Multiple proteins are involved in the protein–DNA complex in the proximal promoter of the human α1(III) collagen gene (COL3A1)

Tomoaki Yoshino\textsuperscript{a,b}, Hideaki Sumiyoshi\textsuperscript{c}, Toshitaka Shin\textsuperscript{c}, Noritaka Matsuo\textsuperscript{c}, Yutaka Inagaki\textsuperscript{d}, Yoshifumi Ninomiya\textsuperscript{a}, Hidekatsu Yoshioka\textsuperscript{c,*}

\textsuperscript{a}Department of Molecular Biology and Biochemistry, Okayama University Graduate School of Medicine and Dentistry, Okayama 700-8558, Japan
\textsuperscript{b}Department of Medicine and Medical Science, Okayama University Graduate School of Medicine and Dentistry, Okayama 700-8558, Japan
\textsuperscript{c}Department of Anatomy, Biology and Medicine, Faculty of Medicine, Oita University, 879-5593 Oita, Japan
\textsuperscript{d}Department of Community Health, Tokai University School of Medicine, Kanagawa 259-1193, Japan

Received 27 September 2004; received in revised form 25 March 2005; accepted 8 April 2005
Available online 27 April 2005

Abstract

We have characterized the proximal promoter of the human α1(III) collagen gene (COL3A1). Transient transfection assays using a series of chimeric constructs linked to the luciferase gene indicated that the segment from −96 to −34 is necessary to activate transcription. Electrophoretic mobility shift assays (EMSAs) showed that the multiple proteins form the DNA–protein complex in different combinations depending on the cell types. A competition assay using mutant oligonucleotides showed that the sequence 5′-GCTCTCATATTTCAGAA-3′ (−79 to −63 bp) is critical for DNA–protein complex formation. This sequence is contained in the B element of mouse α1(III) collagen gene (Col3a1) reported by Ruteshouse and de Crombrugghe (J. Biol. Chem., 1993). In the rhabdomyosarcoma cell line, A204, at least two proteins of 92–118 kDa and 40–52 kDa are involved in the DNA–protein complex bound to this motif.

© 2005 Elsevier B.V. All rights reserved.

Keywords: Type III collagen; Promoter; Transcription; DNA binding protein

1. Introduction

Type III collagen is a member of the fibrillar collagen family [1−5]. The molecule is a homotrimer composed of a single type of α chain. It is widely distributed in soft connective tissues such as fetal skin and blood vessels. This molecule is also detected during the repair process of wound healing, reflecting the immaturity of tissues and increased vascularity. It is thought to play an important role in defining tissue architecture and mechanical properties, although its function is not entirely understood. Mutations in the COL3A1 gene cause type IV Ehlers-Danlos syndrome, a disease leading to aortic rupture in early adult life [6–8]. The absence of type III collagen in a mutant mouse model causes rupture of the blood vessels [9]. Electron microscopic analysis of the vessels revealed that collagen fibrils were missing in the media of the aorta and were irregular in size in the adventitia of the aorta and skin.

Type III collagen is a component of the small argyrophilic collagen fibers that characterize reticular connective tissues [10]. Type III collagen is coexpressed with type I collagen in most tissues. Studies with monoclonal antibodies suggest that type III is a component of the striated fibrils, along with collagen type I [11]. Type III collagen can be present on banded collagen fibrils regardless of fibril diameter [12]. It can also form heterotypic fibrils with type I collagen. The expression of the ratio of type III to type I collagen varies spatially and temporally. It seems likely that transcriptional events are involved in the specific expression of the type III
cell types, were found to form the DNA–protein complex at

Multiple proteins in different combinations, depending on

resistant polypeptide, named the B element binding factor

collagen gene [13]. Ruteshouse and de Crombrugghe

identified two positive

located between

factor related to Jun/AP-1 appeared to bind at the A element

to the B element located between −122 and −106, whereas the factor binding

to the B element located between −83 and −61 was a heat-

resistant polypeptide, named the B element binding factor

(BBF), with a molecular weight of approximately 95 kDa.

In the present study, we characterized the proximal

promoter of the COL3A1 gene. The segment from −96 to

−34 was necessary for the activation of transcription. Multiple proteins in different combinations, depending on

cell types, were found to form the DNA–protein complex at

−79 to −63.

2. Materials and methods

2.1. DNA clone

The COL3A1 clone containing the promoter fragment was provided by Dr. F Ramirez of the Medical College of

Cornell University, New York, NY [16].

2.2. Cells and cell culture conditions

Human rhabdomyosarcoma A204 and mouse NIH3T3

cells were purchased from American Type Culture Collection

(ATCC, Rockville, MD). The cells were cultured in Dulbecco’s modified Eagle’s medium (DMEM) supple-
mented with 10% fetal bovine serum (FBS) at 37 °C in a humidified atmosphere of 5% CO₂ in normal air.

2.3. Construction of chimeric plasmids

The 5′ flanking region of human α1(III) collagen gene of the sequence from −1686 to +68 was subcloned into a pBluescript II vector, referred to as pBL-COL3A1, from the original clone. Constructs containing different promoter lengths but sharing the same 3′ end at +68 were generated as described previously [17]. Briefly, two fragments, −1685/+68Luc and −180/+68Luc, were obtained by using the Deletion kit (Takara, Tokyo, Japan), and the other four fragments, −587/+68Luc, −287/+68Luc, −96/+68Luc and −34/+68Luc, were amplified by PCR from pBL-COL3A1 plasmid DNA. Primers used for the deletion are listed in Table 1. These PCR products were cloned into the pGEM-T Easy vector (Promega, Madison, WI). These six fragments were cut with the appropriate restriction enzyme and linked upstream of the firefly luciferase reporter gene in the pGL3-basic vector (Fig. 1A). Internal mutation constructs used in Figs. 3A and E were generated using inverted PCR methods [18]. The PCR products were self-ligated, cut with the appropriate restriction enzyme and linked to the luciferase vector. The primers are also listed in Table 1. The luciferase constructs used in Fig. 3C were generated from the constructs used in Fig. 3E. Namely, after the 5′ portion of each construct was deleted, the construct was self-ligated. All constructs were sequenced on an ABI 310 sequencer (Applied Biosystem) according to the manufacturer’s protocol.

![Fig. 1. Luciferase assay of the COL3A1 promoter using the 5′-deletion constructs. (A) Schematic illustration of the 5′-deletion constructs of the COL3A1 promoter. (B) Luciferase activity in A204 and NIH3T3 cells. All of the constructs were co-transfected with the pRL-TK vector as an internal control for transfection efficiency. Relative luciferase activities (%) were normalized against the activity of the −1685/+68 Luc construct that was transfected into A204 cells as 100%. Data are the means ± S.D. of four independent experiments.]

![Fig. 2. Comparison of the nucleotide sequences of the proximal promoter of the α1(III) collagen gene among the human, mouse and bovine. The dashes and asterisks indicate the same nucleotides as those of the human and missing nucleotides compared with those of other species, respectively. The horizontal bars labeled A and B indicate cis-elements in the mouse promoter described previously [15].]
2.4. Transient transfection and luciferase assays

Transient transfection was performed by the calcium phosphate precipitation method with some modifications [19]. The cells were plated at a density of $1 \times 10^5$ cells in 24-well dishes with 400 µL of culture medium. After incubation for 1 day at 37 °C, cells were transfected with 1 µg of luciferase plasmid DNA plus 0.1 µg of *Renilla* pRL-SV40 vector (Promega) as an internal control. Six hours after transfection, cells were rinsed with phosphate buffer saline.
PBS, fed with fresh medium and then further cultured for 2 days. The cells were rinsed with PBS, harvested by scraping with passive lysis buffer (Promega) and then centrifuged to pellet the debris.

Luciferase assays were performed using the Dual-Luciferase reporter assay system (Promega) and their activities in the cell extracts were measured using a luminometer (Lumat LB 9507, PerkinElmer Life Sciences) according to the manufacturer’s protocol. The transcription activity of each chimeric construct was evaluated by Renilla luciferase activity to normalize for the transfection efficiency. Results were expressed as means ± S.D. of four to six independent experiments.

2.5. Preparation of nuclear extracts and EMSA

Nuclear extracts were prepared from the cells as described previously [17]. The concentration of protein was determined by Bio-Rad protein assay reagent (Bio-Rad, Hercules, CA) using bovine serum albumin as a standard.

Double-stranded wild-type and mutant oligonucleotides used for EMSA were obtained by PCR amplification using pBL-COL3A1 plasmid DNA as a template [17]. The mutation oligonucleotides in Fig. 6 were generated by site-directed mutagenesis [17]. Primers used for PCR are listed in Table 1. Amplified PCR fragments were subcloned into pGEM-T Easy vector. After digestion with HindIII, DNA was radiolabeled with 32P-dCTP (Amersham Bioscience) using Klenow fragment. EMSAs were carried out as previously described [17]. Briefly, labeled probes (4 × 10⁴ cpm) were incubated with various amounts of nuclear extract (2.5–10 μg) for 30 min at 25 °C in the presence of 3 μg of poly(dI-dC). In the competition experiments, the excess of unlabeled DNA fragments were pre-incubated with nuclear extract for 30 min on ice. In some experiments, nuclear extracts were treated with heating at 95 °C for 5 min, 10% SDS and/or Triton X-100 before incubation with the DNA probe.

The DNA–nucleoprotein complex was separated from the free labeled probe by electrophoresis on a 4.5% polyacrylamide gel. The dried gels were exposed to a bioimaging plate and analyzed with a Fuji Bio Imaging Analyzer (BAS2000 or FAL5000, Fuji Film, Tokyo).

2.6. Estimation of molecular size of nuclear proteins

Nuclear extracts were run on 12.5% or 5% SDS-PAGE gels under reducing conditions and stained with Coomassie
3. Results

3.1. Functional analysis of the COL3A1 promoter

Initially, the transcriptional activity of the −1685/+68 Luc plasmid, containing the human 1.7-kb upstream sequence, was examined in A204 and NIH3T3 (Fig. 1A). Transfection of −1685/+68 Luc plasmid yielded a high level of luciferase enzyme activity in both cells (Fig. 1B). This result suggests that the 1.7-kb promoter sequence contains basic regulatory elements. To narrow down the length of the minimum active promoter sequence, six constructs harboring progressive 5’ deletions of the 1.7-kb promoter fragment were generated (Fig. 1A). An essential loss of transcriptional activity was seen in −34/+68 Luc (Fig. 1B). It suggested that the basic machinery for the transcription of the type III collagen gene is contained in the region from −96 to −34 bp, in which the nucleotide sequences among human [16], mouse [21] and bovine (accession number AB189399) are very similar (Fig. 2).

Next, we generated 20-bp or 30-bp internal deletion constructs of −1685/+68 Luc to characterize the proximal promoter. The deletion construct of −80 to −50 bp showed low activity of the promoter in A204 cells, and those of −100 to −80 bp and −80 to −50 bp also showed low activity in NIH3T3 (Figs. 3A and B). An internal deletion of −50 to −20 showed low activity in both cells. To delineate the mechanism of transcription in detail between −100 and −50 bp, a series of 5’ stepwise deletions with 10 bp internal deletions of −1685/+68 Luc was generated. In the experiment, using the stepwise deletions of 10 bp, the promoter activity decreased at −80 bp in A204, and at −90 bp in NIH3T3 cells (Figs. 3C and D). Similarly, in the experiment using internal deletions of 10 bp, the promoter activity decreased at the region from −80 to −60 bp in A204 cells, and from −90 to −70 in NIH3T3 cells (Figs. 3E and F). Interestingly, the promoter activity was increased in both cells by the internal deletion of 10 bp in the region between −60 and −50 bp. A suppressor factor may be restricted to this region.

3.2. Identification of nuclear factor binding site in the proximal region of COL3A1 promoter

To identify nuclear factors in the proximal promoter, we performed a series of EMSAs. We prepared an oligonucleotide covering the region from −97 to −31 bp. Radiolabeled oligonucleotides were incubated with nuclear extracts from the cells and EMSAs were carried out. The 32P-labeled probe bound several nuclear proteins in the cells (Fig. 4). The pattern of DNA–protein interaction differed in both cell types. GS1 and GS4 were seen in both cells, whereas GS2 and GS3 were seen in NIH3T3 cells, but were hardly seen in A204 cells. GS5 was only seen in A204. The band of GS1 in NIH3T3 cells was seen as a doublet (Fig. 4, lane 10; Fig. 6B, lane 9). Depending on the method of sample preparation, other complexes were sometimes seen. To examine the specificity of the bands, a competition assay was performed.
GS1 and GS5 were competitively inhibited by the competitor, which was an excess of the unlabeled oligonucleotide in A204 (Fig. 4, lanes 3–5). In NIH3T3 cells, the lower band of GS1 was blocked, but not the upper band (Fig. 4, lanes 11–13; Fig. 6C, lane 11). GS4 in both cells and GS2 and GS3 in NIH3T3 cells were only slightly inhibited by the competitor. These results were essentially consistent with those using the short radiolabeled oligonucleotide covering −90 to −60 bp.

Fig. 6. EMSA analysis using 6-bp substituted mutation probes or competitors. (A) The −97/−31 wild-type (Wt) and mutant oligonucleotides (m1 – m5) used for probes and competitors. The underlines in the mouse and human sequences indicate the previously reported B element [15] and the critical site for DNA–protein complexes shown in this study, respectively. (B) The 32P-labeled −97/−31 Wt, m1 – m5 probes were incubated with nuclear extracts from A204 (lanes 2–7) and NIH3T3 cells (lanes 9–14). Lanes 1 and 8 were performed without nuclear extracts. (C) The 32P-labeled −97/−31 Wt probe was incubated with nuclear extracts from A204 and NIH3T3 cells in the presence of 100-fold excess of unlabeled −97/−31 Wt oligonucleotide (lanes 3 and 11) and m1–m5 (lanes 4–8 and 12–16) as competitors. Control assays were performed without nuclear extract and competitor (lanes 1 and 9), and with nuclear extracts but without competitor (lanes 2 and 10).
Fig. 5. However, GS2 and GS3 were not seen in NIH3T3 cells, suggesting that the proteins contained in the GS2 and GS3 complexes bind to the region from −59 to −31 bp.

To identify the binding site for the factors more closely, we prepared 6-bp mutant oligonucleotides in the region from −85 to −57 bp (Fig. 6A). In A204 cells, the mutant oligonucleotides, m2 and m3, failed to form, and m4 weakly formed GS1, GS4 and GS5 protein–DNA complexes (Fig. 6B, lanes 4–6). Similarly, in NIH3T3 cells, m2 and m3 failed to form, and m4 weakly formed protein–DNA complexes of the lower band of GS1 and GS4 (Fig. 6B, lanes 11–13). Next, mutant oligonucleotides were used as cold competitors (Fig. 6C). m2 and m3 hardly, and m4 partially, competed with isotope-labeled wild-type oligonucleotide, resulting in the formation of GS1, GS4 and GS5 in A204 cells (Fig. 6C, lanes 5–7), whereas m1 and m5 significantly, and wild-type completely, inhibited the formation of the GS1 and GS5 complexes. A similar tendency was seen towards blocking the formation of the lower band of GS1 in NIH3T3 cells (Fig. 6C, lanes 10–16). In both experiments, the data were consistent except for GS4. A competition assay showed that the sequence 5′-GCTCTCA-TATTTCAGAA-3′ (−79 to −63 bp) is critical for DNA–protein complex formation.

3.3. Characterization of the DNA-binding proteins

To characterize the DNA-binding proteins, we examined the influence of the concentration of nuclear extract, the heating and the detergent. Depending on the concentration, the pattern of the bands changed. In high concentration, GS1 and GS5 formed complexes in A204 cells (Fig. 7, lane 2–4). The proteins of GS1, GS4 and GS5 in A204 cells and those of the lower band of GS1 and GS4 in NIH3T3 cells were resistant to heat (Fig. 7, lanes 5 and 9; Fig. 8, lane 3). However, following treatment with SDS after heating, the proteins of GS1, GS4 and GS5 in A204 cells completely lost their ability to bind DNA (Fig. 8, lane 4). Again, after adding an excess of Triton X-100, the proteins of GS1 and GS5, but not of GS4, gained the ability to bind. Similarly, proteins of the lower band of GS1 in NIH3T3 cells also gained the ability to bind following renaturation (Fig. 8, lane 6).

Finally, to estimate the molecular sizes of the DNA-binding proteins, the nuclear extract from A204 was fractionated by SDS-PAGE (Figs. 9A and B). After the proteins had been recovered from the gel and renatured, EMSAs were performed. The molecular sizes of GS1 and GS5 in A204 cells were estimated to be 92–118 kDa and 40–52 kDa, respectively (Fig. 9C, lanes 4 and 6). Other DNA binding proteins, designated as the GS6 complex, were also estimated at 36–40 kDa (Fig. 9C, lane 7).

4. Discussion

In the present study, we characterized the proximal promoter of the COL3A1. The 1.7-kb genomic fragment that covers the −1685 to +68 region has strong transcriptional activity. Transient transfection assays using a series of
chimeric constructs linked to the luciferase indicated that the segment from $-96$ to $-34$ is necessary for the activation of transcription. In particular, the region from $-80$ to $-60$ seems to be important for the activation in A204 cells, and that of the $-90$ to $-70$ NIH3T3 cells. A competition assay using normal and mutant oligonucleotides showed that the sequence 5\(-V\)-GCTCTCATATTTCAGAA-3\(V\) (from $-79$ to $-63$ bp) was critical for forming DNA–protein complexes in both cells. In NIH3T3 cells, this sequence was not consistent with the data of the luciferase assay, in that the $-90$ to $-70$ fragment, rather than the $-79$ to $-63$ fragment, was essential. We do not know the reason, but it may reflect the effect of cell type on the complexity of the regulation of this gene. This sequence is contained in the B element of the mouse gene reported by Ruteshouse and de Crombrugghe [15]. The DNA-binding protein of GS1 and GS5 in A204 cells, and the lower band of GS1 in NIH3T3 cells specifically bound to this region (Fig. 10). The proteins of GS1 and GS5 were heat-resistant and their molecular sizes were estimated at 92–118 kDa and 40–52 kDa, respectively (Fig. 9). By fractionation using SDS-PAGE, the small protein of the GS6 complex was also identified. However, we do not know whether it is specific for the sequence. The protein forming the GS1 complex is identical or similar to BBF reported by Ruteshouse and de Crombrugghe in terms of the molecular size and the character of heat resistance [15]. The protein of GS1 itself was found in many cells (data not shown). Therefore, it might be a ubiquitous protein. The complex of proteins, including this protein, could contribute to the specificity of gene expression. They also reported another upstream element, the A element, in the mouse gene. However, we could not find an essential loss of transcriptional activity in the deletion of the element (Fig. 1). Three nucleotides in the A element region are different in the human and the mouse (Fig. 2). Therefore, the element may be important for the mouse gene not the human gene.

Initially, type III collagen was studied with regard to the tissue expression and function using biochemical and immunohistochemical techniques. It occurs in heterotypic fibrils with type I collagen such as fetal tissues and blood vessels.
vessels. Both collagens are coordinately expressed, but the expression of the ratio varies spatially and temporally. It has been thought that the specific transcriptional events are involved in type III collagen gene. More than 10 years ago, Vuorio and de Crombrugghe intensively characterized the basal promoter activity of Col3a1. However, the transcriptional mechanism of the type III collagen gene was poorly characterized. On the other hand, the study of the regulation of type I collagen has progressed during this period. The transcription of type I collagen genes is controlled by a series of complex interactions of positive and negative transcription factors. A number of regulatory elements that are required for constitutive expression have been identified, including the ubiquitous factors such as Sp1/Sp3 and CBF/NF-Y for basal activity of these promoters [22–25]. Other putative transcription factors, such as c-Krox, BFCOL1 and IF-1, are also involved in the transcriptional regulation of type I collagen genes [26–28].

The sequence of the B element is inconsistent with the consensus motif of CBF/NFY and Sp1. Indeed, competition and supershift assays excluded the possibility of CBF/NFY and Sp1 (data not shown). The Ku antigen, which has two components, Ku80 (80 kDa) and Ku (70 kDa), bound to the proximal promoter of COL3A1 gene after induction with lysis oxidase [29]. We tried to show binding of the Ku antigen at the basal level expression, but the protein–DNA complexes of GS1 to GS5 were not supershifted by the specific antibodies of the Ku antigen (data not shown). In addition to the proteins binding to the −79 to −63 region, our experiment suggests that there may be a repressor that binds downstream of this region (Fig. 3F). Transcription of type III collagen genes might be controlled by the interaction of positive and negative transcription factors. It seems to be more complicated than that of the type I collagen gene. The identification of the factors in the −79 to −63 region or the repressor should provide further information as to the regulation of type III collagen expression.

Acknowledgements

We are indebted to Dr. F. Ramirez for providing the COL3A1 clone containing the promoter sequence; Dr. S. Hagiwara for making several constructs; Dr. S. Kusachi for generous support throughout the work; and the staff of the Division of Radioisotope Research, Institute of Scientific Research, Oita University. This work was supported by a Grant-In-Aid for Scientific Research (11470312 and 14370468 to H.Y.) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References


[9] J.C. Brown, E. Vuorio, B. de Crombrugghe, The mouse type-III procollagen gene (Col3a1), clone containing the promoter sequence; Dr. F. Ramirez for providing the


