RESEARCH COMMUNICATION

Cloning, sequencing and characterization of the 5′-flanking region of the human collagenase-3 gene

Ginette TARDIF*, Jean-Pierre PELLETIER*, Martine DUPUIS*, John E. HAMBOR†, and Johanne MARTEL-PELLETIER*‡

*University of Montreal, Osteoarthritis Research Unit, Louis-Charles Simard Research Centre, Notre-Dame Hospital, 1560 Sherbrooke Street East, Montreal, Quebec, Canada H2L 4M1, and †Pfizer Central Research, Eastern Point Road, Groton, CT 06340, U.S.A.

Collagenase-3 (matrix metalloprotease-13) is a recently discovered human collagenase produced in normal articular cartilage chondrocytes and thought to be involved in the pathological process of osteoarthritis. We have sequenced and characterized 1.6 kb of the human collagenase-3 gene 5′-flanking region. The transcription start site was located 22 bp upstream from the ATG start codon. Sequence analysis of the 5′-flanking region revealed the presence of the consensus recognition sites for the TATA and CCAAT DNA-binding proteins, activator protein-1 and E26 transformation specific/polyoma virus enhancer, as well as three core motifs of hormone response elements. Transient transfection assays demonstrated that a small fragment of 133 bp, containing the activator protein-1 and E26 transformation specific/polyoma virus enhancer sites promoted transcription in normal and osteoarthritic human chondrocytes with significantly higher activity than the original 1.6 kb fragment. Nucleotide sequence comparison of the promoter region of human collagenase-3 revealed a stronger similarity to the mouse collagenase-1 promoter than to the human collagenase-1 promoter.

INTRODUCTION

The degradation of native collagen is initiated by the action of one type of metalloprotease, the collagenases. Three collagenases have been found so far in human tissues: collagenase-1 [matrix metalloprotease-1 (MMP-1)], neutrophil collagenase (MMP-8), and a recently discovered one, collagenase-3 (MMP-13). Unlike collagenase-1, which has a widespread distribution in human tissues, the other two appear to have a more restricted pattern of tissue expression [1–4]. In contrast to human cells, rodent cells seem to produce only one type of collagenase [5], the human collagenase-3 being considered the homologue of the rodent one. The rat and mouse genes share 86 % similarity with the human collagenase-3 gene, and only 52 % with the human collagenase-1 gene [3,6,7].

Initially, expression of the human collagenase-3 gene appeared to be restricted to the pathological breast carcinoma cells [3]. Recent studies, however, suggest its involvement in bone formation during human fetal development [8], and in human articular cartilage [4,9] during pathological processes such as in those occurring in osteoarthritis and rheumatoid arthritis [9,10].

To better understand the regulation and tissue-specificity of collagenase-3, we have cloned, sequenced and characterized the 5′-flanking region of its gene. The organization of the proximal promoter region was found to be similar to that of the other MMPs. Its nucleic acid sequence showed stronger similarity to that of the mouse collagenase-1 promoter than to the human collagenase-1 promoter.

MATERIALS AND METHODS

Cloning of the 5′-flanking region of the collagenase-3 gene

The 5′-flanking region of the collagenase-3 gene was obtained using the PromoterFinder DNA Walking kit (Clontech, Palo Alto, CA, U.S.A.), and involved, basically, a primary and a secondary, or nested, PCR reaction. The kit consists of five human genomic libraries, each one being digested with a different restriction enzyme (*Eco*RI, *Sca*I, *Dra*I, *Pvu*II and *Ssp*I). Each DNA fragment of the libraries is ligated at the 5′-end to a linker to which specific primers can hybridize (a primary or outer primer and a secondary or inner primer). The 3′-end primers for the PCR reactions are gene specific. The sequence of the primers for the collagenase-3 gene was 5′-CTGCAAACTGGAGGTC-TTCCTCAGACAA-3′ (primary PCR reaction primer, nt 95–122 [3], and 5′-AGTCCAGCTCAAGAAGAGGAAGGCA-3′ (secondary PCR reaction primer, nt 25–49).

The primary PCR reaction was done with all five DNA libraries as templates, the outer primer specific for the linker and the collagenase-3 outer primer. The reaction was incubated at 95 °C for 1 min followed by 35 cycles, each at 95 °C/30 s and 68 °C/5 min, with a final elongation step at 68 °C/8 min. The secondary, or nested, PCR reaction used as a template 1 μl of the diluted (1/100) primary PCR reaction; the primers were the inner-primer specific for the linker and the inner primer of the collagenase-3. The PCR reaction was for 25 cycles, and the PCR fragments obtained from the secondary reactions were cloned directly into the vector pCRII (Invitrogen, San Diego, CA, U.S.A.). The resulting plasmids were transformed into *Escherichia coli* DH5α and purified by passage through a column (Qiagen, Chatsworth, CA, U.S.A.), as suggested by the manufacturer. This procedure yielded pure plasmid DNA that was used directly in the sequencing reactions and in the transfection assays. The AutoRead Sequencing kit (Pharmacia Biotech, Uppsala, Sweden) was used for the dideoxynucleotide sequencing of double-stranded DNA template with T7 DNA polymerase. Each cloned fragment was sequenced on both strands.

Abbreviations used: AP-1, activator protein-1; β-GAL, β-galactosidase; CAT, chloramphenicol acetyltransferase; Ets, E26 transformation specific; HRE, hormone response element; MMP, matrix metalloprotease; PEA-3, polyoma virus enhancer; RACE, rapid amplification of cDNA ends.

‡ To whom correspondence should be addressed.

The nucleotide sequence data reported in this paper has been deposited in the Genbank database under accession number U52692.

Figure 1 Nucleotide sequence of the 5′-flanking region of the human collagenase-3 gene

The 1.6 kb fragment was obtained by PCR amplification of a human genomic library, and sequenced. The transcription start site, as indicated by +1 and an arrow, was determined by the 5′-RACE method using the interleukin-1-stimulated cartilage RNA as template. The start codon for the gene is shown in bold characters. The CCAAT and TATA boxes are underlined and have lines above and below respectively. Consensus sequences for the binding of the transcription factors AP-1, PEA-3 and the core motifs of HREs are boxed.

Transcription start site

The 5′-rapid amplification of cDNA ends (RACE) method (Gibco-BRL) was used to identify the transcription start site of the collagenase-3 mRNA. This method makes use of an antisense collagenase-3-specific primer (nt 621–643) to extend the polymerization from the primer towards the 5′-end of the message, and was performed as described previously [4] with total RNA from interleukin-1-stimulated human cartilage as template. The amplified fragments were cloned directly into the vector pGEM-T (Promega, Madison, WI, U.S.A.). The plasmids from six independent clones were sequenced in order to determine where the similarity with the collagenase-3 5′-sequence ended. The last similar base was considered to be the transcription start site.

Specimen selection and cell culture

Human chondrocytes from articular cartilage (femoral condyles and tibial plateaus) were obtained from eight osteoarthritic patients (age 69 ± 7 years; mean $\pm$ SD) undergoing total knee arthroplasty. Four normal human cartilage specimens were obtained from adult cadavers (age 64 ± 16 years), within 12 h of death, from the same location as the osteoarthritic cartilage. Chondrocytes were released from articular cartilage by sequential enzymic digestion at 37 °C, as previously described [9], and cultured in Dulbecco's modified Eagle's medium (Gibco-BRL) supplemented with 10 % (v/v) heat-inactivated fetal calf serum (HyClone, Logan, UT, U.S.A.) and an antibiotic mixture (100 units/ml of penicillin base and 100 μg/ml of streptomycin base; Gibco-BRL) at 37 °C in a humidified atmosphere of 5 % CO_2/95 % air.

Analysis of promoter activity

Chloramphenicol acetyltransferase (CAT) reporter plasmids were constructed to test the activity of the human collagenase-3 promoter *in vitro*. Enzymes used for restriction and ligation of DNA were purchased from Pharmacia Biotech (Uppsala, Sweden). The 1.6 kb fragment amplified from the *Ssp*I library was released from pCRII by digestion with *Nsi*I, and subcloned in both orientations [−1599CAT and −1599CAT (3′–5′)] into the *Pst*I site of the promoterless vector pCAT-basic (Promega). The latter was constructed in order to rule out any non-specific transcriptional activity. Four deletions of the −1599CAT plasmid were constructed. The first deletion (−514CAT) was obtained by digesting the plasmid with *Hin*dIII, removing the internal fragments and religating the plasmid, thus leaving 514 bp of the 5′ end of the fragment and 22 bp downstream of the transcription start site. This construct lacked the hormone response element (HRE) core motifs as well as four out of the five CCAAT-binding sites. The −514CAT plasmid obtained was further digested with *Xba*I, the internal fragment (406 bp) was removed and subcloned in the 5′–3′ orientation into the *Xba*I site of the CAT-basic vector (−406CAT). The third deletion (−133CAT) was obtained by restricting the −514CAT plasmid with *Ava*I and *Hin*dIII and religating the plasmid with a high concentration of T4 DNA ligase (10 units/μg of DNA), leaving only 133 bp. The −406CAT and −133CAT retained the TATA box as well as the activator protein-1 (AP-1) and E26 transformation specific/polyoma virus enhancer (Ets/PEA3) sites, but the −406CAT contained a CCAAT motif. The last deletion (−39CAT) was obtained by restricting the −1599CAT plasmid with *Stu*I and *Sma*I, removing the large internal fragment and religating the blunt ends. Only 39 bp of the 5′-flanking sequence were left in this construct which retained only the TATA box. Both strands of each plasmid construct were sequenced.

The transfections were done by the calcium phosphate co-precipitation method [11], with a glycerol shock in order to improve DNA uptake [12]. Briefly, first-passage cultured chondrocytes were seeded at 8.5×10^5 cells per 9.4 cm² well in Dulbecco's modified Eagle's medium with 10 % (v/v) fetal calf serum, and incubated for 24 h. The cells were transfected with the DNA–calcium preparation, which consisted of 125 mM $CaCl_2$, 15 μg of the CAT constructs and 5 μg of the plasmid pSV-β-galactosidase (β-GAL; Promega), the latter plasmid was used as positive control of transfection. After a 48 h incubation period the cells were washed and lysed in Triton X-100/Mops-buffered saline. The cellular extracts served for CAT, β-GAL and total protein level determinations. The CAT and β-GAL levels were measured using ELISA kits from Boehringer Mannheim (Mannheim, Germany), and total protein was measured by the bicinchoninic acid method (Pierce, Rockford, IL, U.S.A.). Each transfection was performed in 5–8 independent osteoarthritic chondrocytes and in 4 independent samples for normal chondrocytes. Results were normalized according to total protein in each well and expressed as pg/μg of total protein. CAT synthesis from the transfection of the −1599CAT plasmid into chondrocytes from each specimen was assigned a value of 100 %; CAT syntheses from the other plasmids were expressed as a percentage of the −1599CAT plasmid.

Statistical analysis

Data are expressed as means ± SEM. Analysis of variance was performed to assess differences between groups. When a significant difference was detected, the groups were then subjected to Student's *t*-tests, and a result of $P < 0.05$ was considered significant.

RESULTS

Characterization of the 5′-flanking region of collagenase-3

The amplification of the *Ssp*I, *Dra*I and *Sca*I libraries yielded fragments of 1600 bp, 600 bp and 400 bp respectively, and the sequencing revealed that they had identical 3′-ends. Figure 1 shows the nucleotide sequence of the 1600 bp fragment. The sequence of the 600 bp *Dra*I and the 400 bp *ScaI* fragments were identical with the last 600 bp and 400 bp of the 1.6 kb fragment respectively.

Analysis of the 5′-flanking sequence (Figure 1) revealed the presence of various consensus recognition sites for DNA-binding proteins. The proximal promoter sequence contained a consensus TATA box at position −37 to −32, an AP-1 consensus sequence at −50 to −56, as well as an Ets/PEA3 binding site at −77 to −83. Further upstream in the sequence were three core motifs (AGGA/TCA) for HREs at positions −590, −891 and −920, and five regions for the binding of the ubiquitous CCAAT-binding proteins at −500, −657, −1218, −1240 and −1541.

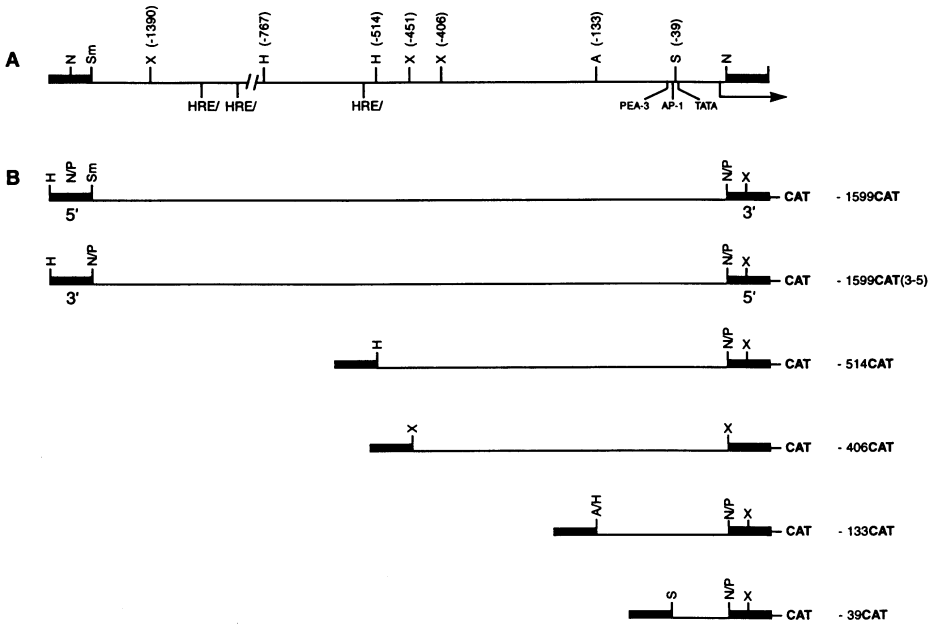
Plasmid DNA from six independent clones obtained by the 5′-RACE method had idential sequences, and their similarity to the cloned 5′-flanking sequence stopped 22 bp upstream from the ATG start codon (Figure 1).

Sequence comparison of the proximal promoter region of the human collagenase-3, mouse collagenase-1 and human collagenase-1 genes revealed that the first two had a very high degree of similarity: 214 out of the 264 bp upstream of the collagenase-3 TATA box were similar to those in the mouse promoter. These regions of similarity encompassed the regions with potential recognition sites for DNA-binding proteins (TATA, AP-1, Ets/PEA3), as well as the −77 to −113 (5 mismatches), −136 to −166 (2 mismatches) and −198 to −235 (3 mismatches) regions. In contrast, sequence comparison of the promoter region of human collagenase-3 and human collagenase-1 revealed very little nucleic acid similarity.

Analysis of promoter activity

To evaluate the promoter activity of the 1.6 kb fragment, constructed deletions of the fragment were transfected into chondrocytes (Figure 2). The restriction map of the original 1.6 kb fragment, cloned into the vector pCRII, is illustrated in Figure 2(A). The 1.6 kb fragment and the CAT-construct derivatives were transfected into normal or osteoarthritic chondrocytes and their relative promoter activities determined (Figure 2B).

The −1599CAT plasmid promoted a similar level of transcription in normal ($n = 4$; 0.28 ± 0.08 pg of CAT/μg of protein) and in osteoarthritic ($n = 8$; 0.25 ± 0.07 pg of CAT/μg of protein) chondrocytes. A similar level of activity was also found for the other plasmids for both chondrocyte types. Transfection of the



RELATIVE PROMOTER ACTIVITY (%)

Normal	Osteoarthritic
100	100
not detectable	not detectable
241.9±48.1 (p<0.03)	183.9±31.4 (p<0.02)
188.4±36.7 (p<0.05)	215.3±46.5 (p<0.02)
262.3±58.8 (p<0.03)	335.5±57.4 (p<0.0002)
not detectable	not detectable

Figure 2 Promoter activity of the 1.6 kb fragment of the 5′-flanking sequence of human collagenase-3 and its deleted derivatives

(**A**) Restriction map of the 1.6 kb fragment. The thick lines represent pCRII vector DNA and the thin lines the collagenase-3 DNA. The transcription start site is indicated by an arrow. The position of the sites for the restriction enzymes *Nsi*I (N), *Sma* I (Sm), *Xba*I (X), *Hin*dIII (H), *Ava*I (A), *Stu*I (S) and *Pst*I (P) are shown. The location of the consensus sequences for the binding of transcription factors TATA protein, AP-1 and PEA-3 and core motifs of HREs, are indicated. (**B**) Schematic representation of the plasmid constructs containing various lengths of the 5′-flanking region of the collagenase-3 gene. The 1.6 kb fragment was cloned in both orientations into the pCAT-basic vector [−1599CAT and −1599CAT (3′–5′)], and the deleted derivatives were constructed as described in the Materials and methods section. The Table illustrates the relative promoter activity of the CAT constructs in normal ($n = 4$ independent individuals) and osteoarthritic ($n = 5$–8) human chondrocytes. Promoter activity was measured as pg of CAT/μg of total protein, and data are expressed as relative units of activity with respect to the −1599CAT, which has been given a value of 100%. *P* values indicate differences from the −1599CAT construct.

1.6 kb fragment cloned in the 3′–5′ orientation [−1599CAT (3–5)] showed no detectable activity, thus indicating that the collagenase-3 5′-flanking DNA had appropriate unidirectional polarity (Figure 2B). A statistically higher level of transcriptional activity was found when the −514CAT, −406CAT and −133CAT plasmid constructs were compared with the −1599CAT construct in both normal and osteoarthritic chondrocytes. Statistical difference was also reached when the −133CAT was compared with the −514CAT ($P < 0.03$) for the osteoarthritic cells. No detectable promoter activity was found for the −39CAT construct in either chondrocyte type.

DISCUSSION

Degradation of cartilage extracellular matrix is thought to result from the combined action of several proteases, among which the MMP family appears to be important group. Human collagenase-3, a new member of the MMPs, is likely to play a crucial role in the cartilage pathophysiological process.

Comparison of the promoter region (up to −349 bp) of human collagenase-3 revealed a very high level of nucleic acid sequence similarity with mouse collagenase-1 and a very low level with human collagenase-1 promoters. These results extend those of Henriet et al. [7] and Freije et al. [3], who reported similarity within the structural sequences of the two genes.

The general organization of the collagenase-3 proximal promoter region (TATA, AP-1, Ets/PEA-3 sites) does not seem to differ greatly from those of the other MMPs, suggesting common regulatory mechanisms of gene transcription. AP-1 sites have been found in all the MMP promoters studied, except in the gelatinase 72 kDa promoter [13,14]. This site, which binds the heterodimer Fos-Jun, has been implicated in the expression of human collagenase-1 [15] and collagenase-3 [16]. The Ets/PEA-3 site is present also in the promoters of human and mouse collagenase-1, stromelysin-1 and matrilysin genes. This site binds the Ets family of oncoproteins (such as PEA-3, Ets-1, Ets-2) that operates at the level of transcription by binding to specific purine-rich regions of DNA. The mouse collagenase-1 promoter has two consensus Ets/PEA-3-binding sequences, AGGAAG, located 46 bp apart [14], while the human collagenase-3 had only one. However, the human sequence has a purine-rich sequence located at about 40 bp upstream of the consensus Ets/PEA-3 site, which could be a potential binding site.

Although not necessary for the expression of the MMP, other transcription factors, such as hormone or retinoic acid receptors, have the ability to modulate gene expression. Our data show that human collagenase-3 has three core motifs of HREs, but they are only half-sites and are not in close proximity. Binding sites for retinoic acid receptors have, until now, only been reported in the 5′-flanking region of the human stromelysin-3 gene [17].

Deleted derivatives of the cloned promoter were constructed to define the region essential for basic transcription. Our transient-transfection results showed that the sole presence of the TATA box was not sufficient to promote transcription in human cells. The −39CAT construct with only the TATA-box could not promote transcription, whereas the smallest plasmid that retained the AP-1 and PEA3 sites (−133CAT) had transcriptional activity, an activity which was even markedly higher than that obtained with the 1.6 kb fragment. This increase in transcriptional activity may result from the removal of a sequence involved in the binding of a potential repressor molecule.

The CCAAT box is a widely distributed sequence found in several promoters. It has been suggested [18] that this sequence increased the activity of neighbouring enhancer motifs. In our study, the CCAAT box did not appear to be essential in enhancing transcription, since both the −406CAT and −133CAT constructs, which lacked the sequence, showed no decrease in transcription. These constructs, in fact, had a higher transcriptional activity than the −1599CAT.

No difference in promoter activity was noted when normal chondrocytes were compared with osteoarthritic ones. This was surprising given that a higher level of collagenase-3 transcription was found in these pathological cells [9]. Although we have no clear explanation at this time, it could be that, because the level of collagenase-3 expression in normal chondrocytes is induced by factors to a level produced by osteoarthritic cells [4,9], the transfection process induced the former cells to synthesize collagenase-3 transcription factors. The osteoarthritic cells, however, might already be maximally stimulated. This hypothesis will be addressed in future experiments.

In summary, we have characterized, by both structure and activity, the human collagenase-3 gene promoter. Furthermore, promoter sequence comparison between collagenases from humans and rodents revealed the high similarity of collagenase-3 to the mouse collagenase-1, suggesting a closely related role and regulation of these enzymes, whereas the low similarity with human collagenase-1 may imply that they have evolved as specialized enzymes with different physiological roles.

This work was supported in part by grants from the Medical Research Council of Canada. The authors also thank Ms. Santa Fiori for her secretarial assistance.

REFERENCES

1 Welgus, H. G., Jeffrey, J. J. and Eisen, A. Z. (1981) J. Biol. Chem. **256**, 9511–9515
2 Hasty, K. A., Jeffrey, J. J., Hibbs, M. S. and Welgus, H. G. (1987) J. Biol. Chem. **262**, 10048–10052
3 Freije, J. M., Diez–Itza, I., Balbin, M., Sanchez, L. M., Blasco, R., Tolivia, J. and Lòpez-Otin, C. (1994) J. Biol. Chem. **269**, 16766–16773
4 Mitchell, P. G., Magna, H. A., Reeves, L. M., Lopresti-Morrow, L. L., Yocum, S. A., Rosner, P. J., Geoghegan, K. F. and Hambor, J. E. (1996) J. Clin. Invest. **97**, 761–768
5 Partridge, N. C., Walling, H. W., Bloch, S. R., Omura, T. H., Chan, P. T., Pearman, A. T. and Chou, W.-Y. (1996) Crit. Rev. Eukaryot. Gene Exp. **61**, 15–27
6 Quinn, C. O., Scott, D. K., Brinckerhoff, C. E., Matrisian, L. M., Jeffrey, J. J. and Partridge, N. C. (1990) J. Biol. Chem. **265**, 22342–22347
7 Henriet, P., Rousseau, G. R. and Eeckhout, Y. (1996) FEBS Lett. **310**, 175–178
8 Stahle-Bäckdahl, M., Sandstedt, B. and Lopez-Otin, C. (1996) Proc. Am. Assoc. Cancer Res. B5 (Abstract)
9 Reboul, P., Pelletier, J. P., Tardif, G., Cloutier, J. M. and Martel-Pelletier, J. (1996) J. Clin. Invest. **97**, 2011–2019
10 Martel-Pelletier, J., Moldovan, F., Hambor, J. E., Cloutier, J. M. and Pelletier, J. P. (1996) Arthritis Rheum. **39**, S211 (Abstract)
11 Graham, F. L. and van der Eb, A. J. (1973) Virology **52**, 456–467
12 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, New York
13 Gaire, M., Magbanua, Z., McDonnell, S., McNeil, L., Lovett, D. H. and Matrisian, L. M. (1994) J. Biol. Chem. **269**, 2032–2040
14 Schorpp, M., Mattei, M.-G., Herr, I., Gack, S., Schaper, J. and Angel, P. (1995) Biochem. J. **308**, 211–217
15 Angel, P., Baumann, I., Stein, B., Delius, H., Rahmsdorf, H. J. and Herrlich, P. (1987) Mol. Cell. Biol. **7**, 2256–2266
16 Borden, P., Solymar, D., Sucharczuk, A., Lindman, B., Cannon, P. and Heller, R. A. (1996) J. Biol. Chem. **271**, 23577–23581
17 Anglard, P., Melot, T., Guérin, E., Thomas, G. and Basset, P. (1995) J. Biol. Chem. **270**, 20337–20344
18 Strähle, U., Schmid, W. and Schütz, G. (1988) EMBO J. **7**, 3389–3395

Received 9 January 1997/3 February 1997; accepted 4 February 1997