The EMBO Journal vol.11 no.12 pp.4281-4290, 1992

Structure of recombinant N-terminal globule of type VI collagen α3 chain and its binding to heparin and hyaluronan

Ulrich Specks, Ulrike Mayer, Roswitha Nischt[1], Thomas Spissinger[2], Karlheinz Mann, Rupert Timpl, Jürgen Engel[2] and Mon-Li Chu[3]

Max-Planck-Institut für Biochemie, D-8033 Martinsried, Germany, [1]Department of Dermatology, University, D-5000 Köln, Germany, [2]Biozentrum, University of Basel, CH-4056 Basel, Switzerland and [3]Departments of Biochemistry and Molecular Biology and Dermatology, Jefferson Institute of Molecular Medicine, Thomas Jefferson University, Philadelphia, PA 19107, USA

Communicated by R.Timpl

A large portion of the N-terminal globule of human collagen VI was prepared from the culture medium of stably transfected human embryonic kidney cell clones. The recombinant product corresponds to sequence positions 1-1586 of the α3(VI) chain that consists of eight homologous ~200 residue motifs (N9 to N2) being similar to the A domain motif of von Willebrand factor. By ultracentrifugation fragment N9-N2 showed a molecular mass of 180 kDa and an asymmetric shape. Elongated structures that consist of eight small globes (diameter ~5 nm) were demonstrated by electron microscopy. The data indicate that each A domain motif represents a separate folding unit which are connected to each other by short protease-sensitive peptide segments. Circular dichroism studies demonstrated about 38% α helix, 14% β sheets and 17% β turns. Fragment N9-N2 showed binding to heparin which could be abolished by moderate salt concentrations. Heparin binding was assigned to domains N9, N6 and N3 which were obtained after partial proteolysis. Domains N7, N5 and N4 lacked affinity for heparin. In addition, N9-N2 showed strong binding to hyaluronan that required exposure to 6 M urea for full dissociation. Ligand binding studies indicated some affinity of N9-N2 for the triple helical region of collagen VI suggesting a role of the N-terminal globule in the self-assembly of microfibrils. No or only little binding was, however, observed to fibril-forming collagens I and III, several basement membrane proteins and other extracellular proteins. Fragment N9-N2 was also an inactive substrate for cell adhesion.

Key words: A domain motif/collagen microfibrils/eukaryotic expression vector/glycosaminoglycans/protein-protein and cell-matrix interactions

Introduction

Collagen VI is a unique member within the large family of collagenous proteins (Van der Rest and Garrone, 1991) that forms abundant microfibrils in a large variety of extracellular matrices (Timpl and Engel, 1987). The self-assembly process is also rather unique and involves the formation of defined dimers and tetramers which by end-to-end aggregation and some lateral associations build up microfibrils with a characteristic 100 nm periodicity (Furthermayr *et al.*, 1983; von der Mark *et al.*, 1984; Bruns *et al.*, 1986; Engvall *et al.*, 1986; Keene *et al.*, 1988). Collagen VI was also identified as a major cell-adhesive protein (Aumailley *et al.*, 1989b; Wayner and Carter, 1987) with an activity comparable with that of laminin, fibronectin and collagen IV. More recent studies showed that collagen VI microfibrils associate with hyaluronan and that this interaction is essential for fibril assembly (McDevitt *et al.*, 1991; Kielty *et al.*, 1992). Electron microscopical evidence also indicated a close association of collagen VI with major fibrils composed of collagens I and III and with the distal surface of basement membranes (von der Mark *et al.*, 1984; Bruns *et al.*, 1986; Keene *et al.*, 1988; Okada *et al.*, 1990).

The various potential functions of collagen VI have so far not been sufficiently addressed at the molecular level reflecting the inherent difficulty of obtaining native material from tissues (Timpl and Engel, 1987). A new approach became feasible by the complete elucidation of the sequence of the three polypeptide chains, α1(VI), α2(VI) and α3(VI), based on cDNA cloning of human (Chu *et al.*, 1988, 1989, 1990) and chick collagen VI (Koller *et al.*, 1989; Bonaldo *et al.*, 1989, 1990; Bonaldo and Colombatti, 1989). The data demonstrated ~1000 amino acid residues for the α1(VI) and α2(VI) and ~3000 residues for the α3(VI) chain. The triple helical collagenous domain contributed by all three chains comprises only 20% of the total mass. All three chains also share three 200 residue non-collagenous domains which are homologous to the A domain motif originally identified in von Willebrand factor (Titani *et al.*, 1986; Titani and Walsh, 1988). Meanwhile it is known that such motifs occur in several more extracellular and receptor proteins (reviewed in Colombatti and Bonaldo, 1991). The extra mass of the α3(VI) chain is primarily contributed by eight more A motifs, referred to as N9 to N2 in human collagen VI (Chu *et al.*, 1990), that are located at the N-terminus. Some of them were recently investigated in the form of fusion proteins indicating that they may bind to fibril-forming collagens (Bonaldo *et al.*, 1990).

In the present study we have prepared a recombinant fragment comprising motifs N9-N2 in a eukaryotic expression system used previously for the production of other extracellular matrix proteins (Fox *et al.*, 1991; Nischt *et al.*, 1991). Structural studies demonstrated that each individual motif represents an independent folding unit that forms a globular structure. The whole fragment possesses several binding sites for heparin and hyaluronan and also binds to the triple helical domain of collagen VI. However, several other extracellular components including fibril-forming collagens and a variety of adherent cells lacked binding activity for fragment N9-N2.

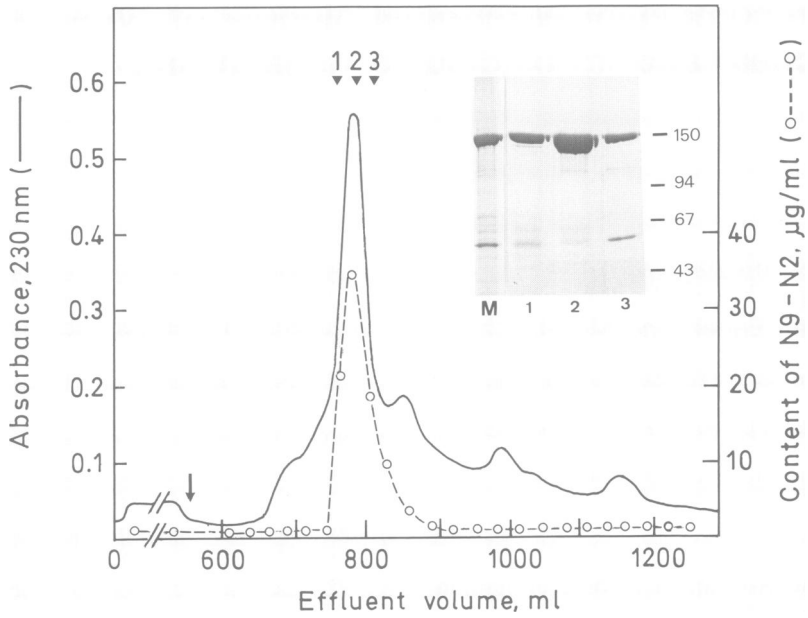

Fig. 1. Purification of recombinant fragment N9–N2 from serum-free medium on DEAE cellulose. The column (2.5 × 25 cm) was equilibrated in 0.05 M Tris–HCl pH 8.6, loaded with 0.5 l medium in the same buffer and eluted with a linear gradient (start denoted by arrow) from 0 to 0.5 M NaCl (500/500 ml). The content of N9–N2 was measured in individual fractions by radioimmuno-inhibition assay (dashed line). The insert shows SDS electrophoresis of total medium (M) and individual fractions (Nos 1–3, arrowheads). The run was calibrated with globular marker proteins (denoted in kDa).

Results

Production, isolation and identification of recombinant fragment N9–N2

A construct was made in a eukaryotic expression vector encoding the signal peptide of human collagen VI α3 chain, a consecutive row of eight homologous 200 residue domains N9–N2 (Chu *et al.*, 1990) and at the C-terminus a tetrapeptide derived from the polylinker. This vector was used to establish several stably transfected clones in the human kidney 293 cell line. In serum-free culture medium of the clones the recombinant product could be easily identified as a major 150 kDa band by electrophoresis (see insert in Figure 1) while this band was lacking in the parental 293 cells. Northern hybridization of the clones demonstrated abundant amounts of a 5.0 kb mRNA which is the size expected from the cDNA construct. No hybridization was observed with 293 cell RNA indicating a low production of endogenous collagen VI α3 chain (Figure 2).

The recombinant product was purified from large volumes (0.5–1 l) of serum-free culture medium of a single clone. Such medium contained 10–15 µg/ml of fragment N9–N2 as shown by radioimmunoassay. Passage over DEAE cellulose showed that most of the material eluted from the column in a single peak with a purity of ~90% (Figure 1). Further purification to >95% electrophoretic homogeneity was then achieved either by molecular sieve chromatography or by heparin affinity chromatography (not shown). The final yields were 5–8 mg/l culture medium.

Purified fragment N9–N2 showed in SDS electrophoresis a single band of 150 kDa that did not change mobility under reducing conditions (Figure 3). The amino acid composition was within the limits of analytical error of that predicted from the sequence. The fragment contained only little glucosamine (0.3 mol/mol protein) and no galactosamine. Edman degradation failed to yield any significant amounts of PTH

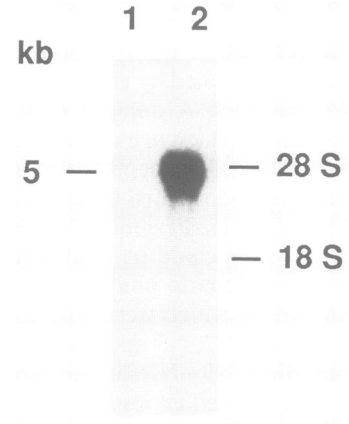

Fig. 2. Identification of mRNA encoding fragment N9–N2 by Northern hybridization. Lanes were loaded with total RNA from the parental kidney cell line 293 (lane 1, 10 µg) and from a 293 cell clone stably transfected with N9–N2 (lane 2, 2 µg). The size of the mRNA (5 kb) and of the positions of rRNA are indicated.

derivatives indicating a blocked N-terminus. Cyanogen bromide cleavage, however, produced a mixture of fragments. Four of the major fragments of ~50, 29, 25 and 21 kDa could be well resolved by electrophoresis and showed after blotting a distinct N-terminal sequence (Figure 3). These sequences fitted perfectly with the observed size of the peptides and that predicted from the position of the adjacent methionines (Chu *et al.*, 1990).

Size, shape and conformation

Fragment N9–N2 showed in neutral buffer a molecular mass of 180 000 ± 5000 as determined by sedimentation equilibrium runs. This value is in excellent agreement with the presence of eight domains each having a molecular mass of ~22 kDa. Sedimentation velocity runs demonstrated that

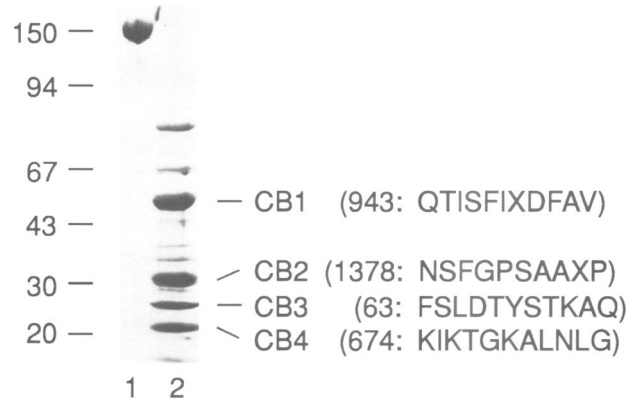

Fig. 3. SDS gel electrophoresis of purified N9−N2 (lane 1) and CNBr cleaved N9−N2 (lane 2). Major CNBr peptides (CB1−CB4) were after blotting subjected to Edman degradation and their N-terminal sequences and starting positions are identified. Electrophoresis was performed under non-reducing conditions and the run was calibrated with globular markers (molecular masses in kDa shown on the left side).

$s_{20,w}$ = 6.5. A frictional ratio f/f_o = 1.51 was calculated from the sedimentation coefficient and the molecular mass (Tanford, 1961). This indicates a rather elongated shape in solution with an axial ratio a/b of ~9.

An elongated shape was also observed for fragment N9−N2 by electron microscopy (Figure 4). Rotary shadowing demonstrated mainly compact rod-shaped or oval particles. A similar gross shape was also shown by negative staining. Here, most of the particles appeared as six to eight individual globular structures of similar diameter. Most characteristic were particles of ~20 nm length with a compact zig-zag arrangement of eight globules. A more extended row-like association of these globules with a total length of 40 nm was also occasionally observed. The individual globes had a diameter of ~5 nm. The calculated diameter for spheres comprising individual N domains with a molecular mass of 22 kDa would be 3.7 nm. The reasonable agreement between both values indicates that each globular structure visualized by negative staining corresponds to an individual N domain predicted from the sequence. The dimension (5 × 40 nm) for the most extended N9−N2 structure is also in good agreement with the axial ratio of 9 determined by ultracentrifugation.

The circular dichroism spectrum of fragment N9−N2 showed in the far UV two distinct peaks of negative ellipticity which are characteristic of α-helix (Figure 5). The spectrum was analyzed in the wavelength range 190−250 nm for the contributions of known secondary structure elements (Hennessey and Johnson, 1981). This indicated that fragment N9−N2 consists of 38% α-helix, 14% β-sheets, 17% β-turns and 30% non-periodic structures.

Immunological studies

A rabbit antiserum was raised against fragment N9−N2 and showed strong binding to the fragment in radioimmunoassays and ELISA (titer ~1:50.000). The antiserum did not react with pepsin-solubilized human placenta collagen VI (titer <1:200) which as previously shown has lost almost all of its N-terminal globular domains (Chu *et al.*, 1987). It was also possible to establish a sensitive radioimmuno-inhibition assay for fragment N9−N2 (Figure 6) where the fragment was a potent inhibitor (IC_{50} = 8 ng/ml). This assay was used to quantitate the fragment in the culture medium of transfected cell clones (~10 μg/ml) and to monitor its purification (see above). Distinct although lower amounts of antigens possessing N9−N2 epitopes (0.02−0.1 μg/ml) were detected in HT1080 cell culture medium. Studies with several other extracellular proteins (collagen XIV, fibronectin, von Willebrand factor) that in part possess domains homologous to N9−N2 demonstrated that they were non-inhibitory even when used at 10 000-fold higher concentrations. A surprising observation was the distinct and relatively strong inhibition by human serum (Figure 6). The inhibition curves for fragment N9−N2 and the serum antigen showed identical slopes indicating that both share most of the epitopes. This allowed determination of the average concentration in sera of eight normal adults as 240 ± 66 ng/ml assuming that the serum antigen has the same size and affinity as fragment N9−N2.

The antiserum was also a potent reagent to detect as little as 100 ng fragment N9−N2 after electrophoresis and immunoblotting indicating its usefulness for the analysis of small samples of tissue-extracted collagen VI. This was demonstrated for 4 M urea-extracted and reduced collagen VI from a rat chondrosarcoma which showed two reactive bands in the range 200−220 kDa after blotting (data not shown). The antiserum also produced strong immunofluorescence staining of extracellular structures on frozen sections of human skin, heart, lung, liver and kidney. The staining patterns were identical to that obtained with an antiserum against pepsin-solubilized human placenta collagen VI (data not shown).

Binding to heparin and hyaluronan and identification of binding domains

Several indications exist that collagen VI may interact with glycosaminoglycans and proteoglycans (Timpl and Engel, 1987; McDevitt *et al.*, 1991; Kielty *et al.*, 1992). Therefore, we examined binding of fragment N9−N2 to heparin and hyaluronan by affinity chromatography. Rather complete binding (98%) as determined by radioimmunoassay was observed to a heparin column equilibrated in 0.1 M NaCl, 0.05 M Tris−HCl pH 7.4. Elution with a linear NaCl gradient resulted in complete displacement of bound N9−N2 at 0.20−0.23 M NaCl indicating a weak interaction with heparin. Binding of fragment N9−N2 to a hyaluronan column was of comparable efficiency (99%). A stepwise increase of NaCl concentrations from 0.15 to 0.4, 1 and 5 M NaCl resulted, however, only in little displacement (1, 8 and 14%, respectively). The major portion (75%) of N9−N2 required 6 M urea in the neutral buffer for complete elution from the column.

Limited cleavage with α-chymotrypsin and elastase followed by affinity chromatography was used to map the heparin binding domains within the N9−N2 structure more precisely. An extensive digestion with α-chymotrypsin for 24 h at 37°C produced five to six protein bands with the molecular mass of 21−28 kDa that very likely correspond to single N domains (see insert lane C in Figure 7). A second major fraction contained several 40−50 kDa bands (presumably N domain dimers) and some larger, apparently

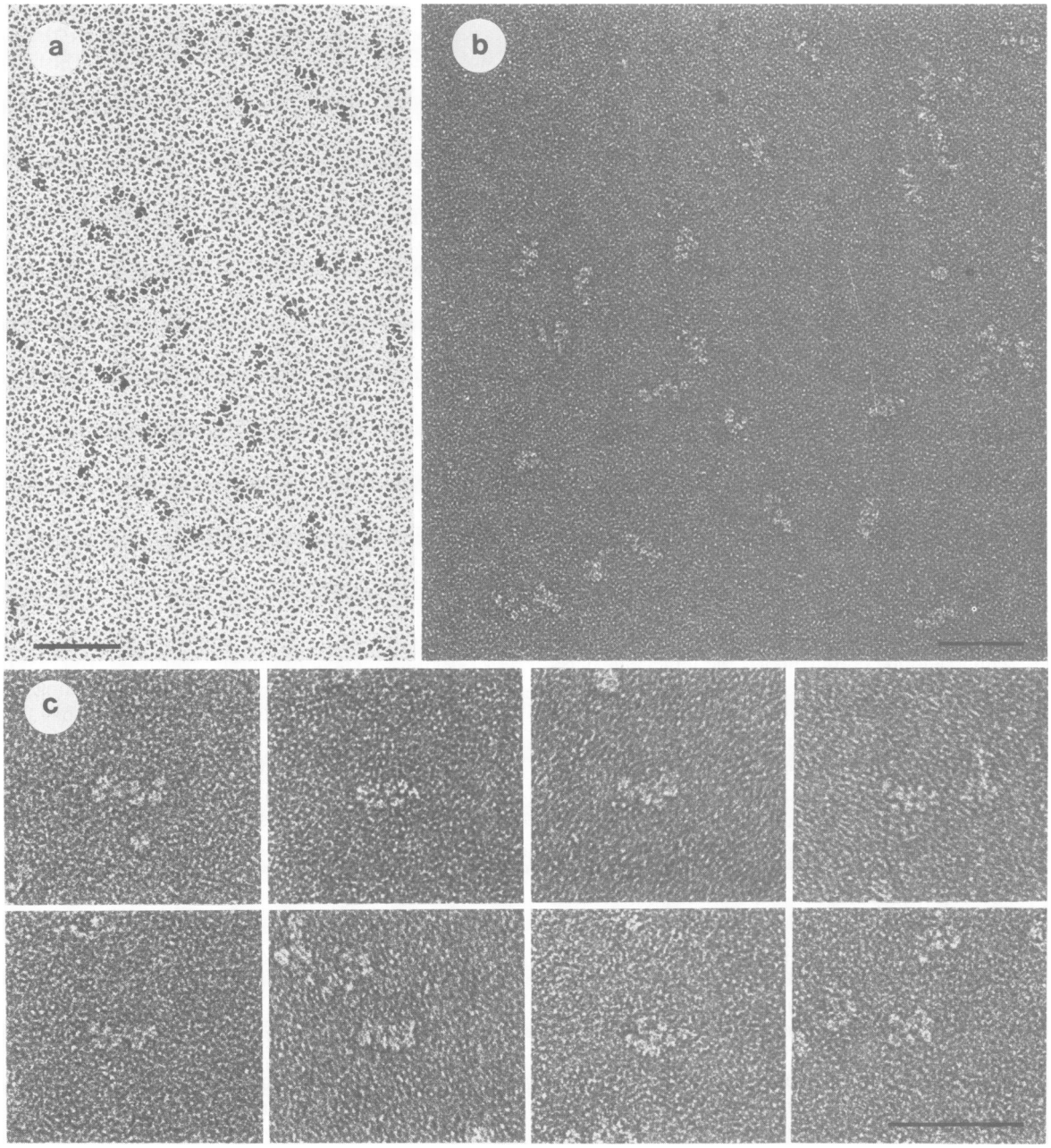

Fig. 4. Visualization of fragment N9−N2 by electron microscopy after rotary shadowing (**a**) or negative staining (**b** and **c**). Representative fields are shown in (a and b) and selected particles in (c). Bars: 50 nm.

less completely cleaved fragments. Passage of the digest over a heparin column showed that ~50% of the total material was not bound. This material consisted mainly of a 26 kDa fragment C1 and a 21 kDa fragment C2. Sequencing after blotting showed for C1 and C2 N-terminal sequences beginning in front of the predicted domains N5 and N7 respectively (Table I). This strongly indicates that C1 and C2 comprise each a single and entire N domain. The heparin binding fragments were partially resolved into three major and several minor peaks by elution with an NaCl gradient (Figure 7). Major bands of the lower molecular weight fragments were C3 (45 kDa) and C4 (28 kDa). Sequence analysis of C3 showed that it contains domains N6 and N5. No sequence was obtained from fragment C4 indicating that it originates from the N-terminus (N9). A less abundant 24 kDa fragment C5 was also obtained in the heparin-bound fraction and shown to correspond to domain N6 (Table I).

The failure to identify all individual N domains in the chymotryptic digest indicated the extensive degradation of some of them. This was underscored by the large amounts of non-binding fragments which could not be accounted for by domains N5 and N7 identified in this fraction (Figure 7). Cleavage with elastase was therefore used as a second approach. This produced a digest comparable in complexity with that obtained with α-chymotrypsin as demonstrated by electrophoresis and heparin affinity chromatography (data not shown). It allowed us to identify a non-binding 28 kDa fragment E1 as domain N4 by sequence analysis. Major binding fragments were E2 (50 kDa), E3 (29 kDa) and E4 (28 kDa). Fragment E2 was shown to correspond to domains

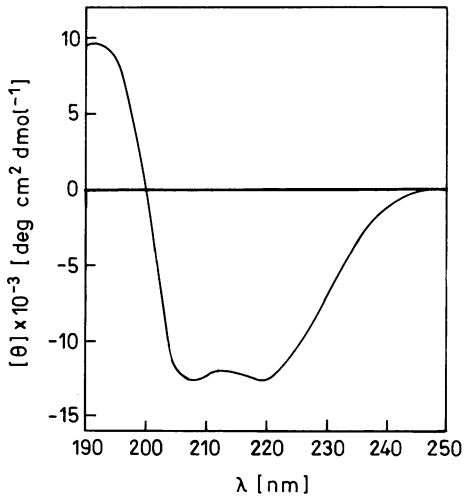

Fig. 5. Circular dichroism spectrum of fragment N9–N2. The spectrum was recorded in 0.05 M Tris–HCl pH 7.5 at 20°C.

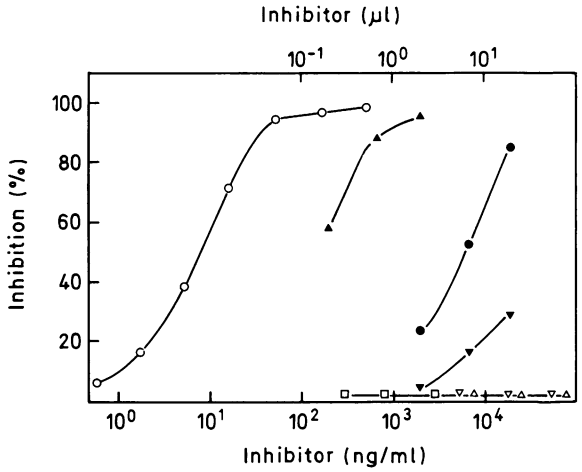

Fig. 6. Radioimmuno-inhibition assay for fragment N9–N2 and comparison with other proteins and biological fluids. The assay system consisted of a rabbit antiserum against N9–N2 and ^{125}I-labeled N9–N2. Non-labeled competitors used were N9–N2 (○), collagen XIV (□) fibronectin (△), von Willebrand factor (▽), normal human serum (●) and serum-free culture medium from HT1080 cells (▼) and a stably transfected 293 cell clone (▲). The scale on top refers to the amounts of biological fluid added in a volume of 0.2 ml.

N3 plus N2 and E3 to domain N3 (Table I). Fragment E4 lacked a recognizable N-terminal sequence and therefore very likely corresponds to domain N9.

Activity in cell adhesion and interactions with extracellular matrix proteins

Since collagen VI is a known cell-adhesive protein we also examined fragment N9–N2 as immobilized substrate in cell attachment and spreading assays. No activity was observed with nine different cell lines: HBL-100 (human mammary epithelium), HT 1080 (human fibrosarcoma), A375 (human melanoma), A431 (human epidermoid), SAOS-2 (human osteosarcoma), 293 (embryonal kidney cells), Rugli (rat glioma), RN22 (rat schwannoma) and B16F10 (mouse melanoma). All these cell lines showed strong adhesion and spreading on pepsin-solubilized human collagen VI, which as found previously (Aumailley *et al.*, 1989b), is mediated by the triple helical collagenous segment of the protein. Lack of significant activity of fragment N9–N2 was not due to insufficient coating capacity that was found to be ~4%. This is within the usual range observed for other cell-adhesive proteins.

Further binding studies were carried out with several extracellular matrix proteins using fragment N9–N2 as soluble and/or immobilized ligand. The most significant binding was observed between pepsin-solubilized collagen VI and immobilized N9–N2 while the binding of soluble N9–N2 to immobilized pepsin-solubilized collagen VI was somewhat weaker (Figure 8). Yet the reaction was completely abolished after reducing most of the disulfide bonds of pepsin-solubilized collagen VI under non-denaturing conditions. Some weaker binding of N9–N2 was also observed to recombinant nidogen, von Willebrand factor and vitronectin. Quite a few other proteins were, however, inactive as ligands. These included various batches of triple helical collagens I and III and pN-collagens I and III which were used under conditions avoiding denaturation of the triple helix. Further inactive ligands were fibronectin and the basement membrane proteins collagen IV, laminin–nidogen complex, BM-40 and a large heparan sulfate proteoglycan. The latter observation was of interest in relation to the binding to heparin demonstrated for fragment N9–N2.

Discussion

The recombinant production of the 180 kDa fragment N9–N2 has for the first time allowed us to study a substantial portion (~32%) of the collagen VI structure not yet available from tissues. It corresponds to position 1–1586 of the human α3(VI) chain sequence for which folding into eight ~200 residue domains (N9–N2 from the N-terminus) was predicted from sequence analysis and exon borders (Chu *et al.*, 1990; Stokes *et al.*, 1991). Homologous protein motifs were initially described as A domains in von Willebrand factor (Titani and Walsh, 1988) and are found in several more extracellular proteins (Colombatti and Bonaldo, 1991). However, collagen VI has the highest abundance of such motifs compared with any other known protein. The analysis of size, composition and of CNBr peptide sequences of fragment N9–N2 demonstrated the correct expression of the designed product. The glucosamine content was negligible despite the presence of three potential NXT/S oligosaccharide acceptor sites. However, two of the sequences are NPS, considered to be an inactive acceptor site. Complete glycosylation of other recombinant proteins was observed when produced by the same expression system (Fox *et al.*, 1991; Nischt *et al.*, 1991). It indicates that the genuine N9–N2 structure may not be significantly glycosylated at all. Antibody studies particularly those based on radioimmuno-inhibition assays demonstrated the same N9–N2 epitopes in cell culture media, tissue-derived collagen VI and serum. This suggested that recombinant N9–N2 was obtained in a properly folded form. Since sufficient quantities of the recombinant product could be obtained, it permitted a more comprehensive study of structural and functional properties in order to support this interpretation.

Negative staining of recombinant N9–N2 showed a maximal number of eight globular domains demonstrating that each of the predicted 200 residue motifs represents an

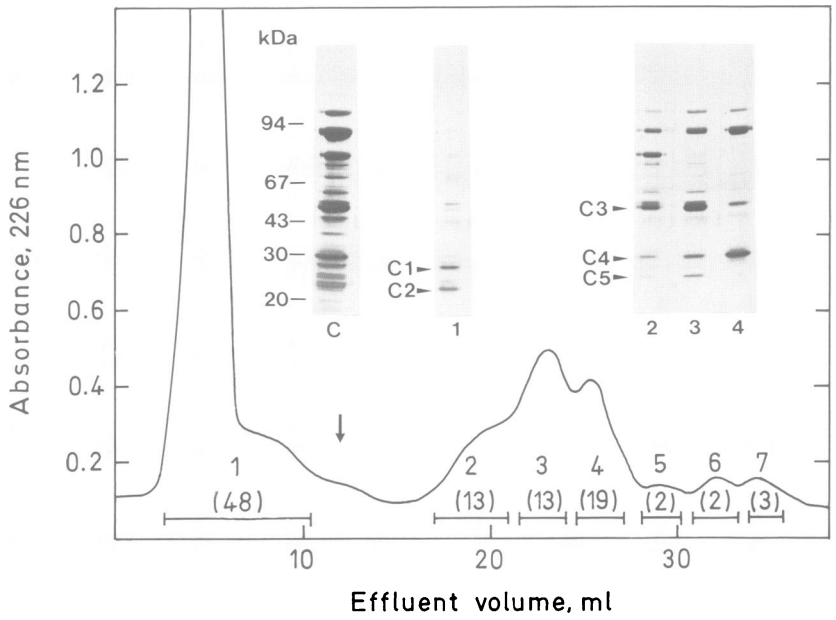

Fig. 7. Heparin affinity chromatography of a chymotryptic digest of N9–N2 and analysis by SDS electrophoresis (inserts). Pools nos 1–7 were collected as indicated by horizontal bars and their relative protein contents (in %) are shown in parentheses. Electrophoresis patterns are shown for the total digest (c) and the major pools (1–4). The runs were calibrated with globular proteins (molecular masses are given in kDa). Peptides denoted C1–C5 were identified by Edman degradation (see Table I). The arrow indicates the start of the NaCl gradient (0–0.4 M NaCl, 30 ml).

independent folding unit. These units may be associated in different ways as indicated in Figure 9a,b. An elongated form of 5 × 40 nm fits the ultracentrifugal data and may comprise the structure of fragment N9–N2 in solution. More compact forms are the most frequently observed electron-microscopical structures indicating a tighter zig-zag structure or the association of the first with the second row of four N domains which bend back on each other (Figure 9b). Since no negative staining data are yet available for native collagen VI, it is impossible to decide between these possibilities. Circular dichroism analysis also demonstrated ~70% α helical and β-sheet/turn structures. Together the data strongly indicate that the recombinant N9–N2 has formed a proper secondary and tertiary structure that is likely to resemble the structure found in native collagen VI. It remains, however, to be determined to what extent the N9–N2 structure, which is unique to α3 chain, is modified by the association with the α1 and α2 chains as in intact collagen VI. These chains share with the α3 chain at the N-terminus only a single, homologous N1 domain. The N1 domains are bordered on each side by short cysteine-rich segments (Chu *et al.*, 1988, 1990) indicating that they may associate with each other by disulfide bridges and thus form a folding and assembly unit independent of N9–N2.

The nature of the connections between the individual N domains of recombinant N9–N2 was studied by cleavage with neutral proteases. The data showed a high susceptibility to cleavage by α-chymotrypsin and elastase just three to ten amino acid residues in front of the start sites of five of the seven folding units as predicted from exon borders (Figure 9c). Each of the eight domains within the N9–N2 structure is encoded by a single exon in chicken and human (Doliana *et al.*, 1990; Stokes *et al.*, 1991). Variation of this structure by alternative splicing is known which, in human collagen VI, involves domains N9 and N7 and a large part of N3 (Stokes *et al.*, 1991). The latter event would very likely interfere with proper folding. Various kinetic studies (unpublished) indicate that several neutral proteases cleave initially between N domains but with different speeds. This suggests that each N domain possesses at its C-terminal end a probably short and flexible peptide link. Once released, the various N domains differ apparently in their protease stability as shown here for extensive (24 h) chymotrypsin and elastase digests (see Table I). Pepsin, however, degrades fragment N9–N2 rapidly to peptides smaller than 10 kDa (unpublished).

Table I. Size and N-terminal sequences of chymotrypsin (C) and elastase (E) fragments obtained from recombinant N9–N2 and correlation to single domains and binding activity for heparin

Fragment	Size (kDa)	Position[a] sequence	Domain	Binding to heparin
C1	26	781, KSV	N5	–
C2	21	385, SGTPEV	N7	–
C3	45	597, XQPESK	N6–N5	+
C4[b]	28	1,	N9	+
C5	24	597, AQPESK	N6	+
E1	28	987, LQPLPS	N4	–
E2	50	1192, XXPPAV	N3–N2	+
E3	29	1192, XYPPXVES	N3	+

[a]Starting position in α3(VI) chain (Chu *et al.*, 1990), X = unidentified residue.
[b]Blocked N-terminus indicating N-terminal position.

It is of interest that a full-length α3 chain of predicted 340 kDa (Chu *et al.*, 1990) has so far not been found in tissue-extracted collagen VI. The α3 chain appeared rather as two to five bands with an electrophoretic migration corresponding to molecular masses of 180–240 kDa (Trüeb and Winterhalter 1986; Wu *et al.*, 1987; Ayad *et al.*, 1989; Colombatti *et al.*, 1989; Kielty *et al.*, 1990; Ronzière *et al.*, 1990). A more uniform band of ~260 kDa that did not get

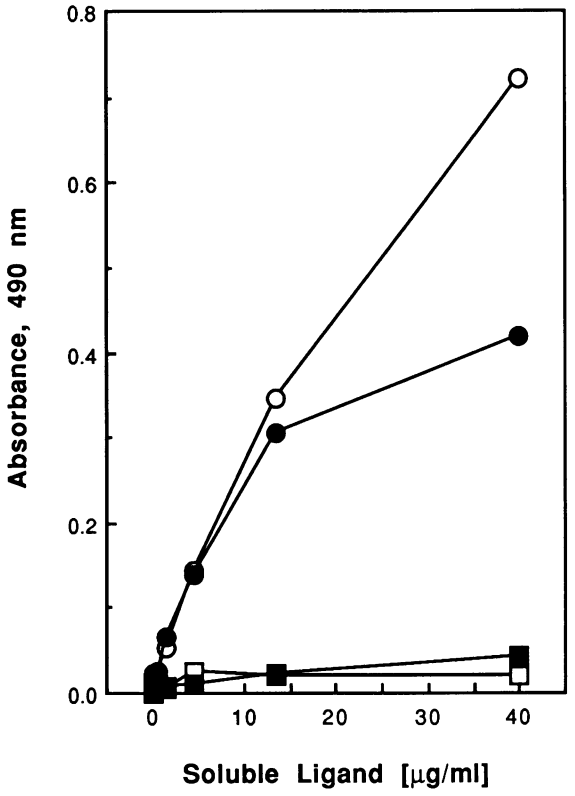

Fig. 8. Binding assay between fragment N9−N2 and pepsin-solubilized collagen VI. Open symbols refer to pepsin-solubilized collagens VI as soluble and N9−N2 as immobilized ligands. Closed symbols refer to N9−N2 as soluble and pepsin-solubilized collagen VI as immobilized ligands. Pepsin-solubilized collagen VI was used prior to (○,●) and after (□,■) reduction of disulfide bonds under non-denaturing conditions.

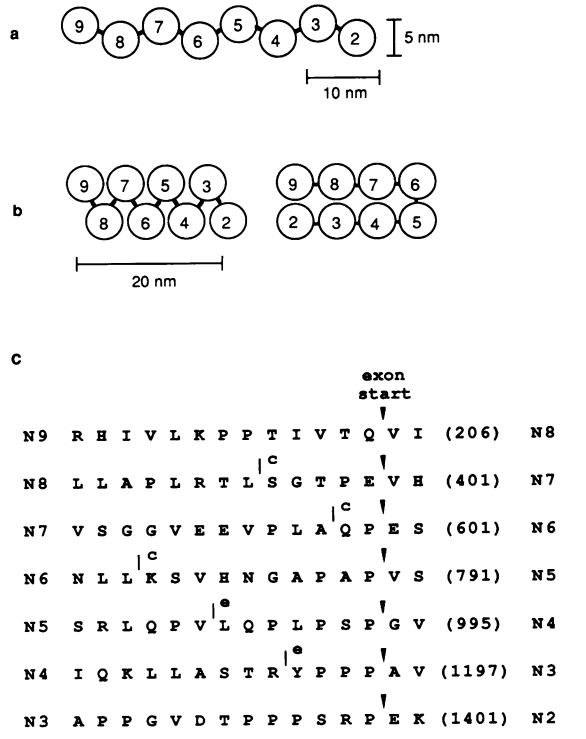

Fig. 9. Models of domain arrangement in fragment N9−N2 (**a** and **b**) and correlation of its exon borders with proteolytic cleavage sites (**c**). The models show a fully extended form (a) as predicted from ultracentrifugation and two ways of compact arrangement (b) as indicated from negative staining. Exon borders (Stokes *et al.*, 1991) are indicated by arrowheads showing preferred cleavage sites by chymotrypsin (c) and elastase (e) by a short vertical line. Numbers in brackets refer to the sequence position (Chu *et al.*, 1990) of the last residue shown in each line.

further processed was, however, found in cell culture media (Colombatti *et al.*, 1987; Kielty *et al.*, 1990). This indicates modification of the α3 structure not only by alternative splicing but also by limited proteolysis which in the latter case should also include domains N9−N2. The recombinant production and structural and immunological approaches described here should therefore be useful to study the structure and function of such truncated α3 chains. It may also reveal whether such processed forms are those actually deposited in tissues.

One of these possible functions of the α3 chain is binding to heparin which has not been described before. The affinity of fragment N9−N2 for heparin is apparently not very strong since a moderate increase in NaCl concentrations above physiological levels dissociates the complex. Yet the calculated isoelectric point pI = 5.4 of the N9−N2 structure indicates the presence of specific heparin binding sequences rather than binding to the affinity column due to a strongly basic nature of the protein. Potential binding structures were mapped to domains N9, N6 and N3 and possibly N2 while domains N4, N5 and N7 were identified as non-binding structures. Two binding domains, N6 with pI = 9.0 and N3 with pI = 10.8, are rather basic. The other potential candidates N9 and N2 have a pI = 4.7−4.8. It remains therefore to be elucidated which of these domains actually contribute to the heparin binding of fragment N9−N2. A heparin sulfate proteoglycan from a basement membrane-producing tumor (Paulsson *et al.*, 1987b) was not an active ligand for recombinant N9−N2. The heparan sulfate chains of this proteoglycan have a low sulfate content and also a low affinity for anti-thrombin III (Pejler *et al.*, 1987). Heparan sulfates obtained from other tissues are richer in sulfate and of higher affinity (Pejler and David, 1987) and may therefore be better candidates for collagen VI binding *in vitro* and *in situ*. Heparin binding has also been shown for the A1 domain of von Willebrand factor (Mohri *et al.*, 1989; Sixma *et al.*, 1991) that is homologous to the N domains. No binding data exist for other proteins possessing similar domains. Our data show in addition that such protein motifs do not necessarily possess obligatory heparin binding sequences.

Binding of fragment N9−N2 to hyaluronan was apparently stronger than that to heparin since it required exposure to denaturing agents for full dissociation. This indicates that the strength of binding does not so much depend on ionic but rather on conformation-dependent carbohydrate−protein interactions. A stable complex between collagen VI microfibrils and hyaluronan was recently obtained and characterized from fetal bovine skin (Kielty *et al.*, 1992). Degradation of hyaluronan caused a dissociation of the microfibrils into collagen VI tetramers, a process that could be reversed by adding hyaluronan. This indicates that hyaluronan directs the assembly of microfibrils which in other tissues then become stabilized by disulfide bridges (Timpl and Engel, 1987). Cartilagenous tissues are also known to possess substantial amounts of collagen VI (Wu

et al., 1987; Ayad *et al.*, 1989) and hyaluronan. It indicates that there they may form similar complexes (McDevitt *et al.*, 1991) suggesting a general importance for such supramolecular structures. It therefore appears likely that the N9−N2 structure is important for such associations.

We have also examined the binding of fragment N9−N2 to several extracellular matrix proteins. The most significant association was observed with pepsin-solubilized collagen VI. This material consists mainly of the triple helix and some remnants of N1 and C1 domains of all three collagen VI α chains (Odermatt *et al.*, 1983; Chu *et al.*, 1987, 1989). Reduction under non-denaturing conditions destroyed binding. Such treatment causes dissociation of collagen VI oligomers by opening cysteine-rich helix sealing segments but leaves the triple helical conformation intact (Furthmayr *et al.*, 1983; Odermatt *et al.*, 1983). It indicates that either oligomeric structures, the sealing segments, the ends of the triple helix or a combination of all of them are important for binding. According to previous models for collagen VI tetramer end-to-end aggregation (Furthmayr *et al.*, 1983; Chu *et al.*, 1989) all these sites should come into close contact with the N9−N2 structure which is consistent with the binding data. No binding was observed for the fibril-forming collagens I and III and their truncated precursor pN-collagens. This is in conflict with distinct binding data reported for collagen I and a fusion protein containing about five N domains of chicken collagen VI α3 chain (Bonaldo *et al.*, 1990). In this study the conformation of these N domains has not been analyzed and the collagen coat was prepared at 37°C which could have caused denaturation. These differences could therefore account for the controversial data. Nevertheless, the close ultrastructural association of collagens I and VI suggested by electron microscopy (von der Mark *et al.*, 1984; Bruns *et al.*, 1986; Kene *et al.*, 1988; Okada *et al.*, 1990) remains a challenge to be approached by further molecular studies.

Other binding data with various basement membrane proteins indicated that nidogen when used in recombinant form shows some weak affinity for N9−N2. Much stronger interactions were previously demonstrated between nidogen, laminin and collagen IV (Fox *et al.*, 1991). Because of the close association of collagen VI microfibrils with certain basement membranes (Keene *et al.*, 1988) we cannot exclude that the low affinity of N9−N2 for nidogen may be of relevance. Further weak affinities may exist for vitronectin and von Willebrand factor which merit further investigations. Fragment N9−N2 was, however, not a substrate for cell adhesion. Domain N4 possesses a LDV sequence (Chu *et al.*, 1990). This sequence, when present in fibronectin, was shown to bind to the integrin receptor α4β1 (Mould *et al.*, 1990; Komoriya *et al.*, 1991). Since several of our cell lines tested also possess α4β1 (HBL-100, A375; see Sonnenberg *et al.*, 1990) it indicates that the LDV sequence present in N4 is not available for integrin recognition.

The data reported here for a recombinant fragment of collagen VI have opened the way for the precise analysis of a particular 200 residue protein motif not too well characterized so far in its structural properties. Since it can be prepared in significant quantities it may also allow X-ray crystallography. Several significant binding properties could be established that may be specific for collagen VI. Their biological relevance remains to be explored.

Materials and methods

Construction of vectors, cell transfection and Northern hybridization

Plasmid F019 with a 5.5 kb cDNA insert encoding a large 5′ segment of the α3(VI) chain (Chu *et al.*, 1990) was restricted with *Sac*I and *Nsi*I. The *Nsi*I site was previously not identified and led to a correction of the sequence from CG to GC at nucleotide position 5089, 5090. The *Nsi*I restriction site was made blunt-ended with mung bean nuclease following a standard protocol. This fragment was then ligated to the *Sal*I−*Sma*I restricted Bluescript vector. The translation stop codon in this construct is provided by the polylinker region. After release of the insert with *Xho*I and *Not*I the resulting fragment was ligated to the *Xho*I−*Not*I restricted eukaryotic expression vector pCis (pCMVXXNH). This expression vector (Eaton *et al.*, 1986; Gorman *et al.*, 1990) contains the transcriptional control region of human CMV. The new plasmid was designated pNVIα3. It encodes the signal peptide and sequence positions 1−1586 of human α3(VI) chain (Chu *et al.*, 1990) and a foreign RGIH sequence at the C-terminus. The corresponding secreted protein product is referred to as N9−N2.

Human embryonic kidney cells 293 (American Type Culture Collection) were cotransfected with plasmid pNVIα3 and pSV2pc, the latter providing puromycin resistance. Stable cell clones were then selected in the presence of puromycin and characterized by Northern hybridization and SDS gel electrophoresis of serum-free culture medium. The probe used for Northern hybridization was the insert released with *Xho*I and *Not*I from the Bluescript vector. Details of these procedures have been previously reported (Fox *et al.*, 1991; Nischt *et al.*, 1991).

Chromatography and cleavage of proteins

Confluent cultures of stable cell clones were incubated for 48 h in serum-free culture medium (Nischt *et al.*, 1991). This medium (0.5−1 l) was then dialyzed at 4°C against 0.05 M Tris−HCl pH 8.6 and passed over a DEAE cellulose column (2.5 × 25 cm) equilibrated in the same buffer. Elution was performed with a linear gradient from 0 to 0.5 M NaCl (500/500 ml). The pool containing the recombinant product was concentrated by ultrafiltration and passed over a Sepharose C1-6B column in neutral buffer (Paulsson *et al.*, 1987a). Alternatively, the pool was dialyzed against 0.1 M NaCl 0.05 M Tris−HCl pH 7.4, 2 mM EDTA and passed over a heparin−Sepharose column or a HiTrap FPLC column (both from Pharmacia). Elution was then carried out with a linear gradient from 0.1 to 0.4 or 0.6 M NaCl of appropriate volume size. Hyaluronan (100 mg, from Sigma) was coupled to EAH−Sepharose (10 ml, from Pharmacia) according to a published procedure (Tengblad, 1979) with a coupling yield of 1−2 mg/ml gel as determined by the carbazol reaction (Butter and Muir, 1962) after digestion with hyaluronidase. The affinity column was then equilibrated in 0.15 M NaCl, 0.05 M Tris−HCl pH 7.4 for binding the proteins. Stepwise elution was done by increasing the NaCl concentration to 0.4, 1 and 5 M followed by 6 M urea, 0.05 M Tris−HCl pH 7.4.

Cleavage with CNBr was performed in 70% formic acid for 4 h at 30°C with a 100-fold excess of CNBr (w/w). Reagents were then removed by lyophilization. Digestions with neutral protease (trypsin, α-chymotrypsin, pancreatic elastase, SV8 protease) was done in 0.2 M ammonium bicarbonate at 37°C for varying periods of time. Pepsin digestion was carried out in 0.1 M acetic acid at 25°C. The enzyme:substrate ratio was in all cases 1:50−100.

Physical methods and electron microscopy

A Spinco model E analytical ultracentrifuge (Beckman Instruments) equipped with a photoelectric scanner was used for sedimentation velocity (at 56 000 r.p.m.) and equilibrium (at 9000 and 1200 r.p.m) runs following previously used methods of calculating molecular mass and sedimentation coefficient (Maurer *et al.*, 1992). The frictional ratio f/f_0 was calculated from both values assuming a degree of hydration of 0.5 g water/g protein. The axial ratio *a*/*b* of prolate ellipsoids as hydrodynamic equivalents was estimated by Perrin's table (Tanford, 1961).

Circular dichroism spectra were measured in a Cary 61 spectropolarimeter (Varian) which was calibrated with α-10-camphor sulfonic acid. They were recorded in a thermostatted quartz cell of optical path length 0.1 mm. The molar ellipticity $[\theta]$ was calculated assuming a mean residue molecular mass of 110 kDa. Secondary structure contributions (Hennessey and Johnson, 1981) were analyzed by a computer-aided curve fitting procedure Provec provided by H.P.Bächinger (Portland, OR).

Electron microscopy of proteins visualized by rotary shadowing or negative staining followed established protocols (Engel and Furthmayr, 1987). Samples (50 μg/ml) dissolved in 0.2 M ammonium bicarbonate and mixed with an equal volume of glycerol were used for rotary shadowing. Other

samples (5 μg/ml) dissolved in 0.05 M Tris−HCl pH 7.4 were adsorbed to carbon film that was rendered hydrophilic by glow discharge and used for negative staining.

Analytical methods

Amino acid and hexosamine compositions were determined after hydrolysis (16 h, 110°C) with 6 or 3 M HCl, respectively, on a LC5001 analyzer (Biotronic). Sequence analyses by Edman degradation were performed in gas phase sequencers 470A and 473A (Applied Biosystems) following the manufacturer's instruction. Samples were purified by SDS gel electrophoresis following established methods and then blotted onto immobilon membranes (Millipore) for sequencing. Isoelectric points (pI) were calculated from sequence data (Chu *et al.*, 1990) by the program isoelectric of the sequence analysis package (version 7.1) from Genetics Computer Group (Devereux *et al.*, 1984).

Immunological methods

An antiserum was generated in a rabbit by three injections of 0.3 mg antigen together with complete Freund's adjuvant. Radioimmuno binding and inhibition (sequential saturation type) assays have been described (Timpl, 1982). ELISA , indirect immunofluorescence and immunoblotting followed standard protocols.

Protein binding and cell adhesion assays

Protein−protein interactions were analyzed with recombinant N9−N2 using it as an immobilized and/or as soluble ligand following a previously described method (Aumailley *et al.*, 1989a; Fox *et al.*, 1991). Binding of soluble ligands was detected by specific antibodies which were adjusted in ELISA to a maximal color yield (OD range 1.0−1.5). The other ligands used included collagen IV from the mouse EHS tumor (Yurchenco and Furthmayr, 1984), mouse laminin−nidogen complex (Paulsson *et al.*, 1987a), mouse recombinant nidogen (Fox *et al.*, 1991), human recombinant BM-40 (Nischt *et al.*, 1991), EHS tumor heparan sulfate proteoglycan (Paulsson *et al.*, 1987b) and human serum vitronectin (Yatohogo *et al.*, 1988). Pepsin-solubilized collagen VI and its reduced but still triple helical derivative (Odermatt *et al.*, 1983) were from human placenta. The preparation of pepsin-solubilized collagens from human placenta (Miller and Rhodes, 1982) and of truncated procollagens pN-collagens I and III (Timpl *et al.*, 1975) has been described. Fibronectin and von Willebrand factor from human plasma were a kind gift of N.Heimburger, Behringwerke, Marburg, Germany.

Cell adhesion assays were performed in microtiter wells by established procedures (Aumailley *et al.*, 1989b). The coating concentrations examined were in the range 0.6−40 μg/ml. The transformed cell lines used were identified previously (Aumailley *et al.*, 1989b; Sonnenberg *et al.*, 1990). Coating efficiency to the wells was determined by adding traces of radio-labeled substrate (Aumailley *et al.*, 1989b).

Acknowledgements

We appreciated the excellent technical assistance of Mischa Reiter, Vera van Delden, Heidi Alberty, Ariel Lustig, Charlotte Fauser and Hanna Wiedemann. We thank Dr C.Gorman for providing the expression vector. The study was supported by grants from the Deutsche Forschungsgemeinschaft (SFB266), the Swiss National Science Foundation (to J.E.), the NIH (grant 38912) and by a fellowship of the Mayo Foundation, Rochester, MN (to U.S.).

References

Aumailley,M., Wiedemann,H., Mann,K.and Timpl,R. (1989a) *Eur. J. Biochem.*, **184**, 241−248.

Aumailley,M., Mann,K., von der Mark,H. and Timpl,R. (1989b) *Exp. Cell Res.*, **181**, 463−474.

Ayad,S., Marriott,A., Morgan,K. and Grant,M.E. (1989) *Biochem. J.*, **262**, 753−761.

Bitter,T. and Muir,H.M. (1962) *Anal. Biochem.*, **4**, 330−334.

Bonaldo,P. and Colombatti,A. (1989) *J. Biol. Chem.*, **264**, 20235−20239.

Bonaldo,P., Russo,V., Bucciotti,F., Bressan,G.M. and Colombatti,A. (1989) *J. Biol. Chem.*, **264**, 5575−5580.

Bonaldo,P., Russo,V., Bucciotti,F., Doliana,R. and Colombatti,A. (1990) *Biochemistry*, **29**, 1245−1254.

Bruns,R.R., Press,W., Engvall,E., Timpl,R. and Gross,J. (1986) *J. Cell Biol.*, **103**, 393−404.

Chu,M.-L., Mann,K., Deutzmann,R., Pribula-Conway,D., Hsu-Chen, C.-C., Bernard,M.P. and Timpl,R. (1987) *Eur. J. Biochem.*, **168**, 309−317.

Chu,M.-L., Conway,D., Pan,T.-C., Baldwin,C., Mann,K., Deutzmann,R. and Timpl,R. (1988) *J. Biol. Chem.*, **263**, 18601−18606.

Chu,M.-L., Pan,T.C., Conway,D., Kuo,H.-J., Glanville,R.W., Timpl,R., Mann,K. and Deutzmann,R. (1989) *EMBO J.*, **8**, 1939−1946.

Chu,M.-L., Zhang,R.-Z., Pan,T., Stokes,D., Conway,D., Kuo,H.-J., Glanville,R., Mayer,U., Mann,K., Deutzmann,R. and Timpl,R. (1990) *EMBO J.*, **9**, 385−393.

Colombatti,A. and Bonaldo,P. (1991) *Blood*, **77**, 2305−2315.

Colombatti,A., Bonaldo,P., Ainger,K., Bressan,G.M. and Volpin,D. (1987) *J. Biol. Chem.*, **262**, 14454−14460.

Colombatti,A., Ainger,K. and Colizzi,F. (1989) *Matrix*, **9**, 177−185.

Deliana,R., Bonaldo,P. and Colombatti,A. (1990) *J. Cell Biol.*, **111**, 2197−2205.

Devereux,J., Haeberli,P. and Smithies,O. (1984) *Nucleic Acids Res.*, **12**, 387−395.

Eaton,D.L., Wood,W.I., Eaton,D., Hass,P.E., Hollingshead,P., Wion,K., Mather,J., Lawn,R.M., Vehar,G.A. and Gorman,C. (1986) *Biochemistry*, **25**, 8343−8347.

Engel,J. and Furthmayr,H. (1987) *Methods Enzymol.*, **145**, 3−78.

Engvall,E., Hessle,H. and Klier,G. (1986) *J. Cell Biol.*, **102**, 703−710.

Fox,J.W., Mayer,U., Nischt,R., Aumailley,M., Reinhardt,D., Wiedemann, H., Mann,K., Timpl,R., Krieg,T., Engel,J. and Chu,M.-L. (1991) *EMBO J.*, **10**, 3137−3146.

Furthmayr,H., Wiedemann,H., Timpl,R., Odermatt,E. and Engel,J. (1983) *Biochem. J.*, **211**, 303−311.

Gorman,C.M., Gies,D.R. and McGray,G. (1990) *Protein Engng. Tech.*, **2**, 3−10.

Hennessey,J.P. and Johnson,W.C. (1981) *Biochemistry*, **20**, 1085−1094.

Keene,D.R., Engvall,E. and Glanville,R.W. (1988) *J. Cell Biol.*, **107**, 1995−2006.

Kielty,C.M., Boot-Handford,R.P., Ayad,S. and Schuttleworth,C.A. (1990) *Biochem. J.*, **272**, 787−795.

Koller,E., Winterhalter,K.H. and Trüeb,B. (1989) *EMBO J.*, **8**, 1073−1077.

Komoriya,A., Green,L.J., Mervic,M., Yamada,S.S., Yamada,K.M. and Humphries,M.J. (1991) *J. Biol. Chem.*, **266**, 15075−15079.

Maurer,P., Mayer,U., Bruch,M., Jenö,P., Mann,K., Landwehr,R., Engel,J. and Timpl,R. (1992) *Eur. J. Biochem.*, **205**, 233−240.

McDevitt,C.A., Marcelino,J. and Tucker,L. (1991) *FEBS Lett.*, **294**, 167−170.

Miller,E.J. and Rhodes,R.K. (1982) *Methods Enzymol.*, **82**, 33−64.

Mohri,H., Yoshioka,A., Zimmerman,T.S. and Ruggeri,Z.M. (1989) *J. Biol. Chem.*, **164**, 17361−17367.

Mould,A.P., Wheldon,L.A., Komoriya,A., Wayner,E.A., Yamada,K.M. and Humphries,M.J. (1990) *J. Biol. Chem.*, **265**, 4020−4024.

Nischt,R., Pottgiesser,J., Krieg,T., Mayer,U., Aumailley,M. and Timpl,R. (1991) *Eur. J. Biochem.*, **200**, 529−536.

Odermatt,E., Risteli,J., van Delden,V. and Timpl,R. (1983) *Biochem. J.*, **211**, 295−302.

Okada,Y., Naka,K., Minamoto,T., Ueda,Y., Oda,Y., Nakanishi,I. and Timpl,R. (1990) *Lab. Invest.*, **63**, 647−656.

Paulsson,M., Aumailley,M., Deutzmann,R., Timpl,R., Beck,K. and Engel,J. (1987a) *Eur. J. Biochem.*, **166**, 11−19.

Paulsson,M., Yurchenco,P.D., Ruben,G.C., Engel,J. and Timpl,R. (1987b) *J. Mol. Biol.*, **197**, 297−313.

Pejler,G. and David,G. (1987) *Biochem. J.*, **248**, 69−77.

Pejler,G., Bäckström,G., Lindahl,U., Paulsson,M., Dziadek,M., Fujiwara,S. and Timpl,R. (1987) *J. Biol. Chem.*, **262**, 5036−5043.

Ronzière,M.C., Ricard-Blum.S., Tiollier,J., Hartmann,D.J., Garrone,R. and Herbage,D. (1990) *Biochim. Biophys. Acta.*, **1038**, 222−230.

Sixma,J.J., Schiphorst,M.E., Verwej,C.L. and Pannekoek,H. (1991) *Eur. J. Biochem.*, **196**, 369−375.

Sonnenberg,A., Linders,C.J.T., Modderman,P.W., Damsky,C.H., Aumailley,M. and Timpl,R. (1990) *J. Cell Biol.*, **110**, 2145−2155.

Stokes,D.G., Saitta,B., Timpl,R. and Chu,M.-L. (1991) *J. Biol. Chem.*, **266**, 8626−8633.

Tanford,C. (1961) *Physical Chemistry of Macromolecules*. John Wiley, New York.

Tengblad,A. (1979) *Biochim. Biophys. Acta*, **578**, 281−289.

Timpl,R. (1982) *Methods Enzymol.*, **82**, 472−498.

Timpl,R. and Engel,J. (1987) In Mayne,R. and Burgeson,R.E. (eds), *Structure and Function of Collagen Types*. Academic Press, Orlando, FL, pp. 105−143.

Timpl,R., Glanville,R.W., Nowack,H., Wiedemann,H., Fietzek,P.P. and Kühn,K. (1975) *Hoppe-Seyler's Z. Physiol. Chem.*, **356**, 1783–1792.
Titani,K. and Walsh,K.A. (1988) *Trends Biochem. Sci.*, **13**, 94–97.
Titani,K., Kumar,S., Takio,K., Ericsson,L.H., Wade,R.D., Ashida,K., Walsh,K.A., Chopek,M.W., Sadler,J.E. and Fujikawa,K. (1986) *Biochemistry*, **25**, 3171–3184.
Trüeb,B.and Winterhalter,K. (1986) *EMBO J.*, **5**, 2815–2819.
Van der Rest,M. and Garrone,R. (1991) *FASEB J.*, **5**, 2814–2823.
Von der Mark,H., Aumailley,M., Wick,G., Fleischmajer,R. and Timpl,R. (1984) *Eur. J. Biochem.*, **142**, 493–502.
Wayner,E.A. and Carter,W.G. (1987) *J. Cell Biol.*, **105**, 1873–1884.
Wu,J.-J., Eyre,D.R. and Slayter,H.S. (1987) *Biochem. J.*, **248**, 373–381.
Yatohogo,T., Izumi,S., Kashiwagi,H. and Hayashi,M. (1988) *Cell Structure Function*, **13**, 281–292.
Yurchenco,P.D. and Furthmayr,H. (1984) *Biochemistry*, **23**, 1839–1850.

Received on July 7, 1992