Functional analysis of human alpha 1(I) procollagen gene promoter. Differential activity in collagen-producing and -nonproducing cells and response to transforming growth factor beta 1.

Sergio A. Jimenez  
*Thomas Jefferson University*

John A. Varga  
*Thomas Jefferson University*

Anne Olsen  
*Thomas Jefferson University*

Liye Li  
*Thomas Jefferson University*

Arturo Diaz  
*Thomas Jefferson University*

Follow this and additional works at: [https://jdc.jefferson.edu/medfp](https://jdc.jefferson.edu/medfp)

Recommended Citation

Jimenez, Sergio A.; Varga, John A.; Olsen, Anne; Li, Liye; Diaz, Arturo; Herhal, Janet; and Koch, Julie, "Functional analysis of human alpha 1(I) procollagen gene promoter. Differential activity in collagen-producing and -nonproducing cells and response to transforming growth factor beta 1." (1994). *Department of Medicine Faculty Papers*. Paper 186.

[https://jdc.jefferson.edu/medfp/186](https://jdc.jefferson.edu/medfp/186)
Authors
Sergio A. Jimenez, John A. Varga, Anne Olsen, Liye Li, Arturo Diaz, Janet Herhal, and Julie Koch

This article is available at Jefferson Digital Commons: https://jdc.jefferson.edu/medfp/186
Functional Analysis of Human α1(I) Procollagen Gene Promoter

Differential Activity in Collagen-Producing and Nonproducing Cells and Response to Transforming Growth Factor β1

(Received for publication, December 30, 1993)

Sergio A. Jimenez†§¶, John Varga†, Anne Olsen§¶, Liye Li‡, Arturo DiazS, Janet Herhalt, and Julie Koch†

From the Departments of Medicine and Biochemistry and Molecular Biology and the Institute of Molecular Medicine, Jefferson Medical College, Thomas Jefferson University, Philadelphia, Pennsylvania 19107

To gain a further understanding of the regulation of human type I collagen gene expression under physiologic and pathologic conditions, we characterized 5.3 kilobase pairs (kb) of the human α1(I) procollagen gene promoter. A series of deletion constructs containing portions of the α1(I) procollagen 5′-flanking region (with end points from −5.3 kb to −84 base pairs (bp)) ligated to the chloramphenicol acetyltransferase (CAT) reporter gene were transiently transfected into NIH/3T3 cells. Maximal CAT activity was observed with constructs having 5′ end points from −804 to −174 bp. A further 5′ deletion to −84 bp caused a marked reduction in CAT activity. Cells transfected with plasmids containing longer 5′-flanking fragments of the α1(I) procollagen gene (−2.3 or −5.3 kb) showed reduced CAT activity compared with the −804 bp construct. The activity of the α1(I) procollagen promoter was much lower in cells that do not normally express type I collagen (HeLa cells) compared with collagen-producing NIH/3T3 cells. The CAT activity of deletion constructs containing longer 5′ regions than −84 bp was increased by ~2-fold in NIH/3T3 cells treated with transforming growth factor β1 (TGFβ1). Electrophoretic mobility shift assays indicated that protein-DNA complex formation with a probe corresponding to the −170 to −90 bp fragment of the α1(I) procollagen promoter was markedly enhanced in nuclear extracts prepared from TGFβ1-treated fibroblasts as compared with untreated fibroblasts. The DNA binding activity stimulated by TGFβ1 was specific for an Sp1-like sequence at positions −164 to −142 bp in the promoter. These results demonstrate that 1) there are both positive and negative cis-acting regulatory elements in the human α1(I) procollagen promoter, 2) these regulatory regions function differently in collagen-producing and -nonproducing cells, 3) the α1(I) procollagen promoter contains TGFβ1-responsive sequences located between −174 and −84 bp from the transcription start site, and 4) TGFβ1 caused marked stimulation of the DNA binding activity of a nuclear factor interacting with an Sp1-like binding site located within a region encompassing −164 to −142 bp of the α1(I) procollagen promoter.

Normal fibroblasts modulate their biosynthetic activity to maintain a precise balance between the synthesis and degradation of their products during dynamic events of tissue remodeling such as development, differentiation, and repair. It has been suggested that abnormalities in these regulatory mechanisms may be responsible for the excessive extracellular matrix accumulation in a variety of fibrotic diseases such as systemic sclerosis and idiopathic pulmonary fibrosis.

The collagens comprise a large family of widely distributed proteins that play a crucial role in the maintenance of the structural properties of the extracellular matrix. Despite the important structural and functional roles that the collagens play in normal tissues, the mechanisms that regulate their production are not completely understood. Variations in net collagen production occurring during growth and differentiation (1–3), viral (4–6) and chemical (7–9) transformation, cytokine and growth factor modulation (10–14), and spontaneous (15–17) and experimentally induced (18, 19) fibrotic processes have been ascribed to fluctuations in the steady-state collagen mRNA levels. The regulatory mechanisms responsible for the maintenance of normal procollagen mRNA levels have not been completely elucidated. Most of the available evidence suggests that the principal mechanisms operate at the level of transcription, although translational control and changes in mRNA processing and stability may also play a role (reviewed in Refs. 20 and 21). The broad spectrum of regulatory signals that can influence collagen gene transcription suggests that the collagen gene promoters are responsive to various trans-acting pathways. Several putative regulatory elements that may determine the transcriptional efficiency of procollagen genes have been identified in their corresponding promoters. These include the consensus TATA and CCAAT motifs as well as additional regulatory elements (22–25), which are the potential targets for the action of promoter-specific transcription factors (26–31). Furthermore, the transcriptional activity of some procollagen gene promoters appears to be modulated by enhancer and silencer elements located 3′ from the transcription start site (32, 33).

Detailed characterization of the cis-acting elements involved in modulation of collagen gene expression is crucial for understanding the physiologic and pathologic regulation of tissue collagen deposition. The purpose of the work reported here was to analyze the human α1(I) procollagen gene promoter in order to identify regulatory regions of the gene that may play a role in the modulation of its expression under normal and pathologic conditions.

MATERIALS AND METHODS

Construction of Plasmids—Several preliminary steps were necessary to prepare fragments with the appropriate restriction sites for ligations. The HindIII site in pSV2 CAT1 was changed to a KpnI site, and a 1.6 kb

1 The abbreviations used are: CAT, chloramphenicol acetyltrans-
KpnI-EamHI fragment was isolated. A 7 kb KpnI procollagen gene KpnI fragment extending from -5.3 to +1.7 kb was isolated from a human cosmid clone and subcloned into pUC19, and the KpnI site at -5.3 kb was changed to a NotI site to give pJH49. A HindIII-ThaI fragment, extending from -804 to +42 bp, was subcloned into the HindIII and Smal sites of pEE19, excised as a HindIII-KpnI fragment, ligated with the KpnI site at -5.3 kb and subcloned into Bluescript KS+ to give p804BS. The 4.5 kb NotI-HindIII fragment from pJH49 (containing alpha1(1) procollagen gene sequences from -5.3 kb to -804 bp) was then subcloned into the NotI and HindIII sites of p804BS to give p5.3kBS.

For the -5.3 and -804 bp constructs, the NotI-KpnI fragments from p5.3kBS and p804a1CAT were ligated with the 1.6 kb KpnI-BamHI coding fragment into the NotI and BamHI sites of Bluescript KS+ to give p6.3kCAT and p804a1CAT, respectively. For the 2.3 kb construct, a 1.5 kb BamHI-HindIII fragment containing sequences from -2.3 kb to -804 bp was converted to a NotI-HindIII fragment and inserted into the NotI and HindIII sites of p804a1CAT to give p2.3kCAT. A series of deletions (from the HindIII site at -804 bp in p804a1CAT toward the start of transcription site) was made following the exonuclease III digestion procedure of Henikoff (34). Exact deletion end points were determined by sequencing, and clones ending at -675, -465, -369, -174, and -84 bp were selected for analysis. A promoterless CAT plasmid, pCAGT, was prepared by removing the 846 bp HindIII-KpnI alpha1(1) procollagen promoter fragment from p804a1 CAT and recloning.

The sequence from -2.3 kb to -804 bp of the alpha1(1) procollagen gene was obtained from deletions made from the NotI site in p2.3k1CAT toward the initiation of transcription site (also following the exonuclease III digestion procedure). DNA sequencing of both strands was performed using the dideoxy chain termination procedure (35) with T7 polymerase (Sequenase, U. S. Biochemical Corp.) following the instructions provided by the supplier.

Cell Transfections-NIH/3T3 cells (obtained from the ATCC) from subconfluent cultures were plated at a density of 3 x 10^5 cells/100-mm dish and cultured in Dulbecco's modified Eagle's medium containing 10% calf serum. Transfections were performed 24 h later employing the calcium phosphate co-precipitation method, as described (36). The NIH/3T3 cells were transfected with 4 μg of alpha1(1) procollagen promoter-CAT plasmid DNA and 0.2 μg of the pSV2AP plasmid DNA containing the SV40 poly(A) signal enhancer. Controls were transfections with a rat alkaline phosphatase promoter (kindly provided by Dr. Kyong Yoon (37)) as an internal standard. After 4 h, the cultures were subjected to a 15% glycerol shock for 3 min at room temperature and then grown in fresh medium for 48 h before harvesting. HeLa cells (obtained from the ATCC) were plated in 60-mm plastic culture dishes (3 x 10^4 cells/dish), and 16 h later they were transfected with the plasmids with the same procedure as that used for NIH/3T3 cells, except that HeLa cells were cultured in minimum Eagle's medium containing 10% fetal calf serum instead of calf serum following the glycerol shock.

To investigate the effects of TGFβ1 on the expression of the various alpha1(1) procollagen promoter constructs, NIH/3T3 cells were transfected with the deletion construct plasmids and the pSV2AP control plasmid as described above. Six hours after the transfections, the culture media were removed, and a fresh medium containing 10% of the serum substitute Serum Plus (Hazelton Biologics) and 12.5 ng/ml human recombinant TGFβ1 (NIH Bureau of Standards, or B & D Systems) was added to duplicate cultures. Control cells incubated in parallel received culture medium without TGFβ1. The control and TGFβ1-treated cells were incubated for an additional 48 h and then harvested for assays of CAT and alkaline phosphatase activities as described below.

Assays of CAT and Alkaline Phosphatase Activities—Cell extracts were prepared by mechanically detaching the cells in 1.5 ml of phosphate-buffered saline and by centrifugation at 5000 x g for 3 min. The cell pellets were resuspended in 100 μl of 0.25 M Tris-HCl, pH 7.8, 0.1% Triton X-100 and sonicated for 15 s. CAT activity was measured on 20-μl aliquots of supernatants following centrifugation of the extracts for 5 min in a Microfuge. The supernatants were heated for 10 min at 60 °C prior to the assay of CAT activity by a liquid scintillation assay using 1 μCi of [14C]acetate-CoA/sample as described (38). In some experiments, CAT activity in the cell extracts was determined employing 1 μCi of [14C]chloramphenicol to acetylated forms was quantified by scraping the corresponding areas from the thin layer chromatography plates and measuring radioactivity by liquid scintillation spectrometry. The two methods of assessing CAT activity yielded comparable results. Alkaline phosphatase activity was assayed on 5 μl of the extracts as described by Yoon et al. (37). CAT activity in each sample was normalized to alkaline phosphatase activity to correct for differences in transfection efficiency.

Electrophoretic Mobility Shift Assays—Nuclear extracts were prepared from confluent cultures of normal human skin fibroblasts that had been incubated with or without TGFβ1 (1 or 10 ng/ml) in the presence of 5% fetal calf serum for 48 h according to the method of Dignam et al. (40). Protein concentrations were determined by a dye binding assay (Bio-Rad), and the nuclear extracts were stored in 50-μl aliquots at -70 °C until use. In order to ensure that fibroblast protein biosynthesis was stimulated by TGFβ1, cultures were labeled with [3H]proline (1.5 μCi/ml) for 24 h prior to harvesting. Only nuclear extracts prepared from cultures exhibiting a greater than 2-fold increase in [3H]proline-labeled protein production over untreated cultures were used. In a typical experiment, a confluent 175-cm2 flask (10^5 cells) yielded ~100 μg of crude nuclear protein.

Probes for electrophoretic mobility shift assays were prepared by enzymatic digestion of the p804a1CAT plasmid with NotI and KpnI. The resulting 431 bp fragment was gel-purified and further digested with NaeI and HindIII. The resulting 219 bp NaeI-HindIII (200-3) and 90 bp NaeI-NaeI (200-5) fragments were purified and 5'-end labeled with [32P]dCTP using the Klenow fragment of DNA polymerase I according to conventional procedures (41). Additional oligonucleotide probes were prepared by automated DNA synthesis (Applied BioSystems), and oligonucleotide Sp1 was obtained from Stratagene. Each oligonucleotide was annealed to its complementary strand and end-labeled with γ[32P]ATP using T4 polynucleotide kinase (Boehringer Mannheim). Competition studies were performed with molar excesses of unlabeled DNA fragments or double-stranded oligonucleotides. Electrophoretic mobility assays were performed using low ionic strength buffers as described (42). Binding reactions contained 5-10 ng of the nuclear extract, 1-2 μg of double-stranded p(dI-dC) (Pharmacia LKB Biotechnology Inc.), and radiolabeled probes containing 50,000 cpm for a total of approximately 0.5 ng. The reaction mixtures were incubated for 30 min on ice in a buffer containing 60 mM KCl, 10 mM HEPES pH 7.9, 1 mM dithiothreitol, 1 mM EDTA, and 5% glycerol in a total volume of 10 μl. Protein-DNA complexes were resolved from the free probes in non-denaturing 5% polyacrylamide gels. Electrophoresis was performed in 1× Tris/glycine buffer (50 mM Tris-HCl, 260 mM glycine) at 120 V for 120 min. The gels were dried under vacuum and exposed to x-ray film with intensifying screens at -70 °C for 16-48 h.

RESULTS

Nucleotide Sequence of -2.3 kb to -804 bp Region of alpha1(1) Procollagen Gene—The nucleotide sequences of the human alpha1(1) procollagen promoter and 5'-flanking regions of 5 kb were determined using 3'-SPAR probes (43). The 3'-end point from -5.3 kb to -84 bp were previously reported (23). Our results extend the upstream sequence of the gene to the BamHI site at -2.3 kb (Fig. 1). The only species for which the alpha1(1) procollagen sequence in this area has been reported to date is rat (43). A comparison of the nucleotide sequences of human and rat alpha1(1) procollagen gene 5'-flanking regions indicates a relatively low overall nucleotide identity (62%). However, there are several regions within the human and rat sequences (underlined in Fig. 1) that are greater than 90% identical. A search of all mammalian DNA sequences in GenBank (release 80.0) revealed no additional sequences of identity or high homology to either strand of the new upstream sequence reported here.

Deletion Analysis of 5.3 kb Upstream Sequence of alpha1(1) Procollagen Gene—To analyze sequences that may be important for regulation of collagen gene expression, a series of CAT constructs with 5' end points from -5.3 kb to -84 bp were prepared (Fig. 2) and examined in transient transfection assays of functional activity. To eliminate the alpha1(1) procollagen gene ATG initiation codon at +120 bp and the two additional upstream ATG codons at +411 bp, a promoterless CAT fragment was +42 bp. Therefore, translation should initiate only at the ATG codon of the CAT gene. The plasmids were transfected into NIH/3T3 cells, and cell extracts were analyzed for CAT activity. To control for differences in transfection efficiency between samples, all results were normalized to a co-transfected reference plasmid (pSV2AP). A summary of these
A further 3' deletion to -84 bp caused a significant reduction in activity of the -804 bp construct was 10-fold lower. In contrast to NIH/3T3 cells, transfection of constructs with longer 5' fragments resulted in about 2-fold greater CAT expression maximal and constant in cells transfected with each construct corrected for alkaline phosphatase activity in the same extracts relative to the CAT activity of the -804 bp construct. Plasmids were assayed in at least two separate transfection experiments with duplicate or triplicate determinations in each, except for the -675 bp construct, which was used in duplicate in a single experiment. Values shown represent the average ± S.E.

Comparison of Activity of α(I) Procollagen Promoter in Collagen-producing and -Nonproducing Cells—In order to assess the activity of α(I) procollagen promoter sequences in a cell line that does not normally exhibit high levels of type I collagen gene production, constructs with end points at -5.3 kb, -2.3 kb, and -804 bp were transfected into HeLa cells. As shown in Fig. 3, the CAT activity driven by these three constructs was markedly lower in HeLa cells than in NIH/3T3 cells. In NIH/3T3 cells, CAT activity driven by the -804 bp construct was about 26-fold higher than the CAT activity observed with the promoterless construct pOCAT, whereas in HeLa cells the relative activity of the -804 bp construct was 10-fold lower. In contrast to NIH/3T3 cells, transfection of constructs with longer 5' sequences into HeLa cells resulted in about 2-fold greater CAT activity than that obtained with the -804 bp plasmid. These results suggest that the positive and negative transcriptional regulatory regions of the α(I) procollagen promoter function differently in collagen-producing and -nonproducing cells.

Localization of α(I) Procollagen Promoter Regions Responsive to TGFβ1—In order to localize regions within the α(I) procollagen promoter that are responsive to stimulation by TGFβ1, the effects of TGFβ1 on the CAT activity of cells transfected with promoter deletion constructs was examined. Incubation with TGFβ1 for 48 h resulted in an approximately 2-fold stimulation of CAT activity in cells transfected with each of the three constructs (Fig. 4), suggesting that TGFβ1-responsive elements were located within the proximal region of the promoter 3' to the -804 bp end point. To further

![Fig. 1. Nucleotide sequence of the 5'-flanking region of the human α(I) procollagen gene. The sequence of the BamHI-HindIII fragment encompassing positions -2292 to -798 bp is shown. Regions with greater than 90% identity to the corresponding rat sequences (43) are underlined.](image)

![Fig. 2. Summary of CAT activity driven by various lengths of the 5'-flanking sequence of the human α(I) procollagen gene. A schematic linear map of p5.3Kal CAT and 5' end points of the deletion constructs are shown on the left. NIW3T3 cells were transiently transfected with the α(I) procollagen-CAT constructs containing various lengths of promoter sequences and the alkaline phosphatase-CAT construct; CAT and alkaline phosphatase activities were determined as described under "Materials and Methods." The hatched bars on the right represent CAT activity driven by each construct corrected for alkaline phosphatase activity in the same extracts relative to the CAT activity of the -804 bp construct. Plasmids were assayed in at least two separate transfection experiments with duplicate or triplicate determinations in each, except for the -675 bp construct, which was used in duplicate in a single experiment. Values shown represent the average ± S.E.](image)
identify TGFβ1-responsive sequences within the proximal region of the α1(I) procollagen promoter, NIH/3T3 cells were transfected with constructs with 5'-flanking regions longer than -84 bp with the promoterless construct pOCAT and incubated with TGFβ1. The results indicate that TGFβ1 caused a greater than 2-fold increase in CAT activity driven by the constructs with 5'-flanking regions longer than -84 bp. In contrast, TGFβ1 did not stimulate the CAT activity driven by the -84 bp promoter. A representative transfection experiment is shown in Fig. 5. These results indicated that TGFβ1-responsive sequences are located between -174 and -84 bp of the α1(I) procollagen promoter.

Nuclear Protein Binding to α1(I) Procollagen Promoter in Control and TGFβ1-treated Fibroblasts—In order to elucidate changes in trans-acting protein-DNA interactions involving the α1(I) procollagen promoter that were associated with TGFβ1-induced stimulation of collagen production, electrophoretic mobility shift assays were performed. For this purpose, nuclear extracts were prepared from untreated and TGFβ1-treated fibroblasts. Two fragments (200-2 and 200-3) of the α1(I) procollagen promoter region spanning the sequences from -389 to -80 bp relative to the transcription start site were used as probes (Fig. 6A). Incubation of the 200-2 probe (-170 to -80 bp) with nuclear extracts from untreated or TGFβ1-treated cells resulted in the formation of two complexes with retarded electrophoretic mobility (labeled R1 and R2), indicating the presence of nuclear DNA binding factor(s) recognizing sequences within the probe (Fig. 7). The intensity of the R1 and R2 complexes determined by laser densitometry of the autoradiograms was increased 18- and 8-fold, respectively, when nuclear extracts from TGFβ1-treated cells were used as probes (Fig. 7). Competition experiments indicated that a 4-fold molar excess of unlabeled 200-2 probe completely prevented the formation of the R1 complex and re-

Fig. 3. Activity of the human α1(I) procollagen gene promoter in collagen-producing (NIH/3T3) and nonproducing (HeLa) cells. Cultured NIH/3T3 and HeLa cells were transiently transfected with various deletion constructs (described in the legend to Fig. 2) under identical conditions. CAT activity in the cell extracts was determined following a 48-h incubation, as described under “Materials and Methods.” CAT activity driven by the individual constructs corrected for transfection efficiency, as determined by alkaline phosphatase activity in the same samples, is shown relative to the CAT activity of the promoterless plasmid pOCAT. The values shown represent the average ± S.E. of three separate transfections with each plasmid, each assayed in duplicate.

Fig. 4. Effects of TGFβ1 on CAT activity driven by various lengths of the 5'-flanking sequence of the human α1(I) procollagen gene. NIH/3T3 cells were transiently transfected with the progressive deletion constructs derived from the p5.3KBS plasmid, as described under “Materials and Methods,” and incubated in Dulbecco's modified Eagle's medium containing 10% Serum Plus with or without TGFβ1 (12.5 ng/ml) for 48 h. CAT activity of the cell extracts, corrected for alkaline phosphatase activity, is shown relative to the -804 bp construct.

Fig. 5. Regulation of the α1(I) procollagen promoter activity by TGFβ1. A, representative CAT assay demonstrating the effect of TGFβ1 on three deletion constructs of plasmid P804α1 CAT. NIH/3T3 cells were transfected with the -369, -174, and -84 bp end point constructs. Following a 48-h incubation with or without TGFβ1, CAT activity in the cell extracts was determined as described under “Materials and Methods.” Values are expressed as percentage of [14C]chloramphenicol that was acetylated. B, histogram showing results of transient transfections of deletion constructs from four separate experiments. CAT activity in TGFβ1-treated cells (mean count/min ± S.E. from duplicate determinations) is expressed relative to the CAT activity in untreated cells (100%).
Binding of Nuclear Proteins from Control or TGFβ1-treated Fibroblasts to Oligonucleotides with Sp1-like Binding Sequences—A detailed analysis of the nucleotide sequence of the −170 to −80 bp region of the human α1(I) procollagen promoter revealed two regions of homology with consensus sequences recognized by the transcription factor Sp1, here designated Sp1.1 and Sp1.2 (44, 45). To determine whether nuclear proteins from human fibroblasts may interact with these Sp-1-like elements in the α1(I) procollagen promoter, electrophoretic mobility shift assays were performed with synthetic double-stranded oligonucleotides corresponding to the Sp1.1 and Sp1.2 sequences as well as with oligonucleotides corresponding to promoter sequences that display homology with the AP-1 and NF-1 consensus binding elements. The location and nucleotide sequences of these potential regulatory elements in the α1(I) procollagen promoter are shown in Fig. 6. Upon incubation of the Sp1.1 probe with nuclear extracts prepared from untreated fibroblasts, two distinct DNA-protein complexes, designated R1 (upper, a duplex) and R2 (lower), were detected (Fig. 8A). In nuclear extracts prepared from fibroblasts treated with TGFβ1, a reproducible 4-fold increase in the intensity of the R1 complex and a less consistent 30% increase in the intensity of the R2 complex were noted compared with the intensity of the corresponding complexes formed with nuclear extracts from untreated fibroblasts. Essentially no complex formation was detected when Sp1.2 was used as the probe (Fig. 8A). To determine whether the formation of the R1 and R2 complexes resulted from specific protein interactions with the Sp1.1 probe, competition experiments with increasing amounts of unlabeled oligonucleotides containing the Sp1.1 or Sp1.2 sequences or the Sp1 consensus binding site were performed. As shown in Fig. 8B, the formation of the R2 complex was completely prevented by the addition of excess unlabeled Sp1.1 oligonucleotide, whereas excess Sp1.2 or consensus Sp1 oligonucleotides failed to compete for binding. In contrast, the formation of the R1 duplex was markedly reduced by each of the three competitor oligonucleotides. Incubation of the nuclear extracts with the AP-1.1 or NF-1.1 oligonucleotide probes resulted in the formation of single DNA-protein complexes (Fig. 8C). However, in contrast to results obtained with the Sp1.1 probe, there were no differences in the intensities of protein-DNA complexes when nuclear extracts prepared from untreated or from TGFβ1-treated fibroblasts were used. These results indicate that a nuclear trans-acting protein binds to the Sp1.1 element at −164 to −142 bp on the antisense strand of the α1(I) procollagen promoter and that this DNA binding activity is increased by treatment of human fibroblasts with TGFβ1. An additional DNA-protein complex (R2) is formed with this Sp1-like element, but this binding activity is less consistently increased by TGFβ1.

**DISCUSSION**

The mechanisms involved in the regulation of collagen production under normal or pathologic conditions are not completely understood (reviewed in Refs. 20 and 21). Although the synthesis of most proteins in eukaryotic cells appears to be regulated at a transcriptional level, post-transcriptional events, such as the regulation of the stability of newly synthesized mRNA, may play an important role under certain conditions (46, 47). Studies of the transcriptional regulation of various collagen genes in human and rodent cells in vitro have demonstrated the presence of regulatory elements located immediately 5′ upstream of the transcription initiation site (22–26). In addition, an enhancer element located within the first intron has been identified in the type I collagen genes (32–34). Sequences located far upstream of the initiation of the transcription site may also be involved in the regulation of expres-
in order to identify upstream elements that may be involved in the regulation of transcription of the human \(\alpha(1)\) procollagen gene, we determined the nucleotide sequence of the promoter region encompassing from -804 to -2292 bp. Comparison of the newly obtained sequence with that of the promoter region in the rat \(\alpha(1)\) procollagen gene (43) indicated less than 65% overall nucleotide identity. However, several regions with highly conserved sequences in the two species were found between -1900 and -1540 bp (underlined in Fig. 1). The high degree of nucleotide sequence conservation between the two species in this region suggests that these sequences may have important regulatory functions. In the rat gene, this region contains two repeats that are variations of the palindromic sequence CCCTCCC. Analysis of the newly obtained sequences of the human gene reported here demonstrated two identical palindromes localized at -1687 and -1063 bp. Other putative regulatory sequences were also identified in the newly obtained sequence. These included two Sp1 binding sites (GGCGCGG) at positions -2168 and -1614 bp and one NF-1 (half-site) binding sequence (GCCCA) at position -830 bp (reverse strand).

To perform a functional analysis of the upstream sequences, several chimeric constructs containing DNA fragments spanning up to -5.3 kb of the \(\alpha(1)\) procollagen promoter ligated to the CAT gene were prepared, and their expression was examined following transient transfection into NIH/3T3 cells. Maximal transcriptional activity was noted when promoter sequences with 5' end points from -176 to -804 bp were included. Constructs containing further upstream sequences showed progressively decreasing activity (Fig. 2). These results are similar to those reported by Rippe et al. (49) for the mouse \(\alpha(1)\) procollagen gene.

In order to examine if there were differences in the regulation of \(\alpha(1)\) procollagen gene transcription in cells that constitutively produce high levels of collagen and cells that normally do not display expression of interstitial collagen genes, we compared the expression of the promoter-CAT constructs following their transfection into NIH/3T3 cells or into HeLa cells. Marked differences in the expression of the promoter were observed between these two types of cells. Transfection of non-collagen-producing HeLa cells resulted in relatively low levels of CAT activity, which, in marked contrast to NIH/3T3 cells, was increased when promoter constructs containing -2.3 and -5.3 kb upstream sequences were examined (Fig. 3). The differences in the expression of the newly obtained sequence with that of the promoter region encompassing -164 to -142 bp or -166 to -142 bp and -85 to -69 bp, respectively, were end-labeled and used as probes in electrophoretic mobility shift assays. Nuclear extracts were prepared from normal human skin fibroblasts that had been treated with TGFβ1 (10 ng/ml) or left untreated for 48 h. Labeled probes were incubated with nuclear extracts under identical conditions, as described in the legend to Fig. 1. The positions of the shifted complexes (R1 and R2) are indicated, along with the position of the unbound DNA. Lane 1, control containing labeled Sp1.2 probe but no protein. C, nuclear extract prepared from untreated fibroblasts; T, nuclear extract prepared from TGFβ1-treated fibroblasts. Panel B, same as panel A, except that only the Sp1 oligonucleotide probe and increasing concentrations (2.5, 100, and 250-fold molar excesses) of unlabeled double-stranded competitor oligonucleotides were used. Nuclear extracts were prepared from TGFβ1 (10 ng/ml)-treated fibroblasts, except in lane 2 (C, untreated fibroblasts). Panel C, same as panel A, except the AP-1.1 and NF-1.1 oligonucleotides, corresponding to the \(\alpha(1)\) procollagen gene sequence from positions -175 to -160 bp and -102 to -85 bp, respectively, were used. The arrowhead denotes the retarded DNA-protein complex (R1).
ferences in expression of the various constructs when trans- 
ferred into collagen-producing and -nonproducing cell types 
suggest that the intracellular milieu plays an important role 
in the regulation of the rates at which the collagen genes 
are transcribed. Furthermore, these results suggest that cell-spe- 
cific transcriptional factors that act on the upstream sequences 
of the a(1) procollagen gene may influence the expression of 
this gene.

To further investigate the functional role that the a(1) pro-
collagen promoter may play during dynamic events requiring 
remodeling of the extracellular matrix, we examined the effects 
of TGFβ1 on the expression of the various promoter constructs. 
It has been suggested that this growth factor acting in auto-
crine and paracrine fashion plays an important role in the 
regulation of extracellular matrix metabolism during develop-
ment, differentiation, and repair (reviewed in Refs. 50 and 51). 
TGFβ causes marked accumulation of collagen in vivo and in 
vitro (52, 53). In animal models of hepatic and pulmonary fibrosis, 
TGFβ mRNA is expressed in high levels in tissues prior to 
the development of fibrosis (54–56). Furthermore, recent 
observations indicate that TGFβ is implicated in the pathogenesis of various 
human diseases characterized by exaggerated fibrosis (57–60).

In the present study, we found that TGFβ1 stimulated CAT 
activity driven by a(1) procollagen promoter constructs with 5′ 
end points distal to −84 bp in transiently transfected NIH/3T3 
cells. These results indicate that TGFβ1-responsive sequences 
in the human a(1) procollagen promoter are located between 
−174 and −84 bp from the transcription start site.

To determine if TGFβ1 induced alterations in DNA binding 
activity in fibroblasts, we examined DNA-protein complex for-
mation in nuclear extracts prepared from TGFβ1-treated and 
untreated cells by electrophoretic mobility shift analysis using 
DNA probes spanning the proximal region of the a(1) procol-
lagen promoter. The results showed a marked increase in DNA 
binding activity that was specific for the probe encompassing 
nucleotides −170 to −80 bp in nuclear extracts from TGFβ1-
treated cells. This region of the human a(1) procollagen 
protein contains a binding site for Sp1 at −87 bp and an element 
with complete identity to the 3′ portion of the canonical NF-1-binding motif (GCCAA), located in reverse orientation at −95 
bp. These overlapping potential binding sites have been shown 
to be important in the basal expression of the mouse a(1) 
procollagen gene (29). In addition, a sequence similar to the 
AP-1-binding site (5′-GAGTCC) is located in reverse orienta-
tion at −165 bp, and a sequence of perfect identity with a 
binding site for Sp1-like factor from the GC2 cis-acting 
element of the rat growth hormone and the F2 element of the 
human growth hormone genes (GGGAGGAG) is found at −148 
bp in reverse orientation (45). When sequences in the −170 to 
−80 bp region of the human a(1) procollagen promoter homolo-
gous to these consensus binding sites were examined in elec-
trophoretic mobility shift assays, only the Sp1.1 probe (corre-
sponding to the sequence identical with GC2 and F2) 
demonstrated increased DNA-protein complex formation with 
nuclear extracts from TGFβ1-treated when compared with un-
treated fibroblasts. The formation of the R1 complex was abol-
ished or decreased in competition assays with all three Sp1-
related oligonucleotides, in contrast to the R2 complex, which 
was prevented only by the Sp1.1 oligonucleotide competitor. 
These observations suggest that at least one component of the 
R1 complex may be a member of the Sp1 family of transcriptional 
factors. Of interest in this regard is the previous demonstration 
that Sp1 interacted with cis-acting elements within the first 
intron (34) or the promoter (61) of the a(1) procollagen gene. 

Sp1, originally described as a factor required for SV40 trans-
scription, binds to “GC boxes,” which are found in the promot-
ers of many mammalian genes (44). Although ubiquitously ex-
pressed, recent evidence points to transcriptional regulation as 
an important role for Sp1 (62). Overexpression of Sp1 was shown 
to inhibit mouse a(1) procollagen promoter activity in transiently transfected NIH/3T3 cells (29). However, when co-
transfected into Drosophila melanogaster Schneider L2 cells, 
which are devoid of homologs of mammalian transcription 
factors, Sp1 overexpression caused potent trans-activation of the 
a(1) procollagen promoter (63).

In the rat a(1) and the mouse a(2) procollagen genes, 
TGFβ-responsive elements resembling the NF-1 consensus 
sequence have been described (64, 65). In the rat a(1) procolla-
genome, the putative “TGFβ activating element” was reported 
to be located 1.6 kb upstream from the transcription start site 
(65). The element contains the 3′ portion of the canonical NF-1 
sequence GCCAAAG also found in the mouse a(2) promoter. 
However, in contrast to the mouse a(2) procollagen gene, 
stimulation of rat a(1) procollagen gene expression by TGFβ1 
does not appear to involve NF-1 binding to its cognate DNA 
element (66). Analysis of the nucleotide sequence of the human 
a(1) procollagen gene indicates that a NF-1-like sequence 
similar to the rat TGFβ activating element is located at −1718 
bp (5′-TGCCCACGGCCAGC). However, our results indicate 
that deletion of a 1.5 kb fragment including this element did 
not prevent stimulation of promoter activity by TGFβ1 (Fig. 4), 
suggesting that in NIH/3T3 cells this NF-1-like element is not 
included in TGFβ1-induced transcriptional activation of the 
a(1) procollagen gene. The nature of trans-acting factors, 
which are implicated in the modulation of the expression of 
diverse TGFβ-responsive genes, is not completely understood. 
Recent studies suggest that, depending on the cell type and the 
gene that is regulated, a variety of distinct nuclear proteins 
may be involved (67). Moreover, transcriptional regulation by 
TGFβ is likely to be a complex process, associated with the 
combinatorial interactions of ubiquitous and inducible trans-
acting factors. The identification of cell-specific DNA binding 
factors and of the stimulatory and inhibitory cis-acting ele-
ments in the promoter of the human a(1) procollagen gene 
that are targets for these factors will permit a better understand-
ing of the complex mechanisms that modulate the transcriptional 
activity of collagen genes during physiologic processes of develop-
ment and repair as well as in a variety of diseases charac-
terized by excessive collagen production.

Acknowledgments—We thank Dr. Kyong Yoon for providing the 
pSV2AP plasmid and Dr. David Hall for thoughtful suggestions. The 
skilled secretarial assistance of O. Ma in the preparation of this manu-
script is gratefully acknowledged.

REFERENCES

1. Moen, R. C., Rowe, D. W., and Palmiter, R. D. (1979) J. Biol. Chem. 254, 3526
2. Tolstoshev, P., Larimer, J. D., and Hall, D. J. (1983) J. Biol. Chem. 258, 5672
3. Merlino, G. T., McKee, C., de Crombrugghe, B., and Pastan, I. (1983) J. Biol. Chem. 258, 1004
4. Adams, S. L., Sobol, M. E., Howard, B. H., Olden, K., Yamada, K. M., de Crombrugghe, B., and Pastan, I. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 3399
5. Rowe, C. M., Racusen, J. C., Raynor, C. H., Bornstein, P., and Palmiter, R. D. (1978) Biochemistry 17, 1561
6. Sandmeier, S., and Bornstein, P. (1979) J. Biol. Chem. 254, 4950
7. Hata, R. I., and Peterkofsky, B. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 2933
8. Sandmeier, S., Smith, R., Kühn, D., and Bornstein, P. (1981) Nature 281, 830
9. Peterkofsky, B., and Prather, W. (1986) J. Biol. Chem. 261, 16818
10. Jimenez, S. A., and Rosenblom, J. (1984) J. Clin. Invest. 74, 112
11. Goldberg, M. B., Sandell, L. J., Stephenson, M. L., and Krane, S. M. (1986) J. Biol. Chem. 261, 9049
12. Goldberg, M. B., and Krane, S. M. (1987) J. Biol. Chem. 262, 16724
13. Varga, J., Rosenblom, J., and Jimenez, S. A. (1987) Biochem. J. 247, 597
14. Varga, J., Olsen, A., Ehrlich, J., Constantine, G., Rosenblom, J., and Jimenez, S. A. (1988) Eur. J. Clin. Invest. 20, 457
15. Jimenez, S. A., Feldman, G., Bashey, R. I., Bienkowski, R., and Rosenblom, J.
Human α1(I) Procollagen Gene Promoter

1986 Biochem. J. 237, 837
16. Jimenez, S. A., Williams, C. J., Myers, J. C., and Bashey, R. I. (1986) J. Biol. Chem. 261, 657
17. Ohita, A., and Uitto, J. (1987) Arthritis Rheum. 30, 404
18. Chojkier, M., Lynch, K. D., and Filip, M. (1988) Hepatology 8, 889
19. Pierce, R. A., Glaug, M. R., Greco, E. S., Mackenzie, J. W., Boyd, C. D., and Deak, S. B. (1987) J. Biol. Chem. 262, 1652
20. Vuorio, E., and de Crombrugghe, B. (1990) Annu. Rev. Biochem. 59, 837
21. Slack, J. L., Liska, D. J., and Bornstein, P. (1993) Am. J. Med. Genet. 45, 140
22. Schmidt, A., Yamasda, Y., and de Crombrugghe, B. (1984) J. Biol. Chem. 259, 7411
23. Bornstein, P., McKay, J., Morishima, J., Devrayalu, S., and Gelinus, R. E. (1987) Proc. Natl. Acad. Sci. U.S.A. 84, 8869
24. Chu, M.-L., deWet, W., Bernard, M., and Ramirez, F. (1985) J. Biol. Chem. 260, 2315
25. deWet, W., Bernard, M., Benson-Chanda, V., Chu, M.-L., Dickson, L., Weil, D., and Ramirez, F. (1985) J. Biol. Chem. 262, 16002
26. Hatamochi, A., Paterson, B., and de Crombrugghe, B. (1986) J. Biol. Chem. 261, 11310
27. Oikarinen, J., Hatamochi, A., and de Crombrugghe, B. (1987) J. Biol. Chem. 262, 11064
28. Marty, S. N., Gelmbek, P. T., Karsenty, G., and de Crombrugghe, B. (1988) Science 241, 802
29. Nehls, M., Rippe, R., Velot, L., and Brenner, D. A. (1991) Mol. Cell. Biol. 11, 4065
30. Karsenty, G., and de Crombrugghe, B. (1990) J. Biol. Chem. 265, 9934
31. Sinkevich, C. P., Thompson, J. P., Poppleton, H., and Raghoff, R. (1992) Biochem. J. 286, 179
32. Bornstein, P., and McKay, J. (1988) J. Biol. Chem. 263, 1603
33. Liska, D. J., Robinson, V. R., and Bornstein, P. (1992) Gene Expression 2, 379
34. Henikoff, S. (1984) Gene (Amst.) 12, 351
35. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 5463
36. Graham, F., and Van der Eb, A. (1973) Virology 52, 456
37. Yoon, K., Thiele, M. A., and Rodan, G. A. (1983) Gene (Amst.) 16, 11
38. Newman, J. R., Morency, C. A., and Rustian, K. D. (1987) BioTechniques 5, 444
39. Gorman, C. M., Moffat, L. F., and Howard, B. M. (1982) Mol. Cell. Biol. 2, 1044
40. Digasam, J. D., Martin, P. L., Shastry, B. S., and Roeder, R. G. (1983) Methods Enzymol. 101, 582
41. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, p. 5.40, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
42. Caceres, J., Glazier, G., Bravo, R., and Weinman, R. (1990) Oncogene 5, 59
43. Lichtler, A., Stover, M. L., Angily, J., Kream, B., and Rowe, D. W. (1989) J. Biol. Chem. 264, 3072
44. Gishon, D., Dyan, W. S., and Tijan, R. (1984) Nature 312, 409
45. Robinoux, S., Gosselin, P., Harvey, M., Lebecq, S., and Guettin, S. L. (1992) Mol. Cell. Biol. 12, 1736
46. Piasek, M. (1988) Nature 334, 863
47. Brawerman, G. (1987) Cell 54, 5
48. Boast, S., Su, M. W., Ramirez, F., Sanchez, M., and Arvedamento, E. V. (1990) J. Biol. Chem. 265, 13501
49. Rippe, R. A., Lorenzen, S.-I., Brenner, D. A., and Breindl, M. (1989) Mol. Cell. Biol. 9, 2224
50. Sporn, M. B., and Roberts, A. B. (1991) J. Cell Biol. 118, 1017
51. Massague, J. (1990) Annu. Rev. Cell Biol. 4, 597
52. Roberts, A. B., Sporn, M. B., Asoian, R. K., Smith, J. M., Roche, N. S., Wakefield, L. M., Heine, U. I., Liotta, L. A., Falanga, V., Kehrl, J. H., and Fauci, A. S. (1986) Proc. Natl. Acad. Sci. U.S.A. 83, 4167
53. Ignott, R. A., and Massague, J. (1986) J. Biol. Chem. 261, 4377
54. Raghoff, R., Irish, P., and Kang, A. H. (1989) J. Clin. Invest. 84, 1836
55. Borunda-Arriendanz, J., Seyer, J. M., Kang, A. H., and Raghoff, R. (1990) FASEB J. 4, 215
56. Craxi, M. J., Weiner, F. R., Flanders, K. C., Giambone, M.-A., Wind, R., Biempica, L., and Zern, M. A. (1989) J. Cell Biol. 108, 2477
57. Connor, T. B., Roberts, A. B., Sporn, M. B., Danielour, D., Dari, L. L., Michels, R. G., delHustros, S., Forger, C., Kato, H., Linsang, M., Hayashi, H., and Glaser, B. M. (1989) J. Clin. Invest. 83, 1661
58. Peltonen, J., Kahara, L., Jaakkola, S., Kahari, V. M., Varga, J., Uitto, J., and Jimenez, S. A. (1990) J. Invest. Dermatol. 94, 965
59. Gruschwitz, M., Muller, P. U., Sepp, N., Hofer, E., Fontana, A., and Wick, G. (1990) J. Invest. Dermatol. 94, 187
60. Broekelmann, T. J., Limper, A. H., Colby, T. J., and McDonald, J. A. (1991) Proc. Natl. Acad. Sci. U.S.A. 88, 6042
61. Nehls, M. C., Rippe, B. A., Velot, L., and Brenner, D. A. (1991) Biochem. Biophys. Res. Commun. 177, 538
62. Saffer, J. D., Jackson, S. J., and Thurston, S. J. (1990) Genes & Dev. 4, 669-666
63. Nehls, M. C., Grapilon, M. L., and Brenner, D. A. (1992) DNA Cell Biol. 11, 443
64. Rossi, P., Karsenty, G., Roberts, A. B., Roche, N. S., Sporn, M. B., and de Crombrugghe, B. (1988) Cell 52, 406
65. Ritsenthaler, J. D., Goldstein, R. H., Fine, A., Lichtler, A., Rowe, D. D. W., and Smith, B. D. (1991) Biochem. J. 280, 157
66. Ritsenthaler, J. D., Goldstein, R. H., Fine, A., and Smith, B. D. (1993) J. Biol. Chem. 268, 1925
67. Riccio, A., Pedone, P. V., Lund, L. R., Olesen, T., Olesen, H. S., and Andresen, P. A. (1992) Mol. Cell. Biol. 12, 1846