Structural and Functional Features of the α3 Chain Indicate a Bridging Role for Chicken Collagen VI in Connective Tissues[†,‡]

Paolo Bonaldo, Vincenzo Russo, Francesco Bucciotti, Roberto Doliana, and Alfonso Colombatti*
Divisione di Oncologia Sperimentale 2, Centro di Riferimento Oncologico, 33081 Aviano, Italy
Received July 6, 1989; Revised Manuscript Received September 19, 1989

ABSTRACT: Type VI collagen is a component of 100 nm long periodic filaments with a widespread distribution around collagen fibers and on the surface of cells. It is an unusual collagen constituted by three distinct chains, one of which (α3) is much larger than the others and is encoded by a 9-kb mRNA. The amino acid sequence of the α3(VI) deduced from the present cDNA clones specifies for a multidomain protein of at least 2648 residues made of a short collagenous sequence (336 residues), flanked at the N-terminus by nine 200 residue long repeating motifs and at the C-terminus by two similar motifs that share extensive identities with the collagen-binding type A repeats of von Willebrand factor. Type VI collagen and α3(VI) fusion proteins bound to insolubilized type I collagen in a specific, time-dependent, and saturable manner. The α3(VI) chain has three Arg-Gly-Asp sequences in the collagenous domain, and cell attachment was stimulated by the triple helix of type VI collagen and by α3(VI) fusion proteins containing Arg-Gly-Asp sequences. This function was specifically inhibited by the Arg-Gly-Asp-Ser synthetic peptide. The type I collagen-binding and the cell-attachment properties of the α3(VI) chain provide direct information for the role of type VI collagen in connective tissues.

The importance of components of the extracellular matrix in several biological functions has been established from studies on morphogenesis, cell differentiation, cell migration, and cell adhesion. Protein chemistry and molecular cloning have provided structural information that suggests a molecular basis for the multiple interactions attributed to the various components. These functions are mediated through specialized and unique domains identified by in vitro studies on a number of extracellular matrix macromolecules. Presumably, the variety of biological activities assigned to numerous components or domains are related to their biological functions in vivo.

For several proteins, however, structure–function relationships are not yet known. Type VI collagen is an example of the latter group. This collagen is the major component of 100 nm long periodic filaments and fibrils that are found on the surface of cells and around or between collagen fibers both in vitro and in vivo (von der Mark et al., 1984; Bruns, 1984; Bruns et al., 1986; Keene et al., 1988). These fibrils occur in both embryo (Bruns et al., 1986) and adult tissues (von der Mark et al., 1984; Keene et al., 1988) and are localized in several tissues including cartilage.

Some peculiar structural and biosynthetic properties of type VI collagen contrast sharply with those of the fibrillar collagens, but also with the characteristics of the other collagens such as type IV, VII, and VIII (Miller & Gay, 1987). The monomers are composed of three genetically distinct chains of about M_r = 260000 (α3), 150000 (α1), and 140000 (α2) (Trüeb & Winterhalter, 1986; Colombatti, et al., 1987). After digestion of the native molecule with pepsin, the intact subunits give rise to three small fragments of apparent M_r = 70000 (α1), 55000 (α2), and 40000 (α3) that represent the collagenous structure (Furuto & Miller, 1980; Jander et al., 1981). Another unique feature is the polymerization process taking place intracellularly soon after synthesis (Engvall et al., 1986; Colombatti et al., 1987) and leading to the formation of disulfide-bonded dimers and tetramers, which constitute the "building blocks" (Wu et al., 1987; Colombatti et al., 1989) of the oligomers and periodic microfibrillar structures. In addition, collagen type VI does not undergo proteolytic processing with removal of large N- and C-propeptides to form a mature collagen molecule: no evidence of processing for any of the individual subunits to shorter polypeptides was found in vitro (Colombatti & Bonaldo, 1987) and, at least for the α1 and α2 chains, also in vivo (Jander et al., 1984; Knight et al., 1984; von der Mark et al., 1984; Gibson & Cleary, 1985; Trüeb & Winterhalter, 1986). Finally, the most distinctive feature of type VI collagen is the unusual length of the α3 chain, which is twice the apparent M_r of the other two chains (Trüeb & Winterhalter, 1986; Colombatti, et al., 1987, 1989). We report here on the isolation and analysis of cDNA clones spanning most of the α3(VI) chain. The deduced amino acid sequence specifies for a multidomain protein that at its carboxyl-terminal part is structurally similar to the smaller α1(VI) (Bonaldo et al., 1989) and α2(VI) (Koller et al., 1989) chains consisting of a short collagenous domain with three Arg-Gly-Asp sequences flanked by a triplicated domain similar to the type A domains in von Willebrand factor. The large amino-terminal portion, unique for the α3(VI) chain, is formed by eight repeating units of about 190–210 amino acids that are also closely related to the type A repeats of von Willebrand factor. In vWF[1] two independent type I collagen-binding sites have been ascribed to the type A region (Pareti et al., 1987). We have tested and were able to demonstrate that type VI collagen has both Arg-Gly-Asp-dependent cell-attachment and type I collagen-binding properties. The information derived from our sequence data and functional assays together with electron microscopic studies reported by others (Furthmayr

[†] This work was supported by Grant 87.01485.44 "P.F.-Oncologia" from Consiglio Nazionale delle Ricerche and by a grant from Associazione Italiana per la Ricerca sul Cancro.

[‡] The nucleotide sequence reported in this paper has been submitted to the GenBank/EMBL Data Bank under Accession Number M24282.

* To whom correspondence should be addressed.

[1] Abbreviations: BSA, bovine serum albumin; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; vWF, von Willebrand factor.

0006-2960/90/0429-1245$02.50/0

et al., 1983; Bruns et al., 1986; Keene et al., 1988) provide a structural–functional basis for the specific bridging role of type VI collagen molecules in connective tissues.

Experimental Procedures

Isolation of cDNA Clones. Poly(A+) RNA was purified from 2-day-old chick aortas according to Chirgwin et al. (1979) and cDNA synthesized by use of a random hexanucleotide as a primer. The library was constructed in the expression vector pEX1 by use of adaptors (Haymerle et al., 1987). Screening of the library was performed by established procedures using mouse monoclonal antibodies for the α3 chain (Colombatti et al., 1988) or rabbit polyclonal antibodies affinity purified onto a nitrocellulose-bound pepsin-resistant fragment of α3(VI). To obtain other cDNA clones, the library was screened by hybridization with nick-translated DNA fragments obtained by digestion of clones with restriction enzymes. All the clones were rescreened twice, and positive clones that cross-hybridized to each other were sequenced.

Polyacrylamide Gel Electrophoresis and Protein Blotting. SDS–PAGE was performed with a 5% slab gel. The procedure for performing transfer of peptides from the polyacrylamide gels to nitrocellulose filters and incubation of the blots was essentially the same as described by Towbin et al. (1982). The filter was incubated at room temperature with a 1:100 dilution of different antisera, followed by peroxidase-labeled rabbit IgG to mouse immunoglobulins.

Northern Blotting. Poly(A+)-enriched RNA was prepared from chick gizzard on an oligo(dT)–cellulose affinity column. Electrophoresis of the RNA was performed on a 1% (w/v) agarose gel containing 2.2 M formaldehyde in MOPS buffer. RNA was then transferred onto nitrocellulose and hydridized with cDNA probes labeled to high specific activity (Feinberg & Vogelstein, 1983).

DNA Sequencing and Analysis. Sequencing was performed by the dideoxynucleotide chain termination method (Sanger et al., 1977) as modified by Biggin et al. (1983) and by the use of modified bacteriophage T7 DNA polymerase (Tabor & Richardson, 1987) after subcloning of cDNA inserts into the phage vectors M13 mp 18 and M13 mp 19 (Messing, 1983). DNA and deduced amino acid sequence analyses were done as in Bonaldo and Colombatti (1989).

Binding to Collagen Type I. Fibrillar type I collagen (1–2 μg/well) (Vitrogen, Collagen Corp., Palo Alto, CA) was adsorbed overnight at 37 °C to polystyrene microtiter plates after dilution in 0.1 M carbonate buffer solution, pH 8.6. The wells were saturated with 1% BSA for 3 h and were then incubated with ^{125}I-labeled type VI collagen or expressed protein of clone pB10 in 50 μL of Hank's buffered (pH 7.6) solution containing 0.01% SDS for a period of 1 h at 4 °C. In the competition experiments the labeled and cold ligands were added at the same time in the wells containing type I collagen. Saturability was analyzed in dose–response experiments, and calculations of binding parameters were carried out by the method of Scatchard. At the end of the incubation the wells were washed extensively with Hank's solution, and the amount of radioactivity bound was assessed in a γ-counter. Determinations were undertaken in duplicate, and results are expressed as the mean of the two. Values have been corrected for a blank value obtained by binding ^{125}I-labeled type VI collagen or pB10 protein to adsorbed BSA. This value was usually less than 1% of the total radioactivity added. ^{125}I-Labeled type VI collagen or pB10 protein was prepared by the iodogen procedure (Fraker & Speck, 1978). Type VI collagen used in these experiments was purified from chick gizzard by a differential extraction procedure as described in detail elsewhere

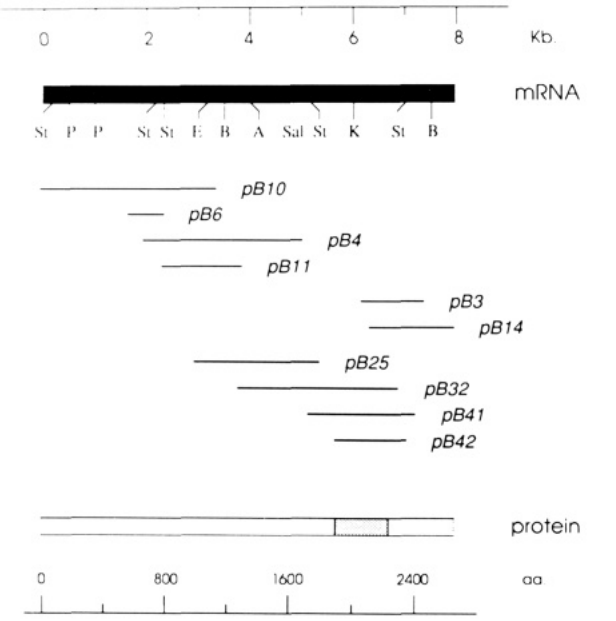

FIGURE 1: Overlapping cDNA clones used for sequence analysis. A schematic diagram and a partial restriction map of the α3(VI) chain mRNA is shown on the top. Restriction sites are as follows: St, *Sst*I; P, *Pst*I; E, *Eco*RI; B, *Bam*HI; K, *Kpn*I; Sal, *Sal*I. Solid lines in the central portion show representative clones. The shaded box in the diagram of the protein at the bottom indicates the triple helix (COL domain).

(Colombatti et al., 1989). The pepsin-resistant fragment of type VI collagen (triple-helical domain) was obtained from chick gizzard following the procedure described by Trüeb et al. (1987). vWF was a gift from Dr. L. De Marco (1983). The procedure to obtain the pB10 and TE7 proteins is briefly outlined: the clones pB10 coding for α3 (VI) and TE7 coding for chick tropoelastin (Bressan et al., 1987) were inserted in an intermediate form of the pUEX1 expression vector, which lacks the β-galactosidase gene and maintains only a deleted cro gene (Bressan & Stanley, 1987). Therefore, the specific expressed protein is fused only to 18 amino acids, 9 derived from the cro sequence and 9 from the adaptor (Haymerle et al., 1987). The pB10 and TE7 proteins, obtained from lysated *Escherichia coli* grown at 42 °C, were further purified by SDS–PAGE. At the end of the run the bands corresponding to the two proteins were visualized by incubation in ice-cold 1 M potassium acetate, excised, electroeluted, and used for binding experiments. Specific binding ranged between 8 and 20% depending on the labeling conditions and the purified preparation of type VI collagen.

Cell Adhesion Assay. Cell adhesion assays were performed with metabolically labeled aorta smooth muscle cells from chick embryos. Cell monolayers were collected by trypsinization, washed in methionine-free Dulbecco's modified Eagle's medium (DMEM) containing 1% dialyzed fetal calf serum (FCS), and metabolically labeled with [^{35}S]methionine (800 Ci/mmol; Amersham) at 100 μCi/mL. Aliquots (2.0×10^4) of the cell suspension in DMEM containing 1% BSA were plated onto protein-coated 96-well polystyrene microtiter plates and incubated at 37 °C for different lengths of time. At the end of the incubation the medium was removed, and the adherent cells were washed twice. Bound radioactivity was detached with 1 M NaOH and counted in a β-counter. Background values of cells attached to BSA-coated wells were subtracted from the values obtained in the presence of the other proteins. Under our experimental conditions maximal binding ranged between 15 and 30% of the added cells. Experiments carried out with cells maintained in the presence of 20 μg/mL

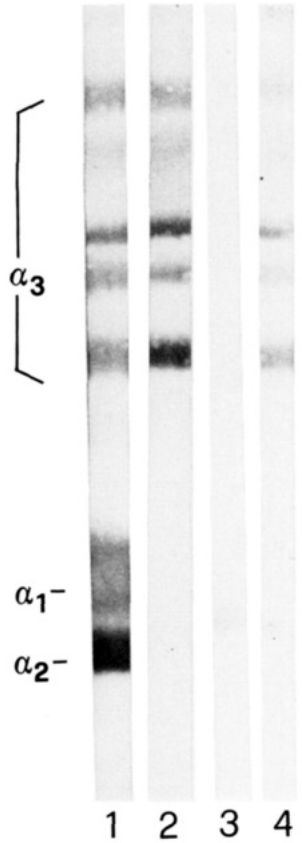

FIGURE 2: Immunoblotting of type VI collagen. Purified tissue-form type VI collagen was electrophoresed under reducing conditions on an SDS–5% polyacrylamide gel and then transferred onto nitrocellulose, and strips were incubated with polyclonal antibodies. Specificity of the antibodies: anti-pepsin-resistant type VI collagen (lane 1); anti-pepsin-resistant type VI collagen affinity purified onto pB32 fusion protein (lane 2); normal rabbit serum (lane 3); anti-pB10 and -pB4 fusion protein (lane 4). α1–α3 indicate the chains of the tissue form of type VI collagen.

cycloheximide for 30 min before and during the adhesion assay gave similar results.

Coating of the microtiter plates was done by air-drying protein solutions overnight at 37 °C. Adsorption of the proteins was followed by blocking with 1% BSA for few hours before performing the cell adhesion assay.

Results

Isolation of cDNA Clones and Northern Blot Analysis. A random-primed library made in the pEX1 expression vector from chick smooth muscle cell mRNA was screened by monoclonal antibodies to the tissue form of the α3(VI) chain and by polyclonal antibodies to the pepsin-resistant form of the α3(VI) chain. Two nonoverlapping open-reading frames were obtained. The library was then rescreened with ^{32}P-labeled probes (clone pB3 and the *Acc*I–*Sal*I restriction fragments of clone pB4). After several rescreenings 32 clones were isolated altogether that cross-hybridized with each other and covered a total of 8.0 kb. A partial restriction map and representative clones covering the sequence are shown in Figure 1.

Authenticity of the cDNA clones was established by the identification of the α3(VI) chain in a retroblot assay with antibodies affinity purified onto fusion proteins (Figure 2, lane 2). Furthermore, a polyclonal serum raised against fusion proteins of clones pB10 and pB4 (Figure 2, lane 4) recognized the α3(VI) chain in immunoblotting assay.

Northern blotting analysis of poly(A+) RNA isolated from chick gizzard and probed with several different cDNA clones showed a broad band corresponding to an mRNA of approximately 9 kb (Figure 3). No additional bands of smaller size were detected in the present experiments, but with one of the α3(VI) cDNA probes a smear of probable partial degradation was also detected. The insufficient resolution of the present gel did not allow the identification of the multiple bands described by others in human cells (Chu et al., 1987). In other experiments to be published elsewhere, the 9-kb band was shown to be composed of several alternatively spliced mRNAs (Doliana et al., unpublished results).

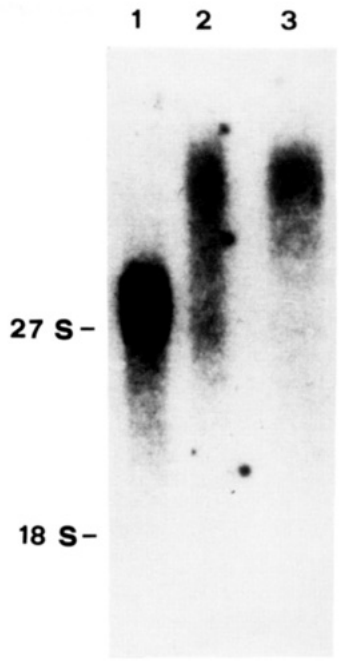

FIGURE 3: Autoradiogram of a Northern blot analysis. Each lane contained 3 μg of poly(A+) RNA extracted from chick gizzard. Inserts from clone pB4 (lane 2), clone pB10 (lane 3), and clone pA13 specific for the α1(VI) chain (lane 1) (Bonaldo et al., 1989) were ^{32}P labeled and hybridized to the filter. On the left the migration of ribosomal RNA is indicated.

Nucleotide and Predicted Amino Acid Sequences. Twenty-eight clones were characterized by restiction enzyme analysis and DNA sequencing. The coding sequence and the deduced amino acid sequence of the overlapping cDNA clones studied are shown in Figure 4. The sequence contains a single large open-reading frame of 7944 bases coding for a polypeptide of 2648 amino acids (M_r = 287 035). No translation initiation codon was found among the clones isolated to date, suggesting that the mature protein is even larger than predicted from previous studies by ourselves (Colombatti et al., 1987; Colombatti & Bonaldo 1987) and others [see Miller and Gay (1987) for a review]. The carboxyl end of the sequence including the stop codon is reported in a separate paper (Bonaldo & Colombatti, 1989).

The deduced sequence contains a very extended noncollagenous N-terminal portion of 1819 amino acids residues, followed by a collagenous (COL) C-terminal portion of 336 residues and by a noncollagenous carboxyl sequence of 493 residues.

The sequence contains six potential N-glycosylation sites (Asn-Xaa-Ser/Thr) one of which is within the COL domain. If one assumes an average molecular mass of 2500 daltons per N-linked carbohydrate side chain, the addition of six such moieties to the mature protein core would result in an estimated M_r = 302 035. At the junctions between the COL and the noncollagenous domains there are two clusters of cysteines that are likely candidates for interchain disulfide bonds that prevent digestion of the nonreduced triple helix by pepsin and

```
   1 CAGGCTGGAGCCAAAGGACTGGTTAAAGACATCACAGCTCAAGAGTCTGCTGACCTTATTTTCCTTATTGACGGATCAGACAACATCGGAAGTGTCAATTTTCAAGCCATACGCGATTTC
     Q  A  G  A  K  G  L  V  K  D  I  T  A  Q  E  S  A  D  L  I  F  L  I  D  G  S  D  N  I  G  S  V  N  F  Q  A  I  R  D  F    40
 121 CTTGTAAATTTGATTGAAAGCCTCAGAGTTGGAGCTCAGCAGATACACATTGGGGTGGTGCAGTATAGTGATCAACCCAGAACTGAGTTCGCTTTGAACAGCTACTCCACAAAAGCGGAT
     L  V  N  L  I  E  S  L  R  V  G  A  Q  Q  I  H  I  G  V  V  Q  Y  S  D  Q  P  R  T  E  F  A  L  N  S  Y  S  T  K  A  D    80
 241 GTTCTGGATGCCGTGAAGGCACTCAGTTTCCGCGGTGGCAAGGAAGCTAACACTGGTGCAGCACTGGAGTATGTGGTGGAGAACCTCTTCACCCAGGCAGGAGGCAGCAGGATAGAGGAA
     V  L  D  A  V  K  A  L  S  F  R  G  G  K  E  A  N  T  G  A  A  L  E  Y  V  V  E  N  L  F  T  Q  A  G  G  S  R  I  E  E   120
 361 GCGGTGCCTCAGATCTTAGTGCTGATAAGTGGTGGAGAGTCTAGCGATGATATCCGAGAGGGCTTGTTAGCGGTGAAGCAAGCTAGTATATTTTCATTTAGCATTGGGGTCCTGAATGCT
     A  V  P  Q  I  L  V  L  I  S  G  G  E  S  S  D  D  I  R  E  G  L  L  A  V  K  Q  A  S  I  F  S  F  S  I  G  V  L  N  A   160
 481 GACAGTGCAGAGCTGCAGCAGATTGCTACTGATGGAAGCTTCGCATTTACCGCCCTGGATATCCGTAACCTTGCTGCTCTTCGAGAATTATTACTGCCCAACATTGTTGGAGTGGCCCAA
     D  S  A  E  L  Q  Q  I  A  T  D  G  S  F  A  F  T  A  L  D  I  R  N  L  A  A  L  R  E  L  L  L  P  N  I  V  G  V  A  Q   200
 601 AGACTCATTTTGTTAGAGGCCCCAACCATTGTGACTGAAGTTATTGAAGTGAACAAGAAGGATATAGTCTTCCTGATAGATGGCTCAACAGCTCTGGGGACTGGCCCCTTTAACTCAATT
     R  L  I  L  L  E  A  P  T  I  V  T  E  V  I  E  V  N  K  K  D  I  V  F  L  I  D  G  S  T  A  L  G  T  G  P  F  N  S  I   240
 721 CGTGATTTCGTTGCCAAAATCGTCCAAAGGCTGGAGGTTGGACCTGACCTGATCCAAGTTGCGGTGGCTCAGTATGCAGACACTGTGAGGCCAGAGTTTTATTTTAACACCCATCAGAAC
     R  D  F  V  A  K  I  V  Q  R  L  E  V  G  P  D  L  I  Q  V  A  V  A  Q  Y  A  D  T  V  R  P  E  F  Y  F  N  T  H  Q  N   280
 841 AGAAAGGATGTGATGGCTAATGTGAAGAAAATGAAGCTCATGGGTGGTACGGCCCTCAACACTGGCTCGGCACTGGACTTTGTGAGGAACAACTTCTTCACCAGTGCTGCCGGGTGTAGG
     R  K  D  V  M  A  N  V  K  K  M  K  L  M  G  G  T  A  L  N  T  G  S  A  L  D  F  V  R  N  N  F  F  T  S  A  A  G  Ⓒ  R   320
 961 ATGGAGGAAGGGGTGCTCCCCATGCTGGTACTCATCACTGGTGGTAAGTCCATGGATGCCGTTGAGCAAGCTGCAGCAGAGATGAAAAGGAACAGGATAGTGATCCTTGCTGTTGGGTCG
     M  E  E  G  V  L  P  M  L  V  L  I  T  G  G  K  S  M  D  A  V  E  Q  A  A  A  E  M  K  R  N  R  I  V  I  L  A  V  G  S   360
1081 AGAAATGCTGACGTGGCAGAACTCCAAGAAATTGCTCATGAGAGAGATTTTGTGTTCCAACCAAATGATTTCCGCCTTCAGTTCATGCAAGCCATTCTTCCTGAGGTGCTGTCACCCATC
     R  N  A  D  V  A  E  L  Q  E  I  A  H  E  R  D  F  V  F  Q  P  N  D  F  R  L  Q  F  M  Q  A  I  L  P  E  V  L  S  P  I   400
1201 CGGACACTCTCCGGAGGAATGGTCATCCACGAAACACCATCAGTTCAAGTAACCAAAAGGGATATCATCTTTCTTTTGGATGGATCACTCAACGTCGGAAATGCCAACTTCCCCTTTGTG
     R  T  L  S  G  G  M  V  I  H  E  T  P  S  V  Q  V  T  K  R  D  I  I  F  L  L  D  G  S  L  N  V  G  N  A  N  F  P  F  V   440
1321 CGGGACTTTGTTGTGACTTTAGTTAACTACCTTGATGTCGGAACCGACAAAATCCGAGTTGGCTTAGTGCAATTTAGTGACACTCCTAAAACCGAGTTCTCCCTATACTCATACCAAACC
     R  D  F  V  V  T  L  V  N  Y  L  D  V  G  T  D  K  I  R  V  G  L  V  Q  F  S  D  T  P  K  T  E  F  S  L  Y  S  Y  Q  T   480
1441 AAATCAGACATAATTCAACGTTTGGGGCAGCTGAGGCCCAAGGGTGGGTCAGTGCTGAACACTGGCTCTGCACTGAACTTTGTGCTTTCGAATCACTTCACTGAAGCTGGTGGGAGCAGA
     K  S  D  I  I  Q  R  L  G  Q  L  R  P  K  G  G  S  V  L  N  T  G  S  A  L  N  F  V  L  S  N  H  F  T  E  A  G  G  S  R   520
1561 ATAAATGAACAAGTGCCGCAGGTCCTAGTCCTGGTGACGGCAGGGAGGTCGGCCGTACCCTTCCTGCAAGTTTCCAATGACTTAGCTCGGGCCGGAGTGCTGACTTTTGCTGTTGGAGTT
     I  N  E  Q  V  P  Q  V  L  V  L  V  T  A  G  R  S  A  V  P  F  L  Q  V  S  N  D  L  A  R  A  G  V  L  T  F  A  V  G  V   560
1681 AGGAATGCGGATAAGGCAGAGCTTGAACAGATTGCATTTAATCCAAAAATGGTATATTTTATGGATGATTTCAGTGATCTGACAACATTACCTCAGGAGTTAAAAAAGCCTATAACAACT
     R  N  A  D  K  A  E  L  E  Q  I  A  F  N  P  K  M  V  Y  F  M  D  D  F  S  D  L  T  T  L  P  Q  E  L  K  K  P  I  T  T   600
1801 ATAGTTAGTGGAGGTGTGGAGGAGGTTCCACTCGCCCCAACAGAAAGCAAGAAAGACATTTTATTCCTGATTGATGGTTCAGCCAACCTCTTGGGTAGCTTTCCTGCTGTCCGGGACTTT
     I  V  S  G  G  V  E  E  V  P  L  A  P  T  E  S  K  K  D  I  L  F  L  I  D  G  S  A  N  L  L  G  S  F  P  A  V  R  D  F   640
1921 ATACACAAAGTCATTTCTGACCTGAACGTGGGTCCCGACGCCACGCGAGTAGCTGTGGCTCAGTTCAGTGACAACATCCAAATTGAATTTGACTTTGCTGAACTCCCATCTAAGCAAGAC
     I  H  K  V  I  S  D  L  N  V  G  P  D  A  T  R  V  A  V  A  Q  F  S  D  N  I  Q  I  E  F  D  F  A  E  L  P  S  K  Q  D   680
2041 ATGCTTCTGAAAGTGAAAAGGATGAGGTTAAAAACTGGGAAGCAGCTGAATATCGGAGTTGCGCTCGATGAAGTCATGAGGAGGCTGTTCGTGAAGGAAGCTGGAAGCAGGATTGAAGAA
     M  L  L  K  V  K  R  M  R  L  K  T  G  K  Q  L  N  I  G  V  A  L  D  E  V  M  R  R  L  F  V  K  E  A  G  S  R  I  E  E   720
2161 GGCATTCCTCAGTTCTTGGTCCTCCTGGCTGCTGGGAGGTCAACCGATGAGGTGGAGCGACCATCGGGAGCTCTGAAGGAAGCTGGAGTTGTGACCTTTGCTATCAAAGCCAAAAATGCT
     G  I  P  Q  F  L  V  L  L  A  A  G  R  S  T  D  E  V  E  R  P  S  G  A  L  K  E  A  G  V  V  T  F  A  I  K  A  K  N  A   760
2281 GATTTATCAGAGCTGGAAAGGATAGCTTACGCCCCACAGTTCATTCTGAATGTCGAATCCCTTCCTCGGATTTCGGAGCTCCAGGCAAACATAGTGAACCTACTGAAAACTATCCAGTTT
     D  L  S  E  L  E  R  I  A  Y  A  P  Q  F  I  L  N  V  E  S  L  P  R  I  S  E  L  Q  A  N  I  V  N  L  L  K  T  I  Q  F   800
2401 CAGCCAACAGTAGTTGAGAGAGGTGAGAAGAAGGATGTGGTGTTTCTAATCGATGGCTCGGATGGTGTCAGAAGAGGTTTCCCTCTCCTGAAAACCTTTGTGGAAAGAGTTGTTGAAAGC
     Q  P  T  V  V  E  R  G  E  K  K  D  V  V  F  L  I  D  G  S  D  G  V  R  R  G  F  P  L  L  K  T  F  V  E  R  V  V  E  S   840
2521 CTTGATATTGGCCGTGACAAGGTCCGTGTTGCCATTGTGCAGTACAGCAACGCCATACAACCTGAGTTCTTGCTGGATGCCTACGAAGACAAAGCCGATTTAGTCAGTGCCATCCAGGCG
     L  D  I  G  R  D  K  V  R  V  A  I  V  Q  Y  S  N  A  I  Q  P  E  F  L  L  D  A  Y  E  D  K  A  D  L  V  S  A  I  Q  A   880
2641 CTGACAATTATGGGAGGATCTCCTCTGAACACCGGAGCAGCCCTCGACTATCTCATCAAGAACGTGTTCACCGTGTCGAGTGGCAGCAGGATAGCTGAGGGCGTCCCGCAGTTCCTGATC
     L  T  I  M  G  G  S  P  L  N  T  G  A  A  L  D  Y  L  I  K  N  V  F  T  V  S  S  G  S  R  I  A  E  G  V  P  Q  F  L  I   920
2761 CTGCTCACTGCCGACCGATCCCAGGATGATGTGAGGAGGCCCTCAGTGGTCCTCAAGACGAGTGGCACAGTGCCCTTCGGCATCGGGATTGGAAATGCCGACCTCACAGAGCTTCAGACG
     L  L  T  A  D  R  S  Q  D  D  V  R  R  P  S  V  V  L  K  T  S  G  T  V  P  F  G  I  G  I  G  N  A  D  L  T  E  L  Q  T   960
2881 ATCTCCTTCCTCCCGGATTTTGCTATCTCTGTACCCGACTTCAGCCAGCTGGATTCAGTGCAACAGGCCGTGTCCAACAGAGTCATCCGGCTGACCAAGAAAGAAATAGAGTCTCTTGCC
     I  S  F  L  P  D  F  A  I  S  V  P  D  F  S  Q  L  D  S  V  Q  Q  A  V  S  N  R  V  I  R  L  T  K  K  E  I  E  S  L  A  1000
3001 CCTGATCTGGTTTTTACATCACCCAGCCCAGTGGGTGTGAAGAGGGATGTTGTGTTCCTGGTCGATGGCTCCCGCTACGCCGCCCAGGAATTCTACCTCATCCGTGATCTCATTGAGAGG
     P  D  L  V  F  T  S  P  S  P  V  G  V  K  R  D  V  V  F  L  V  D  G  S  R  Y  A  A  Q  E  F  Y  L  I  R  D  L  I  E  R  1040
3121 ATAGTGAACAACCTGGATGTGGGCTTTGACACCACTCGGATTTCGGTGGTCCAGTTCAGCGAGCATCCCCATGTAGAGTTTCTTCTCAATGCCCACTCCACCAAGGATGAGGTGCAAGGT
     I  V  N  N  L  D  V  G  F  D  T  T  R  I  S  V  V  Q  F  S  E  H  P  H  V  E  F  L  L  N  A  H  S  T  K  D  E  V  Q  G  1080
3241 GCAGTGAGGCGGCTGCGGCCACGGGGTGGCCAGCAGGTGAACGTGGGGGAAGCCCTTGAATTTGTGGCAAAAACCATCTTTACCCGTCCGTCTGGGAGCCGCATAGAAGAGGGCGTCCCT
     A  V  R  R  L  R  P  R  G  G  Q  Q  V  N  V  G  E  A  L  E  F  V  A  K  T  I  F  T  R  P  S  G  S  R  I  E  E  G  V  P  1120
3361 CAGTTTTTGGTTATCCTCTCCTCCCGCAAGTCCGACGATGACTTAGAATTCCCATCGGTACAAGTCAAACAAGTAGGAGTGGCACCTATGGTCATAGCAAAGAACATGGATCCTGAGGAG
     Q  F  L  V  I  L  S  S  R  K  S  D  D  D  L  E  F  P  S  V  Q  V  K  Q  V  G  V  A  P  M  V  I  A  K  N  M  D  P  E  E  1160
3481 ATGGTGCAGATTTCCCTTAGCCCTGACTATGTGTTCCAAGTCTCCAGCTTCCAGGAACTGCCCAGCCTTGAACAGAAGCTGCTGGCTCCTATTGAAACCCTGACTGCAGACCAAATCAGA
     M  V  Q  I  S  L  S  P  D  Y  V  F  Q  V  S  S  F  Q  E  L  P  S  L  E  Q  K  L  L  A  P  I  E  T  L  T  A  D  Q  I  R  1200
3601 CAACTCCTGGGAGATGTGACAACAATTCCAGATGTTTCTGGTGAGGAAAAGGACGTAGTCTTCCTTATTGATAGCTCCGACAGCGTTAGGTCTGATGGGCTTGCTCACATTCGGGATTTC
     Q  L  L  G  D  V  T  T  I  P  D  V  S  G  E  E  K  D  V  V  F  L  I  D  S  S  D  S  V  R  S  D  G  L  A  H  I  R  D  F  1240
3721 ATCAGCAGAATTGTCCAGCAGCTTGATGTTGGGCCAAACAAAGTGCGGATTGGAGTGGTGCAATTCAGCAATAATGTCTTCCCTGAGTTCTACTTGAGAACCCACAAGTCTAAGAATGCC
     I  S  R  I  V  Q  Q  L  D  V  G  P  N  K  V  R  I  G  V  V  Q  F  S  N  N  V  F  P  E  F  Y  L  R  T  H  K  S  K  N  A  1280
3841 GTACTGCAAGCCATCCGCCGCTTACGGCTTAGAGGAGGGTACCCCGTGAACGCTGGCAAAGCCCTGGACTACGTGGTGAAGAACTACTTCATCAAGTCTGCAGGGAGCAGGATAGAAGAT
     V  L  Q  A  I  R  R  L  R  L  R  G  G  Y  P  V  N  A  G  K  A  L  D  Y  V  V  K  N  Y  F  I  K  S  A  G  S  R  I  E  D  1320
3961 GGAGTGCCCCAGCATCTAGTTGTCATCCTTGGAGATCAGTCCCAGGATGATGTCAACAGGCCTGCTAATGTGATCTCTTCAACGAGTATTCAACCTCTGGGTGTAGGAGCCAGGAATGTA
     G  V  P  Q  H  L  V  V  I  L  G  D  Q  S  Q  D  D  V  N  R  P  A  N  V  I  S  S  T  S  I  Q  P  L  G  V  G  A  R  N  V  1360
4081 GACAGGAACCAGCTGCAGGTCATCACCAATGATCCTGGCCGCGTGCTTGTAGTGCAGGACTTCACAGGACTGCCCACTTTAGAAAGAAAGGTGCAGAACATTCTTGAAGAGCTTACAGTA
     D  R  N  Q  L  Q  V  I  T  N  D  P  G  R  V  L  V  V  Q  D  F  T  G  L  P  T  L  E  R  K  V  Q  N  I  L  E  E  L  T  V  1400
4201 CCTACAACAGAAGGCCCTGTATATCCAGGACCAGAAGGCAAGAAGCAGGCTGACATTGTGTTTTTGCTGGATGGCTCCATCAACCTGGGAAGAGATAACTTCCAAGAGGTTCTCCAGTTC
     P  T  T  E  G  P  V  Y  P  G  P  E  G  K  K  Q  A  D  I  V  F  L  L  D  G  S  I  N  L  G  R  D  N  F  Q  E  V  L  Q  F  1440
4321 GTGTATTCTATAGTGGATGCCATCTATGAGGATGGTGACTCCATCCAGGTAGGACTGGCACAGTACAACTCTGATGTCACTGATGAGTTTTTCCTCAAGGATTACTCCAGCAAACCTGAG
     V  Y  S  I  V  D  A  I  Y  E  D  G  D  S  I  Q  V  G  L  A  Q  Y  N  S  D  V  T  D  E  F  F  L  K  D  Y  S  S  K  P  E  1480
4441 ATTTTAGACGCCATCAACAAGGTGATCTACAAAGGTGGCCGAGTGGCAAACACTGGTGCAGCCATCAAGCACCTTCAGGCAAAACACTTTGTGAAGGAGGCTGGCAGCCGAATTGACCAG
     I  L  D  A  I  N  K  V  I  Y  K  G  G  R  V  A  N  T  G  A  A  I  K  H  L  Q  A  K  H  F  V  K  E  A  G  S  R  I  D  Q  1520
4561 AGGGTGCCTCAGATCGCCTTCATCATCACAGGGGGGAAATCTTCCGATGATGGGCAAGGTGCCTCCATGGAAGTTGCACAGAAGGGTGTCAAGGTGTTTGCAGTTGGTGTAAGGAACATC
     R  V  P  Q  I  A  F  I  I  T  G  G  K  S  S  D  D  G  Q  G  A  S  M  E  V  A  Q  K  G  V  K  V  F  A  V  G  V  R  N  I  1560
4681 GACCTGGAGGAGGTCAGCAAGTTGGCCAGCGAGAGTGCCACGAGTTTCAGAGTGTCGACGGCTCAGGAGTTGTCTGAACTGAATGAGCAGGTCTTGGTTACATTGGCAGCTGCTATGAAG
     D  L  E  E  V  S  K  L  A  S  E  S  A  T  S  F  R  V  S  T  A  Q  E  L  S  E  L  N  E  Q  V  L  V  T  L  A  A  A  M  K  1600
4801 GAGAAACTGTGCCCAGGAACGACAGATGTGACAAGGGACTGCGACTTAGACGTGATACTTGGCTTCGATGTGTCAGATGTCGGAGCCGGCCAAAACATCTTCAATTCCCAGAGAGGGCTG
     E  K  L  Ⓒ  P  G  T  T  D  V  T  R  D  Ⓒ  D  L  D  V  I  L  G  F  D  V  S  D  V  G  A  G  Q  N  I  F  N  S  Q  R  G  L  1640
```

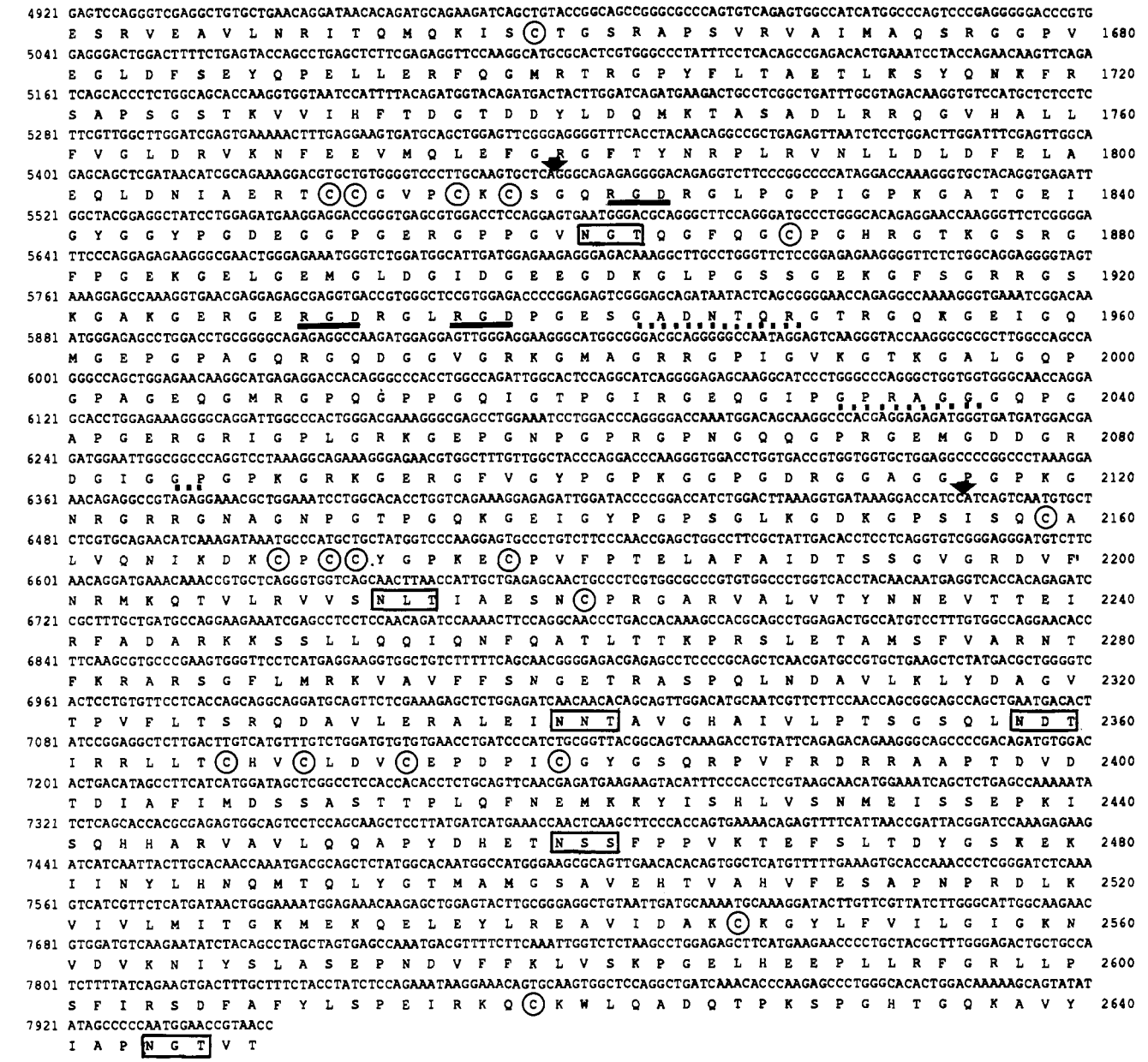

FIGURE 4: Nucleotide and predicted amino acid sequence of the α3(VI) chain cDNA: first line, nucleotide sequence; second line, deduced amino acid sequence. Nucleotides and amino acids are numbered on the left- and right-hand side, respectively. The start and the end of the triple helix are marked by an arrow. Six potential N-attachment sites for oligosaccharides are boxed, and cysteine residues are circled. Three potential Arg-Gly-Asp (RGD) cell-binding sequences are underlined by solid lines, and three interruptions in the triple helix are underlined by dotted lines.

collagenase (Furuto & Miller, 1980; Jander et al., 1981; Odermatt et al., 1983). Very similar clusters are found in the α1 (VI) (Bonaldo et al., 1989) and α2 (VI) (Koller et al., 1989) chains. One additional cluster of cysteines is contained in the carboxyl noncollagenous portion of the α3(VI) sequence. The collagenous sequence presents three imperfections of the triplet repeats that may explain the supertwisting between type VI dimers in the formation of a tetramer as suggested by the analysis of the molecule after rotary shadowing (Furthmayr et al., 1983). The triple helix contains also one cysteine and three potential cell-binding Arg-Gly-Asp sequences. No other Arg-Gly-Asp sequences were found throughout the rest of the α3(VI) sequence.

Domain Structure. The present 2648 residue long sequence of the α3(VI) chain is very unusual for a collagen: the triple helix encompasses only about one-eighth of the still incomplete sequence; the N-terminal portion is much more extended compared with the C-terminal portion; and finally, several features of this chain such as clusters of cysteines and potential N-linked glycosylation sites are confined to the carboxyl end of the sequence. To further investigate the domain organization of the α3(VI) chain, internal amino acid sequence similarities were searched by comparing the sequence with itself by the diagonal dot matrix (Figure 5). This analysis revealed the existence of a pattern of multiple repeats that covered all of the sequence. Apart from the triplet repeats of the COL domain at about two-thirds of the sequence, the most striking characteristic is the presence of 11 repeats composed of about 200 residues each and comprising all of the remaining sequence. The similarity is higher for the eight repeats at the N-terminal end, whereas the three repeats separated by the COL domain show a lower degree of similarity comparable to that exhibited by the three repeats found in the α1(VI) chain (Bonaldo et al., 1989).

A search of the NBRF Protein Sequence Data Base using the Lipman and Pearson (1985) FASTP program revealed a significant identity of the eleven 200 residue long repeats with the type A repeats of vWF (Shelton-Inloes et al., 1986; Titani

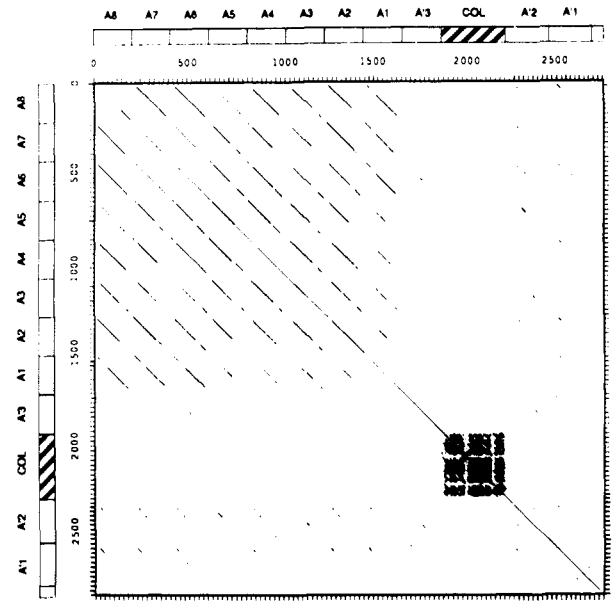

FIGURE 5: Dot-matrix analysis of the deduced amino acid sequence of the α3(VI) chain. The amino acid sequence of the α3(VI) chain is compared to itself (Dayoff, 1979). Sequences of 36 contiguous residues are compared, and a dot is placed when the total score for the comparison is 50% or more.

et al., 1986). The eight repeats at the N-terminal side of the α3(VI) sequence exhibit a higher overall degree of similarity with the repeats of vWF (range between 26 and 36%); on the other hand, the repeats at the C-terminal side of the α3(VI) molecule have a lower similarity with vWF (range between 18 and 26%). Consequently, we designated the former group as type "A" repeats and the three repeats at the two sides of the COL domain as type "A'". In both types of repeats there is shared clustering of sequence identities in several regions (Figure 6).

A schematic domain structure of the α3(VI) chain compared with that of α1(VI) chain (Bonaldo et al., 1989) is shown in Figure 7.

Binding to Collagen Type I. The type A domains of vWF are three tandem repeats between residues 496 and 1111 of the mature protein. Two independent collagen-binding sites have been localized to the same region (Pareti et al., 1987). The present finding that the α3(VI) chain contains several repeats homologous to the type A repeats of vWF together with the notion that also the α1(VI) (Bonaldo et al., 1989) and α2(VI) chains (Koller et al., 1989) have three similar types of repeats predicts that type VI collagen would be functionally able to interact with type I collagen. We first investigated the ability of the tissue form of type VI collagen to bind to type I collagen fibrils adsorbed onto plastic surfaces. The results demonstrated that type VI collagen bound efficiently to type I collagen and that the binding was inhibited to a great extent by the expression protein of clone pB10, which codes for about 5.5 type A repeats (Figure 8A). The ability of purified vWF to inhibit most of the binding activity of labeled type VI collagen for type I collagen and the absence of any competition by other proteins that lack type A repeats such as tropoelastin, bovine serum albumin, and also the pepsin-resistant form of type VI collagen (Figure 8A) extended the significance of the phenomenon observed. Type VI collagen is purified from the tissue as a tetramer with an approximate M_r = 2 000 000, and the α3(VI) chain is present as multiple bands [see Figure 2 and Colombatti et al. (1989)]. Consequently, to further characterize binding parameters, we used clone pB10. The gel-eluted expression protein of clone pB10 exhibited type I collagen binding, and this binding was time dependent (Figure 8B), probably saturable (Figure 8C), and

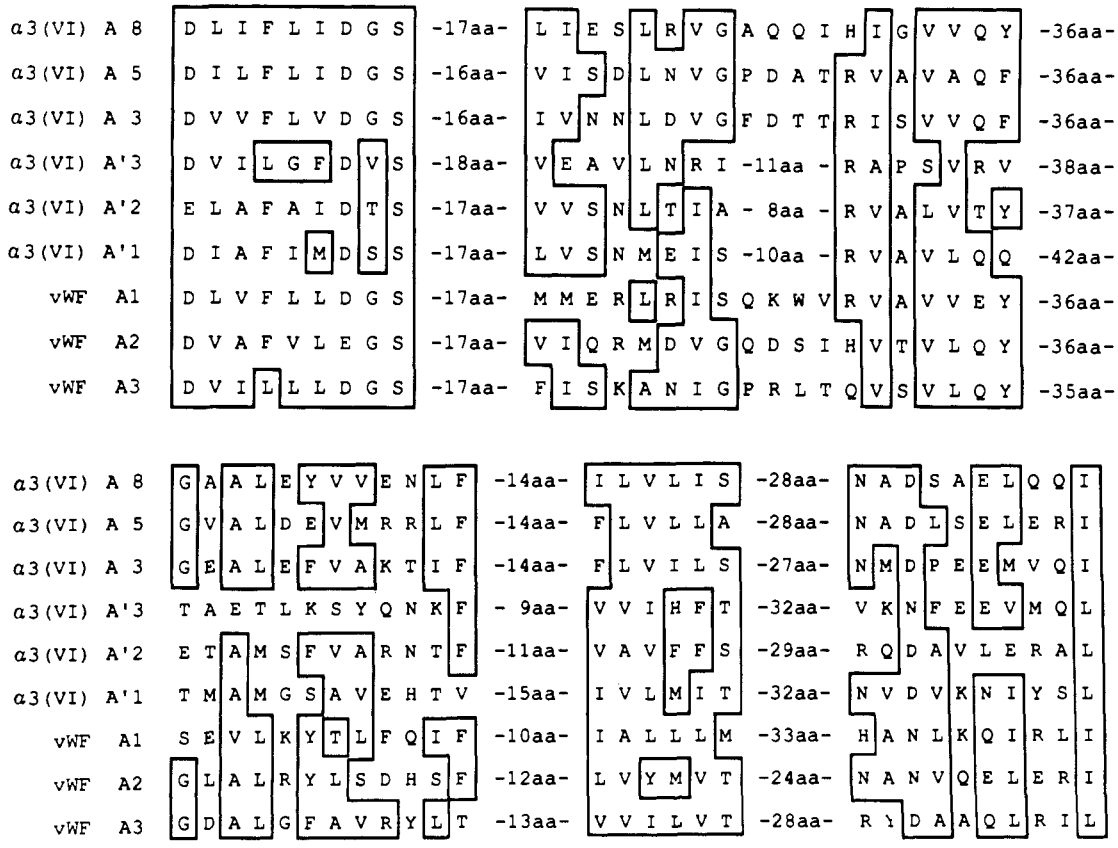

FIGURE 6: Alignment of chick α3(VI) and human vWF. Clusters of higher shared similarities are shown. Identical or similar residues shared by at least five repeats are boxed. Only three out of eight type A repeats are shown.

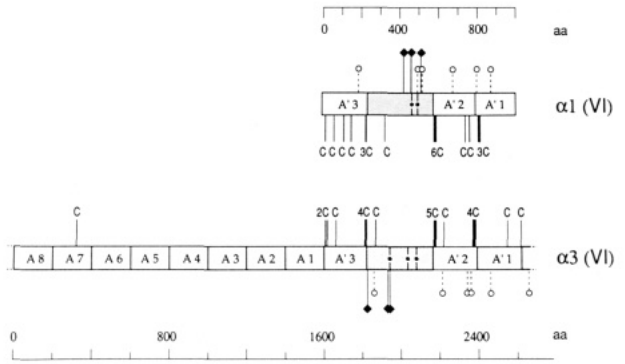

FIGURE 7: Schematic domain structure of the α3(VI) chain compared with the α1(VI) chain. The shaded box indicates the triple helix (COL domain). Cysteines, interruptions in the triple helix, potential cell-attachment sites, and glycosylation sites are indicated by C, closed circles, closed diamonds, and open circles, respectively.

inhibited by an excess of the respective cold ligand (Figure 8D). These assays were repeated several times with different labelings and different preparations of type VI collagen and expression protein pB10 and confirmed the extent of the binding observed.

Cell Adhesion. The deduced sequence of the α3(VI) chain has indicated that there are three potential Arg-Gly-Asp cell-binding sequences in the triple helix. Interestingly enough, two of them are located in a region of the sequence where several Gly-Xaa-Yaa triplets are lacking a proline in the X or Y position. This could probably endow the triple helix with less rigidity, and the Arg-Gly-Asp sequences could be available and functionally interact with integrin-like receptors present at the cell surface (Ruoslahti & Pierschbacher, 1987). To investigate the possibility that type VI collagen might promote cell adhesion, the pepsin-resistant form of type VI collagen was adsorbed to plastic and assayed for its ability to mediate the attachment of embryo arterial smooth muscle cells. The pepsin-resistant form of type VI collagen stimulated cells attachment in a dose- and time-dependent fashion. The expression protein of clone pB32 that codes for the whole triple-helical domain of α3(VI) could also mediate cell adhesion (data not shown). Cell adhesion was inhibited by the addition of increasing amounts of the synthetic peptide Gly-Arg-Gly-Asp-Ser-Pro (GRGDSP) but not of the peptide Gly-Arg-Gly-Glu-Ser-Pro (GRGESP) (Figure 9). Laminin, used as a positive control for cell adhesion, gave a higher adhesion in terms of number of smooth muscle cells attached. However,

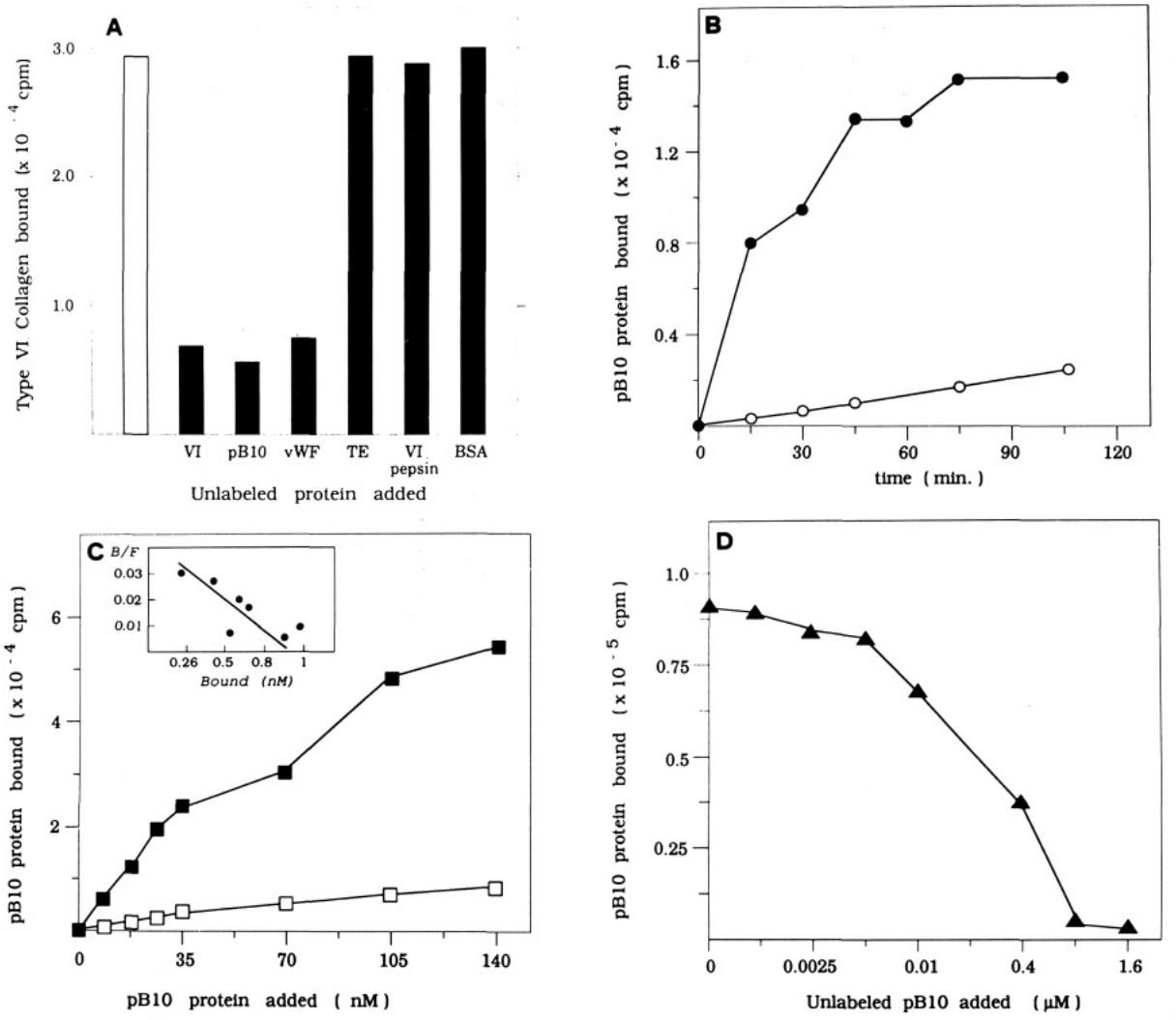

FIGURE 8: Binding of ^{125}I-labeled type VI collagen and pB10 protein to fibrillar type I collagen. (A) Binding was carried out at 4 °C for 60 min. Shaded boxes indicate the binding values obtained in the presence of a 100-fold excess of cold ligand. VI, tissue form of type VI collagen; vWF, von Willebrand factor; TE, fusion protein of tropoelastin; VI pepsin, pepsin-resistant form of type VI collagen; BSA, bovine serum albumin; pB10, fusion protein of pB10 clone. Kinetics (B), saturability (C), and reversibility (D) of binding of pB10 clone in the absence (●, ■, ▲) or in the presence (○, □) of a 100-fold molar excess of cold ligand. The inset in panel C shows a Scatchard-type plot of the experimental data using the computer-assisted program LIGAND (Munson, 1983). By assuming an M_r = 123 000 for pB10 fusion protein, a K_d of 2.66 × 10^{-8} M was calculated. In panels A, B, and D about 90 ng of labeled ligand (4100 cpm/ng) was used.

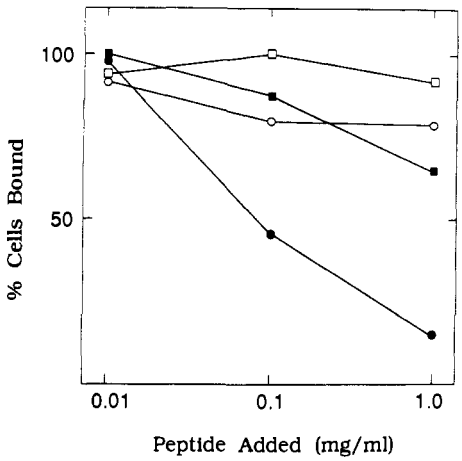

FIGURE 9: Inhibition of cell adhesion to type VI collagen. Ten micrograms per well of laminin (■, □) and pepsin-resistant type VI collagen (●, ○) was air-dried onto a multiwell dish. [^{35}S]-Methionine-labeled embryo smooth muscle cells were centrifuged (100*g*, 1 min) onto the plated proteins and incubated for 75 min in the presence of the varying concentrations of the GRGDSP (■, ●) or GRGESP (□, ○) peptides indicated on the abscissa. After extensive washing, cells bound to the wells were quantitated by counting of the remaining cell-associated radioactivity. Binding is expressed as a percent of that observed in the absence of competing peptides.

in this case cell binding was not affected by the addition of either peptide.

Discussion

The results presented here establish most of the primary sequence of the α3 chain of type VI collagen and provide functional data that help to understand the role of type VI collagen in connective tissues.

Confirmation of the authenticity of the cDNA clones isolated was obtained by several methods: antibodies from the polyclonal serum were affinity purified onto fusion proteins and in a retroblot experiment recognized the tissue form of α3(VI); we elicted also antibodies against fusion proteins and demonstrated reactivity with the α3(VI) chain. Northern blot analysis of chick mRNA gave a band of about 9–10 kb, compatible with the size of the α3(VI) chain; finally, the sequence of the COL domain matched the partial sequences recently published for the α3 chain of human type VI collagen (Chu et al., 1987, 1988).

The deduced amino acid sequence specifies for a 2648 amino acid polypeptide with a molecular weight of 286 892. If one adds posttranslational modifications such as N- and O-linked glycosylation, hydroxylation of proline and lysine in the third position of the Gly-Xaa-Yaa tripeptides, and disaccharide addition to hydroxylysines, the molecular weight might be higher. Considering that other sequences are still lacking from the present reading frame at both the 5′- and 3′-ends, the total mass of the α3(VI) chain should be at least on the order of 310 000 daltons, suggesting that the value of M_r = 260 000 implied by several authors including ourselves from the apparent migration after SDS–PAGE was underestimated. With this information deduced from the sequence, the actual size of the α3(VI) chain is about three times larger compared with the size of the α1(VI) and α2(VI) chains. This unusual length of the α3(VI) chain, which is a unique characteristic among collagens discovered to date, is due to the very extended amino-terminal region of about 1600 residues comprising at least eight repeating units.

The most distinctive feature that emerged from the analysis of this sequence is the finding that about 85% of the α3(VI) chain is represented by two types of similar repeating motifs, designated domains A and A′, that share extensive identities with the type A repeats of vWF. Identity with the type A repeats of vWF has been detected also for the noncollagenous regions of the α1(VI) and α2(VI) chains of chick (Bonaldo et al., 1989; Koller et al., 1989) and human molecules (Chu et al., 1989). Similar sequences are present in several other apparently unrelated proteins such as the chicken cartilage matrix protein (Argraves et al., 1987), the complement factors B and C2 (Mole et al., 1984; Bentley, 1986), and the α chain of the integrin molecules p150,95 (Corbi et al., 1987), Mac-1 (Pytela et al., 1988; Arnaout et al., 1988; Corbi et al., 1988), LFA-1 (Larson et al., 1989), and VLA-2 (Takada & Hemler, 1989). At least in vWF (Pareti et al., 1986, 1987) and cartilage matrix protein (Argraves, 1987) it is very likely that the type A domains are responsible for the interaction with type I collagen. The data presented here demonstrate that type VI collagen interacts with type I collagen immobilized on plastic and that this type I collagen-binding activity resides in the noncollagenous part of the α3(VI) chain containing the type A repeats. vWF, which shares with type VI collagen only the type A repeats, inhibits to a great extent the interaction between type VI and type I collagens, whereas the triple helix of type VI collagen fails to compete for the binding to collagen type I. The expressed fusion protein of the clone pB10, encoding about 5.5 type A repeats and no other sequences, bound to type I collagen in a saturable, time-dependent, and reversible manner, supporting the concept that the binding functions are limited to these repeats. This does not exclude the possibility of additional sites with affinity for collagen I.

In this context, although not proven by our data, the type A′ repeats of the α3(VI), of the α1(VI) (Bonaldo et al., 1989), and of the α2(VI) chains (Koller et al., 1989) might also share these binding properties with the α3(VI) type A repeats. It could be speculated that one function of multiple (17 for the three chains altogether) type A like domains within the sequence of type VI collagen is to augment the overall affinity of the interaction with type I collagen by acting in concert. This hypothesis can now be tested by producing fusion proteins with a different number of repeats and measuring the relative binding affinities. A recent observation seems pertinent to this discussion: it has been found that the α3(VI) chain is present in the tissue as a ladder of polypeptides that differ from each other by about 20 000 daltons (Colombatti et al., 1989), which is the approximate size of the type A domain. If this variation is not only the consequence of proteolytic degradation, it must imply, for one thing, the existence of a mixed population of type VI molecules with local variations in the number of type A repeats. It is tempting to speculate on the role of these repeats in protein–protein interactions since they could affect the molecular structure of the extracellular matrix. By varying the length of the interacting region, changes in the overall affinity of the interaction between type I collagen and type VI molecules could ensue and act to modulate the plasticity in the connective tissues. It is conceivable that alternative splicing of exons for the type A repeats (Doliana et al., unpublished results) could also participate in modulating the processes of interaction as has been postulated for other multimodule extracellular matrix proteins such as fibronectin (Kornblihtt et al., 1984). The finding of multiple forms of the M_r = 260 000 α3(VI) subunit in immunoprecipitates from lysates of cells metabolically labeled by very short pulse–chase labeling (Colombatti & Bonaldo, 1987) is in agreement with the above suggestion. The fine tissue distribution of type VI collagen is consistent with the observations reported above. Microfilaments of type VI collagen often appear closely ap-

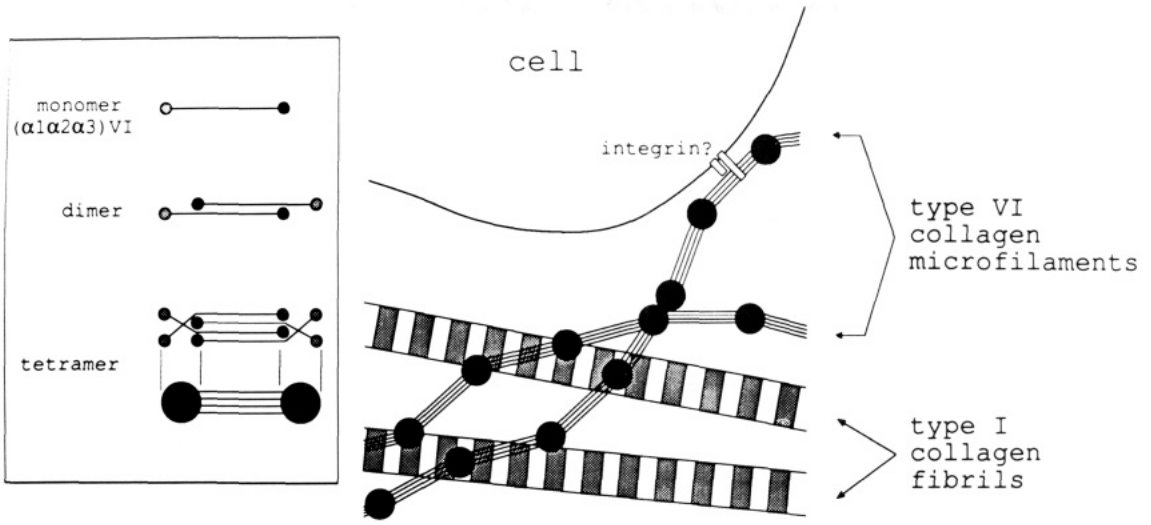

FIGURE 10: Schematic drawing of type VI collagen assembly and tissue organization. The functions ascribed to type VI collagen on the basis of the present findings are illustrated. The inset shows a model of assembly of type VI collagen based on electron microscopy data by Furthmayr et al. (1983).

posed to large striated collagen fibers or oriented perpendicular to the major collagenous fibers and forming an interconnecting meshwork (von der Mark et al., 1984; Bruns, 1984; Bruns et al., 1986; Keene et al., 1988). Physiological interactions between type VI collagen and the major fibrous collagens are very likely to occur at multiple sites. The resolution of electron microscopy is not sufficient to indicate whether the interaction with collagen fibers is involving either the "beads" of the type VI microfibrils, which correspond to the globular ends of the tetramers and are composed of the type A and A′ repeats, or the "rods" containing the triple helices. The present results by showing that type A repeats bind to immobilized type I collagen in vitro suggest that the structural components of the type VI microfilaments that interact with type I collagen fibers in vivo could be the beads. Our data demonstrate the existance of a heterophilic interaction between type I collagen and type A repeats of collagen VI. They do not exclude the possibility of other interactions: for instance, the type A repeats of one collagen VI molecule could interact with the collagenous portion of the antiparallel monomer of collagen VI and contribute to the formation of the dimer.

Given the wide participation of the Arg-Gly-Asp sequence in cell–matrix interactions (Ruoslahti & Pierschbacher, 1987), it was of great interest to determine whether the multiple Arg-Gly-Asp sequences present in the α3(VI) chains contained sites active in promoting cell adhesion. Multiple cell receptors of the integrin type and of a different type have been found on the cell membrane and described to interact with collagens (Wayner & Carter, 1987; Wayner et al., 1988). Our data demonstrate that the pepsin-resistant form of type VI collagen and a fusion protein of an α3(VI) clone coding for the collagenous domain have cell-attachment properties. This function is inhibited by synthetic peptides of the Arg-Gly-Asp type, but at the moment it is not known which Arg-Gly-Asp sequences in the type VI collagen chains are active in this function. A cell-attachment activity that can be similarly inhibited by synthetic RGDS-containing peptides has been recently demonstrated for the triple helix of human type VI collagen (Aumailley et al., 1989). It is worth mentioning that Arg-Gly-Asp sequences are found only in the COL domain in all three chains of type VI collagen, thus restricting the cell adhesion function to the short triple helix of this molecule. An RGDS-independent cell adhesion activity also has been described for type VI collagen (Wayner & Carter, 1987), and therefore, this collagen could interact with the cell membrane via multiple molecules, or different receptors could be present on different cell types. The close proximity of type VI collagen to cells (Bruns et al., 1986; Keene et al., 1988) suggests that cell adhesion could be one of the main functions of this molecule. Support to this notion comes from the recent finding that chick type VI collagen has 13 Arg-Gly-Asp sequences [three in the α1 chain (Bonaldo et al., 1989), seven in the α2 chain (Koller et al., 1989), and three in the α3 chain (this paper)].

In summary, the deduced amino acid sequence presented here establishes the basic structure of the α3 subunit of type VI collagen. The data show that this unique collagenous component is a multidomain polypeptide made of an Arg-Gly-Asp-containing triple helix and of several repeating units similar to the type A repeats of vWF. The cell-attachment and the type I collagen-binding properties of type VI collagen as demonstrated in this paper provide a functional correlate to the structural and tissue localization data (see Figure 10 for a schematic representation).

ACKNOWLEDGMENTS

We are grateful to Dr. G. M. Bressan for the cDNA library cloned in the pEX1 vector and for the pUEX1 intermediate vector and to Dr. G. Tarone for the gift of synthetic peptides. We thank Maria Teresa Mucignat for technical assistance and Elisabetta Montagner for typing the manuscript.

Registry No. von Willebrand factor, 109319-16-6.

REFERENCES

Argraves, W. S., Deak, F. Sparks, J. J., Kiss, I., & Goetinck, P. F. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 464–468.

Arnaout, M. A., Gupta, S. K., Pierce, M. W., & Tenen, D. G. (1988) *J. Cell Biol. 106*, 2153–2158.

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

Bentley, D. R. (1986) *Biochem. J. 239*, 339–345.

Biggin, M. D., Gibson, T. J., & Hong, G. F. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 3963–3965.

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

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

Bressan, G. M., & Stanley, K. K. (1987) *Nucleic Acids Res. 15*, 10056.

Bressan, G. M., Argos, P., & Stanley, K. K. (1987) *Biochemistry 26*, 1497–1503.
Bruns, R. (1984) *J. Ultrastruct. Res. 89*, 136–145.
Bruns, R., Press, W., Engvall, E., Timpl, R., & Gross, J. (1986) *J. Cell Biol. 103*, 394–404.
Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., & Rutter, W. J. (1979) *Biochemistry 18*, 5294–5298.
Chu, M. L., Mann, K., Deutzmann, R., Pribula-Conway, D., Hsu-Chen, C. C., Bernard, M. P., & Timpl, R. (1987) *Eur. J. Biochem. 168*, 309–317.
Chu, M. L., Conway, D., Pan, T., Baldwin, C., Mann, K., Deutzmann, R., & Timpl, R. (1988) *J. Biol. Chem. 263*, 18601–18606.
Chu, M. L., Pan, T., Conway, D., Kuo, H. J., Glanville, R. W., Timpl, R., Mann, K., & Deutzmann, R. (1989) *EMBO J. 8*, 1939–1946.
Colombatti, A., & Bonaldo, P. (1987) *J. Biol. Chem. 262*, 14461–14466.
Colombatti, A., Bonaldo, P., Ainger, K., Bressan, G. M., & Volpin, D. (1987) *J. Biol. Chem. 262*, 14454–14460.
Colombatti, A., Ainger, K., Mucignat, M. T., & Bonaldo, P. (1988) *Collagen Relat. Res. 8*, 85–88.
Colombatti, A., Ainger, K., & Colizzi, F. (1989) *Matrix 9*, 177–185.
Corbi, A., Miller, L. J., O'Connor, K., Larson, S. R., & Springer, T. A. (1987) *EMBO J. 6*, 4023–4028.
Corbi, A. L., Kishimoto, T. K., Miller, L. J., & Springer, T. A. (1988) *J. Biol. Chem. 263*, 12403–12411.
Dayhoff, M. D. (1979) in *Atlas of Proteins Sequence and Structure* (Schwartz, R. M., & Dayhoff, M. D., Eds.) Vol. 5, Suppl. 3, National Biomedical Research Foundation, Washington, DC.
Engvall, E., Hessle, H., & Klier, G. (1986) *J. Cell Biol. 102*, 703–710.
Feinberg, A. P., & Vogelstein, B. (1983) *Anal. Biochem. 132*, 6–13.
Fraker, D. J., & Speck, J. C. (1978) *Biochem. Biophys. Res. Commun. 80*, 849–857.
Furthmayr, H., Wiedemann, H., Timpl, R., Odermatt, E., & Engel, J. (1983) *Biochem. J. 211*, 303–311.
Furuto, D. K., & Miller, E. J. (1981) *J. Biol. Chem. 255*, 290–295.
Gibson, M. A., & Cleary, E. G. (1985) *J. Biol. Chem. 260*, 1149–1159.
Haymerle, H., Herz, J., Bressan, G. M., Frank, R., & Stanley, K. K. (1986) *Nucleic Acids Res. 14*, 8615–8624.
Jander, R. D., Rauterberg, J., Voss, B., & von Bassewitz, D. B. (1981) *Eur. J. Biochem. 114*, 17–25.
Jander, R. D., Troyer, D., & Rauterberg, J. (1984) *Biochemistry 23*, 3675–3681.
Keene, D. R., Engvall, E., & Glanville, R. W. (1988) *J. Cell Biol. 107*, 1995–2006.
Knight, K. R., Ayad, S., Shuttleworth, A., & Grant, M. E. (1984) *Biochem. J. 220*, 395–403.
Koller, E., Winterhalter, K. H., & Trüeb, B. (1989) *EMBO J. 8*, 1073–1077.
Kornblihtt, A. R., Vibe-Pedersen, K., & Baralle, F. E. (1984) *EMBO J. 3*, 221–226.
Larson, R. S., Corbi, A. L., Berman, L., & Springer, T. (1989) *J. Cell Biol. 108*, 703–712.
Lipman, D. J., & Pearson, W. R. (1985) *Science 227*, 1435–1441.
Messing, J. (1983) *Methods Enzymol. 101*, 20–78.
Miller, E. J., & Gay, S. (1987) *Methods Enzymol. 144*, 3–41.
Mole, J. E., Anderson, J. ., Davison, E. A., & Woods, D. E. (1984) *J. Biol. Chem. 259*, 3407–3412.
Munson, P. J. (1983) *Methods Enzymol. 92*, 542–576.
Odermatt, E., Risteli, J., Van Delden, V., & Timpl, R. (1983) *Biochem. J. 211*, 259–302.
Pareti, F. I., Fujimura, Y., Dent, J. A., Holland, L. Z., Zimmerman, T. S., & Ruggieri, Z. M. (1986) *J. Biol. Chem. 261*, 15310–15315.
Pareti, F. I., Niiya, K., McPherson, J. M., & Ruggieri, Z. M. (1987) *J. Biol. Chem. 262*, 13835–13841.
Pytela, R. (1988) *EMBO J. 7*, 1371–1378.
Ruggeri, Z. M., De Marco, L., Gatti, L., Bader, P., & Montgomery, R. R. (1983) *J. Clin. Invest. 72*, 1–12.
Ruoslahti, E., & Pierschbacher, M. D. (1987) *Science 238*, 491–497.
Sanger, F., Micklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.
Shelton-Inloes, B. B., Titani, K., & Sadler, J. E. (1986) *Biochemistry 25*, 3164–3171.
Tabor, S., & Richardson, C. C. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 4767–4771.
Takada, Y., & Hemler, M. E. (1989) *J. Cell Biol. 109*, 394–407.
Titani, K., Kumar, S., Takio, K., Ericsson, L. H., Wade, R. D., Ashida, K., Walsh, K. A., Chopek, M. W., Sadler, J. E., & Fujikawa, K. (1986) *Biochemistry 25*, 3171–3184.
Towbin, H., & Martin, G. R. (1982) *Proc. Natl. Acad. Sci. U.S.A. 76*, 4350–4354.
Trüeb, B., & Winterhalter, K. H. (1986) *EMBO J. 5*, 2815–2819.
Trüeb, B., Schreier, T., Bruckner, P., & Winterhalter, K. H. (1987) *Eur. J. Biochem. 166*, 699–703.
von der Mark, H., Aumailley, M., Wick, G., Fleishmajer, R., & Timpl, R. (1984) *Eur. J. Biochem. 142*, 493–502.
Wayner, E. A., & Carter, W. G. (1987) *J. Cell Biol. 105*, 1873–1884.
Wayner, E. A., Carter, W. G., Piotrowicz, R. S., & Kunicki, T. J. (1988) *J. Cell Biol. 107*, 1881–1891.
Wu, J., Eyre, D. R., & Slayter, H. S. (1987) *Biochem. J. 248*, 373–381.