Multiple Forms of Chicken α3(VI) Collagen Chain Generated by Alternative Splicing in Type A Repeated Domains

Roberto Doliana, Paolo Bonaldo, and Alfonso Colombatti

Divisione di Oncologia Sperimentale 2 Centro Di Riferimento Oncologico, 33081 Aviano, Italy

Abstract. Type VI collagen is a structurally unique component widely distributed in connective tissues. Its molecular structure consists of monomers that have the potential to assemble intracellularly into dimers and tetramers which, once secreted, can form microfilaments by end-to-end association. Individual monomers are composed of chains of M_r = ~140,000 (α1 and α2) and >300,000 (α3). Type VI collagen molecules contain a short triple helix with large globular domains at both ends. These domains are made for their greatest part of repetitive units similar to type A repeats of von Willebrand Factor. The α 3(VI) chain, contributing most of the mass of the NH_2-terminal globule, appeared heterogenous both at the mRNA and protein level. Several α3(VI)-specific clones that lack the sequences corresponding to repeats A8 and A6 were isolated from a chicken aorta cDNA library. Northern blot hybridization of poly (A+)-enriched RNA from chicken gizzard with cDNA fragments corresponding to several individual type A repeats showed that A8- and A6-specific probes did not hybridize to the lower M_r transcripts. Clones spanning ~20 kb of the 5'-end of the α 3(VI) gene were isolated from a chicken genomic library and subjected to analysis by restriction mapping, Southern blotting, and selective sequencing of the intron–exon boundaries. At the most 5'-end of the gene an additional type A repeat (A9), previously undetected in cDNA clones, was identified. Furthermore, it was determined that the presumed signal peptide and repeats A9 through A6 are encoded within individual exons. Reverse transcription and polymerase chain reaction of aorta RNA suggested that a mechanism of alternative mRNA splicing by a phenomenon of exon skipping generates α3(VI) isoform variants that contain different numbers of type A repeats. Immunohistochemistry of frozen sections of chicken embryo tissues with repeat-specific mAbs showed that an antibody directed against a conditional exon has a more restricted tissue distribution compared to an antibody against a constitutive exon.

TYPE VI, one of the major collagens of connective tissues, is a component of 100-nm-long periodic microfilaments that are found at the surface of cells and around or between collagen fibers (von der Mark et al., 1984; Bruns, 1984; Bruns et al., 1986; Keene et al., 1988). The widespread occurrence of these thin fibrils in embryo (Bruns et al., 1986) and adult tissues (von der Mark et al., 1984; Keene et al., 1988) and the diversity in localization, ranging from cartilage to soft tissues (Burgeson, 1988), are characteristic features of this collagen. The molecular mechanisms of microfilament formation are presently unknown but electron microscopic (Furthmayr et al., 1983) and biosynthetic studies (Engvall et al., 1986; Colombatti et al., 1987; Colombatti and Bonaldo, 1987) have provided evidence that the polymerization process takes place intracellularly soon after synthesis and leads to the formation of disulfide-bonded dimers and tetramers. Furthermore, the individual chains do not seem to undergo proteolytic processing with removal of the large N- and C- propeptides that do not represent precursor structures. The tetramers associate extracellularly by end-to-end to form the oligomeric microfilaments (Furthmayr et al., 1983).

Recently, we (Bonaldo et al., 1989, 1990) and others (Koller et al., 1989; Chu et al., 1989) provided evidence that a major portion of the constituent chains of chicken and human type VI collagen consists of repeating units of ~200 residues that are closely related to the type A repeats of von Willebrand Factor (Shelton-Inloes et al., 1986). The most distinctive feature that emerged from the analysis of these sequences was the finding that ~85% of the α3 (VI) chain is represented by two types of similar repeating motifs, designated domains A and A' (Bonaldo et al., 1990). In a previous study Engvall et al. (1986) described the heterogeneity of the α3 (VI) chain present as three or more closely spaced bands in SDS-PAGE. The possibility was put forward that this heterogeneity was the consequence of posttranslational events. Similar discrete bands were detected by us after a short 7-min pulse (Colombatti et al., 1987) and even after immunoprecipitation of tunicamycin and α,α'-dipyridyl–treated chicken embryo cells (Colombatti and Bonaldo, 1987).

Moreover, by hybridization of mRNA obtained from human cell lines with α3 (VI)-specific cDNA probes multiple messages were detected (Chu et al., 1987). It appears more likely then that alternative splicing of mRNA is generating protein diversity through multiple forms of α3 (VI) transcripts.

We report here that alternative splicing in the chicken α3 (VI) gene generates several mRNAs that differ by one or more type A repeated domains. As a result of this mechanism different α3(VI) polypeptides are produced that have a specific tissue distribution and may be important in tissue-specific functions.

Materials and Methods

Isolation of cDNA Clones

The construction of a chicken aorta cDNA library in the expression vector pEX1 (Bressan et al., 1987) and the isolation of several cDNA clones encoding the α3(VI) chain have previously been described (Bonaldo and Colombatti, 1989; Bonaldo et al., 1990). A 538-bp-long Pst I restriction fragment from the most 5'-end clone pB10 was purified, nick translated to a specific activity of 7×10^5 cpm/ng, and used to rescreen the cDNA library.

Northern Blotting

Total RNA and poly(A+)-enriched RNA were prepared from chicken gizzard using standard procedures (Maniatis et al., 1982). Electrophoresis of the RNA was performed on 0.7% (wt/vol) agarose gel containing 2.3 M formaldehyde in MOPS buffer for 8 h at 150 V using 20-cm-long plates. RNA was then transferred onto nitrocellulose filters and hybridized with [α-^{32}P] CTP-labeled cDNA probes derived from clone pB10 and specific for different type A repeats (see Figs. 1 and 2).

The filters were hybridized at 68°C overnight in 6× SSC and 10× Denhardt's solution. After washing in 0.2× SSC and 0.1% SDS at room temperature the filters were exposed to β-max Hyperfilms (Amersham International, Amersham, UK).

Isolation of Genomic Clones

A chicken genomic library in EMBL-3 (Clontech Laboratories, Inc., Palo Alto, CA) was plated and the plaques transferred to nitrocellulose filters. The filters were hybridized with synthetic oligonucleotide probes prepared in a DNA synthesizer (Applied Biosystems, Inc., Foster City, CA) and the 5'-end was labeled with [γ-^{32P}]ATP (Amersham International) and T4 polynucleotide kinase (Boehringer Mannheim, GmbH, FRG). The synthetic oligonucleotides were derived from the cDNA sequences encoding the α3(VI) signal peptide sequence and the repeat A8 (see Fig. 4). Four clones were isolated and one (λ gen 5) was further studied and is reported here.

Restriction Enzyme Mapping and DNA Sequence Analysis

Plasmid DNA and lambda phage DNA were isolated by standard procedures (Maniatis et al., 1982). Restriction enzyme digestions were performed as described by the manufacturers. Phage DNA fragments were separated by electrophoresis on 0.7% agarose gels, transferred to nitrocellulose, and hybridized with synthetic oligonucleotides specific for the α3(VI) cDNA clone pB10. Positive fragments were subcloned into the M13-derived vectors, mp18 and mp19 (Messing, 1983), and the nucleotide sequence was obtained by the dideoxy chain termination method (Sanger et al., 1977) as modified by Biggin et al. (1983) using modified bacteriophage T7 DNA polymerase (Tabor and Richardson, 1987). Some sequences were determined directly on caesium chloride-purified lambda DNA using synthetic oligonucleotide primers and Taq DNA polymerase (Promega Biotec, Madison, WI).

Reverse Transcription/Polymerase Chain Reaction

Reverse transcription (RT)[1]/polymerase chain reaction (PCR) was slightly modified from the method described by Rappolee et al. (1988). Total RNA (0.8 μg) was heated at 95°C for 5 min and quickly cooled on ice. The reaction (20 μl of PCR buffer: 50 mM KCl, 10 mM Tris-HCl, pH 8.3, 1.5 mM $MgCl_2$, 0.01% [wt/vol] gelatin) contained 20 U of AMV reverse transcriptase (Promega Biotec), 1 mM dNTPs (each), 20 U RNasin (Promega Biotec), and 50 pmol of α3(VI)-specific oligonucleotide antisense primer. The reaction mixture was incubated for 10 min at room temperature, 60 min at 42°C, 5–10 min at 95°C, and then chilled on ice. The resulting cDNA was amplified by using the DNA amplification reagent kit (Perkin-Elmer/Cetus, Norwalk, CT). 2.5 U of Thermus aquaticus (Taq) polymerase and 50 pmol of α3(VI)-specific oligonucleotide sense primer were added and the reaction was carried out through 40 cycles of amplification. Aliquots of the PCR mixture were electrophoretically separated in agarose gel and were visualized with ethidium bromide staining. The oligonucleotides used and their position within the sequence are the following: sense primers A (nucleotides 267–283), F (947–970), and D (2022–2045); antisense primers B (1200–1229), C (1788–1817), G (2388–2417), and E (2956–2985).

Immunoperoxidase Staining

Two α3(VI) chain-specific mAbs were selected according to their reactivity with hybrid proteins. In brief, hybrid proteins, obtained from lysates of *E. coli* transformed with different cDNA clones and grown at 42°C as detailed elsewhere (Bonaldo et al., 1987), were plated onto polystyrene microtiter plates. mAbs were then assayed for their binding activity for the different hybrid proteins by an ELISA type of assay. Antibody 111C10, that recognized only the pB10 protein and mapped in the spliced repeat A8, and antibody 111A3, that mapped in a constitutive region of the α3(VI) chain, were then selected and used for immunoperoxidase staining.

Tissues from 15-d-old chicken embryos were quickly dissected, embedded in OTC (Miles Laboratories Inc., Naperville, IL), and snap-frozen in liquid nitrogen. Sections (5–8 μm) were cut, air dried, and fixed for 5 min in a 1:1 acetone/chloroform solution. Specimens were rehydrated with PBS, and after incubation with normal horse serum (1:50 dilution), the sections were incubated with the primary antibody (10–20 μg/ml) for 30 min at room temperature, followed by biotin-labeled second antibody (1:200 dilution), 30 min at room temperature, and finally the avidin-biotin complex (ABC, kits PK-4001 and PK-4002; Vector Labs, Burlingame, CA) was applied for 45 min at room temperature. Brown staining was produced by 5-min treatment with 3-3'diaminobenzidine (50 mg in 100 ml of PBS, pH 7.4, containing 0.01% hydrogen peroxide and 10 mM imidazole). Specimens were counterstained with Mayer's hematoxylin. Negative controls were performed by treating sections with an antiricin mAb.

Results

Isolation of cDNA Clones

We reported previously most of the sequence of the chicken α3(VI) chain deduced from several overlapping cDNA clones (Bonaldo and Colombatti, 1989; Bonaldo et al., 1990). The missing upstream sequences were obtained from the same library by screening with a 538-bp-long Pst I restriction fragment of the most 5'-end clone pB10 (Bonaldo et al., 1990). Several positive clones (pB101–pB112) were isolated, purified, and characterized by restriction enzyme analysis and DNA sequencing.

Nucleotide and Amino Acid Sequences of cDNA Clones

Five clones have additional sequences that were absent from clone pB10, whereas one clone (pB112) overlaps over all its sequence with clone pB10 (Bonaldo et al., 1990). The addi-

1. *Abbreviations used in this paper*: PCR, polymerase chain reaction; RT, reverse transcription.

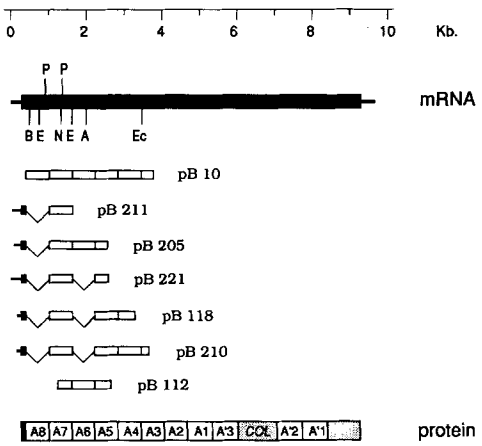


Figure 1. Diagram of chicken α3(VI) collagen cDNA clones. Schematic representation of chicken α 3(VI) cDNA clones approximately to scale beginning with the most upstream clone pB221. In the diagram of the mRNA at the top a partial restriction map is shown. Solid lines indicate the nontranslated sequences. The closed box indicates the sequences coding for the putative signal peptide; ⌣ indicates contiguity. The diagram of the protein at the bottom follows the designations reported previously (Bonaldo et al., 1990) where A and A' indicate type A repeats and COL the triple helix. Domains unrelated to type A repeats are shown as shaded boxes. *A*, Ava I; *B*, Ban II; *E*, Eco RV; *Ec*, Eco RI; *N*, Nco I; *P*, Pst I.

tional sequences contain 5'-untranslated regions of different lengths followed by a short sequence resembling a signal peptide (see below). Following the presumed signal peptide all five clones overlap with clone pB10 and they extend into different type A repeats. Surprisingly, all five clones lack the sequences corresponding to repeat A8 and three clones lack also the sequences corresponding to repeat A6 (Fig. 1).

Northern Blot Analysis

To determine whether the differences in the cDNA clones reflected differences in the respective mRNAs and not cloning artifacts, poly(A+)-enriched RNA from chicken gizzard was examined by Northern blot hybridization under stringent conditions with cDNA restriction fragments specific for sequences found in repeats A8, A7, A6, and A3 (Fig. 1). As seen in Fig. 2, all probes hybridized with a complex pattern of closely spaced multiple bands in the range of α3(VI) transcripts (9–10 kb) (Bonaldo et al., 1990). The uppermost band was poorly resolved and resulted in a broader signal with all these cDNA probes, probably signifying more than one mRNA species. However, the probes encompassing part of repeats A8 or A6 did not hybridize to the lower M_r message. Ethidium bromide staining of the agarose RNA gel revealed equivalent amounts of intact rRNA in all the samples. The presence of several transcripts hybridizing with α3(VI) probes suggests that multiple α3(VI) RNA species may arise by a mechanism of alternative splicing.

Genomic Clones Reveal an Additional Type A Repeat and Show Individual Exons Coding for Repeats A9 through A6

To clarify the genetic basis for the mRNA variants, the intron/exon structure of the 5'-end of chicken α3(VI) gene was investigated by screening a chicken genomic library with synthetic oligonucleotide probes specific for sequences found in the presumed signal peptide and repeat A8 (oligonucleotide I and oligonucleotide III, see Fig. 4). Overlapping genomic clones spanning a total of ~20 kb were isolated and a partial restriction map was constructed. A more detailed analysis was performed for a 14-kb-long clone (λ gen 5). Restriction fragments were isolated from this clone by hybridization to oligonucleotide probes specific for the cDNA clone pB10 (oligonucleotides I–V, see Fig. 4). By a combination of restriction mapping, Southern analysis, and selective sequencing, the exon structure and the intron–exon boundaries of the 5'-terminal part of the α3(VI) gene were deduced (Fig. 3). Five exons were found, four of which code exactly for one type A repeat each (Fig. 3 *A*). The precise intron–exon boundaries were determined and the splice donor and acceptor sequences are shown in Fig. 3 *B*. Each donor and acceptor site is conventional and is in good agreement with the standard consensus motifs (Padgett et al., 1986; Krainer and Maniatis, 1988). All splice junctions are in frame and introns lie between the first and second nucleotide of a codon (phase 1 introns) (Sharp, 1981). Fig. 4 reports a composite nucleotide sequence and deduced amino acid sequence derived from the different cDNA clones and the genomic clone λ gen 5. The sequence starts with a short (266 bp) 5'-untranslated region followed by a sequence that codes for 25 amino acids characteristic of a signal peptide (van Heijne, 1986).

The NH_2-terminal sequence of α3(VI) is not known, therefore, we assume from the deduced sequence that the mature protein initiates with a glutamine as has been reported both for the chicken (Koller et al., 1989; Bonaldo et al., 1989) and the human (Chu et al., 1989) α1(VI) and α2(VI) chains. Restriction mapping, subcloning, and sequencing of the λ gen 5 genomic clone showed the existence of an additional open reading frame of 625 bp upstream to the sequences that completely matched with those of the re-

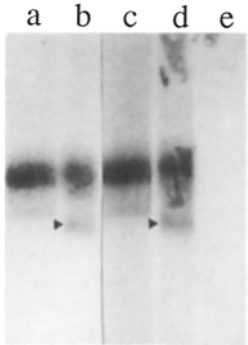
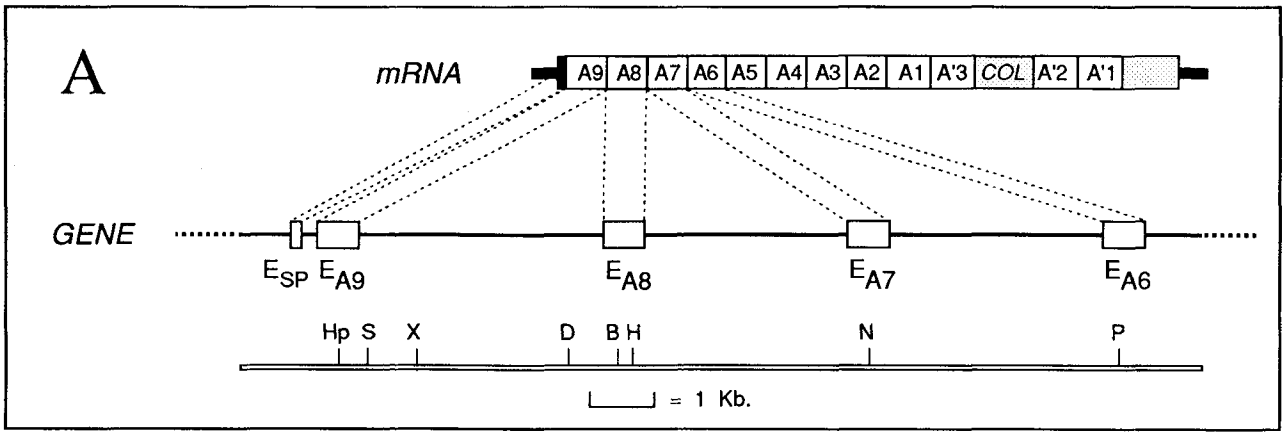


Figure 2. Northern blot analysis of chicken gizzard showing α3(VI) collagen-specific mRNA bands. Each lane was loaded with 7 μg of poly (A+)-enriched RNA. Individual strips were hybridized to [α-^{32}P] dCTP-labeled cDNA fragments derived from clone pB10 (Bonaldo et al., 1990) and containing sequences specific for different type A repeats. (lane *a*) A8-specific probe (255-bp-long Ban II–Eco RV fragment); (lane *b*) A7-specific probe (253-bp-long Nco I–Eco RV fragment); (lane *c*) A6-specific probe (383-bp-long Eco RV–Ava I fragment); (lane *d*) A3-specific probe (318-bp-long Eco RI–Bam HI fragment); (lane *e*) unrelated probe (pEX1 vector). The arrowheads indicate the migration of the mRNA that is not detected by A8 and A6 probes. On the right the migration of molecular weight markers is indicated in kb. Only the relevant part of the autoradiogram is shown.



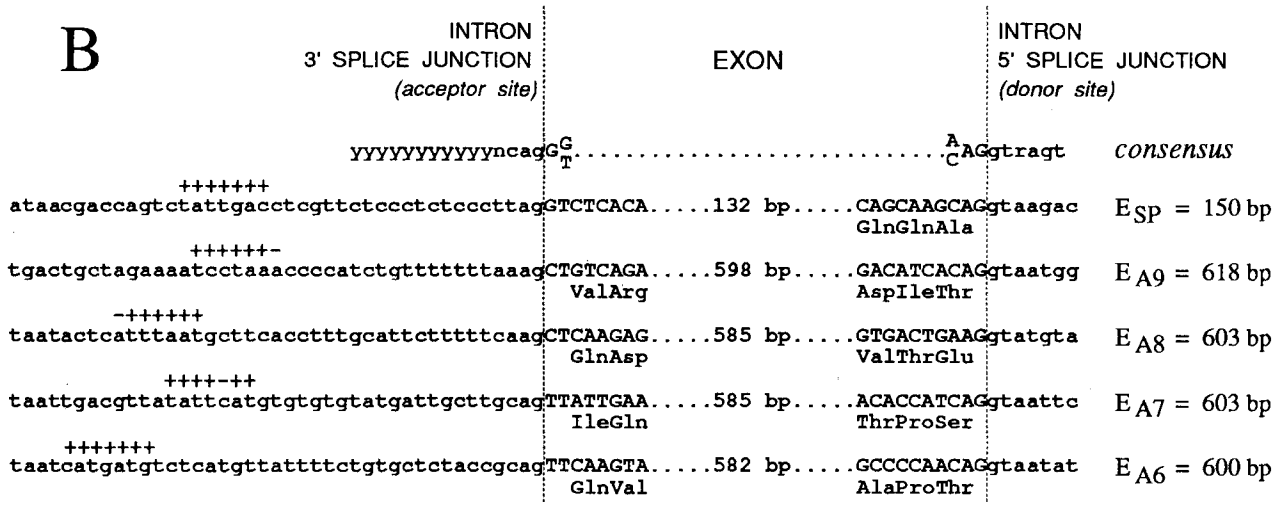


Figure 3. Sequences and physical map of the 5'-end of the chicken α3(VI) collagen gene. (*A*) Diagram of the physical map and partial restriction sites of clone λ gen 5. Exons are indicated by open rectangles and introns by thick lines. *B*, Bgl I; *D*, Dra I; *H*, Hind III; *Hp*, Hpa I; *N*, Nco I; *P*, Pvu II; *S*, Sma I; *X*, Xba I. (*B*) Nucleotide sequences at the exon–intron boundaries. The splice junctions of five introns are aligned and compared to the splice consensus sequences for eukaryotic genes. Intron sequences are indicated by lowercase and exon sequences by capital letters. Deduced amino acids are indicated by the three letter code. The lariat branchpoint consensus sequence *ynytray* (Padgett et al., 1986) is shown by + and −. At the right the definition of the exons and their length is shown. *r*, purine; *y*, pyrimidine; *n*, purine or pyrimidine.

peat A8. Comparison of the deduced amino acid sequences with the sequences of the eight type A repeats of α3(VI) previously identified (Bonaldo et al., 1990) revealed that this open reading frame is an exon precisely encoding a full type A repeat. Except for a few residues present in clone pB10 (Bonaldo et al., 1990) this repeat was not detected in any of our previous cDNA clones. With the addition of this extra type A repeat (A9), residue 1 of clone pB10 in our previous report (Bonaldo et al., 1990) becomes residue 199. Fig. 4 also shows that repeats A9, A8, A7, and A6 are encoded within single exons (E_{A9}, E_{A8}, E_{A7}, and E_{A6}), whereas the presumed signal peptide is encoded together with 62 bp of 5'-untranslated mRNA sequence by a separate exon (E_{SP}).

α3(VI) mRNA Heterogeneity Is Due to Exon Skipping of Type A Repeats

In view of the heterogeneity of the α3(VI) mRNAs, the selective hybridization with the cDNA probe specific for different repeats, the isolation of cDNA clones lacking individual type A repeats, and the demonstration that the repeats from A9 through A6 are encoded within single exons, we applied the RT/PCR amplification assay to analyze further this complex transcription unit. Using this approach together with α3(VI)-specific primers we examined the region of probable isoform variation comprised between the signal peptide and repeat A5 (Fig. 5).

Evidence both for spliced and unspliced transcripts was obtained. Amplified fragments from transcripts missing E_{A9} and E_{A8} were detected using the sense primer A in the signal peptide and the downstream antisense primers B and C in the repeats A8 and A7, respectively. Similarly, using the sense primer D in the repeat A7 and the antisense primer E in the repeat A5, an amplified fragment of 364 bp missing E_{A6} was detected (Fig. 5 *B*, *left side*). Our assay conditions favor the amplification of short sequences, consequently higher M_r cDNA including the spliced exons are not visible

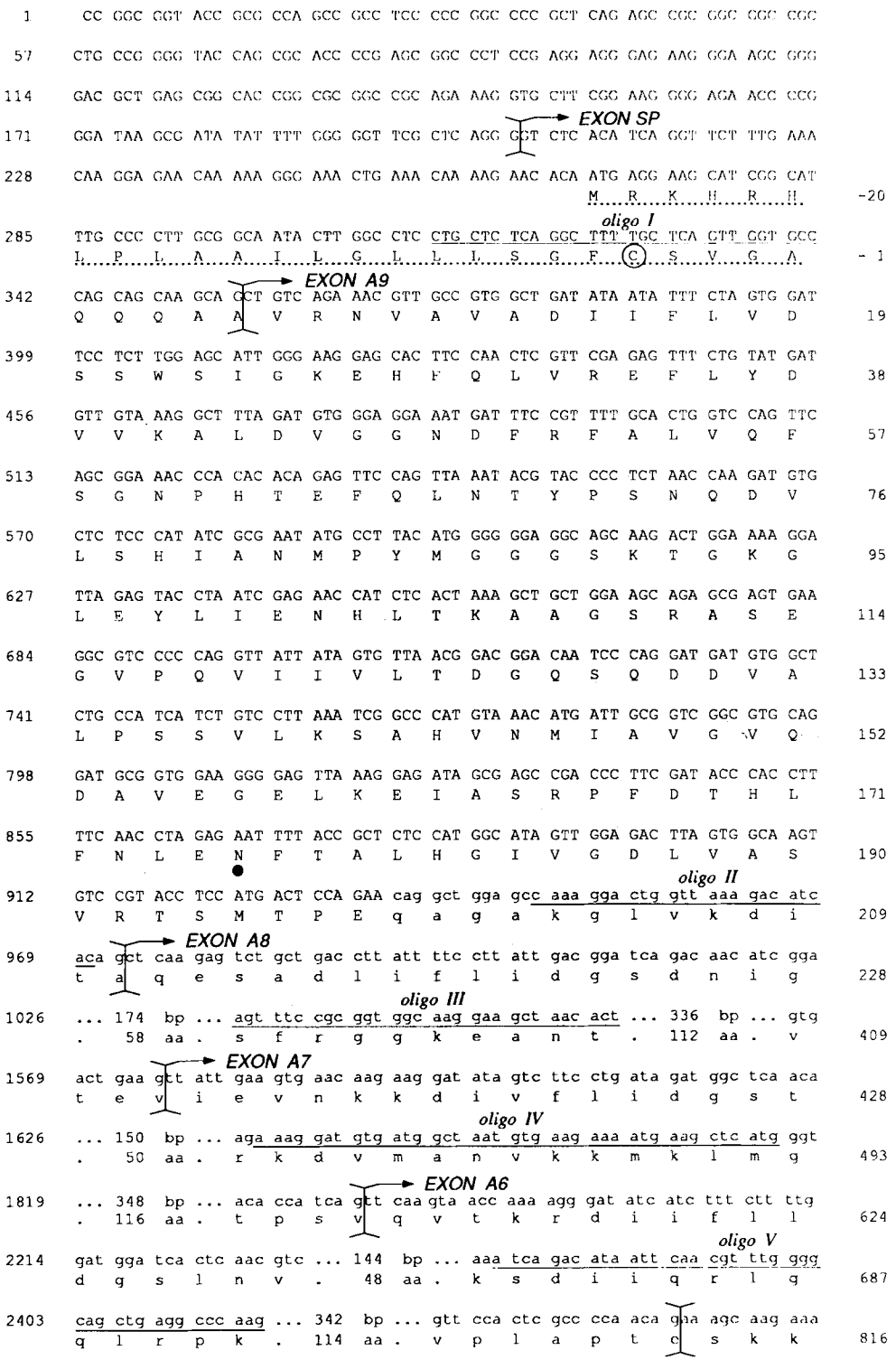


Figure 4. Composite nucleotide sequence of the 5′ end of α3(VI) collagen gene. The sequence is derived in part from cDNA clones pB211, pB205, pB221, pB118, and pB210 (see Fig. 1) and in part (positions 352–935) from the genomic clone λ gen 5. First line, nucleotide sequence; second line, deduced amino acid sequence (one letter code). Nucleotides and amino acids are numbered on the left and right-hand sides, respectively. Lower case letters indicate the nucleotide and deduced amino acid sequence already reported (Bonaldo et al., 1990). The putative signal peptide sequence of chicken α 3(VI) collagen is underlined by a dotted line. Vertical lines separate the different exons. The amino acid position 199 marks the start of clone pB10 and corresponds to position 1 in the previous report. One potential N-attachment site for oligosaccharide is marked by a dot, and a cysteine residue is circled. Synthetic oligonucleotides used to screen the EMBL-3 genomic library (*oligo I* and *III*) and to map the exons (*oligo II*, *IV*, and *V*) are underlined. These sequence data are available from EMBL/GenBank/DDBJ under accession number M24282.

in our gel and the lower efficiency of synthesis of larger transcripts is not informative of the different α3(VI) mRNAs relative abundance. However, amplified fragments from transcripts containing E_{A9}, E_{A8}, and E_{A6} were detected using the sense primers F and D in the repeats A7 and A9 and the antisense primers B and G in the repeats A8 and A6 (Fig. 5 *B*, *right side*). All the amplified fragments seem genuine since their size is in accord with the size expected from the primary sequence.

Negative controls in which amplification was performed omitting the reverse transcriptase or the RNA ensured that amplification arose directly from α3(VI) mRNA rather than

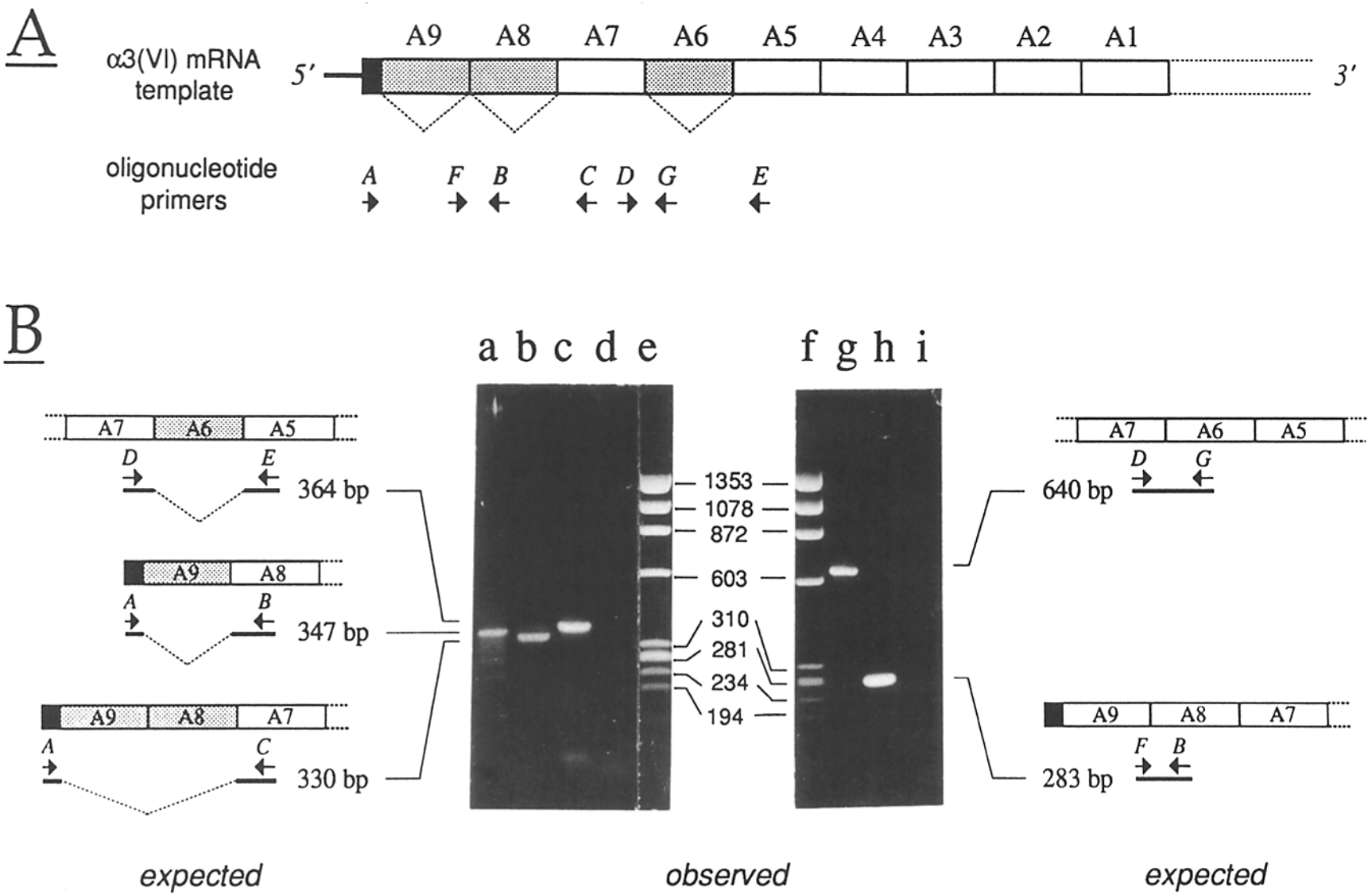


Figure 5. Detection of mRNA transcripts coding for the α3(VI) isoforms by the RT/PCR method. (*A*) Linear model of the α3(VI) mRNA showing the position of the alternatively expressed sequences (*shaded boxes*) and of the synthetic oligonucleotide primers used in the amplifications. (*B*) Amplification showing the expected and observed DNA fragments obtained with the following combinations of oligonucleotides: primers A and B (lanes *a* and *d*); primers A and C (lane *b*); primers D and E (lane *c*); primers D and G (lanes *g* and *i*); primers F and B (lane *h*). In negative control reaction mixtures the reverse transcriptase (lane *d*) or the template RNA (lane *i*) were omitted. A Hae III digest of φX174 DNA was used as standard M_r marker (lanes *e* and *f*).

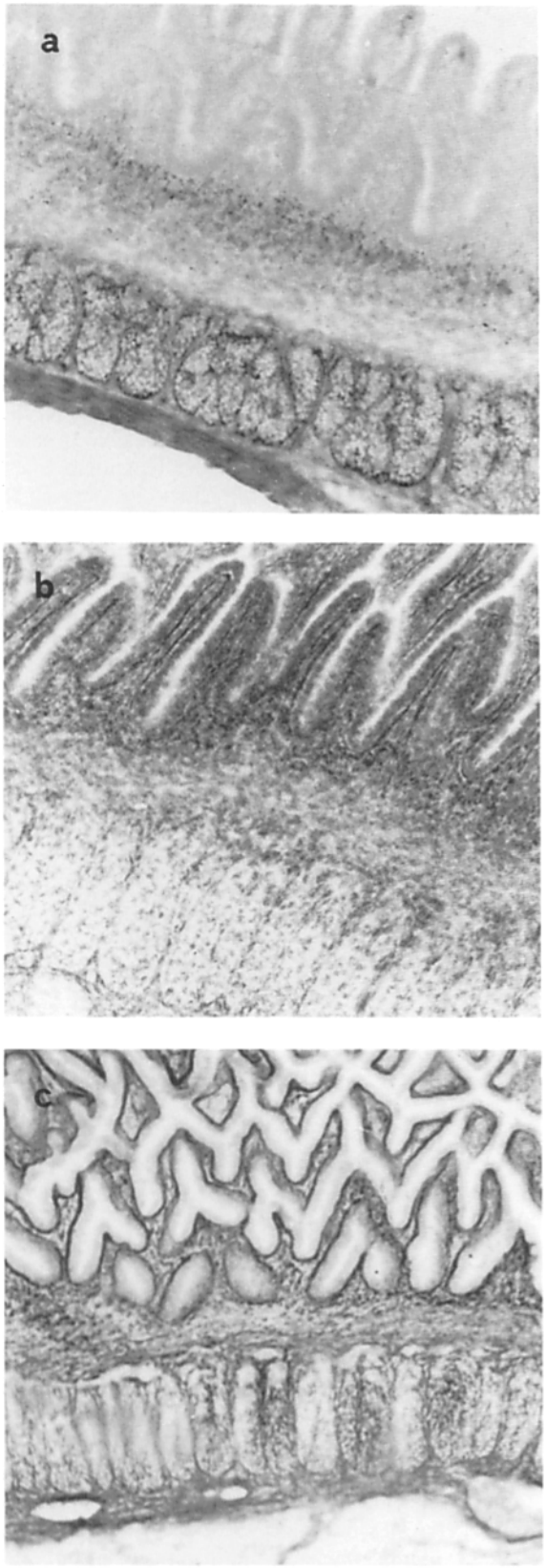


Figure 6. Distribution of α3(VI) collagen isoforms in chicken intestine by indirect immunoperoxidase staining. (*a*), mAb 111C10; (*b*) mAb 111A3; and (*c*) polyclonal antibody to type VI collagen. mAb 111C10 does not react with the mucosa and submucosa whereas mAb 111A3 shows a very strong reaction.

from minute amounts of contaminating plasmid cDNA since no DNA bands were detected in this case (Fig. 5 *B*, lanes *d* and *h*). The size of the amplified fragments excluded also the possibility that amplification arose from unprocessed RNA or genomic DNA which might contaminate the reaction.

Distribution of α3(VI) Isoforms in Chicken Embryo Tissues

The above observations on the possible existence of isoforms of the α3(VI) chain containing a variable number of type A repeats together with previous studies that showed heterogeneity of the α3(VI) polypeptides in SDS-PAGE (Engvall et al., 1986; Colombatti et al., 1987, 1989) suggest that there might be some specific function provided by the individual repeats. A first step toward the understanding of the biological meaning of these isoforms is the study of the tissue distribution. Using an mAb (111C10, to be reported elsewhere) specific for repeat A8 and an mAb (111A3, Colombatti et al., 1988) specific for constitutive sequences, we performed indirect immunoperoxidase staining of frozen sections to analyze the pattern of reactivity of embryo tissues. The mAb 111A3 showed a strong positive reaction with all tissues examined (Figs. 6 and 7). On the other hand, mAb 111C10, which recognizes α3(VI) molecules containing the repeat A8, showed a strong reactivity with the extracellular matrix of the intestinal muscular layers, whereas the reactivity with the mucosa and submucosa was very weak or absent (Fig. 6). In several other tissues the reactivity of the two mAbs was superimposable (data not shown).

The observed restriction in the expression of A8-positive α3(VI) isoforms could result either from a cell- or tissue-specific mRNA splicing difference or from a quantitative difference in the level of all variants of α3(VI) mRNA so that an undetectable level of A8-positive mRNA would become sufficient in some sites to produce enough protein to give a

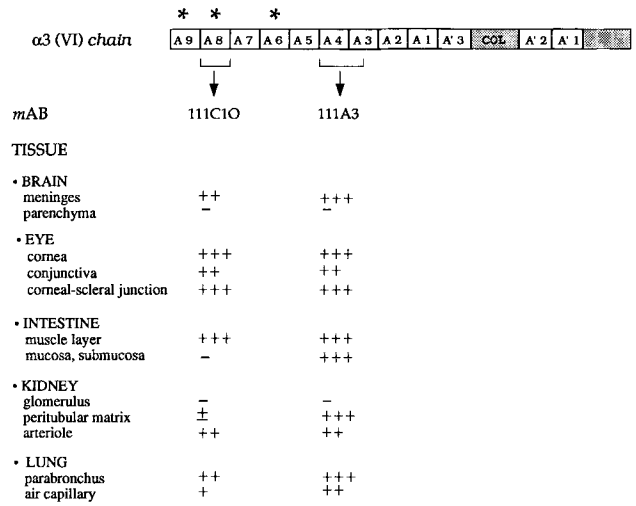


Figure 7. Distribution of mAbs reactivity in chicken embryo tissues. The localization of mAb 111C10 and 111A3 epitopes within the sequence of the α3(VI) polypeptide chain is shown in the diagram at the top. Asterisks indicate the position of the alternatively spliced exons in the mRNA. Reactivity: (+++) very strong, (++) strong, (+) weak, (±) very weak, (−) lack of reactivity.

positive immunoperoxidase signal. To exclude this latter possibility we incubated adjacent sections with several dilutions of mAb 111A3 and 111C10. In these experiments the distinct pattern of the reactivity of the two mAbs was still unchanged.

Discussion

We have used cDNA and genomic clones to study exon/intron organization of the 5'-end of chicken α3(VI) collagen. The demonstration that the pattern of multiple mRNAs of this chain is at least in part the result of a mechanism of multiple alternative splicing of exons encoding type A repeats is a major finding of this study. In addition, the exon structure suggests that the α3 chain of type VI collagen evolved by multiple processes of gene shuffling and amplification. Alternative splicing of α3(VI) gene transcripts was first suggested by the observation that several cDNA clones lacked one or more sequences of ~600 bp coding for individual type A repeats that were present in other cDNA clones from the same library. The α3(VI) gene constitutes a complex transcription unit and the size of the transcripts (~10 kb) does not allow a fine resolution of the different messages. Nevertheless, by hybridization of Northern blots with type A-specific cDNA probes, the presence of variant mRNAs that differed in size was initially demonstrated in this study. The finding that individual type A repeats are encoded within single exons and the possible correspondence between the size difference of the various mRNAs and the size of type A-coding exons were highly suggestive that the different transcripts might be the result of an alternative splicing mechanism involving type A repeats. Furthermore, the pattern of hybridization of the Northern blot with the type A-specific cDNA probes corresponding to the spliced A8 and A6 domains suggests that certain mRNA molecules not only exclude both the exons but might involve skipping of additional exons (A9 and maybe other exons). Given the high M_r of the mRNA and the nearly identical size of the different type A repeats it is conceivable that each band of the Northern blot represents a mixture of comigrating mRNA species that have skipped different exons. The RT/PCR amplification analysis, using various primers specific for the signal sequence and for presumed constitutive exons, yielded fragments with sizes expected if mRNA isoforms missing either one or at least two type A-encoding exons were expressed. Appropriate controls excluded the possibility that cDNA clones missing individual repeats could contaminate the reaction mixture and serve as templates. Evidence for unspliced fragments also was obtained by the RT/PCR amplification. It is conceivable that the ratio between the various α3(VI) RNA transcripts may change depending on specific sites or physiological and pathological conditions as has been already shown for the different isoforms of another extracellular matrix glycoprotein, namely fibronectin (Zardi et al., 1987; ffrench-Constant and Hynes 1988, 1989; ffrench-Constant et al., 1989; Carnemolla et al., 1989).

Alternative splicing is an important mechanism of gene regulation and it is well documented for several proteins (Andreadis et al., 1987), including the extracellular matrix constituents fibronectin (Kornblihtt et al., 1984; Schwarzbauer et al., 1987; Gutman and Kornblihtt, 1987), elastin (Indik et al., 1987), tenascin/cytotactin (Jones et al., 1989; Gulcher et al., 1989), and link protein (Rhodes et al., 1988). Among collagens, apart from the α3(VI), there is evidence that transcripts of the human α2(VI) (Chu et al., 1989), of the human α1(XIII) (Tikka et al., 1988), and of the chicken α1(IX) (Nishimura et al., 1989) and α2(I) (Bennett et al., 1989), undergo alternative splicing. The use of alternative promoters by the α1(IX) collagen gene results in protein products with different sequence domains and specific tissue distribution. Regarding the other collagen genes it is not known at the moment whether the different mRNAs are translated in different proteins. At least for the α2(I) this is not the case. Little is presently known about the mechanisms involved in the determination and regulation of the alternative splicing, mainly because of the lack of suitable in vitro systems that preserve cell-specific features (Padgett et al., 1986). Only in few instances has it been possible to study the expression and processing of cell-specific splicing pathways by transfecting different cell lines as shown for fibronectin (Baron et al., 1989). In this case it was demonstrated that all the information necessary to induce tissue-specific alternative splicing is in *cis* respective to the exons undergoing splicing and that *trans*-acting factors differentially expressed in the various cell lines confer the tissue-specific expression. It was beyond the purpose of the present study to obtain information on the intron sequences further upstream and downstream from the exon–intron boundaries, but from the short intron sequences available and the limited number of repeats analyzed (three conditional A9, A8, and A6 and two constitutive A7 and A5) we could not identify specific sequences that would distinguish alternatively spliced exons from constitutive exons.

Splicing of A9, A8, and A6 exons of α3(VI) is similar to the optional skipping of ED-A and ED-B exons of fibronectin (Kornblihtt et al., 1984; Gutman and Kornblihtt, 1987; Schwarzbauer et al., 1987). Through this mechanism several functionally appropriate α3(VI) polypeptides can be generated that have a different number of type A repeats. Indeed, heterogeneity at the protein level had been reported previously for the α3(VI) chain both in in vitro biosynthetic studies (Engvall et al., 1986; Colombatti et al., 1987; Colombatti and Bonaldo, 1987) and in vivo (Jander et al., 1984; Wu et al., 1987; Colombatti et al., 1989). There is no direct evidence at the moment to relate the different mRNAs to the different polypeptides, although it is tempting to speculate that the ladder of multiple polypeptides derives from messages that have skipped one or more exons coding for complete type A repeats. The finding that an mAb with specificity for conditional exons has a different tissue distribution than an mAb with specificity for constitutive exons is consistent with this notion and with the α3(VI) polypeptide heterogeneity detected in tissue extracts (Colombatti et al., 1989). Given the fact that A8-specific mAb 111C10 detects a single epitope, it is in principle still possible that the lack of reactivity of the intestinal mucosa and submucosa with this mAb might not be the result of the local synthesis of type VI molecules with α3 chains devoid of repeat A8, but only the consequence of epitope masking and/or interaction with different constituents of the extracellular matrix.

We have reported that type VI collagen and recombinant fusion proteins of the NH_2-terminal portion of α3(VI) chain have the potential to interact under physiological conditions in vitro with type I collagen fibrils (Bonaldo et al.,

1990; Russo, V., A. Appi, and A. Colombatti, manuscript in preparation). Given the widespread distribution of type VI collagen and its close vicinity to the cells, it seems reasonable to imply that the presence of multiple type A repeats modulates the interaction of type VI collagen with type I collagen and also with other potential ligands in the extracellular matrix and at the cell surface.

We thank Dr. Sandro Poletti for his help with the immunoperoxidase technique and Elisabetta Montagner for typing the manuscript.

This work was supported by a grant from the Associazione Italiana per la Ricerca sul Cancro.

Received for publication 26 February 1990 and in revised form 18 July 1990.

References

Andreadis, A., M. E. Gallego, and B. Nadal-Ginard. 1987. Generation of protein isoform diversity by alternative splicing: mechanistic and biological implications. *Annu. Rev. Cell Biol.* 3:207-242.

Barone, M. V., C. Henchcliffe, F. E. Baralle, and G. Paolella. 1989. Cell type specific trans-acting factors are involved in alternative splicing of human fibronectin pre-mRNA. *EMBO (Eur. Mol. Biol. Organ.) J.* 8:1079-1085.

Bennett, V. D., I. M. Weiss, and S. L. Adams. 1989. Cartilage-specific 5' end of chick alpha-2(I) collagen messenger RNAs. *J. Biol. Chem.* 264:8402-8409.

Biggin, M. D., T. J. Gibson, and G. F. Hong. 1983. Buffer gradient gels and [^{35}S]-label as an aid to rapid DNA sequence determination. *Proc. Natl. Acad Sci. USA.* 80:3963-3965.

Bonaldo, P., and A. Colombatti. 1989. The carboxyl terminus of the chicken α3 chain of collagen VI is a unique mosaic structure with glycoprotein Ib-like, fibronectin type III, and Kunitz modules. *J. Biol. Chem.* 264:20235-20239.

Bonaldo, P., F. Bucciotti, and A. Colombatti. 1987. Isolation of cDNA clones corresponding to the Mr = 150,000 subunit of chick type VI collagen. *Biochem. Biophys Res. Commun.* 149:347-354.

Bonaldo, P., V. Russo, F. Bucciotti, G. M. Bressan, and A. Colombatti. 1989. α1 chain of chick type VI collagen: the complete cDNA sequence reveals a hybrid molecule made of one short collagen and three von Willebrand factor type A-like domains. *J. Biol. Chem.* 264:5575-5580.

Bonaldo, P., V. Russo, F. Bucciotti, R. Doliana, and A. Colombatti. 1990. Structural and functional features of the chicken α3 chain indicate a bridging role for collagen VI in connective tissues. *Biochemistry.* 29:1245-1254.

Bressan, G. M., P. Argos, and K. K. Stanley. 1987. Repeating structure of chick tropoelastin revealed by complementary DNA cloning. *Biochemistry.* 26:1497-1503.

Bruns, R. 1984. Beaded filaments and long-spacing fibrils: relation to type VI collagen. *J. Ultrastruct. Res.* 89:136-145.

Bruns, R., W. Press, E. Engvall, R. Timpl, and J. Gross. 1986. Type VI collagen in extracellular 100-nm periodic filaments and fibrils: identification by immunoelectron microscopy. *J. Cell Biol.* 103:394-404.

Burgeson, R. E. 1988. New collagens, new concepts. *Annu. Rev. Cell Biol.* 4:551-577.

Carnemolla, B., E. Balza, A. Siri, L. Zardi, M. R. Nicotra, A. Bigotti, and P. G. Natali. 1989. A tumor-associated fibronectin isoform generated by alternative splicing of messenger RNA precursors. *J. Cell Biol.* 108:1139-1148.

Chu, M. L., K. Mann, R. Deutzmann, D. Pribula-Conway, C. C. Hsu-Chen, M. P. Bernard, and R. Timpl. 1987. Characterization of three constituent chains of collagen type VI by peptide sequences and cDNA clones. *Eur. J. Biochem.* 168:309-317.

Chu, M. L., T. Pan, D. Conway, H. J. Kuo, R. W. Glanville, R. Timpl, K. Mann, and R. Deutzmann. 1989. Sequence analysis of α1(VI) and α2(VI) chains of human type VI collagen reveals internal triplication of globular domains similar to the A domains of von Willebrand factor and two α2(VI) chain variants that differ in the carboxyl terminus. *EMBO (Eur. Mol. Biol. Organ.) J.* 8:1939-1946.

Colombatti, A., and P. Bonaldo. 1987. Biosynthesis of chick type VI collagen. II. Processing and secretion in fibroblasts and smooth muscle cells. *J. Biol. Chem.* 262:14461-14466.

Colombatti, A., P. Bonaldo, K. Ainger, G. M. Bressan, and D. Volpin. 1987. Biosynthesis of chick type VI collagen. I. Intracellular assembly and molecular structure. *J. Biol. Chem.* 262:14454-14460.

Colombatti, A., K. Ainger, M. T. Mucignat, and P. Bonaldo. 1988. Monoclonal antibodies for the different chains of chick type VI collagen. *Collagen Relat. Res.* 8:331-377.

Colombatti, A., K. Ainger and F. Colizzi. 1989. Type VI collagen: high yields of a molecule with multiple forms of α3 chain from avian and human tissues. *Matrix.* 9:177-185.

Engvall, E., H. Hessle, and G. Klier. 1986. Molecular assembly, secretion, and matrix deposition of type VI collagen. *J. Cell Biol.* 102:703-710.

ffrench-Constant, C., and R. O. Hynes. 1988. Patterns of fibronectin gene expression and splicing during cell migration in chicken embryos. *Development (Camb.).* 104:369-382.

ffrench-Constant, C., and R. O. Hynes. 1989. Alternative splicing of fibronectin is temporally and spatially regulated in the chicken embryo. *Development (Camb.).* 106:375-388.

ffrench-Constant, C., L. Van De Water, H. F. Dvorak, and R. O. Hynes. 1989. Reappearance of an embryonic pattern of fibronectin splicing during wound healing in the adult rat. *J. Cell Biol.* 109:903-914.

Furthmayr, H., H. Wiedemann, R. Timpl, E. Odermatt, and J. Engel. 1983. Electro-microscopical approach to a structural model of intima collagen. *Biochem. J.* 211:303-311.

Gulcher, J. R., D. E. Nies, L. S. Marton, and K. Stefansson. 1989. An alternatively spliced region of the human hexabrachion contains a repeat of potential N-glycosylation sites. *Proc. Natl. Acad. Sci. USA.* 86:1588-1592.

Gutman, A., and A. R. Kornblihtt. 1987. Identification of a third region of cell-specific alternative splicing in human fibronectin mRNA. *Proc. Natl. Acad. Sci. USA.* 84:7179-7182.

Indik, Z., H. Yeh, N. Ornstein-Goldstein, P. Sheppard, N. Anderson, J. C. Rosenbloom, L. Peltonen, and J. Rosenbloom. 1987. Alternative splicing of human elastin mRNA indicated by sequence analysis of cloned genomic and complementary DNA. *Proc. Natl. Acad. Sci. USA.* 84:5680-5684.

Jander, R., D. Troyer, and J. Rauterberg. 1984. A collagen-like glycoprotein of the extracellular matrix is the undegraded form of type VI collagen. *Biochemistry.* 23:3675-3681.

Jones, F. S., S. Hoffman, B. A. Cunningham, and G. M. Edelman. 1989. A detailed structural model of cytotactin: protein homologies, alternative RNA splicing, and binding regions. *Proc. Natl. Acad. Sci. USA.* 86:1905-1909.

Keene, D. R., E. Engvall, and R. W. Glanville. 1988. Ultrastructure of type VI collagen in human skin and cartilage suggests an anchoring function for this filamentous network. *J. Cell Biol.* 107:1995-2006.

Koller, E., K. H. Winterhalter, and B. Trüeb. 1989. The globular domains of type VI collagen are related to the collagen-binding domains of cartilage matrix protein and von Willebrand factor. *EMBO (Eur. Mol. Biol. Organ.) J.* 8:1037-1077.

Kornblihtt, A. R., K. Vibe-Petersen, and F. E. Baralle. 1984. Human fibronectin cell specific alternative mRNA splicing generates polypeptide chains differing in the number of internal repeats. *Nucleic Acids Res.* 12:5853-5868.

Krainer, A. R., and T. Maniatis. 1988. RNA splicing. *In* Transcription and Splicing. B. D. Hames and D. M. Glover, editors. IRL Press, Oxford. 131-143.

Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Press, Cold Spring Harbor, NY. 545 pp.

Messing, J. 1983. New M13 vectors for cloning. *Methods Enzymol.* 101:20-78.

Nishimura, I., Y. Muragaki, and B. R. Olsen. 1989. Tissue-specific forms of type IX collagen-proteoglycan arise from the use of two widely separated promoters. *J. Biol. Chem.* 264:20033-20041.

Padgett, R. A., P. J. Grabowski, M. M. Konarska, S. Seiler, and P. A. Sharp. 1986. Splicing of messenger RNA precursors. *Annu. Rev. Biochem.* 55: 1119-1150.

Rappolee, A. D., D. Mark, M. J. Banda, and Z. Werb. 1988. Wound macrophages express TGF-alpha and other growth factors in vivo: analysis by mRNA phenotyping. *Science (Wash. DC).* 241:708-712.

Rhodes, C., K. Doege, M. Sasaki, and Y. Yamada. 1988. Alternative splicing generated two different mRNA species for rat link protein. *J. Biol. Chem.* 263:6063-6067.

Sanger, F., S. Micklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74:5463-5467.

Schwarzbauer, J. E., R. E. Patel, D. Fonda, and R. O. Hynes. 1987. Multiple sites of alternative splicing in the rat fibronectin gene transcript. *EMBO (Eur. Mol. Biol. Organ.) J.* 6:2573-2580.

Sharp, P. A. 1981. Speculation on RNA splicing. *Cell.* 23:643-646.

Shelton-Inloes, B. B., K. Titani, and J. E. Sadler. 1986. cDNA sequences for human von Willebrand factor reveal five types of repeated domains and five possible protein sequence polymorphisms. *Biochemistry.* 25:3164-3171.

Tabor, S., and C. C. Richardson. 1987. DNA sequence analysis with modified bacteriophage T7 DNA polymerase. *Proc. Natl. Acad. Sci. USA.* 84:4767-4771.

Tikka, L., T. Pihlajaniemi, P. Henttu, D. J. Prockop, and K. Tryggvason. 1988. Gene structure for the α1 chain of a human short-chain collagen (type XIII) with alternatively spliced transcripts and translation termination codon at the 5' end of the last exon. *Proc. Natl. Acad. Sci. USA* 85:7491-7495.

van Heijne, G. 1986. A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res.* 14:4683-4690.

von der Mark, H., M. Aumailley, G. Wick, R. Fleishmajer, and R. Timpl. 1984. Immunochemistry, genuine size and tissue localization of collagen VI. *Eur. J. Biochem.* 142:493-502.

Wu, J. J., D. Eyre, and H. S. Slayter. 1987. Type VI collagen of the intervertebral disc. *Biochem. J.* 248:373-381.

Zardi, L., B. Carnemolla, A. Siri, T. E. Petersen, G. Paolella, G. Sebastio, and F. E. Baralle. 1987. Transformed human cells produce a new fibronectin isoform by preferential alternative splicing of a previously unobserved exon. *EMBO (Eur. Mol. Biol. Organ.) J.* 6:2337-2342.