THE JOURNAL OF BIOLOGICAL CHEMISTRY
© 1995 by The American Society for Biochemistry and Molecular Biology, Inc.
Vol. 270, No. 1, Issue of January 6, pp. 418–424, 1995
Printed in U.S.A.

Structural and Functional Analysis of the Promoter of the Human α1(XI) Collagen Gene*

(Received for publication, June 8, 1994, and in revised form, September 12, 1994)

Hidekatsu Yoshioka‡, Patricia Greenwel§, Kazuhito Inoguchi‡, Sharada Truter§, Yutaka Inagaki§¶, Yoshifumi Ninomiya‡, and Francesco Ramirez§||

From the ‡Department of Molecular Biology and Biochemistry, Okayama University Medical School, Okayama 700, Japan and the §Brookdale Center for Molecular Biology, Mt. Sinai School of Medicine, New York, New York 10029

In order to eventually elucidate the mechanisms regulating α1(XI) collagen expression in cartilaginous and non-cartilaginous tissues, we performed an initial analysis of the structural-functional features of the promoter of the human gene (*COL11A1*). After cloning and sequencing the 5′ portion of *COL11A1*, primer extension and nuclease protection assays identified several minor transcriptional start sites clustered around a major one located 318 base pairs from the ATG codon. Consistent with this finding, analysis of the upstream sequence revealed the absence of a TATA motif and the presence of several GC boxes. Transient transfection experiments delineated the smallest promoter sequence directing relatively high expression of a reporter gene in a cell type-specific manner. Nine nuclear protein-bound areas were located within this promoter sequence of the *COL11A1* gene. Sequence homologies suggested that the majority of the footprints correspond to potential binding sites for ubiquitous nuclear proteins, such as AP2 and Sp1. Additional experimental evidence indicated that one of the protected areas may bind a transcriptional complex that is identical or closely related to the one that regulates tissue specificity in the coordinately expressed α2(V) collagen gene.

The fibrillar group of collagens includes five structurally related trimers which are traditionally divided into major (types I–III) and minor (types V and XI) types, based on their relative abundance (1). Each of the minor collagens is an integral component of the fibrils of one of the major types (2–4). This and experiments of fibril reconstitution *in vitro* suggest that the minor collagens regulate types I and II collagen fibrillogenesis in cartilaginous (type XI) and non-cartilaginous (type V) matrices (5–7). In addition to cartilage, the α1 subunit of type XI collagen is also produced by a large variety of non-cartilaginous tissues and cultured cells. They include bone, vitreous, skin, heart, sterna, arterial smooth muscle cells, and two rhabdomyosarcoma cell lines (8–15). The non-cartilaginous synthesis of α1(XI) collagen has been often associated with the production of hybrid types V/XI trimers (9, 13, 14).

Based on this evidence and contrary to the previous belief, it is now hypothesized that the minor collagen types represent a distinct subclass of fibrillar molecules, which perform comparable functions in a large variety of tissues and molecular associations (1). A corollary to this postulate is that tissue-specific expression of minor collagen genes must be regulated in coordination with that of the major collagen genes and, in some cases, with each other. For example, distinct transcriptional programs are expected to control α1(XI) collagen production in cartilaginous and non-cartilaginous tissues and in coordination with types II and V collagen expression, respectively. Consistent with this hypothesis, recent work has shown that co-expression of the α1(XI) and α2(V) collagen genes (*COL11A1* and *COL2A5*)[1] in arterial smooth muscle cells is modulated in a similar manner by serum deprivation and transforming growth factor-β1 (15).

Because of their potential relevance to connective tissue physiopathology, we are interested in elucidating the function and regulation of the minor collagen types. Accordingly, we recently began to examine the factors and mechanisms responsible for coordinated and tissue-specific expression of the human *COL11A1* and *COL5A2* genes. Our first report showed that 152 bp of the *COL5A2* promoter are sufficient to sustain cell type-specific transcription of a reporter gene in transient transfection experiments (16). Various DNA assays demonstrated that the activity of the *COL5A2* promoter is under the positive control of two neighboring *cis*-acting elements, termed FP-A and FP-B. These two *COL5A2* promoter sequences apparently contain novel nuclear protein-binding sites which are conserved in the corresponding region of the mouse gene (17).

As an extension of this work, we have now cloned and sequenced the 5′ portion of the human *COL11A1* gene. We also delineated the composition and extent of the minimal *COL11A1* promoter and defined its nuclear protein-binding pattern. Interestingly, a protein-bound area of the minimal *COL11A1* promoter seems to bind a transcriptional complex identical or closely related to the one that recognizes the FP-B element of the *COL5A2* gene.

MATERIALS AND METHODS

Cells and Culture Conditions—Most of the cell lines used in this study were purchased from the American Type Culture Collection; they include: HT-1080 (CCL 121), human fibrosarcoma; RD (CCL 136), human embryonal rhabdomyosarcoma; and A-204 (HTB 82), human rhabdomyosarcoma. In addition, Dr. M. Taubman (Mt. Sinai School of Medicine) kindly provided 1120, a primary line of smooth muscle cells from rat vascular tissue. All of the cell lines were grown in Dulbecco's modified essential

* This work was supported by National Institutes of Health Grant AR-38648 and by Grant-in-aid for Scientific Research 03670178 from the Japanese Ministry of Education, Science and Culture. This is article 154 from the Brookdale Center for Molecular Biology. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) U12139.

¶ Present address: First Dept. of Internal Medicine, Kanazawa University School of Medicine, Kanazawa, Japan.

|| To whom reprint requests should be addressed: Brookdale Center for Molecular Biology, Mt. Sinai School of Medicine, One Gustave L. Levy Place, New York, NY 10029.

[1] The abbreviations used in this paper are: *COL5A2* and *COL11A1*, the genes coding for α2(V) and α1(XI) collagen chains; bp, base pair(s); CAT, chloramphenicol acetyltransferase; kb, kilobase(s); nt, nucleotide(s); RACE, rapid amplification of cDNA ends.

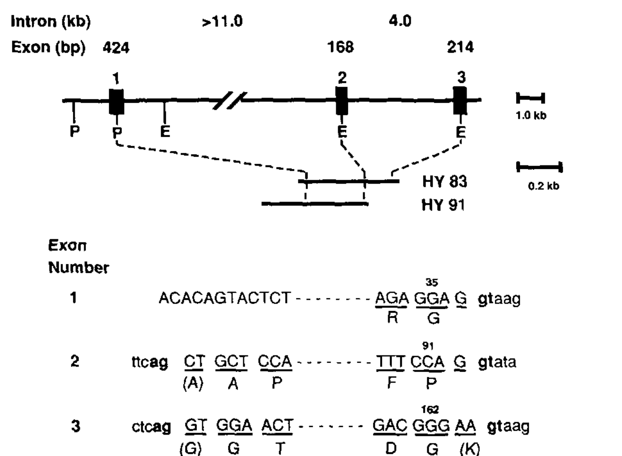


FIG. 1. **Genomic and cDNA clones at the 5′ portion of the human *COL11A1* gene.** *Top*, map of portion of the two genomic inserts with indicated exon number, exon and intron length, and the position of *Eco*RI (*E*) and *Pst*I (*P*) sites. Note that the 5′ foremost *Pst*I fragment is the one used to derive the promoter/CAT constructs. Below the gene map are depicted the two cDNA clones discussed in the text with the indication of the restriction sites common to the genomic clones. *Bottom*, sequences of the exon/intron junctions with the encoded amino acids numbered according to Yoshioka and Ramirez (10).

medium supplemented with 10% fetal calf serum.

Isolation and Characterization of cDNA and Genomic Clones—The human rhabdomyosarcoma cell line RD was used to purify poly(A)$^+$ RNA using the guanidinium thiocyanate method, followed by elution through an oligo(dT)-cellulose column (18). The RNA was then used as a template to generate a λgt10 cDNA library by random priming, and using a commercially available cloning kit (Amersham Corp.). Two genomic libraries were utilized for the isolation of the 5′ portion of the *COL11A1* gene; the first was a kind gift of Dr. M. Mori (Kumamoto Medical School), while the other was purchased from Clontech Laboratories Inc. (Palo Alto, CA). Screening of cDNA and genomic libraries, as well as identification and purification of positive clones, were performed as described previously (8, 10). Sequencing of fragments subcloned into the plasmid pBluescript II SK (Stratagene, La Jolla, CA) was carried out using the dideoxy chain termination procedure on double-stranded DNA (19). Sequences were analyzed with the aid of the computer program MacVector (International Biotechnologies Inc., New Haven, CT).

Determination of the Start Site of Transcription—For primer extension experiments, about 1 ng of a 248-bp *Pst*I/*Eco*RI fragment of clone HY 83 (see Fig. 1) was end-labeled, annealed to approximately 2 μg of poly(A)$^+$ RNA, and extended with reverse transcriptase following the previously described protocol (10). Alternatively, about 100 pg of an end-labeled synthetic oligonucleotide were annealed to approximately 10 μg of total RNA and treated as above. In both cases, RNA was derived from A-204 and HT-1080 cells. For nuclease protection, a 719-bp genomic fragment, whose 3′ end corresponds to the *Pst*I site common to HY83 and located 50 nt upstream of the ATG codon (see Fig. 2), was subcloned into the expression vector pBluescript II SK. About 1 μg of DNA was then linearized at a unique restriction site and used as a substrate to generate uniformly ^{32}P-labeled antisense RNA using the T7 RNA polymerase. The labeled riboprobe was then annealed to 10 μg of total RNA from A-204 cells and subjected to RNase digestion according to the published protocol (20). In all cases, products of the reaction were visualized by autoradiography after electrophoresis in a denaturing 5% polyacrylamide gel. Finally, the procedure of rapid amplification of cDNA ends (RACE) was carried out according to the method described by Frohman *et al.* (21). Briefly, 20 μg of total RNA from RD cells was reverse transcribed using a *COL11A1*-specific primer. Following tailing, a 10-μl aliquot of the double-stranded cDNA molecules was amplified by the polymerase chain reaction technique using a 5′ adapter primer and a 3′ *COL11A1*-specific primer located just upstream of the reverse transcriptase primer. Amplifications were carried out for 40 cycles using a step cycle of 94 °C, 40 s; 55 °C, 40 s; 72 °C, 3 min, followed by a final extension of 15 min at 72 °C. About 5 μl of the RACE products were purified by agarose gel electrophoresis and subcloned using the TA cloning kit (Invitrogen, San Diego, CA).

Construction of Chimeric Plasmids and Transfection Experiments—Different portions of the *COL11A1* promoter sharing the same 3′ end

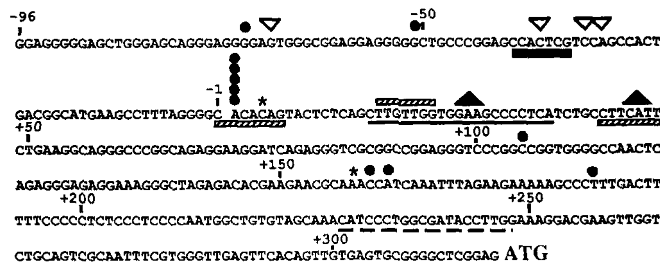


FIG. 2. **Composition of the sequence around the start sites of transcription.** The sequence extends from the ATG codon to nucleotide −96, relative to the major start site of transcription. *Continuous* and *discontinuous underlinings* indicate the positions of the oligonucleotides used in the primer extension and RACE experiments, respectively. The *Pst*I site of the 248-nt fragment used in the primer extension experiments is highlighted by *double underlining*. The 5′ ends of the extension products are signified by the *black triangles* (using the *Pst*I-*Eco*RI fragment as a primer), and by the *white triangles* (using the synthetic oligonucleotide as a primer). The approximate 5′ ends of the RNase protected bands are indicated by the *stippled boxes*, with the *black box* indicating the position of the minor ~300-nt product (see Fig. 3*C*). The 5′ ends of the RACE products are instead shown by the *black dots*. The *asterisks* highlight the 5′ of the cDNAs isolated from this (HY 91, nucleotide +5) and previous (HY 83, nucleotide 142) library screenings (10).

(+269, see Fig. 2) were generated by progressive 5′ deletion of the exon 1-containing 1.7-kb *Pst*I genomic fragment (see Fig. 1) subcloned upstream of the chloramphenicol acetyltransferase (CAT) gene in the pBLCAT3 vector (22). Plasmid DNA, purified as described previously (23), was transfected into various human and murine cells using the calcium phosphate precipitation method followed by a 15% glycerol shock treatment (24, 25). After 48 h, transfected cells were lysed and analyzed for CAT activity as described previously (23). The transcriptional activity of each construct was normalized against the activity of a co-transfected plasmid containing the luciferase gene under the control of the SV40 promoter (25). Luciferase activity was measured in a luminometer with the aid of a commercial assay system (Promega, Madison, WI). The statistical value of the transfection data was evaluated using the Mann-Whitney *U* test. The following cells were used in the transfection assays: the human fibrosarcoma line HT-1080, the human rhabdomyosarcoma line A-204, and the rat smooth muscle cells 1120.

DNA: Protein Binding Assays—Crude nuclear extracts were prepared from 1120 cells according to our published modification of the protocol of Morris *et al.* (25, 26). DNase I footprinting was performed as described previously (25). Briefly, approximately 10^5 cpm of end-labeled DNA was incubated for 30 min at 4 °C with various amounts of nuclear extracts in the presence of 1 μg of poly(dI-dC). The incubation buffer contained 10 mM Tris, pH 7.5, 80 mM NaCl, 1 mM EDTA, 4% glycerol, and 10 mM β-mercaptoethanol. After incubation, buffer composition was adjusted to 2.5 mM $CaCl_2$ and 5 mM $MgCl_2$, and samples were digested with a range of 0.1–1 unit of DNase I (Boehringer Mannheim) for 60 s at room temperature. Following addition of 5 mM EDTA and carrier tRNA, samples were phenol-extracted, ethanol-precipitated, resuspended in a denaturing buffer, and electrophoresed in a 5% polyacrylamide −7 M urea sequencing gel. The binding reaction conditions used for the DNase I footprinting test were also employed for the gel mobility-shift assay. After incubation, samples were electrophoresed in a polyacrylamide gel (30:1 acrylamide/bisacrylamide ratio) and subsequently analyzed by autoradiography. The sequences of the wild-type and mutant *COL5A2* oligonucleotides used in this study have been described in the report of Truter *et al.* (16).

RESULTS AND DISCUSSION

Structural Analysis of the 5′ Portion of COL11A1—Sequencing of several overlapping cDNAs has established the primary structure of the α1(XI) chain (8, 10). Aside from demonstrating the fibrillar nature of the α1(XI) polypeptide, this work has also provided the first evidence for an unrestricted pattern of expression of the corresponding gene (8, 10). In order to gather information relevant to the regulation of *COL11A1* expression, we used the 5′ foremost cDNA clone, HY 83, to screen two human genomic libraries (10). This led to the isolation of two not overlapping phage clones, each containing an insert of about 18 kb. Upon Southern blot and sequencing analyses, the

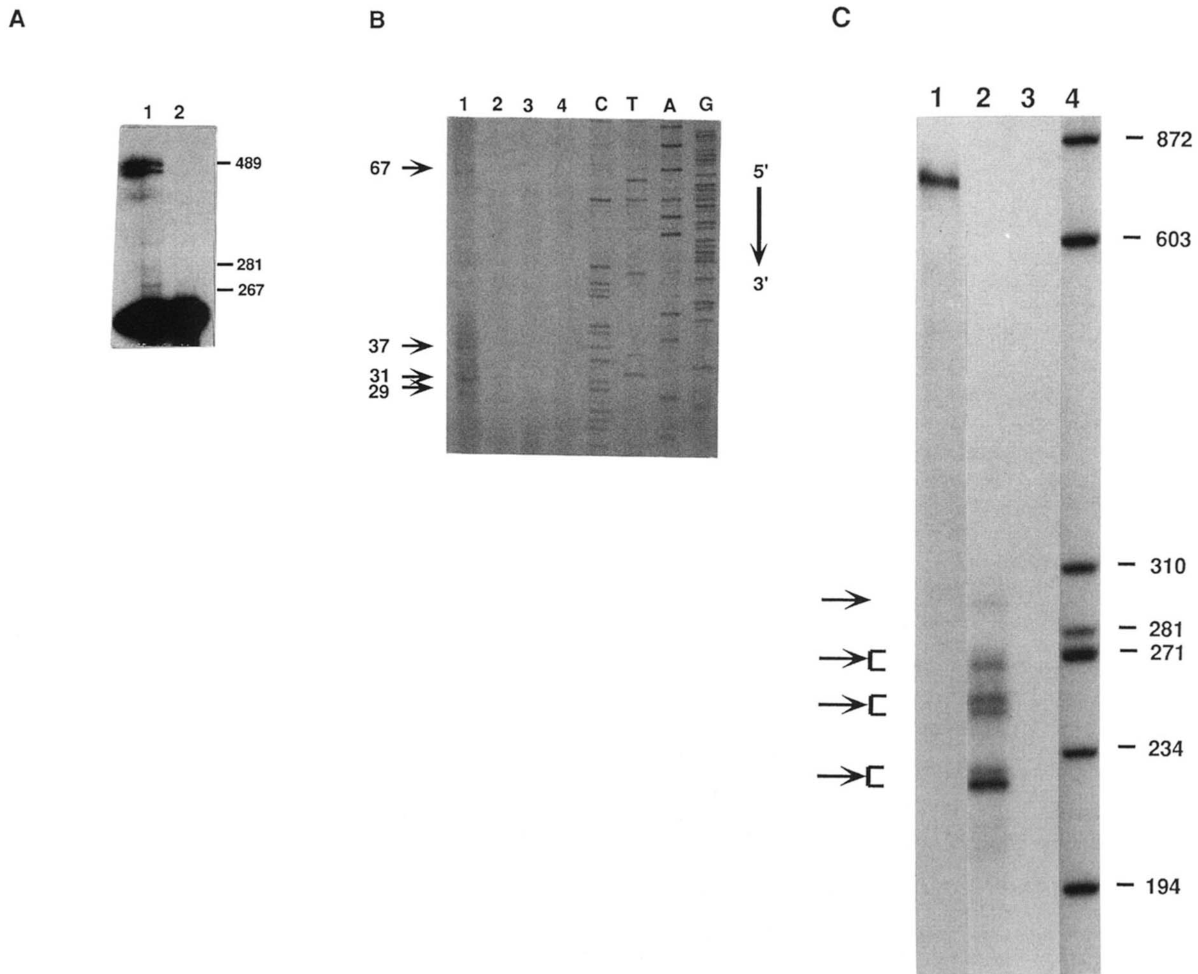


FIG. 3. **Determination of the start site of *COL11A1* transcription.** *Panel A*, primer extension experiment using the *Pst*I-*Eco*RI fragment (see Fig. 2) and RNA from A-204 (*lane 1*) and HT-1080 (*lane 2*) cells. Size markers (in bp) are on the right side of the autoradiogram. *Panel B*, primer extension experiment using the synthetic oligonucleotide primer (see Fig. 2) and RNA from A-204 (*lane 1*), HT-1080 (*lane 2*), and yeast (*lane 3*); *lane 4* is the control sample without addition of RNA. Positions of the various extension products are indicated by the *arrows* with the numbers referring to the nucleotides shown in Fig. 2. Sequencing reactions using the same primer are shown in the last four lanes (*C*, *T*, *A*, and *G*). *Panel C*, RNase protection experiment of the 719-nt riboprobe using RNA from A-204 (*lane 2*) and HT-1080 (*lane 3*) cells. The undigested riboprobe is in *lane 1*, whereas size markers (in bp) are indicated on the *right* of *lane 4*. The *arrows* on the *left* highlight the protected products whose 5′ ends are shown in Fig. 2.

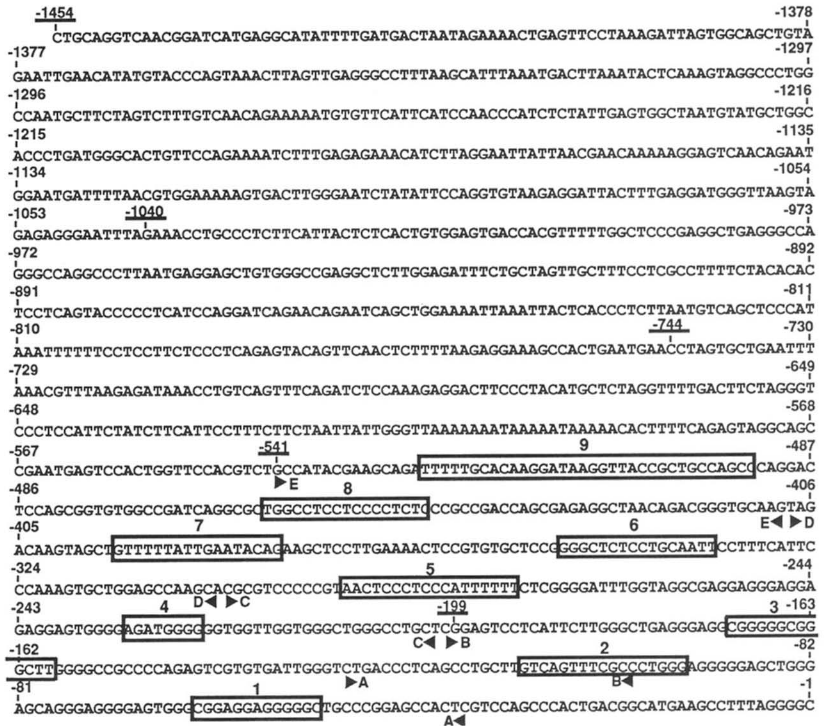
clones were found to harbor the first 488 bp of *COL11A1* coding sequence divided into three separate exons (Fig. 1). Like other fibrillar collagen genes, the first exon codes for the putative signal peptide (27). The other two exons code for portion of the so-called PARP (proline/arginine-rich protein)-like globular domain that distinguishes some of the minor collagen aminopropeptides (10, 28–31).

In addition to identifying these exons, the analysis also determined the composition of about 1.8 kb of sequence lying immediately upstream of the ATG codon (Figs. 2 and 4). Previous work has shown that the 5′ end of clone HY 83 is located 161 bp upstream of the ATG codon (Figs. 1 and 2) (10). In order to identify possible transcripts initiating further upstream of HY 83, we first screened an unamplified cDNA library from RD rhabdomyosarcoma cells with a 5′ fragment of HY 83. Analysis of the resulting positive clone, HY 91, revealed that this cDNA contains an additional 154 bp of non-coding sequence (Figs. 1 and 2). Based on these data and in order to establish the 5′ boundary of exon 1, extension experiments were performed using two different primer sequences and RNA purified from the α1(XI) collagen-producing A-204 cells and non-producer HT-1080 fibrosarcoma cells (32, 33).

The first primer is the 248-nt *Pst*I/*Eco*RI internal fragment of HY 83, whose 5′ is positioned 49 bp upstream of the ATG codon (Figs. 1 and 2) (10). This primer sequence yielded two major extension products with estimated sizes of approximately 470 and 490 nt (Fig. 3*A*). After subtracting the length of the primer sequence, the extension products placed the end of the transcripts about 20 and 40 bp downstream of the 5′ of clone HY 91 (Fig. 2). The second primer extension experiment utilized a 21-mer complementary to a sequence located in the 5′ portion of HY 91 (Fig. 2). This primer was chosen to ascertain the possible presence of transcripts extending further upstream of clone HY 91. The second primer extension experiment gave rise to four identifiable products ending 30–70 nt further 5′ of HY 91 (Figs. 2 and 3*B*). Incidentally, both sets of experiments yielded extension products only with RNA isolated from the α1(XI) collagen-producing cell line A-204.

In order to have an independent estimate of the relative representation of the two groups of extension products, we employed the RACE protocol using primers located downstream of the 5′ end of clone HY 83 (Fig. 2). Following subcloning, 11 positive RACE clones were randomly chosen and sequenced. Four of the RACE products were found to end downstream of HY 91, with two of them nearly coinciding with the last nucleotide of HY 83 (Fig. 2). Two of the RACE products extended upstream of HY 91, approximately within the region where the ends of the second group of primer extension prod-

FIG. 4. **Composition of the *COL11A1* promoter sequence.** The sequence extends from immediately 5′ of the major start site of transcription (−1, see also Fig. 2) to the 5′ *Pst*I site (−1454) of the 1.7-kb genomic clone mentioned under "Materials and Methods." The *underlined* nucleotides indicate the 5′ ends of the promoter/CAT constructs shown in Fig. 5. The *arrows* signify the boundaries of the probes used in the DNase I footprinting assays, while the *boxed* sequences delineate the approximate extent of the nuclear protein-bound areas.



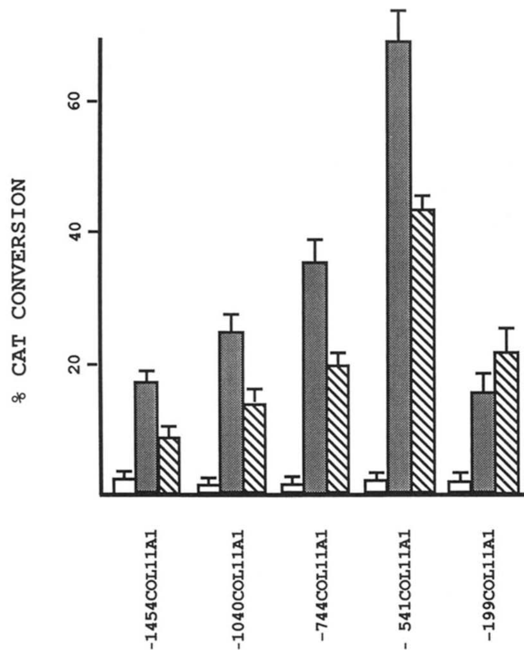


FIG. 5. **Deletion analysis of the COL1A1 promoter.** The histograms indicate the percentage of CAT conversion of the chimeric constructs normalized for the internal control and relative to pSVCAT. Plasmids were transfected in HT-1080 (*white histogram*), A-204 (*gray histogram*), and 1120 (*stippled histogram*) cells. The data represent an average of three to five independent tests ± S.D.

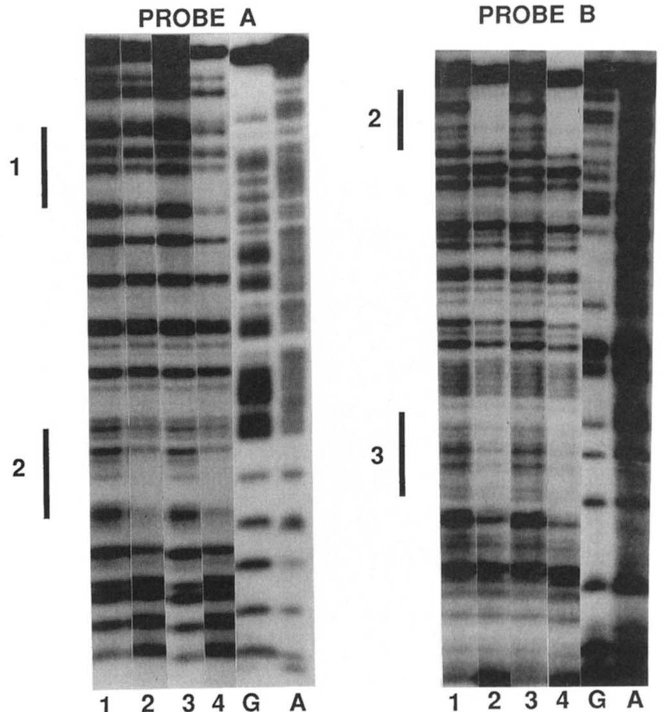


FIG. 6. **DNase I footprinting analysis of the basal promoter sequence.** The −199 to −36 region was analyzed using the overlapping probes A and B of Fig. 4. In each test, the DNA was incubated without nuclear extract (*lanes 1*), and with 60 μg of nuclear extract (*lanes 2*) in the presence of 100-fold molar excess of specific (*lanes 3*) or unspecific (*lanes 4*) competitor DNA. *Lanes G* and *A* are Maxam and Gilbert sequencing reactions used as markers. *Vertical bars* indicate the approximate extent of the nuclease-protected areas (see also Fig. 4).

ucts had been mapped (Fig. 2). Finally, 5 of the 11 race clones ended 3 nucleotides 5′ of HY 91, suggesting that this represents the major start site of *COL11A1* transcription (Fig. 2).

The location of the major start site of transcription was independently assessed by a nuclease protection experiment. To this end, we utilized a 719-nt riboprobe spanning from a *Pst*I site located upstream of the ATG codon and common to both the HY 83 and HY 91 clones (Figs. 1 and 2). After RNase digestion, we observed three major groups of resistant products, ranging in size from 220 to 260 nt, and a minor product of about 300 nt (Fig. 3*C*). The largest of the three major groups of protected bands placed the end of the transcripts at the same location as the 5′ of HY 91 and the majority of the RACE products (Fig. 2). The other two major groups of protected products were found to be nearly superimposable to the 5′ end of the first primer

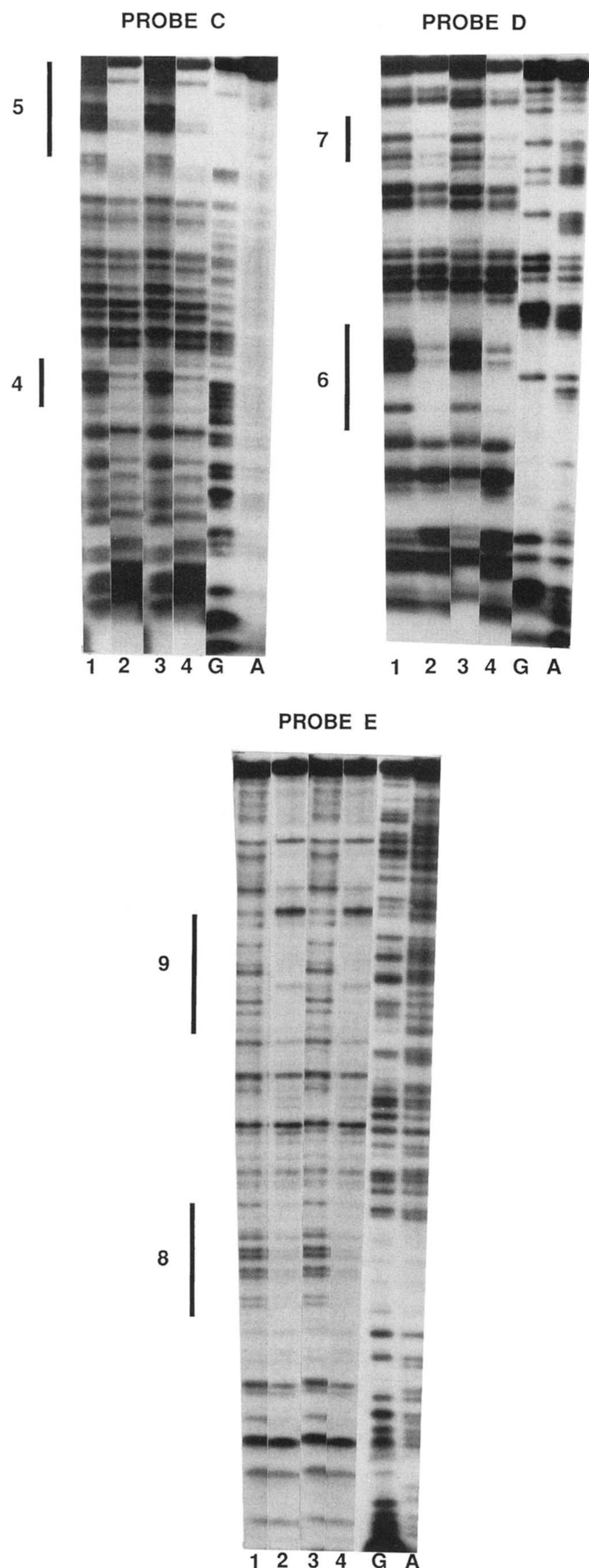


FIG. 7. **DNase I footprinting analysis of the minimal promoter sequence.** The −541 to −200 region was analyzed using the three contiguous probes C, D, and E of Fig. 4. In each test, the DNA was incubated without nuclear extract (*lanes 1*), and with 60 μg of nuclear extract (*lanes 2*) in the presence of 100-fold molar excess of specific (*lanes 3*) or unspecific (*lanes 4*) competitor DNA. *Lanes G* and *A* are Maxam and Gilbert sequencing reactions used as markers. *Vertical bars* indicate the nuclease-protected areas (see also Fig. 4).

extension products (Fig. 2). Finally, the 300-nt resistant species mapped a minor transcript coincident within the region of the second primer extension products (Fig. 2). Also in this case, specificity was supported by the negative data obtained with RNA from the α1(XI) collagen non-producer cell line HT-1080.

Based on this evidence, we concluded that the *COL11A1* gene contains multiple transcriptional start sites with a major one (+1 in Fig. 2) located 318 nt upstream of the translational initiation codon. The presence of multiple start sites of transcription was indirectly supported by the observation that the sequence immediately upstream of +1 lacks a TATA box and contains several potential GC boxes (Fig. 4).

Functional Analysis of the COL11A1 Promoter Sequence—A series of chimeric constructs containing different amounts of the *COL11A1* promoter linked to the CAT gene were generated and tested transiently by cell transfection. The aim of this analysis was to delineate the shortest promoter sequence still exhibiting detectable cell type-specific transcription. These functional tests were carried out in the α1(XI) collagen producer cells A-204 and 1120, and in the negative control line HT-1080.

The transcriptional activity of the *COL11A1*/CAT plasmid containing 1.4-kb upstream sequence was compared in A-204 and 1120 cells *versus* HT-1080 cells. In contrast to HT-1080, transfection of the −1454 *COL11A1*/CAT plasmid in the α1(XI) collagen-producing cells yielded some levels of CAT enzyme activity (Fig. 5). This result suggested that the 1.4-kb promoter sequence contains cell type-specific regulatory elements (Fig. 4). To narrow down the length of the most active promoter sequence with tissue specific expression, five constructs harboring progressive 5′ deletions of the 1.4-kb sequence were assayed in the three cell lines. The results revealed a substantial loss of transcriptional activity when the sequence between −541 and −200 was omitted from the construct (Fig. 5). The −541 and −199 promoter sequences are both expressed in a cell type-specific manner, but at significantly different levels; because of this and in order to distinguish them, they will herein be referred to as the minimal and the basal promoter, respectively. The experiments also suggested the possible presence of negative *cis*-acting elements in the promoter region spanning form −744 to −542 (Fig. 5).

Identification of Nuclear Factor-binding Sites in the Minimal COL11A1 Promoter—In the next set of analyses, the minimal promoter sequence was subjected to the DNase I footprinting analysis in order to identify possible sites of interactions with nuclear proteins. To this end, the promoter segment extending from nucleotide −541 to −36 was divided into five different fragments (*A–E* in Fig. 4). Each of these subfragments was then used as a separate probe in the DNase I footprinting analysis with nuclear proteins purified from 1120 cells. It should be noted that the definition of the approximate boundaries of the footprints was based on results obtained with both DNA strands (Fig. 4); however, here we show only the data of one of the two strands as illustrative examples (Figs. 6 and 7).

The overlapping probes A and B cover the most proximal region of the *COL11A1* promoter (−199 to −36); this is the segment that we have arbitrarily defined as the basal promoter because significantly less active than the −541 sequence. Three different footprints were identified within the 163-bp-long basal promoter (Fig. 6). They correspond to GC-rich boxes (*numbered 1–3* in Fig. 4) highly homologous to the consensus recognition sequences for the binding of transcription factors

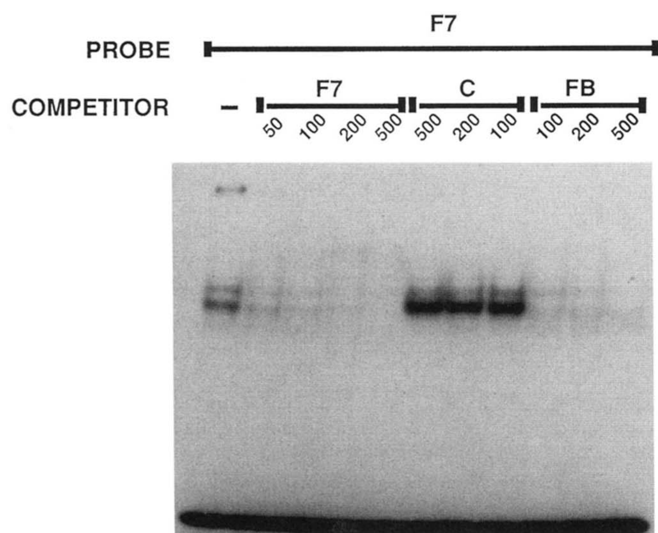


FIG. 8. **Complex competition between *COL11A1* and *COL5A2* promoter elements.** DNA binding was analyzed by the gel mobility shift assay using the footprint 7 (*F7*) probe. Binding was challenged with increasing molar amounts of the same sequence, and of oligonucleotides for FP-B (*FB*) and the sequence of oligonucleotide C (*C*, see also Fig. 4). *Numbers* above each lane indicate the -fold excess of the competitors.

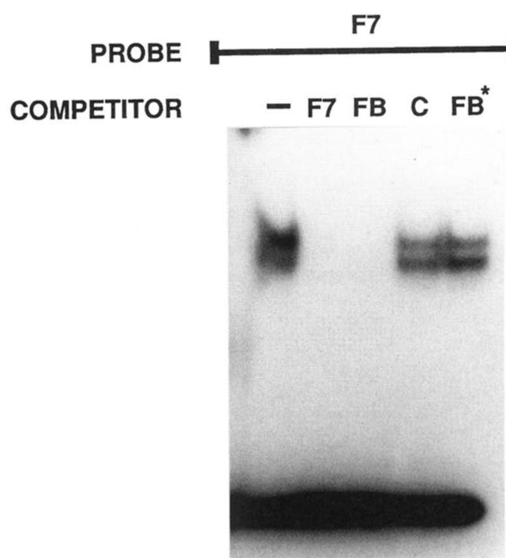


FIG. 9. **Further evidence for a common type V-XI nuclear protein binding site.** The gel mobility shift assay was used to analyze the effect of challenging nuclear protein binding to footprint 7 with 50-fold molar excess of the same sequence (*F7*), FP-B (*FB*), oligonucleotide C (*C*), and the mutated FP-B (*FB**) sequence.

AP2 and Sp1 (34). The presence of these particular binding sites is consistent with the multiple transcriptional start sites and the TATA-less nature of the *COL11A1* promoter (35).

The other three probes cover the −541 to −200 segment of the minimal *COL11A1* promoter. The DNase I footprinting analysis revealed that this region is characterized by the presence of at least six protected areas (Fig. 7). A computer-aided search identified possible homologies to known nuclear factor binding sites (34). They include the ubiquitously expressed factors NF-κB (footprint 6), CF1 (footprints 4 and 5), AP3 (5′ third of footprint 9), and AP2 (3′ third of footprint 9). In addition, a potential GATA-like recognition sequence was identified in the middle of footprint 9 (36). Footprint 8 scored the highest homology with a regulatory sequence (GGG*X*GGPuPu) of human polyoma virus JC (37). The same element was also noted in the 5′ half of footprint 5. Finally, the search identified a good candidate for the binding of the transcriptional complex that interacts with one of the aforementioned *cis*-acting elements of the *COL5A2* promoter. This initial prediction was solely based on the homology between the 3′ half of footprint 7 (TTGAATACAG) and the core sequence of FP-B (ATCAATCAG) (16). FP-B is the major *cis*-acting sequence necessary for high and tissue specific *COL5A2* gene expression (16). Indeed, 5-nucleotide substitution of the CAATC motif in FP-B was shown to result in the loss of protein binding *in vitro* and of transcriptional activity in transfection assays (16).

To confirm the above hypothesis, we performed a competition experiment using the gel mobility shift assay. To this end, a radiolabeled oligonucleotide (−395 to −379) encompassing the sequence of footprint 7 was incubated with nuclear proteins purified from 1120 nuclei. Binding to footprint 7 was challenged by increasing amounts of unlabeled oligonucleotides for footprint 7, FP-B, or an unrelated sequence. With the exception of the last one, the other two oligonucleotides competed binding with comparable effectiveness (Fig. 8). The results therefore suggested that the DNA elements of the *COL11A1* and *COL5A2* promoters bind identical or closely related nuclear proteins. Consistent with this conclusion, the FP-B oligonucleotide containing the aforementioned 5-nucleotide substitution (ACCGA for CAATC) failed to compete binding of nuclear proteins to footprint 7 (Fig. 9).

Summary—The principle scope of this work was to lay the ground for future studies aimed at characterizing the mechanisms that orchestrate the spatiotemporal expression of the *COL11A1* gene, and its modulation in response to environmental stimuli. The results revealed that the *COL11A1* is TATA-less promoter which, like other TATA-less promoters, exhibits multiple start sites of transcription. They also suggested that some of the GC-rich sequences lying immediately upstream to the major start site of transcription are probably occupied by the Spl and AP2 nuclear proteins. There is also strong preliminary evidence for a common regulatory element in the *COL11A1* and *COL5A2* promoters. This nuclear protein binding site resides within functionally critical segments of the promoters. Conceivably, it may represent the element that the two collagen genes share in order to coordinate the production of type V/XI trimers in some non-cartilaginous tissues.

Experiments in progress are testing the validity of this hypothesis. They are also examining the specificity of the *COL11A1* promoter in the more physiological environment of the transgenic mouse. Additional work is mapping the cytokine-responsive elements of the *COL11A1* and *COL5A2* genes using vascular smooth muscle cells as a model. Aside from elucidating how a specific set genes are modulated by the same cytokines, this information may provide new insights into the etiopathogenesis of diseased vascular tissues.

Acknowledgments—We thank Drs. M. Mori and M. Taubman for the generous gifts of the genomic library and the rat cell line, Dr. M. DiLiberto for helpful comments and suggestions, A. Fukutomi and H. Zhou for the excellent technical assistance, and M. Sozomenu and L. Apelis for typing the manuscript.

REFERENCES

1. van der Rest, M., and Garrone, R. (1991) *FASEB J.* **5,** 2814–2823
2. Birk, D. E., Fitch, J. M., Babiarz, J. P., and Linsenmayer, T. F. (1988) *J. Cell Biol.* **106,** 999–1008
3. Mendler, M., Eich-Bender, S. Q., Vaughan, L., Winterhalter, K. H., and Bruckner, P. (1989) *J. Cell Biol.* **108,** 191–197
4. Linsenmayer, T. F., Gibney, E., Igoe, F., Gordon, M. K., Fitch, J. M., Fessler, L. I., and Birk, D. E. (1993) *J. Cell Biol.* **121,** 1181–1189
5. Adachi, E., and Hayashi, T. (1986) *Connect. Tissue Res.* **14,** 257–266
6. Adachi, E., Hayashi, T., and Hashimoto, P. H. (1989) *Matrix* **9,** 232–237
7. Birk, D. E., Fitch, J. M., Babiarz, J. P., Doane, K. J., and Linsenmayer, T. F. (1990) *J. Cell Sci.* **95,** 649–657
8. Bernard, M., Yoshioka, H., Rodriguez, E., van der Rest, M., Kimura I.,

Ninomiya Y., Olsen, B. R., and Ramirez, F. (1988) *J. Biol. Chem.* **263,** 17159–17166
9. Niyibizi, C., and Eyre, D. R. (1989) *FEBS Lett.* **242,** 314–318
10. Yoshioka, H., and Ramirez, F. (1990) *J. Biol. Chem.* **265,** 6423–6426
11. Brown, K. E., Lawrence, R., and Sonenshein, G. E., (1991) *J. Biol. Chem.* **266,** 23268–23273
12. Nah, H. D., Barembaum, M., and Upholt, W. B. (1992) *J. Biol. Chem.* **267,** 22581–22586
13. Kleman, J. P., Hartmann, D. J., Ramirez, F., and van der Rest, M. (1992) *Eur. J. Biochem.* **210,** 329–335
14. Mayne, R., Brewton, R. G., Mayne, P. M., and Baker, J. R., (1993) *J. Biol. Chem.* **268,** 9381–9386
15. Lawrence, R., Hartmann, D. J., and Sonenshein, G. E. (1994) *J. Biol. Chem.* **269,** 9603–9609
16. Truter, S., Di Liberto, M., Inagaki, Y., and Ramirez, F. (1992) *J. Biol. Chem.* **267,** 25389–25395
17. Truter, S., Andrikopoulos, K., Di Liberto, M., Womack, L., and Ramirez, F. (1993) *Connect. Tissue Res.* **29,** 51–59
18. Chirwin, J. M., Przybyla, A. E., MacDonald, R. J., and Rutter, W. J. (1979) *Biochemistry* **18,** 5294–5299
19. Hattori, M., and Sakaki, Y. (1986) *Anal. Biochem.* **152,** 232–238
20. Exposito, J. Y., D'Alessio, M., Di Liberto, M., and Ramirez, F. (1993) *J. Biol. Chem.* **268,** 5249–5254
21. Frohman, M. A., Dush, M. K., and Martin, G. R. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85,** 8998–9002
22. Lackow, B., and Schutz, G. (1987) *Nucleic Acids Res.* **15,** 5490
23. Boast, S., Su, M. W., Ramirez, F., Sanchez, M., and Avvedimento, E. V. (1990) *J. Biol. Chem.* **265,** 13351–13356
24. Graham, F., and Van der Eb, A, (1973) *Virology* **52,** 456–457
25. Inagaki, Y., Truter, S., and Ramirez, F. (1994) *J. Biol. Chem.* **269,** 14828–14834
26. Morris, G. F., Price, D. H., and Marzluff, W. F. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83,** 1–5
27. Vuorio, E., and de Crombrugghe, B. (1990) *Annu. Rev. Biochem.* **59,** 837–872
28. Neame, P. J., Young, C. N., and Treep, J. T. (1990) *J. Biol. Chem.* **265,** 20401–20408
29. Takahara, K., Sato, Y., Okazawa, K., Okamoto, N., Noda, A. Yaoi, Y., Kato, I. (1991) *J. Biol. Chem.* **266,** 13124–13129
30. Greenspan, D. S., Cheng, W., and Hoffman, G. G. (1991) *J. Biol. Chem.* **266,** 24727–24733
31. Zhidkova, N., Brewton, R. G., and Mayne, R. (1993) *FEBS Lett.* **326,** 25–28
32. Alitalo, K., Myllyl, R., Sage, H., Pritzl, P., Vaheri, A., and Bornstein, P. (1982) *J. Biol. Chem.* **257,** 9016–9024
33. Alitalo, K., Vaheri, A., Krieg, T., and Timpl, R. (1980) *Eur. J. Biochem.* **260,** 11216–11222
34. Faisst, S., and Meyer, S. (1992) *Nucleic Acids Res.* **20,** 1–36
35. Dynan, S. W. (1986) *Trends Genet.* **2,** 196–197
36. Orkin, S. H. (1990) *Cell* **63,** 665–672
37. Martin, J. D., King, D. M., Slauch, J. M., and Frisque, R. J. (1985) *J. Virol.* **53,** 306–311