Vol. 177, No. 1, 1991 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
May 31, 1991 Pages 538-544

CONSERVATION OF BINDING SITES FOR REGULATORY FACTORS IN THE COORDINATELY EXPRESSED α1(I) AND α2(I) COLLAGEN PROMOTERS

Gerard Karsenty and Benoit de Crombrugghe

Department of Molecular Genetics
The University of Texas
M. D. Anderson Cancer Center
1515 Holcombe Blvd.
Houston, Texas 77030

Received April 18, 1991

To examine possible mechanisms for the coordinate control of the α1(I) and α2(I) collagen genes, we have searched for DNA binding factors that are common to both genes. We have recently identified in the proximal part of the α1(I) promoter a functional binding site for CBF, a heteromeric transcriptional activator which binds to certain CCAAT sequences, and also functional binding sites for two different transcriptional repressors, designated IF1 and IF2. CBF was previously also shown to bind and activate the α2(I) collagen promoter. We now present evidence that a factor with similar binding characteristics as IF1 binds to the α2(I) promoter at approximately the same distances from the start site of transcription as in the α1(I) collagen promoter. A three bp substitution mutation in the IF1 binding site which abolishes IF1 binding increases the activity of the α2(I) promoter 4-fold as with the α1(I) promoter. We propose that the coordinate regulation of these two genes is at least in part mediated by these common elements.

Type I collagen, the most abundant fibrillar collagen in a number of connective tissues, is composed of two α1(I) chains and one α2(I) chain. Its expression changes in a number of physiological and pathological conditions that include the appearance of type I collagen in selective tissues during embryonic development and its increased synthesis in different forms of fibrosis. Numerous experiments in cultured fibroblasts show that in most circumstances, the two polypeptide chains of type I collagen are synthesized with a 2:1 stoichiometry (1). The same 2:1 stoichiometry is also observed for the steady state levels of the corresponding mRNAs in human, mouse and chick fibroblasts (2) and for their rates of synthesis in nuclear run-on experiments with chick fibroblasts nuclei (2). Although the mechanism(s) of this coordinate regulation is not

0006-291X/91 $1.50

yet understood, one possibility could be that common transcriptional factors control both promoters.

Our previous studies on the gene for the α2 chain of mouse type I collagen have identified several different cis-acting elements upstream of the start of transcription between +1 and -500. These included a functional binding site for CBF, a heteromeric CCAAT binding protein (3). In the segment of the α1(I) collagen promoter between -80 and -200 four DNA binding sites for three different trans-acting factors have been identified (see Fig. 1-A). CBF binds to the proximal of two CCAAT elements (4). A factor, designated IF1 binds to two adjacent upstream sites (4) (-190 to -170 and -160 to -130, marked A and B in Fig. 1-A). Another factor designated IF2 makes major contacts with the distal unit of a 12 bp repeat that brackets the proximal CCAAT motif (this segment is marked as C + D in Fig. 1-A) (4). Upstream of the CCAAT motif in the α2(I) promoter, between -165 and -155, there is a sequence with a high degree of homology (10 out of 11 bp) with the B element in the α1(I) promoter. This region of the α2(I) promoter was called B by analogy with this sequence in the α1(I) promoter. Given this homology we have tested whether IF1 also binds to the α2(I) collagen promoter and whether a mutation in this site has a similar effect on promoter activity as in the α1(I) promoter.

MATERIALS AND METHODS

DNA Constructions and Mutagenesis. Plasmid pR40, which contains sequences between -2000 and +54 of the mouse α2(I) collagen gene (5), was used for site-specific mutagenesis according to the method of Kunkel (6). Briefly, a 240-bp Bgl II - Sma I fragment containing sequences between -350 and -110 in the α2 (I) collagen was inserted in the replicative form of M13 mp19 phage and further processed as previously described (7). After mutagenesis the mutated fragment was substituted for the wild-type fragment in the original pR40 vector to generate plasmids with appropriate mutations (Fig. 1, panel A). To ascertain that a mutation had been inserted in the vector, recombinant plasmids were analyzed by DNA sequencing. Plasmid preparations used in transfection studies were prepared by cesium chloride centrifugation and consisted of monomeric supercoiled DNA as tested by agarose gel electrophoresis.

Gel Retardation. One microliter of nuclear extract of mouse NIH-3T3 fibroblast (between 3 and 5 μg of total protein) was incubated with approximately 10,000 cpm (5 fmole) of end-labeled double-stranded oligonucleotides in a final volume of 10μl. Incubations were carried out at room temperature for 15 min. All binding reactions contained 12 mM HEPES-NaOH (pH 7.9), 12% glycerol, 0.12 M KCl, 0.12 mM EDTA, 0.2 mM dithiothreitol, 0.2 mM PMSF, 1μg/ml pepstatin, 1 μg/ml leupeptin, and 4 μg of poly(dI-dC). Following electrophoresis in a 5% polyacrylamide gel (acrylamide: bis at 30:1) in 0.25 X T BE (45mM Trisborate, 1 mM EDTA), gels were dried and then autoradiographed at -80° C.

DNA Transfections. Mouse NIH-3T3 cells were maintained in Dulbecco's modified Eagle medium supplemented with 10% calf serum, penicillin, and streptomycin. The cells were incubated in an 8% CO_2 atmosphere and plated approximately 15 h before transfection at a density of 0.5 x 10^6 cells in 10 cm diameter plastic dishes. Transfections were performed as previously described (8) using 10 μg of plasmid DNA in 1 ml of $CaPO_4$ coprecipitate and placing the cells in a 4% CO_2 incubator. Cells were washed twice with PBS 14 to 18 h after DNA addition, refed with the above medium, and reincubated in an 8% CO_2 atmosphere. Cells were harvested 40-42 h after transfection and assayed for chloramphenicol acetyl transferase (9). Activity was determined by counting spots cut from thin layer chromatography plates. All transfections were repeated at least three times with at least two independent plasmid DNA preparations. Transfection experiments included 5 μg of pRSVβGal as an internal standard, a plasmid in which the Escherichia coli Lac Z gene is fused to the promoter of the long terminal repeat of Rous sarcoma virus. β-Galactosidase assays served to normalize chloramphenicol acetyl transferase activities between different experiments.

RESULTS AND DISCUSSION

The segment of the α1(I) collagen promoter between -80 and -200 contains four DNA binding sites for three different trans-acting factors (see Fig. 1-A). CBF, a heteromeric CCAAT-binding protein, binds to the proximal of two CCAAT elements and activates transcription of the α1(I) promoter (10). A factor, designated IF1, binds to two adjacent upstream sites (-90 to -170 and -160 to -130, marked A and B in Fig. 1-A) and another factor designated IF2 makes major contacts with the distal unit of a 12-bp repeat that brackets the proximal CCAAT motif (This segment is marked as C + D in Fig. 1-A) (4).

A comparison of the α2(I) promoter sequence upstream of the CCAAT motif with that of the α1(I) promoter (Fig. 1) shows partial homologies with the IF1 binding site in this promoter: Between -165 and -155 in the α2(I) promoter there is a sequence with a high degree of homology (10 bp out of 11) with the B element in the α1(I) promoter. This region of the α2(I) promoter was called B by analogy with this sequence in the α1(I) promoter. Given these homologies we have tested whether IF1 also binds to the α2(I) collagen promoter and whether mutations in these sites have similar effects on promoter activity.

DNA binding studies of the α2(I) promoter. To test whether IF1 binds to the sequence from -165 to -155 in the α2(I) promoter, a double-stranded oligonucleotide containing sequences between -173 and -143 was used in gel retardation assays. This segment is located at a similar distance from the start site of transcription as the proximal IF1 binding site in the α1(I) promoter (-160 to -130). Fig 2-A, lane 1 shows that a factor present in NIH-3T3 nuclear extracts binds to this oligonucleotide. This binding is specific since it can be inhibited by excess amounts of

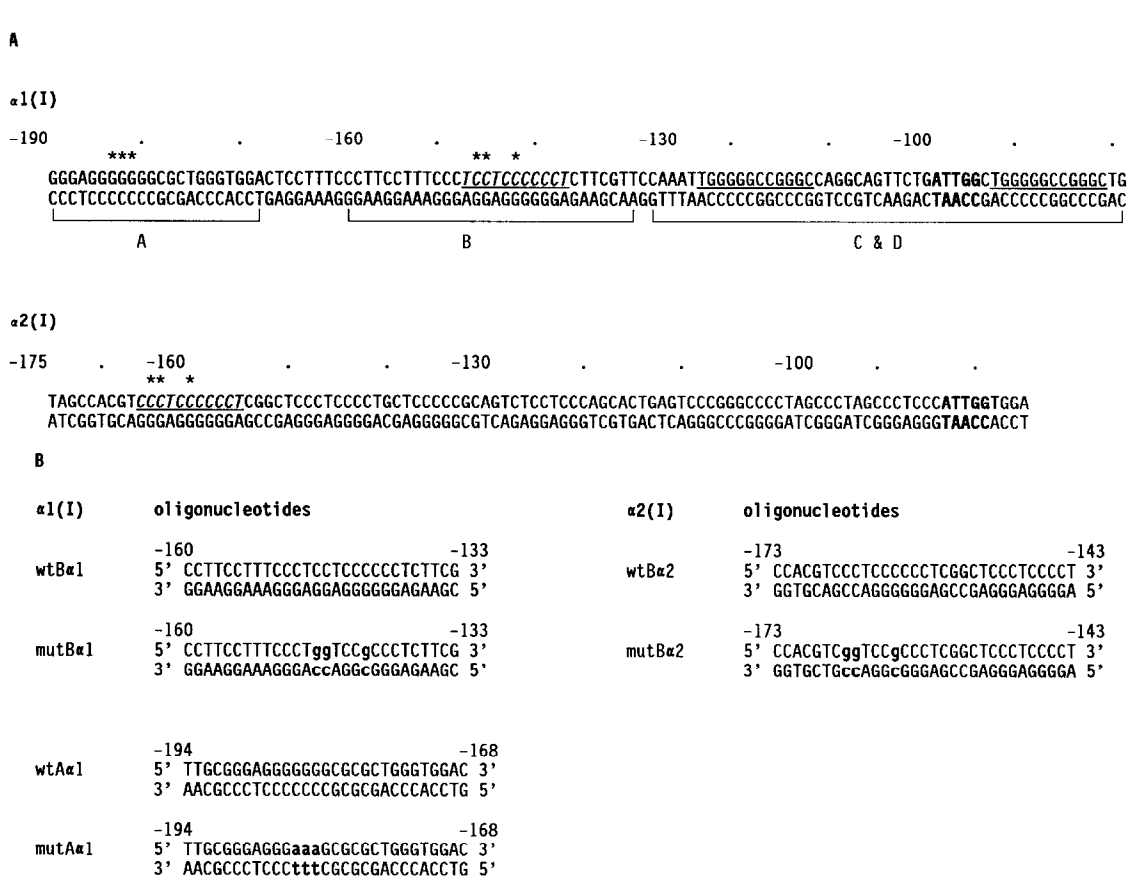

Fig. 1. Sequence and location of binding sites for DNA binding proteins in the α1(I) and the α2(I) collagen promoters. Panel A. The brackets under the α1(I) promoter sequence correspond to domains in the promoter identified by DNase I footprints (7). In the α1(I) promoter the two 12-bp repeats in the C + D segment are underlined in the upper strand. The proximal CCAAT motif and complementary ATTGG in the α1(I) promoter is shown in bold, as is the CCAAT motif in the α2(I) promoter. The position of mutations generated by site-specific mutagenesis are indicated by asterisks above the wild-type sequence. A sequence in the B elements of α1(I) and a homologous sequence in the α2(I) are italicized and underlined. In the α1(I) promoter the 3-bp mutation in the A element changes the three G residues to three A residues; in the B element the 3-bp mutation changes the three marked C residues to three A residues. In the α2(I) promoter the 3-bp mutation in the B element changes the three marked C residues to three A residues. Panel B. Sequence of the double-stranded oligonucleotides used in this work corresponding to the A and B element of the α1(I) and the α2(I) collagen promoters. The mutations are indicated in small bold characters.

the same oligonucleotide but not by a mutant oligonucleotide that contains the same 3-bp substitution mutation (lane 3) that abolishes IF1 binding in the α1(I) promoter. This mutant oligonucleotide does not bind any factor present in NIH-3T3 nuclear extracts (data not shown).

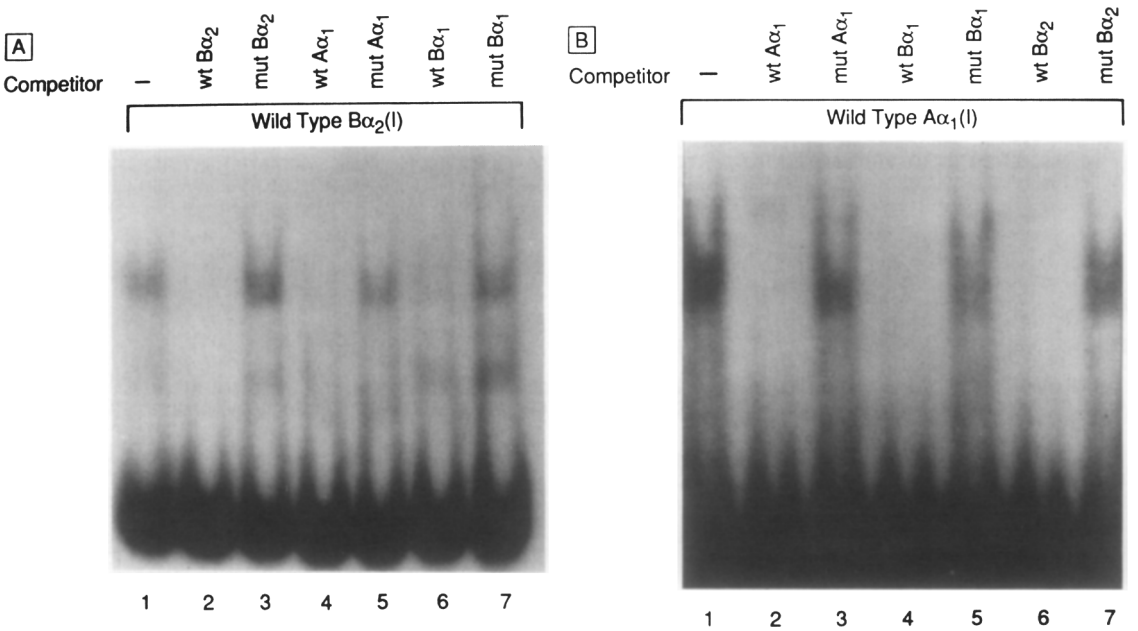

Fig. 2. Binding of IF1 to the B element of the mouse α2(I) collagen promoter. DNA binding was analyzed by gel retardation assay. Panel A. A labeled double-stranded oligonucleotide containing the sequence of the B element of the mouse α2(I) promoter was incubated with nuclear extracts of NIH-3T3 fibroblasts and the reaction subsequently fractionated by electrophoresis on a 5% nondenaturing polyacrylamide gel. Competitions were performed with 100-fold molar excess of either wild-type (w.t.) or mutant (mut) oligonucleotides containing sequences of the α1(I) promoter or the α2(I) promoter. See Fig. 1-B for the oligonucleotides. Panel B. A labeled double-stranded oligonucleotide containing the sequence of the A element of the mouse α1(I) promoter was incubated with nuclear extracts of NIH-3T3 fibroblasts and the reaction fractionated by eletrophoresis on a 5% nondenaturing polyacrylamide gel. Competitions were performed with 100-fold molar excess.

Binding to the wild-type oligonucleotide is also inhibited by an excess of unlabeled wild-type oligonucleotide containing the sequence of either the A or B element of the α1(I) promoter (Fig. 2-A, lanes 4 and 6), but not by oligonucleotides that contain mutations in these sites that inhibit the binding of IF1 (lanes 5 and 7). These experiments suggest that a similar factor binds to homologous sequences in both promoters.

To further demonstrate than an identical factor binds to the IF1 binding sites in the α1(I) promoter and to a similar sequence in the α2(I) promoter, we used an oligonucleotide containing the A element in the α1(I) promoter, which by gel retardation assay was shown to bind IF1 better than oligonucleotide containing the B element (4). With this oligonucleotide the retarded band had the same mobility as the one observed when the labeled oligonucleotide contained the sequence of the B element of the α2(I) promoter (compare Fig. 2-A, lane 1 and Fig. 2-B, lane

Table I. Effect of Small Substitution Mutations on α2(I) Promoter Activity

Plasmids	IF1 Binding Site (B element)	IF2 Binding Site (C element)	Relative CAT Activity
pR40	w t	w t	100
pG56	mut	w t	410 ± 53

DNA transfections were performed as indicated in Materials and Methods. The mutation in pG56 as introduced in pR40 as described in Materials and Methods. CAT activities are expressed relative to the wild-type promoter (pR40) as mean ± S.E.M.

1). This binding was inhibited by an excess of unlabeled oligonucleotide containing the wild-type sequence of the A and B element of the α1(I) promoter (Fig. 2-B, lanes 2 and 4) or the B element of the α2(I) promoter (Fig. 2-B, lane 6) but not by oligonucleotides containing mutations in any of these three element (lanes 3,5 and 7). In summary the competition experiments of Fig. 2-A and B, as well as the equal mobility of the complexes, strongly suggest that the factor that binds to the B element of the α2(I) promoter is the same factor that binds to the A and B elements of the α1(I) promoter. This factor was previously designated IF1 (4).

Effect of mutations in the B and C elements on α2(I) promoter activity. To examine the function of IF1 in the α2(I) promoter, transient DNA transfection experiments of NIH-3T3 fibroblasts were performed using a α2(I) collagen promoter -CAT chimeric construction carrying mutations in the IF1 binding site. As shown in Table I the same 3-bp substitution mutation in the B element that abolished binding of IF1 had a pronounced effect on the activity of the promoter, increasing this activity 4-fold (pG56). This effect is similar to the one observed with the same mutation in the α1(I) promoter (7).

It thus appears that the two closely coregulated α1(I) and α2(I) collagen genes contain in their promoters two sites for common DNA binding factors that have been evolutionary conserved (see Fig. 3 for a schematic representation). These elements are located at approximately the same distances from the transcription start site and display qualitatively similar functions. One element is a binding site for CBF, a heteromeric transcription activator. One other elements is a binding site for an inhibitory factor called IF1. It will be interesting to determine whether additional common functional elements have been conserved in other segments of these two coordinately regulated genes, which have probably diverged for over 500 million years.

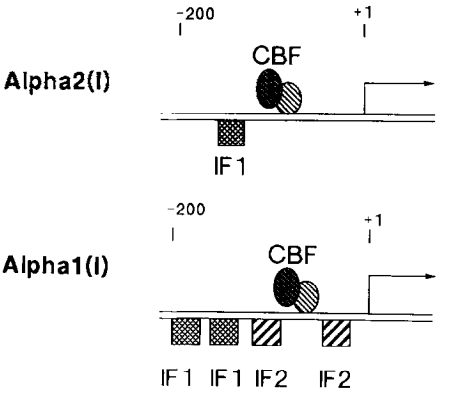

Fig. 3. Schematic representation of the different DNA binding sites for transcription factors identified in the α1(I) promoter and the α2(I) promoter. The upper panel represents the α2(I) promoter, with binding sites for CBF,and IF1. The lower panel represents the α1(I) promoter with binding sites for CBF, IF1, and IF2.

ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health grant, R01-HL41264 and the National Cancer Institute grant, R01-CA49515 awarded to Benoit de Crombrugghe. Gerard Karsenty was a recipient of a Research Investigator Award from the National Arthritis Foundation.

REFERENCES

1. deWet, W.J., Chu, M.L., and Prockop, D.J. (1983) J. Biol. Chem. 258, 14385-14389.
2. Vuurst, J., Sobel, M.E., and Martin, G.R. (1985) Eur. J. Biochem. 151, 449-453.
3. Hatamochi, A., Golumbek, P.T., Van Schaftingen, E., and de Crombrugghe, B. (1988) J. Biol. Chem. 263, 5940-5947.
4. Karsenty, G., and de Crombrugghe, B. (1990) J. Biol. Chem. 265, 9934-9942.
5. Rossi, P., and de Crombrugghe, B. (1987) Proc. Natl. Acad. Sci. USA 84, 5590-5594.
6. Kunkel, T.A. (1985) Proc. Natl. Acad. Sci. USA 82, 488-492.
7. Karsenty, G., Golumbek, P.T., and de Crombrugghe, B. (1989) J. Biol. Chem. 263, 13909-13915.
8. Chen, C., and Okayama, H. (1987) Mol. Cell. Biol. 7, 2745-2752.
9. Gorman, C.M., Moffat, L.F., and Howard, B.H. (1982) Mol. Cell. Biol. 2, 1044-1051.
10. Maity, S.N., Golumbek, P.T., Karsenty, G., and de Crombrugghe, B. (1988) Science 241, 582-585.