17 serine–threonine clusters and 10 potential *N*-glycosylation sites are located in this domain.

Carbohydrate-binding and protein interaction domains at the versican carboxy-terminus

Versican has previously been found to contain two EGF-like repeats (residues 2103–2178) and a lectin-like domain (residues 2179–2305) (Krusius *et al.*, 1987). Further analysis of the published amino acid sequence reveals a previously unrecognized domain in the carboxy-terminal portion of the core protein (residues 2306–2366). This domain is similar to the repeats in the complement regulatory proteins murine factor H (Kristensen and Tack, 1986) and human C4 binding protein (Chung *et al.*, 1985) as well as complement factor B (Mole *et al.*, 1984) (Figure 5). The similarities between the versican domain and the most homologous repeats in these complement proteins vary between 25 and 34% identity. The carboxy-terminal end of the cartilage proteoglycan contains the same motif. It shows the highest similarity with versican (61% over 61 residues). The complement regulatory-like sequences in LEC-CAMs (Lectin, EGF, Complement–Cellular Adhesion Molecules) (Stoolman, 1989), which include GMP-140 (Johnston *et al.*, 1989), ELAM-1 (Bevilacqua *et al.*, 1989) and MEL-14 (Lasky *et al.*, 1989; Siegelman *et al.*, 1989), are also related to this versican domain.

Discussion

Based on our sequence analysis of the versican cDNA clones, we propose a structural model for this proteoglycan, in which the core protein contains seven distinct domains (Figure 6). Among these domains are four potential binding elements that may be involved in interactions among extracellular matrix molecules and between cells and extracellular matrices. These domains are a link protein-like sequence at the amino-terminal end and EGF-like, lectin-like and complement regulatory protein-like domains at the carboxy-terminal portion of the protein core. The intervening 1750-amino acid stretch includes the glycosaminoglycan attachment domain.

Our sequence analysis shows that versican contains attachment sites for 12–15 chondroitin sulfate side chains, 20 *N*-linked oligosaccharide chains and a number of *O*-linked oligosaccharide chains. The presence of *O*- and *N*-linked oligosaccharides could explain the discrepancy between the theoretical M_r of the core protein, which is 262 744 daltons after signal peptide cleavage, and the reported size of 290–400 kd after treatment with chondroitinase ABC (Johansson *et al.*, 1985; Krusius *et al.*, 1987).

Versican contains 12 repeats of a sequence that are likely to serve as one type of chondroitin sulfate attachment sites. The consensus sequence of these repeats is E/DGSGE/D, in which one amino acid residue (usually one of the glycines) is variable. This consensus sequence is similar to the attachment site of the single chondroitin sulfate chain in collagen type IX (Huber *et al.*, 1988). Three sequences in the chondroitin–heparan sulfate proteoglycan syndecan (Saunders *et al.*, 1989) also follow this pattern. Another two sites of versican belong to the SGXG glycosaminoglycan attachment signal previously established in a group of small proteoglycans (Bourdon *et al.*, 1987). Whether both types of consensus sequences are recognized by the same set of glycosylation enzymes has yet to be determined. In versican, attachment of glycosaminoglycan side chains might be crucial for maintaining an extended structure that would ensure the separation of the specific functional elements located at the ends of the polypeptide. Such a structure has been established for the cartilage proteoglycan by electron microscopy (Paulsson *et al.*, 1987).

The globular domain at the NH_2-terminus includes a complete link protein-like structure. The presence of such a domain in versican agrees with the hyaluronic acid binding activity reported for what seems to be the same proteoglycan (Johansson *et al.*, 1985). Versican contains only one link protein-like domain, whereas the cartilage proteoglycan includes a second partial link protein-like domain (Doege *et al.*, 1987). However, electron microscopical data have

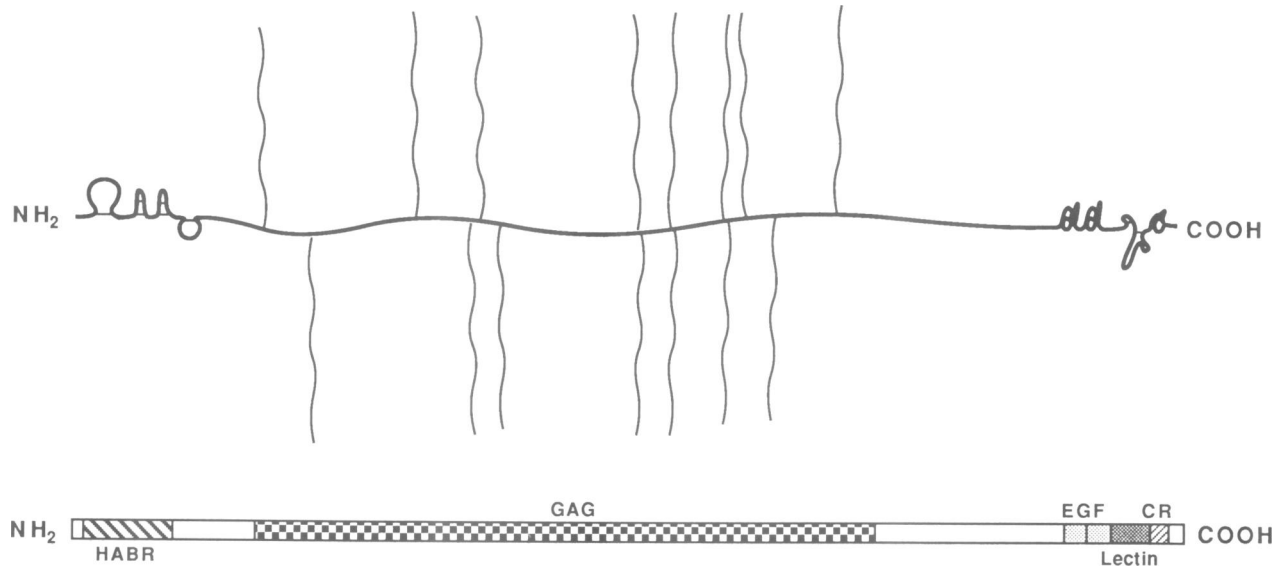

Fig. 6. Structural model of versican. The thicker line represents the core protein and wavy lines illustrate chondroitin sulfate side chains. The disulfide bonds were assigned based on structures of similar proteins. The organization of the link protein-like hyaluronic acid binding region (HABR) and glycosaminoglycan attachment (GAG), EGF-like (EGF), lectin-like (lectin) and complement regulatory protein-like (CR) domains of versican is shown schematically.

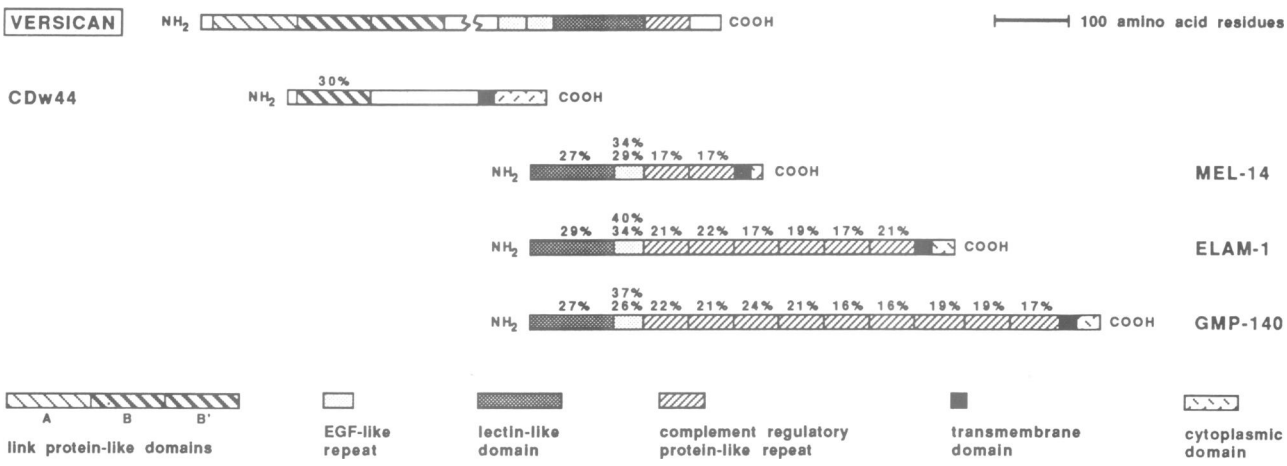

Fig. 7. Comparison of the binding elements in amino- and carboxy-terminal domains of versican with the structural organization of LEC-CAMs and with CDw44/gp90^{Hermes}. The percent identities to corresponding domains of versican are noted above the depicted domains.

demonstrated that this second domain is not involved in the interactions of the large cartilage proteoglycan with hyaluronic acid and link protein (Mörgelin *et al.*, 1988). The B-loops of the link protein-like domain in versican are also similar to a sequence in the lymphocyte adhesion protein CDw44/GP90^{Hermes}, which may also bind to hyaluronic acid (Goldstein *et al.*, 1989; Stamenkovic *et al.*, 1989).

While the link protein-like domains of the proteins discussed above are only related to the corresponding versican domain, one published link protein-like sequence may be identical to it. Short peptide sequences have been recently reported for a glial hyaluronate-binding protein (GHAP) isolated from human brain white matter (Perides *et al.*, 1989). These sequences are virtually identical with sequences within the amino-terminal domain of versican. The few differences probably reflect deamidation reactions and/or sequencing errors. These similarities suggest that GHAP is a proteolytically processed form of versican and that it includes the link protein-like portion and the adjacent glutamic acid rich structure of versican.

The versican sequence adjacent to the link protein-like domain includes an unusual poly-glutamic acid stretch. Similar sequences occur in a number of DNA binding proteins, where they are thought to interact with histones (LaPeyre *et al.*, 1987; Grönemeyer *et al.*, 1985; Kessel *et al.*, 1987). A cluster of glutamic acids seems to be responsible for the binding of the cell adhesion protein bone sialoprotein to hydroxyapatite (Oldberg *et al.*, 1988). Whether also versican is present in bone is not known.

The carboxy-terminal portion of versican consists of two EGF-like repeats, a lectin-like domain and one complement regulatory protein-like sequence. EGF-like repeats of some other proteins are known to participate in a number of protein–protein interactions (for review see Appella *et al.*, 1988). The two versican EGF-like repeats are closely related to similar sequences in coagulation factor IX. The factor IX

repeats are probably involved in the binding to a factor IX specific cell surface receptor (Rimon *et al.*, 1987) and it seems possible that versican might also interact with a cell surface receptor through its EGF-like repeats. A further possibility is that cell growth may be promoted by these EGF-like sequences as has been reported for laminin (Panayotou *et al.*, 1989).

The lectin-like domain in versican may interact with carbohydrates, as the similar region in the cartilage proteoglycan has been demonstrated to bind fucose and galactose (Halberg *et al.*, 1988). However, the galactose-containing collagens type I or II did not interact with the lectin-like domain of the cartilage proteoglycan, suggesting that the affinities of these proteoglycans may be selective for more complex oligosaccharides.

Both versican and the cartilage proteoglycan contain at their carboxy-terminal end a complement regulatory protein-like domain. This structural element, identified in a number of proteins, is thought to mediate protein−protein interactions (reviewed in Müller-Eberhard, 1988). However, versican and the cartilage proteoglycan contain only one of these structures, whereas other proteins usually include several repeats of this sequence element.

A proteoglycan similar to versican has been isolated from developing chicken limb buds (Kimata *et al.*, 1986). The expression of this proteoglycan, termed PG-M, closely follows condensation of mesenchymal cells. Binding experiments suggest that it interacts with collagen type I and fibronectin (Yamagata *et al.*, 1986). These molecules, therefore, are candidates for ligands of either the glycosaminoglycan chains or the various binding elements of versican core protein.

The carboxy terminal portion of versican contains the same structural motifs, EGF-like, lectin-like and complement regulatory protein-like units as a group of cell adhesion molecules termed LEC-CAMs. The same motifs, albeit in different order than in versican, are present in the lymphocyte homing and adhesion molecules ELAM-1 (Bevilacqua *et al.*, 1989), MEL-14 (Siegelman *et al.*, 1989; Lasky *et al.*, 1989) and GMP-140 (Johnston *et al.*, 1989).

Figure 7 shows the structural relations of versican, LEC-CAMs and CDw44/GP90^{Hermes}. The similarity to adhesive proteins in the carboxy-terminal portion and the occurrence of a hyaluronic acid-binding element at its amino-terminus suggest that versican could also connect hyaluronic acid with the cell membrane. Hyaluronate forms a coat around many types of cells, profoundly affecting their adhesive and migratory properties (reviewed in Laurent and Fraser, 1986). The sequence information presented here will now allow exploration of the role of versican in these processes.

Materials and methods

Construction of the random primed cDNA library

Poly(A) RNA was isolated from human placenta by the guanidine isothiocyanate method and subsequent chromatography on oligo(dT)−cellulose (Davis *et al.*, 1986). mRNA (10 μg) was primed with 0.5 μg random hexanucleotides (Pharmacia). cDNA synthesis was carried out by using the Amersham cDNA synthesis system. Prior to the first strand synthesis mRNA and primer were heated for 1 min at 70°C and immediately chilled on ice. Reverse transcription was done at 25°C for 45 min followed by incubation at 42°C for 1 h. The second strand was synthesized by the RNase H−DNA polymerase I procedure (Gubler and Hoffman, 1983). Internal *Eco*RI sites were protected by methylation with *Eco*RI methylase (Davis *et al.*, 1986). Phosphorylated *Eco*RI linkers (decamer, Pharmacia) were blunt-end ligated to the cDNA. The cDNA was restricted with *Eco*RI and size-fractionated on Sephacryl S-1000; fragments larger than 0.5 kb were ligated to *Eco*RI-restricted and dephosphorylated λgt11 arms (Promega). The ligated DNA was packaged with phage extracts (Gigapack gold, Stratagene). The resulting library contained 9×10^6 recombinants.

Plaque screening and subcloning

The unamplified cDNA library was screened with radiolabeled 5′-fragments (random oligolabeling kit, Boehringer) of the previously characterized clones. DNA of positive clones were isolated, cut with *Eco*RI and subcloned into pBluescript KS(+) (Stratagene) . Recombinant plasmids were purified by alkaline lysis (Davis *et al.*, 1986) and CsCl gradient centrifugation.

Double stranded cDNA sequencing

The DNA sequence was determined by the dideoxynucleotide chain termination method by using Sequenase sequencing kit (United States Biochemical Corporation) with Bluescript T3 and T7 primers (Stratagene) or specific 17mer oligonucleotides. The plasmid strands were separated by alkaline denaturation (1 M NaOH, 1 mM EDTA; 5 min) prior to the reaction. The entire cDNA was sequenced from both strands and all sequences were confirmed by unidirectional sequencing of a second overlapping clone. Internal *Eco*RI sites were verified by sequencing parts of the clones in λDNA (Manfioletti and Schneider, 1988). For this purpose, 10 μg λDNA was heat denatured prior to reaction. Annealing of specific primers (50 ng) was done for 5 min at room temperature.

Sequence analysis and databank comparisons

Editing and analysis of the sequence was done using MicroGenie (Beckman) and PCGene (Intelligenetics) software. Sequence comparisons with the databases Genbank (release 58.0), EMBL (17.0), NBRF/PIR (18.0) and SWISS-PROT (9.0) were performed on Bionet (Intelligenetics) using the Pearson−Lipman algorithm (Pearson and Lipman, 1988) with ktup values of 4 and 1 for DNA and protein analysis, respectively.

Acknowledgements

We thank Drs Scott Argraves, Kurt Amrein and Thomas Manes for helpful technical advice, and Dr Eva Engvall for critical reading of the manuscript. This work was supported by Grant CA42507 to E.R., Cancer Center support Grant CA 30199 from the National Cancer Institute and Stipend 86PZ48 of the Swiss National Foundation to D.R.Z.

References

Appella,E., Weber,I.T. and Blasi,F. (1988) *FEBS Lett.*, **231**, 1−4.
Bevilacqua,M.P., Stengelin,S., Gimbrone,M.A.,Jr and Seed,B. (1989) *Science*, **243**, 1160−1165.
Bourdon,M.A., Krusius,T., Campbell,S., Schwartz,N.B. and Ruoslahti,E. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 3194−3198.
Chung,L.P., Bentley,D.R. and Reid,K.B.M. (1985) *Biochem. J.*, **230**, 133−141.
Davis,L.G., Dibner,M.D. and Battey,J.F. (1986) *Basic Methods in Molecular Biology*. Elsevier, Amsterdam.
Deák,F,. Kiss,I., Sparks,K.J., Argraves,W.S., Hampikian,G. and Goetinck,P.F. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 3766−3770.
Doege,K., Hassell,J.R., Caterson,B. and Yamada,Y. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 3761−3765.
Doege,K., Sasaki,M., Horigan,E., Hassell,J.R. and Yamada,Y. (1987) *J. Biol. Chem.*, **262**, 17757−17767.
Goetinck,P.F,. Stirpe,N.S., Tsonis,P.A. and Carlone,D. (1987) *J. Cell Biol.*, **105**, 2403−2408.
Goldstein,L.A., Zhou,D.F.H., Picker,L.J., Minty,C.N., Bargatze,R.F., Ding,J.F. and Butcher,E.C. (1989) *Cell*, **56**, 1063−1072.
Grönemeyer,H., Turcotte,B., Quirin-Stricker,C. and Bocquel,M.T. (1985) *Nucleic Acids Res.*, **13**, 3985−3994.
Gubler,U. and Hoffman,B. (1983) *Gene*, **25**, 263−269.
Halberg,D.F., Proulx,G., Doege,K., Yamada,Y. and Drickamer,K. (1988) *J. Biol. Chem.*, **263**, 9486−9490.
Heinegård,D. and Sommarin,Y. (1987) *Methods Enzymol.*, **144**, 305−319.
Huber,S., Winterhalter,K.H. and Vaughan,L. (1988) *J. Biol. Chem.*, **263**, 752−756.
Johansson,S., Hedman,K., Kjellén,L., Christner,J., Vaheri,A. and Höök,M. (1985) *Biochem. J.*, **232**, 161−168.
Johnston,G.I., Cook,R.G. and McEver,R.P. (1989) *Cell*, **56**, 1033−1044.

Kessel,M., Schulze,F., Fibi,M. and Gruss,P. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 5306–5310.
Kimata,K., Oike,Y., Tani,K., Shinomura,T., Yamagata,M, Uritani,M. and Suzuki,S. (1986) *J. Biol. Chem.*, **261**, 13517–13525.
Kiss,I., Deák,F., Meštric,S., Delius,H., Soos,J., Dékány,K., Argraves,W.S., Sparks,K.J. and Goetinck,P.F. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 6399–6403.
Kozak,M. (1984) *Nucleic Acids Res.*, **12**, 857–872.
Kristensen,T. and Tack,B.F. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 3963–3967.
Krusius,T. and Ruoslahti,E. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 7683–7687.
Krusius,T., Gehlsen,K.R. and Ruoslahti,E. (1987) *J. Biol. Chem.*, **262**, 13120–13125.
LaPeyre,B., Bourbon,H. and Amalric,F. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 1472–1476.
Lasky,L.A., Singer,M.S., Yednock,T.A., Dowbenko,D., Fennie,C., Rodriguez,H., Nguyen,T., Stachel,S. and Rosen,S.D. (1989) *Cell*, **56**, 1045–1055.
Laurent,T.C. and Fraser,J.R.E. (1986) *CIBA Foundation Symp.*, **124**, 9–29.
Manfioletti,G. and Schneider,C. (1988) *Nucleic Acids Res.*, **16**, 2873–2884.
Mole,J.E., Anderson,J.K., Davison,E.A. and Woods,D.E. (1984) *J. Biol. Chem.*, **259**, 3407–3412.
Mörgelin,M., Paulsson,M., Hardingham,T., Heinegård,D. and Engel,J. (1988) *Biochem. J.*, **253**, 175–185.
Müller-Eberhard,H.J. (1988) *Annu. Rev. Biochem.*, **57**, 321–347.
Oldberg,Å., Antonsson,P. and Heinegård,D. (1987) *Biochem. J.*, **243**, 255–259.
Oldberg,Å., Franzén,A. and Heinegård,D. (1988) *J. Biol. Chem.*, **263**, 19430–19432.
Panayotou,G., End,P., Aumailley,M., Timpl,R. and Engel,J. (1989) *Cell*, **56**, 93–101.
Paulsson,M., Mörgelin,M., Wiedemann,H., Beardmore-Gray,M., Dunham,D., Hardingham,T., Heinegård,D., Timpl,R. and Engel,J. (1987) *Biochem. J.*, **245**, 763–772.
Pearson,W.R. and Lipman,D.J. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 2444–2448.
Perides,G., Lane,W.S., Andrews,D., Dahl,D. and Bignami,A. (1989) *J. Biol. Chem.*, **264**, 5981–5987.
Poole,A.R. (1986) *Biochem. J.*, **236**, 1–14.
Rimon,S., Melamed,R., Savion,N., Scott,T., Nawroth,P.P. and Stern,D.M. (1987) *J. Biol. Chem.*, **262**, 6023–6031.
Ruoslahti,E. (1988) *Annu. Rev. Cell Biol.*, **4**, 229–255.
Sai,S., Tanaka,T., Kosher,R.A. and Tanzer,M.L. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 5081–5085.
Saunders,S., Jalkanen,M., O'Farrell,S. and Bernfield,M. (1989) *J. Cell Biol.*, **108**, 1547–1556.
Siegelman,M.H., van de Rijn,M. and Weissman,I.L. (1989) *Science*, **243**, 1165–1172.
Stamenkovic,I., Amiot,M., Pesando,J.M. and Seed,B. (1989) *Cell*, **56**, 1057–1062.
Stoolman,L.M. (1989) *Cell*, **56**, 907–910.
Tomita,M., Furthmayr,H. and Marchesi,V.T. (1978) *Biochemistry*, **17**, 4756–4770.
von Heijne,G. (1986) *Nucleic Acids Res.*, **14**, 4683–4690.
Yamagata,M., Yamada,K.M., Yoneda,M., Suzuki,S. and Kimata,K. (1986) *J. Biol. Chem.*, **261**, 13526–13535.

Received on May 2, 1989; revised on June 13, 1989