

A Potent Far-Upstream Enhancer in the Mouse *Pro α 2(I)* Collagen Gene Regulates Expression of Reporter Genes in Transgenic Mice

George Bou-Gharios,^{*‡} Lee Ann Garrett,^{*} Jerome Rossert,^{*} Karen Niederreither,^{*} Heidi Eberspaecher,^{*} Chad Smith,^{*} Carol Black,[‡] and Benoit de Crombrughe^{*}

^{*}Department of Molecular Genetics, The University of Texas M. D. Anderson Cancer Center, Houston, Texas 77030; and

[‡]Department of Rheumatology, Royal Free Hospital, London NW3, United Kingdom

Abstract. We have identified three DNase I-hypersensitive sites in chromatin between 15 and 17 kb upstream of the mouse *pro α 2(I)* collagen gene. These sites were detected in cells that produce type I collagen but not in cells that do not express these genes. A construction containing the sequences from -17 kb to +54 bp of the mouse *pro α 2(I)* collagen gene, cloned upstream of either the *Escherichia coli* β -galactosidase or the firefly luciferase reporter gene, showed strong enhancer activity in transgenic mice when compared with the levels seen previously in animals harboring shorter promoter fragments. Especially high levels of expression of the reporter gene were seen in dermis, fascia, and the fibrous layers of many internal organs. High levels of ex-

pression could also be detected in some osteoblastic cells. When various fragments of the 5' flanking sequences were cloned upstream of the 350-bp proximal *pro α 2(I)* collagen promoter linked to the *lacZ* gene, the *cis*-acting elements responsible for enhancement were localized in the region between -13.5 and -19.5 kb, the same region that contains the three DNase I-hypersensitive sites. Moreover, the DNA segment from -13.5 to -19.5 kb was also able to drive the cell-specific expression of a 220-bp mouse *pro α 1(I)* collagen promoter, which is silent in transgenic mice. Hence, our data suggest that a far-upstream enhancer element plays a role in regulating high levels of expression of the mouse *pro α 2(I)* collagen gene.

THE genetic programs that control the differentiation of fibroblasts and osteoblasts are still poorly understood. We are using the type I collagen genes as markers for these cell lineages to better understand the mechanisms by which they express their phenotypes. The type I collagen genes are expressed at high levels in osteoblasts, in odontoblasts, and in fibroblasts of tendons and skin, whereas in other tissues, their expression is considerably lower and is due mainly to the presence of fibroblasts and mesenchymal cells in these tissues. Type I collagen belongs to the fibrillar class of collagens and is composed of two α 1(I) chains and one α 2(I) chain (Cole, 1994; Philanjaniemi and Rehn, 1995, and references therein for review of collagen types and functions). Excessive amounts of these proteins along with other extracellular matrix proteins are found in fibrotic diseases including cirrhosis, glomerulosclerosis, and scleroderma.

Transcription of the two genes for the type I collagen polypeptides is coordinately regulated (Vuust et al., 1985), suggesting that similar transcriptional mechanisms probably regulate both genes. Our long-term objective is to define the sequence elements in the type I collagen genes necessary for correct expression in different cell lineages and at various times in development and, subsequently, to identify the DNA-binding proteins that interact with these sequences. Earlier studies with transgenic mice, using promoter sequences of the mouse *pro α 2(I)* collagen gene from -2,000 to +54 bp linked to a CAT reporter gene, showed that CAT enzyme activity was present in many tissues where high levels of the endogenous gene are expressed, such as tail, skin, and bone (Khillan et al., 1986). However, using CAT assays of tissue extracts, it was not possible to localize the precise cells that expressed the transgene. Therefore, transgenic mice containing a 2,000-bp promoter linked to a *lacZ* reporter gene were generated; these mice expressed the transgene in many cells that express the corresponding endogenous gene (Niederreither et al., 1992). Similar qualitative results were obtained with a 350-bp *pro α 2(I)* collagen proximal promoter, except that the levels of expression were lower than with the 2,000-bp promoter and expression in osteoblasts was either absent or occurred at very low levels (D'Souza et al., 1993).

Address all correspondence to Dr. Benoit de Crombrughe, Department of Molecular Genetics, Box 11, 1515 Holcombe Blvd., The University of Texas M.D. Anderson Cancer Center, Houston, TX 77030. Tel.: (713) 792-2590. Fax: (713) 794-4295.

Jerome Rossert's current address is INSERM U64, Hopital Tenon, 75020, Paris, France.

Karen Niederreither's current address is IGBMC, Illkirch Cedex 67404, France.

The first two authors contributed equally to this work.

Although these *pro α 2(I)* collagen promoter-containing transgenes were expressed in a specific manner and at the correct stage of development, the levels of expression of the reporter gene appeared to be lower than the levels of the endogenous gene, suggesting that additional important enhancer elements might be present in other parts of the gene. Earlier transient expression experiments in tissue culture cells provided evidence for the presence of a transcriptional enhancer in the first intron between +418 and +1524 (Rossi and de Crombrughe, 1987; Pogulis and Freytag, 1993), but this sequence did not appear to affect the levels or pattern of expression of the 2,000-bp promoter in transgenic mice (Goldberg et al., 1992). We have examined the DNA segments upstream of the mouse *pro α 2(I)* collagen gene in order to find sequences that enhance expression of the proximal promoter in transgenic mice. We report here the identification of a far-upstream enhancer that substantially increases the activity of the reporter gene, confers a high degree of position independence on the transgene, and can drive a 220-bp mouse *pro α 1(I)* promoter in a cell-specific manner.

Materials and Methods

Isolation of Nuclei and DNase I Digestion

Nuclei were isolated by a modification of the procedure of Bellard et al. (1989). For all the solutions, 1 mM DTT and protease inhibitors (phenylmethylsulfonyl fluoride at 1 mM and pepstatin, leupeptin, and aprotinin at 10 μ g/ml) were used. Four large plates (25 \times 25 cm) of each cell line, NIH/3T3 (American Type Culture Collection, Rockville, MD), MC3T3-E1 (Dr. Farach-Carson, The University of Texas Dental School), EL4 (Dr. Platsoukas, The University of Texas M. D. Anderson Cancer Center, Houston, TX), and NMuLi (American Type Culture Collection), were washed twice with ice-cold PBS. The cells were scraped in 3 ml of PBS and centrifuged at 1,500 rpm for 10 min in a tabletop centrifuge (model GLC-2B; DuPont/Sorvall Instruments, Wilmington, DE). Mouse tissues were dissected from newborn B6D2F1 mice and washed in ice-cold PBS. Brain and liver tissues were then minced with scissors and treated as described below. Skin from newborn mice was cut into small strips and floated on a solution of Dispase (Collaborative Research, Inc., Waltham, MA) at 37°C for 30 min. The epidermis was then stripped from the dermis and discarded. The dermis was chopped into small pieces with a razor blade and digested for 1 h at 37°C in a solution of 3 mg/ml collagenase (Sigma Immunochemicals, St. Louis, MO) in DME. The digested dermis was filtered through cheesecloth. The filtrate was centrifuged at 1,500 rpm for 10 min in a Sorvall GLC-2B tabletop centrifuge. The cell pellet was either used immediately to prepare nuclei or plated out in DME with 10% fetal calf serum at 37°C in 8% CO₂ for 36 h and harvested as described above for cultured cells.

The cell pellet or fragments of tissue were resuspended in 10 ml of buffer 1 (15 mM Tris HCl, pH 7.5; 15 mM NaCl; 60 mM KCl; 1 mM EDTA, pH 8.0; 0.5 mM EGTA, pH 8.0; 1.9 M sucrose; 0.1% Triton X-100; 0.5 mM spermidine; and 0.15 mM spermine) and transferred to a Dounce homogenizer for lysis (Kontes Glass Co., Vineland, NJ). For tissues, the lysate was filtered through cheesecloth to remove any fibrous elements. 10 ml of buffer 2 (buffer 1 without Triton X-100) was added to the lysed cells followed by enough buffer 3 (buffer 1 without Triton X-100 or sucrose) to bring the refractive index of the solution to 1.40–1.42. The suspension was centrifuged at 4°C in an SS34 rotor at 10,000 rpm for 10 min, and the pellet was resuspended in 5 ml of buffer 4 (buffer 1 containing 0.34 M sucrose without EDTA, EGTA, or Triton X-100). The absorbance of 5 μ l of resuspended nuclei was measured at 260 nm. The nuclei were then pelleted again by centrifugation for 10 min at 10,000 rpm in an SS34 rotor, and the pellet was resuspended in buffer 4 at an absorbance of \sim 0.2 OD unit/ μ l.

For each digestion, 15 OD units of nuclei were used. The volume of the nuclei was brought to 80 μ l by the addition of buffer 4. 10 μ l of DNase I assay buffer (400 mM Tris HCl, pH 7.5; 60 mM MgCl₂) and 10 μ l of diluted DNase I (Pharmacia LKB Biotechnology, Inc., Piscataway, NJ) (concentrations 0–6 U/ μ l) were added. The nuclei were incubated at 37°C

for 15 min. The reaction was stopped by the addition of 200 μ l of stop buffer (50 mM Tris HCl, pH 8.0; 100 mM EDTA, pH 8.0; 100 mM NaCl; and 1% SDS) and 40 μ l of Proteinase K (20 mg/ml). The nuclei were then incubated overnight at 55°C to digest proteins. The DNA was extracted by phenol/chloroform, precipitated with ethanol, and resuspended in 100 μ l of Tris EDTA, pH 8.0. Aliquots from the DNase I-digested nuclei of each cell line were digested with the indicated restriction endonucleases, followed by electrophoresis through a 1% agarose gel. The DNA fragments were transferred to Zeta-probe GTG membrane (BioRad Labs, Hercules, CA), and the blot was hybridized according to the manufacturer's recommendation.

DNA Constructs

The 5' flanking sequences of the mouse *pro α 2(I)* collagen gene were derived from two sources: a cosmid clone (Cosmid 1) or a λ FIX II genomic clone (λ 3-3). The cosmid library contained genomic DNA from B6D2F1 mice partially digested with BamHI and cloned into the BamHI site of the vector pWE15 (Stratagene, La Jolla, CA). Cosmid 1 contained sequences of the *pro α 2(I)* collagen gene from the BamHI site at -17 kb to the BamHI site at $+2.5$ kb. The λ FIX II library (Stratagene) contained genomic DNA from 129 mice partially digested with Sau3AI and cloned into the XhoI site of the vector. The clone λ 3-3 contained 5' flanking sequences of the *pro α 2(I)* collagen gene from about -23.5 to -11 kb. Fig. 1 A contains a partial restriction map of these two clones, and Fig. 2 shows diagrams of the constructs derived from these sequences.

The following two reporter gene vectors were modified for use in generating transgenic mice: pLacF (Mercer et al., 1991) and pA₃Luc (Wood et al., 1989). The lacZ vector was modified by the addition of short linkers containing unique restriction enzymes in the polylinker regions to generate pLacH and pLacRM; the luciferase vector was similarly modified to generate pLuc32 and pLucG. Various fragments of the mouse *pro α 2(I)* collagen upstream sequences were cloned into the polylinker regions of these four modified vectors as described below.

A 4-kb fragment from SphI (+54) to HpaI (-3.8 kb) was blunt ended with T4 polymerase and cloned into the SmaI site of pLacH (*lacZ*) and pLuc32 (luciferase) to generate pGB4lac and pGB4luc, respectively. A 6.5-kb fragment from SphI (+54) to BamHI (-6.5 kb) was cloned blunt end using a similar strategy to generate pGB6.5lac and pGB6.5luc. For the 17-kb constructs, PCR was used to convert the SphI site at +54 into a unique XhoI restriction site. The sequences between the XmaI site (-108 bp) and the XhoI site (+54 bp) were subcloned in pBluescript KS (Stratagene) between XhoI and XmaI to yield pBS54. The sequences between -17 kb (NotI) and -108 bp (XmaI) were isolated from Cosmid 1 and cloned into pBS54 between NotI and XmaI to yield pBS17 from -17 kb (NotI) to +54 bp (XhoI). This fragment was excised and cloned into pLacRM and pLucG between XhoI and NotI to yield pGB17lac and pGB17luc, respectively.

For the remaining constructs, only *lacZ* vectors were used. Various fragments of the 5' flanking sequences were cloned upstream of a proximal *pro α 2(I)* collagen promoter. The proximal promoter (-350 to +54 bp) was inserted blunt end into the XbaI site of pLacRM to yield plac350. Four different fragments isolated from subclones of Cosmid 1 were inserted upstream of this 350-bp promoter by blunt-end ligation into the XhoI site to yield pGB11.5/6.5, pGB17/11.5, pGB13.5/11.5, and pGB17/13.5. The insert from λ 3-3 was isolated and cloned in both orientations into the NotI site of plac350 to generate pGB23.5/11 and pGB23.5/11-rev.

Two additional constructs were made by subcloning into a modified form of pBluescript containing a BglII site. The 6-kb BglII fragment of λ 3-3 (-13.5 to -19.5 kb) was cloned into this modified pBluescript between BglII and BamHI. The insert was removed by XhoI and NotI and cloned into plac350 to yield pGB19.5/13.5 or removed by BglII and NotI and cloned into m.p.lacZ (Rossert et al., 1996) to yield pGB α 1(I)19.5/13.5.

Generation of Transgenic Mice

The transgenic mice were generated using mouse *pro α 2(I)* collagen gene constructs as shown in Fig. 2. For constructs in the pLacH vector (pGB4lac and pGB6.5lac), vector sequences were removed with XhoI and SphI. For constructs in the pLuc32 vector (pGB4luc and pGB6.5luc), vector sequences were removed with AatII and XhoI. For constructs in the pLacRM and pLucG vectors (pGB17lac, pGB17luc, pGB11.5/6.5, pGB17/11.5, pGB13.5/11.5, pGB17/13.5, pGB23.5/11, pGB23.5/11-rev, and pGB19.5/13.5), vector sequences were removed with SacII and NruI. For the pGB α 1(I)19.5/13.5 construct, vector sequences were removed with

XhoI and NarI. The digested DNA constructs were gel purified, resuspended at a concentration of 2 ng/ μ l and microinjected into the pronuclei of B6D2F2 fertilized eggs as described previously (Hogan et al., 1986). The injected eggs were then transferred into CD1 foster mothers. The integration of the transgenes in founder mice was assessed by Southern blot analysis of DNA derived from either the tails of 1-mo-old animals or the placentas of embryos.

Assays for the Expression of Reporter Genes

β -Galactosidase activity was assayed on whole embryos as previously described (Mercer et al., 1991). For embryos older than E15.5, the skin was removed before fixation and staining. Embryos were fixed for 45 min and then stained overnight at room temperature. For histological analysis, 8- μ m sections were cut and counterstained with eosin. Some newborn pups were partially clarified in potassium hydroxide (McLeod, 1980).

Luciferase activity was assayed in tissues of 4-wk-old mice containing the transgene as previously described (Niederreither et al., 1992). The results are expressed as luciferase light units per μ g of protein.

Results

DNase I-Hypersensitive Sites in the *pro α 2(I)* Collagen Gene

DNase I-hypersensitive sites in chromatin are frequently associated with control regions that confer high levels of expression or that provide position-independent expression in transgenic mice. For several tissue-specific genes, DNase I-hypersensitive sites have been found far upstream of the promoter sequences (Grosveld et al., 1993; Kaltschmidt et al., 1994; Zhong et al., 1994). These sites are thought to play a role in regulating expression of linked genes through the binding of sequence specific transcription factors. An earlier study in our laboratory showed that a cell-specific DNase I-hypersensitive site was present at about -100 bp in the promoter of the mouse *pro α 2(I)* collagen gene (Liau et al., 1986). A similar hypersensitive site has also been detected in the chick *pro α 2(I)* collagen gene (Beck et al., 1991), as well as other sites in the body of the gene. Hypersensitive sites have also been detected in the mouse and human *pro α 1(I)* promoter region and gene (Barsh et al., 1984; Breindl et al., 1984). To find elements that might enhance expression of the mouse *pro α 2(I)* gene, we searched for additional DNase I-hypersensitive sites in chromatin upstream of the proximal promoter.

A cosmid clone (Fig. 1 A, Cosmid 1) containing 5' flanking sequences of the mouse *pro α 2(I)* collagen gene was isolated from a library of mouse genomic DNA partially digested with BamHI. Restriction mapping of this cosmid clone showed that it contained sequences from -17 kb upstream to +2.5 kb downstream of the start of transcription. For sequences further upstream, we isolated a λ FIX II genomic clone (λ 3-3), which contained sequences from -23.5 to -11 kb upstream of the start of transcription. We used fragments from these two clones to probe Southern blots containing DNA from nuclei treated with DNase I and subsequently digested with the appropriate restriction enzymes (Fig. 1 A).

The first probe, P1 (a 1.4-kb BamHI/EcoRI fragment), hybridized to an 8-kb BamHI fragment (-6.5 to +1.5 kb) (Fig. 1, A and B). In DNA derived from DNase I-treated NIH/3T3 fibroblast nuclei, two additional species, 6.4 and 4.4 kb, could be detected; these species corresponded to

hypersensitive sites at approximately -100 bp and -2.1 kb, respectively (Fig. 1 B). These hypersensitive sites were not detectable in DNA derived from DNase I-treated nuclei of EL4 cells, a T cell line which does not produce type I collagen. The site located at -100 bp, designated HS1, has been previously described (Liau et al., 1986). The other site (HS2), located at -2.1 kb, was significantly less hypersensitive, which could be deduced by the relative intensities of the bands. The presence and location of both hypersensitive sites were confirmed by rehybridizing the blot with a probe homologous to the other end of the 8-kb BamHI fragment (data not shown).

The same blot was stripped and reprobed with fragment P2 (a 900-bp Asp718/BamHI fragment), which binds to a 10.5-kb BamHI fragment (-17 to -6.5 kb) (Fig. 1 A). Two hypersensitive sites were seen in NIH/3T3 chromatin between -15 and -17 kb upstream of the transcription start site, but not in EL4 chromatin (data not shown). No hypersensitive sites were detected between -6.5 and -15 kb in either cell line. To further characterize the hypersensitive sites between 15 and 17 kb upstream, we isolated nuclei from two additional cell lines, MC3T3-E1 (an osteoblastic cell line derived from mouse calvaria) and NMuLi (a liver cell line). MC3T3-E1 cells express type I collagen (Sudo et al., 1983), whereas NMuLi cells do not (data not shown). After DNase I treatment of nuclei, the DNA from the four cell lines was digested with BglII and probed with P3 (1.3-kb AccI/BglII fragment), which hybridized to a 6-kb BglII fragment (-19.5 to -13.5 kb) (Fig. 1 A). Three hypersensitive bands could be seen in DNA from NIH/3T3 and MC3T3-E1 cells but not in DNA from EL4 or NMuLi cells (Fig. 1 C). These sites were mapped to approximately -15.5, -16.2, and -17.1 kb and designated HS3, HS4, and HS5, respectively. The presence and location of these hypersensitive bands was confirmed by rehybridizing with a probe homologous to the other end of the 6-kb BglII fragment (data not shown). The same blot was stripped and rehybridized to fragment P4 (a 600-bp BamHI/BglII fragment), which binds to an 8.5-kb BglII fragment from -28 to -19.5 kb. No major hypersensitive sites were seen with this probe (data not shown).

We also isolated nuclei from tissues of newborn mice to determine whether the hypersensitive sites that were detected with cultured cells were also present in mouse tissues. We used newborn mouse skin, which expresses high levels of type I collagen as a positive tissue, and mouse brain, which expresses very little type I collagen as a negative tissue. Nuclei derived from brain and skin were treated with DNase I and subsequently digested with BamHI and probed with P1 to show the presence of HS1 and HS2. These hypersensitive sites were only evident in DNA derived from skin fibroblasts (data not shown). We then digested the DNase I-treated DNAs with BglII and probed with P3 to show HS3-5. Fig. 1 D shows that these hypersensitive sites were detected in chromatin from skin fibroblasts but not in chromatin from brain cells. The location of the hypersensitive sites in chromatin of cultured cells and mouse tissues was identical. Moreover, we did not detect any new hypersensitive sites confirming the results obtained with cultured cell lines.

Five major hypersensitive sites were shown to be present in discrete locations in the region between -28 kb

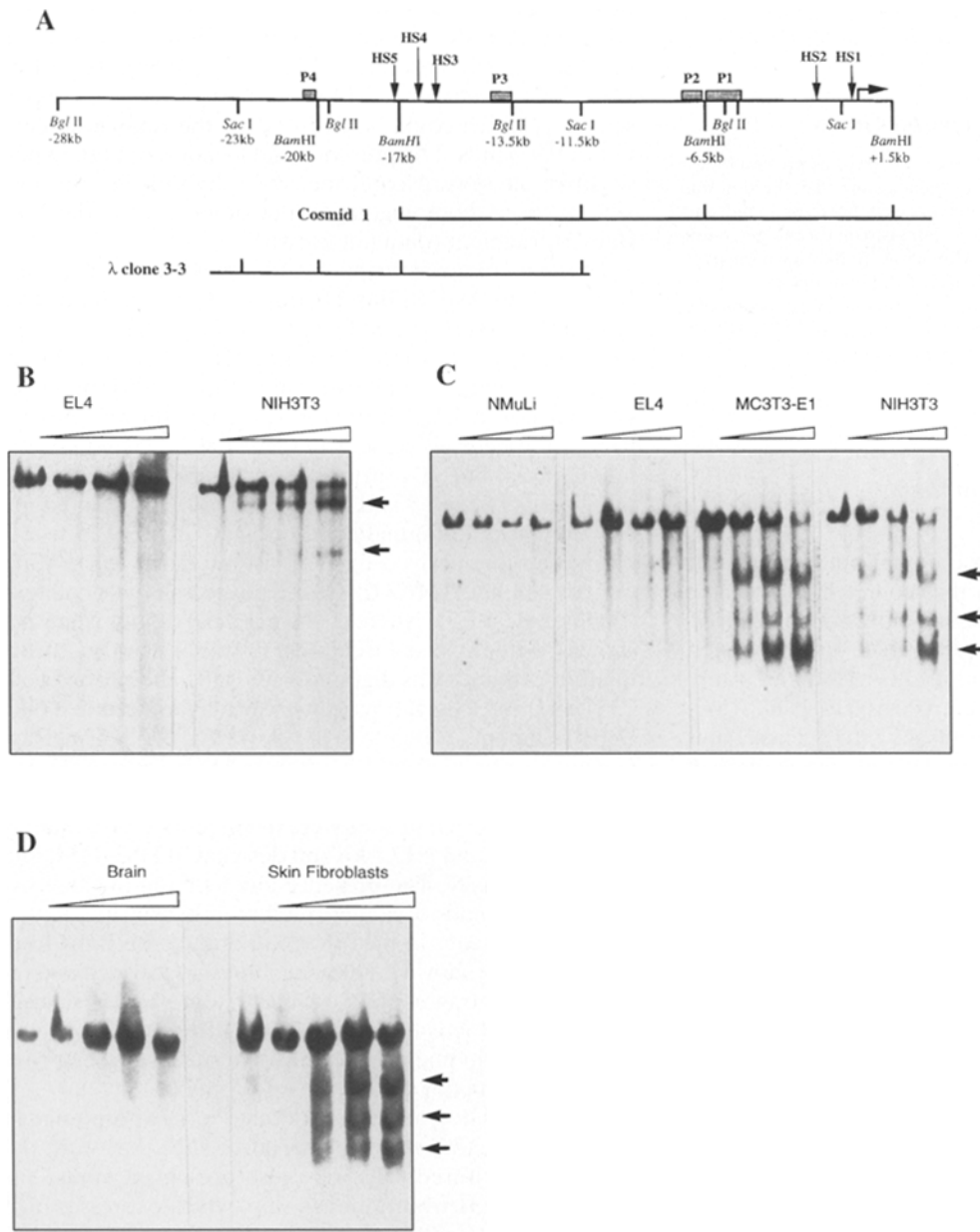


Figure 1. DNase I-hypersensitive sites in chromatin of the 5' flanking region of the mouse *proa2(I)* collagen gene. (A) A partial restriction map of the mouse *proa2(I)* collagen gene from -28 to +1.5 kb with respect to the start site of transcription. The locations of the five hypersensitive sites (*HS1-HS5*) are shown as well as the probes (*P1-P4*) used in hybridization. Diagrams of the two genomic clones used in this study, Cosmid 1 and λ clone 3-3, are also shown. In B-D, nuclei of the indicated cell lines and tissues were digested with increasing concentrations of DNase I. (B) DNA from fibroblasts (*NIH/3T3*) and T cells (*EL4*) was digested with *Bam*HI and hybridized with probe P1. The arrows indicate the position of the species derived from DNase I-digested DNA (6.4 and 4.4 kb). (C) DNA from fibroblasts (*NIH/3T3*), osteoblasts (*MC3T3-E1*), T cells (*EL4*), and liver cells (*NMuLi*) was digested with *Bgl*II and hybridized with probe P3. The arrows indicate the position of the species derived from DNase I-digested DNA (3.6, 2.7, and 2.0 kb). (D) DNA from brain cells and skin fibroblasts was digested with *Bgl*II and hybridized with probe P3. The arrows indicate the position of the species derived from DNase I-digested DNA (3.6, 2.7, and 2.0 kb).

and +1.5 kb of the mouse *proa2(I)* collagen gene. These sites were present in type I collagen-producing cells, but they were not found in cell types that do not produce type I collagen. The existence of these hypersensitive sites suggested that the DNA in chromatin at or around these sites was accessible to DNA-binding proteins in cells that synthesize type I collagen but not in other cell types. To test the effects of the segments containing the hypersensitive sites on transcriptional activity, we made several promoter-reporter gene constructs with 5' flanking sequences and introduced them into transgenic mice.

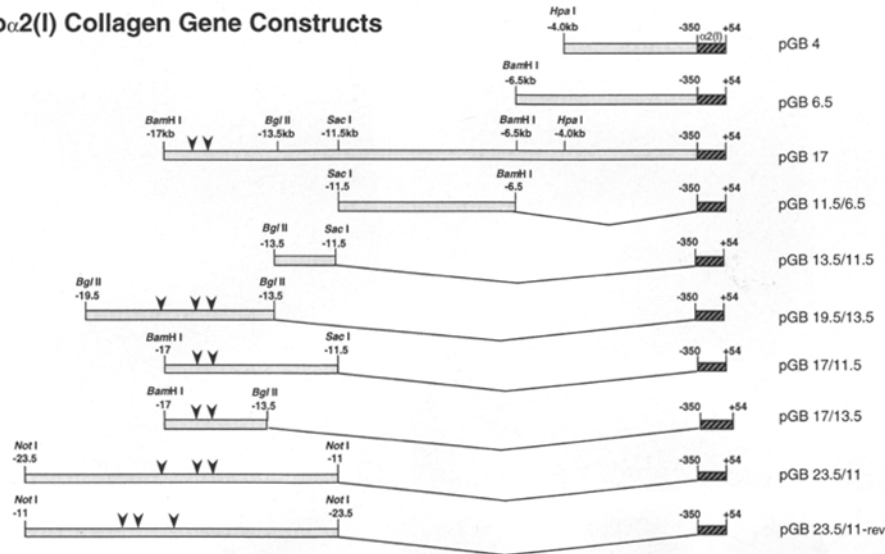
Generation of Transgenic Mice

Transgenic mouse lines were generated using the mouse

proa2(I) collagen gene constructs shown in Fig. 2 and described in detail in the Materials and Methods section. Mice harboring the pGB4, pGB6.5, and pGB17 constructs were coinjected with two separate DNA constructs: one containing the promoter fragment fused to the *lacZ* reporter gene and one containing the promoter fragment fused to the luciferase reporter gene. Animals found to be positive by Southern blot analysis were always positive for both reporter genes. For the other constructs shown in Fig. 2, only *lacZ* reporter genes were used.

For most of the DNA constructs that were microinjected, we sacrificed the foster mothers at 15.5 d of gestation and stained the founder animals for *lacZ* activity. However, for the pGB17, pGB6.5, pGB17/11.5, and pGB17/13.5 constructs, a small number of pups were allowed to

A pro α 2(I) Collagen Gene Constructs



B pro α 1(I) Collagen Gene Construct

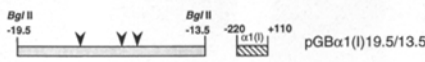


Figure 2. Diagrams of the various constructs introduced into transgenic mice. The locations of the three upstream hypersensitive sites HS3–HS5 are indicated by the arrowheads. (A) Constructs in which the mouse pro α 2(I) collagen far-upstream enhancer sequences are linked to the mouse pro α 2(I) collagen 350-bp proximal promoter. (B) A construct in which the mouse pro α 2(I) collagen far-upstream enhancer sequences are linked to the mouse pro α 1(I) collagen 220-bp proximal promoter.

come to term and positive founder animals were used to establish transgenic lines.

Expression of a pro α 2(I) Collagen lacZ Transgene in Whole Embryos

After overnight staining for β -galactosidase, embryos harboring either the 2,000-bp construct (–2,000lacZ) or the 350-bp construct (–350lacZ) showed relatively faint blue staining in various mesenchymal cells (Niederreither et al., 1992). LacZ staining was detectable in about 40% of the lines that were positive by Southern blot analysis (Table I). In the 12 embryos positive by Southern blot analysis for a construct that contained 3.8 kb upstream of the transcription start site (pGB4), expression of the lacZ transgene was not detectable (Table I). This finding was surprising because these constructs include the 2,000-bp promoter sequences, which direct expression of β -galactosidase in transgenic mice, albeit at relatively low levels. Similarly, among 13 embryos positive by Southern blot analysis for a construction containing 6.5 kb upstream of the transcription start site (pGB6.5), expression of the lacZ gene was not detectable (Table I). These data suggested that there were silencing elements located in the promoter upstream of 2,000 bp that were capable of strongly suppressing the transcriptional activity of this promoter. The pGB4 and pGB6.5 constructs contained DNA segments in which HS1 and HS2 were identified, whereas the previously injected 350- and 2000-bp promoter constructs contained DNA segments in which only HS1 was identified (Fig. 1). It is possible that the sequences around HS2 mediate the silencing effects seen in pGB4 and pGB6.5.

In contrast to the complete absence of X-gal staining in embryos transgenic for the 3.8- and 6.5-kb promoter sequences, six of the eight embryos harboring sequences

from –17 kb to +54 bp (pGB17), which were found to be positive by Southern blot analysis, expressed the transgene (Table I). Moreover, β -galactosidase staining was very strong in these embryos. In Fig. 3, A and B, we compare the intensities of X-gal staining at E15.5 in representative embryos of the –2000lacZ and the pGB17 transgenic mice. Embryos harboring the pGB17 construct were stained very strongly in the skin and appeared almost completely blue in whole-mount views, although not all cells in these mice were stained as revealed by histology (see below).

The pGB17 construct contained a DNA segment in which four hypersensitive sites, HS1 to HS4, were identified (Fig. 1). In this construct, the silencing effects seen with sequences in pGB4 and pGB6.5 were overcome by

Table I. Percent of Mice Expressing the lacZ Transgene

Construct	Positive by Southern	β -gal expressors	Percent expressors	Intensity of staining*
–350lacZ	8	3	38	+
pGB4	12	0	0	–
pGB6.5	13	0	0	–
pGB17	8	6	75	+++++
pGB11.5/6.5	2	0	0	–
pGB13.5/11.5	4	0	0	–
pGB19.5/13.5	6	4	67	+++++
pGB17/11.5	10	9	90	+++++
pGB17/13.5	5	4	80	+++++
pGB23.5/11	7	7	100	+++++
pGB23.5/11/11-rev	6	5	83	+++++
pGB α 1 (I)19.5/13.5	10	7	70	+++++

* Average levels of expression after overnight staining.

–, no detectable staining.

+, weak staining.

++++, strong staining.

+++++, very strong staining.

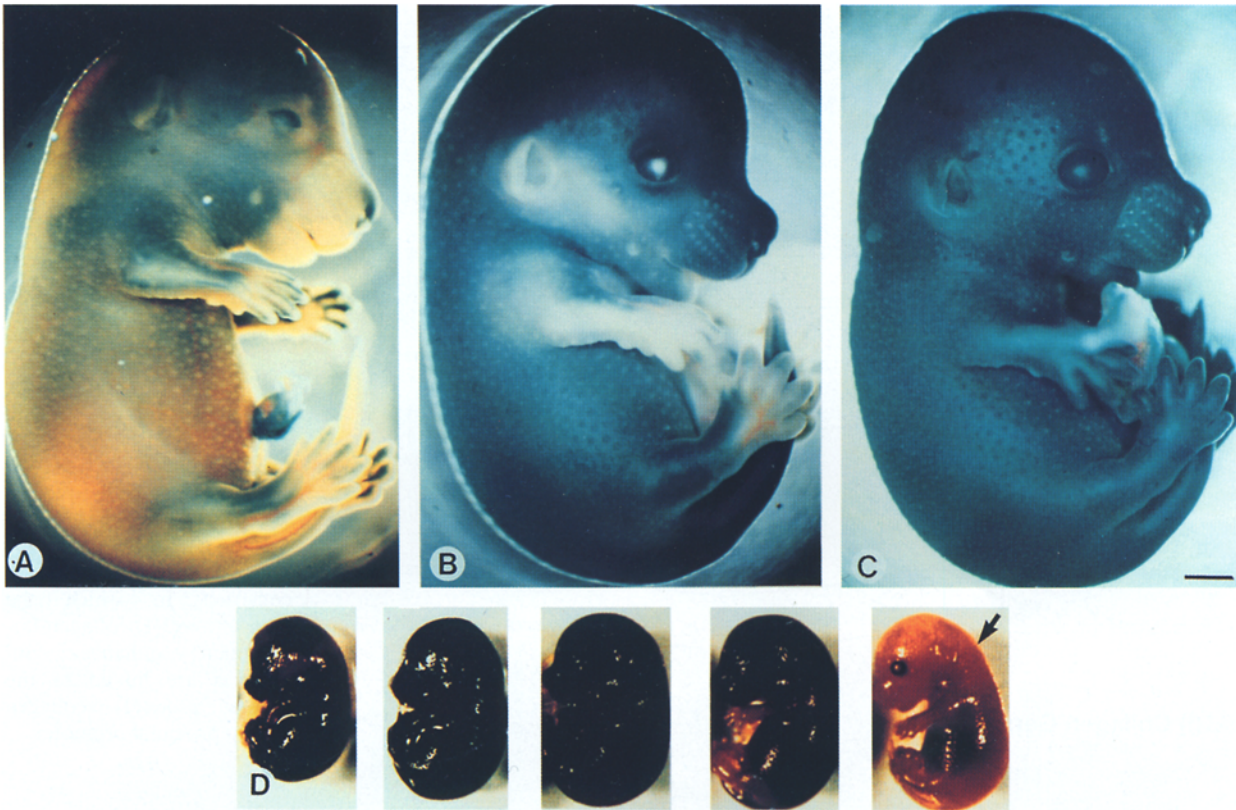


Figure 3. β -Galactosidase staining in typical whole-mount embryos at E15.5 harboring $-2,000lacZ$ (A), pGB17 (B), and pGB17/11.5 (C). Note the increased intensity of X-gal staining in embryos harboring constructs containing the far-upstream enhancer. The embryos shown in B and C are both transgenic founders, whereas the embryo shown in A is from a transgenic line. Differences in the staining pattern of embryos B and C could be due to chimerism in the founder animals. (D) The intensity of staining in founder animals harboring the pGB17/11.5 construct after 4 h of staining. The arrow points to a negative control. Bar, 1 mm.

sequences located further upstream. The presence of these upstream regions also resulted in a considerable increase in the activity of the *lacZ* reporter gene over the levels seen in mice harboring either 2,000 or 350 bp of the promoter, suggesting that a strong enhancer is present.

To further delineate the sequences necessary for enhancing activity, we cloned various fragments of the far-upstream region 5' to the 350-bp *pro α 2(I)* collagen promoter and generated transgenic mice. In positive E15.5 embryos, β -galactosidase activity correlated with the presence of the DNA segments in which the hypersensitive sites had been identified. Constructs with DNA segments in which the three upstream hypersensitive sites (HS3 to HS5) or only HS3 and HS4 had been identified gave strong *lacZ* expression in transgenic mice (Table I). These constructs included pGB17/11.5, pGB17/13.5, pGB19.5/13.5, pGB23.5/11, and pGB23.5/11-rev. As shown with the pGB23.5/11-rev construct, the far-upstream region was active in the reverse orientation as well as in the normal orientation. On the other hand, embryos transgenic for constructs, with DNA segments lacking HS3 to HS5, showed no X-gal staining. Specifically, two embryos positive by Southern blot analysis for the pGB11.5/6.5 construct and four embryos positive by Southern blot analysis for the pGB13.5/11.5 construct did not express *lacZ* (Table I). These results strengthen the notion that a silencer region

lies upstream of $-2,000$ bp and could extend upstream to approximately -13.5 kb. The effect of this silencer region was overcome by the presence of DNA segments in which some or all of the far-upstream hypersensitive sites had been identified.

The strongest levels of staining were seen in embryos harboring the pGB17/11.5 construct (Fig. 3 C). The X-gal staining of some of these embryos became visible in 60 min and was complete by 4 h (Fig. 3 D). Other embryos were only partially stained at 4 h but stained completely after overnight incubation with X-gal. Of 10 embryos positive by Southern blot analysis, nine expressed the transgene (Table I). A shorter construct, pGB17/13.5, was expressed in four of five mice positive by Southern blot analysis, whereas the pGB19.5/13.5 construct was expressed in four of six mice, the pGB23.5/11 construct was expressed in seven of seven mice, and the pGB23.5/11-rev construct was expressed in five of six mice (Table I).

Although the pGB19.5/13.5, pGB23.5/11, and the pGB23/5/11-rev constructs all contain DNA segments in which the three far-upstream hypersensitive sites (HS3 to HS5) were identified, the level or pattern of staining in embryos harboring this DNA was similar to that seen in embryos transgenic for the pGB17 construct or transgenic for the pGB17/11.5 or pGB17/13.5 constructs, which contain DNA segments in which only HS3 and HS4 were identi-

fied (data not shown). This may indicate that there is functional redundancy in these sites and that the presence of a segment containing the sequence in which two DNase I-hypersensitive sites were identified is adequate for optimal expression.

These results indicated that an enhancing region is located between -13.5 and -19.5 kb of the mouse *proa2(I)* collagen gene. The same region was shown to contain three cell-specific DNase I-hypersensitive sites, HS3 to HS5 (Fig. 1). We have denoted this region the far-upstream enhancer of the mouse *proa2(I)* collagen gene.

Luciferase Activity in Tissues of Transgenic Mice

Luciferase activity was assayed in homogenates of tissues from 4-wk-old mice that were shown by Southern blot analysis to harbor the transgene. Table II shows the levels of luciferase in tissue extracts of three strains harboring the pGB17 construct compared with two strains with the 350-bp luciferase construct. The much higher levels of luciferase activity seen in mice carrying the 17-kb promoter confirms the presence of an enhancer element in the larger construct. In one of the strains transgenic for the 17-kb promoter, the levels of luciferase activity in many tissues were considerably higher than in the two other strains. These differences were perhaps due to the site of integration of the transgene, but they did not appear to depend on the number of copies of the transgene integrated.

The highest levels of expression were seen in tail, followed by bone (calvarium) and skin, which are tissues that express high levels of the endogenous gene. Since all organs, including soft tissues such as liver and brain, contain some mesenchymal cells or fibroblasts, which produce type I collagen, it was not surprising to see expression of the transgene in every tissue tested, but the values of luciferase activities in tissues which express relatively little type I collagen were at least 20–100 times lower than the values in tail.

The luciferase activities in tissues of mice transgenic for the pGB6.5 construct were much lower than the activities in tissues of mice harboring the pGB17 construct. These activities were also lower than those in mice carrying the 350-bp proximal promoter. Indeed, the expression in several tissues became undetectable in mice transgenic for the 6.5-kb construct (Table II). Overall, the results of luciferase assays in different tissues confirmed the *lacZ* staining patterns, which indicated that there was no detectable staining in transgenic embryos carrying pGB6.5, weak staining in embryos harboring -350*lacZ*, and strong staining in embryos transgenic for pGB17, thereby demonstrating that there was a potent enhancer element present in the pGB17 construct and a silencer element included in the pGB6.5 construct.

Histochemical Localization of β -Galactosidase Activity in E15.5 Embryos

X-gal-positive embryos were examined at 15.5 d of embryonic development because many bone primordia begin to ossify at this stage, but the skin has not yet fully developed, allowing penetration of the X-gal substrate in intact embryos. In all embryos transgenic for the constructs con-

Table II. Luciferase Levels in Tissues of 4-wk-old Transgenic Mice*

Luciferase copy number	350-bp lines		6.5-kb lines		17-kb lines		
	1	1	1	3	3	5	1
Tissue							
Tail	140	220	80	3.2	1,900,000	71,000	34,000
Skin	260	350	12	0.54	47,000	5,500	6,300
Calvarium	0.14	0.88	0	0.57	42,000	8,500	6,900
Liver	0	0	0	0	610	45	44
Heart	1.1	1.1	0	0	1,900	400	250
Lung	0	0	0	0	7,900	2,500	1,300
Brain	0.28	65	0	0.1	720	200	250
Kidney	0	6.6	0	0	3,300	420	300
Thymus	0	0.44	5	2.9	3,900	120	63
Intestine	0	0.88	0	0	3,500	1,200	480

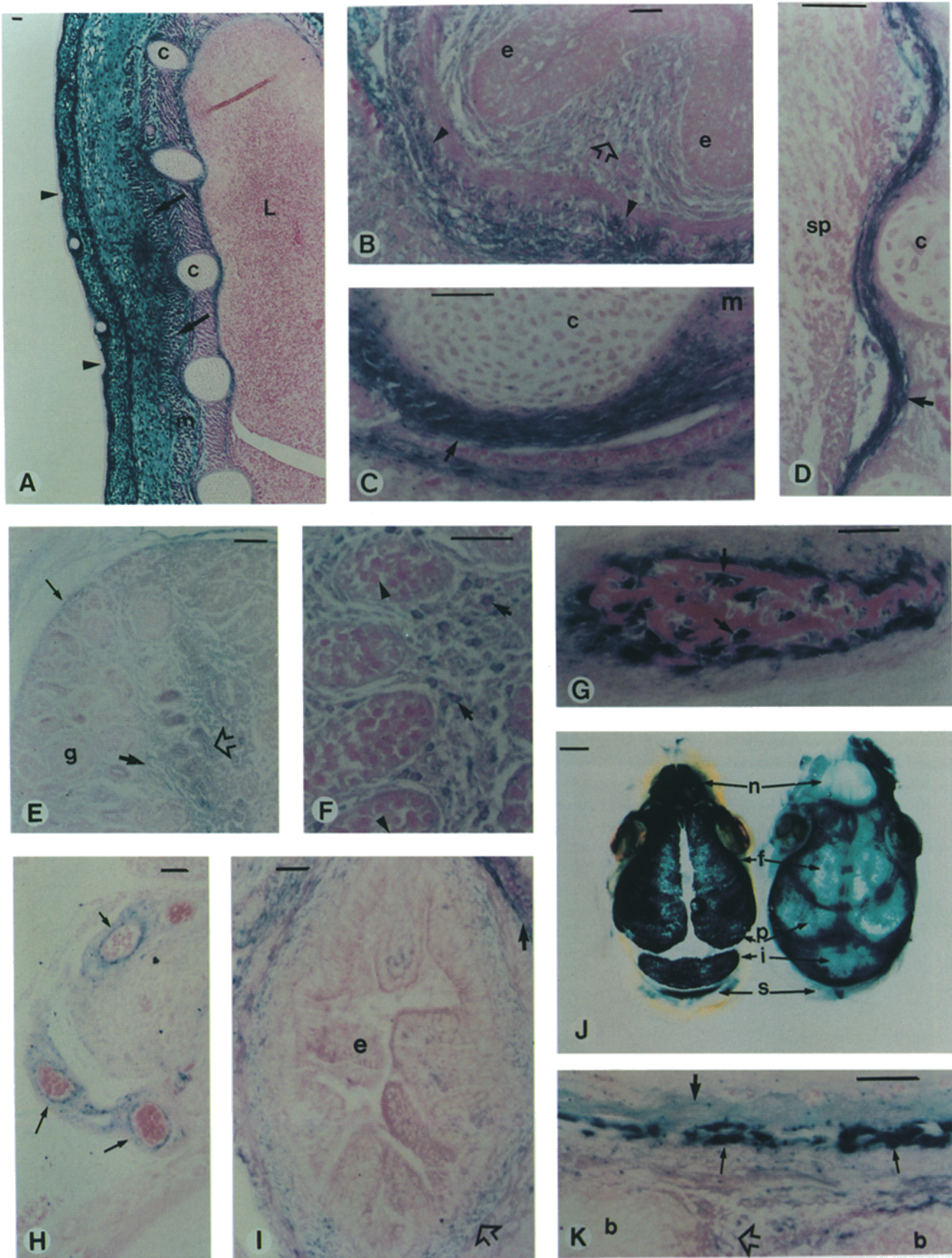
*Luciferase values are expressed as Light Units per μ g protein.

taining the far-upstream enhancer sequences, the *lacZ* gene was expressed in a very similar pattern (Table III).

E15.5 transgenic embryos showed expression of *lacZ* in many but not all cell types in which the endogenous gene is active. In all embryos examined, fibroblast-specific X-gal staining was detected in the connective tissue layers of the dermis and fascia. Fig. 4 A shows a parasagittal section of an embryo in which very intense staining was evident in the connective tissue layers between the epidermis and the muscle. Expression also occurred in fibroblasts between the muscle bundles, whereas the liver cells and the chondrocytes of the ribs were completely negative, as were the epidermal cells. Expression in fibroblasts throughout the gastrointestinal tract was also seen. X-gal staining was present in the fibrous elements of the submucosa and muscularis layers of the stomach but not in the epithelial mucosal cells (Fig. 4 B). Fibroblasts in many tendons were also stained with X-gal (Fig. 4 C). Fibroblasts in the meninges (Fig. 4 D) surrounding the spinal cord and the brain were positive for X-gal. The kidney capsule and con-

Table III. Expression Pattern of the *LacZ* Transgene

	E15.5					Newborn	
	pGB17	pGB 17/11.5	pGB 17/13.5	pGB 19.5/13.5	pGB α 1 19.5/13.5	pGB17	pGB 17/13.5
OSTEOBLASTS							
Frontal/Parietal	+	+	+	+	+	+	+
Clavicle	+	+	+	+	+	+	+
Mandible	+	+	+	+	+	+	+
Other bones	-	-	-	-	-	-	-
Odontoblasts	ND	ND	ND	ND	ND	-	-
FIBROBLASTS							
Dermis/Fascia	+	+	+	+	+	+	+
Tendon	\pm	\pm	\pm	\pm	\pm	\pm	\pm
Meninges	+	+	+	+	+	+	+
GI Tract	+	+	+	+	+	+	+
Kidney	+	+	+	+	+	+	+
Testis	+	+	+	+	+	+	+
Blood vessels	-	-	-	-	-	+	+



nective tissue cells in the stroma were positive for X-gal staining (Fig. 4 E), whereas the cells in the glomeruli were not. In the testis, the stromal fibroblasts were X-gal positive but not the primitive Sertoli cells and spermatogonia in the testicular cords that mature into the seminiferous tubules (Fig. 4 F).

Bone ossification starts at E13.5 to E14.5 in the clavicle, followed at later times by other bones (Rugh, 1990). At E15.5, X-gal staining of embryos was seen in some ossification centers (Table III), i.e., in osteoblasts of the frontal and parietal bones of the skull, in the clavicle (Fig. 4 G), and in the maxillary and mandibular bones, but very little in other bones. Occasional blue-stained cells, which appeared to be osteoblasts by their morphology, were associated with regions of subperiosteal ossification of long bones in some embryos; however, there was no X-gal staining detected in regions of endochondral ossification. In human embryos, the frontal and parietal bones of the skull are formed entirely by intramembranous ossification, whereas the maxilla, mandible, and clavicle are formed partially by intramembranous ossification and partially by endochondral ossification (Gardner, 1971). Hence, the pattern of X-gal staining in forming bones of E15.5 embryos appeared to largely correlate with the presence of osteoblasts in ossification centers that were generated by intramembranous ossification.

In conclusion, X-gal staining of embryos transgenic for constructs containing the far-upstream enhancer was much stronger than in embryos harboring either a 2,000- or a 350-bp promoter. Although the intensity of staining varied somewhat from embryo to embryo, staining was specific for cells of mesenchymal origin, including fibroblasts and some osteoblasts. No staining was detected in various cell types that did not express the endogenous type I collagen gene, such as epithelial cells, neuronal cells, or the parenchymal cells of the spleen, liver, etc. Clearly, however, not all type I collagen-producing cells expressed the transgene. Possible reasons for the lack of transgene expression in many osteoblasts are discussed in more detail later.

Histochemical Localization of β -Galactosidase Activity in Newborn Pups

Three newborn pups harboring the pGB17 construct and two newborn pups harboring the pGB17/13.5 construct were examined. We observed significant X-gal staining surrounding blood vessels in newborn mice that was not seen in embryos at E15.5. Fig. 4 H shows the staining around some of the blood vessel walls of newborn pups. After removal of the skin and clarification (McLeod, 1980), it was possible to see the overall pattern of X-gal staining in bones of the pups. Fig. 4 J shows the heads of two different newborn pups, one harboring the pGB17/13.5 construct and one harboring the 3.2-kb *pro α 1(I)* collagen promoter (3,200lacZ). It was previously shown that pups containing either 2.3 or 3.2 kb of the mouse *pro α 1(I)* collagen promoter strongly expressed a *lacZ* transgene in osteoblasts of all bones, including those of the skull (Rossert et al., 1995). As expected, the pup harboring the 3.2-kb *pro α 1(I)* collagen promoter showed very strong expression in the various skull bones; however, the pup harboring pGB17/13.5 revealed a much more limited and less intense pattern of staining (Fig. 4 J). Strong staining was seen mainly in the sutures where there are many fibroblastic cells and recently differentiated osteoblasts. Histological sectioning confirmed that cells that appeared to be osteoblasts by their morphology were staining as well as fibroblasts in the connective tissue (data not shown). Similar patterns of expression of the transgene in the skull were seen with the other newborn pups examined (data not shown). These data suggested that the majority of transgene expression in the bones of these pups was confined to areas of new bone growth where osteoblasts were in the process of differentiating from their mesenchymal precursors. As with E15.5 embryos, occasional expression was also detected weakly in the periosteal region of long bones, including the ribs (data not shown).

To determine whether odontoblasts were positive, sections through the jaw were made from several newborn

Figure 4. Micrographs of histological sections showing the localization of β -galactosidase staining. All embryos harboring constructs containing the far-upstream enhancer showed similar patterns of expression. The following micrographs were chosen to illustrate the general patterns seen. (A–G) Sagittal sections of embryos at E15.5. (A) An embryo harboring pGB17/11.5 with staining in connective tissue cells of the dermis, the fascia, and the fibroblasts (arrows) of the skeletal muscle (*m*). Note the lack of staining in the cells of the epidermis (arrowheads), the liver (*L*), and the chondrocytes (*c*) of the ribs. (B–F) Sections from embryos harboring the pGB17 construct. (B) Staining is seen in the smooth muscle and fibroblastic cells in the muscularis externa (arrowheads) and in the submucosal (open arrow) layers of the stomach wall, whereas the epithelial cells (*e*) of the rugal folds are not stained. (C) Staining in the tendon (arrow) of the elbow merging into the surrounding skeletal muscle (*m*). The chondrocytes (*c*) of the developing humerus are negative. (D) The meninges (arrow) surrounding the spinal cord (*sp*) were stained, while the neural tissues and the chondrocytes of the ribs (*c*) were negative. (E) Staining in the fibroblastic cells of the kidney capsule (small arrow) and connective tissue (large arrow). There is also some staining associated with tubular cells (open arrow) but not with the glomeruli (*g*). (F) The testis in which the cytoplasm of the interstitial cells (arrows) were stained, but the primitive spermatogonia and Sertoli cells (arrowheads) were not. (G) The clavicle of an embryo harboring the pGB19.5/13.5 construct with blue-stained osteoblastic cells (arrows) associated with the osteoid. (H) Histological section of a newborn pup harboring pGB17/13.5 showed staining in several of the small arteries of the body wall (arrows), while the chondrocytes of a rib and surrounding muscle were negative. (I) Expression in the fibrous submucosal (open arrow) and muscular layers (arrow) of the intestine of an E15.5 pup harboring the pGB α 1(I)19.5/13.5 construct. No staining is seen in the epithelial cells (*e*). (J) Whole-mount staining of newborn pups from two transgenic lines after removal of the skin showed the overall expression pattern in the bones of the skull. Note that the 3.2-kb *pro α 1(I)* promoter (left) was expressed strongly throughout the skull bones, whereas the pGB17/13.5 construct (right) was most strongly expressed in the sutures. The bones of the skull are indicated as nasal (*n*), frontal (*f*), parietal (*p*), interparietal (*i*), and supraoccipital (*s*). (K) In the frontal bone of the skull of an E15.5 pup harboring the pGB α 1(I)19.5/13.5 construct, osteoblasts are strongly stained (small arrows) and the overlying fibroblasts are stained less intensely (large arrow). There is also staining in meningeal fibroblasts (open arrow) between lobes of the brain but not in the brain itself (*b*). (A–I, K) Bars, 100 microns. (J) Bar, 1 mm.

pups. These sections clearly showed that odontoblasts did not stain with X-gal, whereas osteoblasts in the surrounding mandibular bone were positive (data not shown). Again, this contrasts with the pattern seen with newborn pups harboring 2.3 or 3.2 kb of the mouse *pro α 1(I)* collagen promoter in which the odontoblasts are strongly stained (Rossert et al., 1995). Table III summarizes the tissues in which mesenchymal and fibroblastic cells were X-gal positive in newborn pups and embryos at E15.5.

Activity of the Enhancer Cloned Upstream of the Mouse *pro α 1(I)* Collagen Minimal Promoter

The short mouse *pro α 1(I)* collagen promoter from -220 to +110 bp linked to a *lacZ* transgene has been shown to be essentially silent in transgenic mice (Rossert et al., 1996). We asked whether the far-upstream enhancer segment of the *pro α 2(I)* collagen gene from -19.5 to -13.5 kb could activate this promoter. The rationale was that any expression detected would likely be dependent on the presence of the *pro α 2(I)* collagen enhancer region.

As with constructs linked to the *pro α 2(I)* collagen proximal promoter, staining in whole-mount embryos was very strong (data not shown). Histological analysis of several pups containing the chimeric construct (pGB α 1(I)19.5/13.5) showed that staining in these embryos was specific to type I collagen-producing cells with a pattern of expression identical to that seen in mice transgenic for the pGB17 construct. Fig. 4 I shows expression of the pGB α 1(I)19.5/13.5 construct in intestine at E15.5. Fig. 4 K shows expression in the osteoblasts of the calvaria as well as fibroblasts of the fascia and the meninges of the brain. These data suggest that the far-upstream enhancer of the mouse *pro α 2(I)* collagen gene is capable of conferring tissue-specific expression to a linked promoter.

Discussion

We have identified an enhancer in the mouse *pro α 2(I)* collagen gene located in the region between 13.5 to 19.5 kb upstream of the start site of transcription, which greatly increases expression of reporter genes in transgenic mice. This enhancer exerts its effects at various distances from the promoter and in either orientation. Moreover, the enhancer directs high levels of cell-specific expression of the *pro α 1(I)* collagen minimal promoter, which has been shown to be virtually silent in transgenic mice. In chromatin, this element was characterized by the presence of three cell-specific DNase I-hypersensitive sites located 15 to 17 kb upstream of the start site of transcription. The level of expression of luciferase and *lacZ* reporter genes in mice with constructs containing the far-upstream enhancer was significantly higher than the levels seen in constructs lacking the enhancer.

We believe that this sequence is important in controlling expression of the endogenous gene since the hypersensitive sites were found only in cells that produce type I collagen, suggesting that these sites play a role in the transcription of a gene expressed only by these cell types. DNase I digestion of chromatin derived from skin fibroblasts and cell lines producing type I collagen showed the presence of five major hypersensitive sites in the region

between the BamHI site at +1.5kb and the BglII site at -28 kb. None of these hypersensitive sites was detected in the chromatin of brain cells or in cell lines that did not produce type I collagen. We designated these sites HS1 to HS5, beginning with the site previously described at -100 bp. All mice transgenic for constructions containing DNA segments in which all three far-upstream hypersensitive sites (HS3 to HS5) were identified or segments in which only HS3 and HS4 were identified expressed high levels of the reporter genes. In contrast, embryos containing constructions lacking these sequences showed low or no measurable activity. Hence, a direct correlation exists between DNA segments with enhancer activity and those in which the far-upstream DNase I-hypersensitive sites were identified. Moreover, the enhancer region was able to activate transcription of a reporter gene in transgenic mice when it was placed in its natural position at 15 to 17 kb upstream of the start site of transcription with silencer sequences located between the enhancer and the promoter.

Regions of DNase I-hypersensitivity sites are frequently associated with transcriptional control of tissue-specific genes as in the human β globin gene cluster, which is controlled by a strong cell-specific enhancer known as the locus control region (LCR)¹ containing four hypersensitive sites approximately 15 kb upstream (Grosveld et al., 1993). Like the proximal promoters of the γ and β globin genes, which can direct erythroid-specific expression of a transgene at low levels in the absence of the LCR (Townes et al., 1985; Chada et al., 1986), a short proximal promoter (-350 to +54 bp) of the mouse *pro α 2(I)* collagen gene can drive low levels of fibroblast-specific expression of a reporter gene in transgenic mice (Goldberg et al., 1992; Niederreither et al., 1992). In this study, we show that the addition of an upstream sequence (-19.5 to -13.5 kb) to a minimal mouse *pro α 2(I)* collagen promoter greatly enhanced reporter gene expression. Furthermore, when constructs lacking the far-upstream enhancer were used to generate transgenic mice, they occasionally showed promiscuous expression in tissues that do not express the endogenous gene, such as in the brain (Goldberg et al., 1992; Niederreither et al., 1992). However, there was no promiscuous expression in embryos harboring constructs that contain the far-upstream enhancer fragment. Moreover, the ratio of mice that expressed the transgene over mice that were positive by Southern hybridization also increased in mice harboring constructions containing the far-upstream enhancer, suggesting that the constructs were less likely to be suppressed by the effects of flanking sequences. However, unlike the situation with the β -globin LCR, there was no correlation between the number of copies of the transgene and its level of expression (Table II).

Histological analysis of 15.5-d embryos and newborn mouse pups containing constructions with the far-upstream enhancer fragment showed that only mesenchymally derived cells, such as fibroblasts of the skin, fascia, and tendon and osteoblasts of the calvaria, clavicle, and mandible, express the transgene. In this study we show that mesenchymal cells in tissues such as the abdominal viscera were also stained for X-gal in mice harboring the far-up-

1. Abbreviation used in this paper: LCR, locus control region.

stream enhancer, whereas in mice harboring 2,000 bp of the promoter, staining in these cells was not detectable, perhaps due to the much lower level of promoter activity. In newborn animals, transgene expression could also be detected in smooth muscle cells in the walls of blood vessels; no staining in blood vessels was detected at E15.5. It has been shown in sheep that type I collagen levels increase in the aorta in the immediate perinatal period (Bendeck and Langille, 1991). A similar phenomenon in mice might explain the observation of significant X-gal staining in blood vessel walls only in newborn pups. Bedalov et al. (1994) have recently shown by RNA analysis that the *pro α 1(I)* collagen gene is highly expressed in the aorta of 7-d-old mice. The same group showed that a human minigene construct that contains 2.3 kb of the *pro α 1(I)* collagen promoter linked to the first five and last six intron/exon units and 2 kb of 3' flanking sequences was expressed in the aorta of transgenic mice by Western blot analysis of tissue extracts, whereas a construct containing 3.5 kb of the rat *pro α 1(I)* collagen promoter fused to the CAT gene had no activity in smooth muscle cells (Bedalov et al., 1994). These data suggest that elements in the *pro α 1(I)* collagen gene that control expression in vascular smooth muscle cells might be present in the gene itself or in the 3' flanking sequences, whereas similar elements for the *pro α 2(I)* collagen gene appear to be located in a far-upstream enhancer.

Only a subset of osteoblastic cells expressed the *lacZ* transgene. In E15.5 embryos and newborn pups, histological analysis indicated that some osteoblasts in calvarium, mandible, and clavicle were stained with X-gal. In human embryos, these bones are known to be formed either completely or in part by intramembranous ossification (Gardner, 1971). In contrast, osteoblasts in other bones, including the long bones of the limbs, were largely X-gal negative. Occasional expression of the transgene was seen in long bones in regions of periosteal ossification, but no expression was seen in areas of endochondral ossification. The most strongly stained regions of the bones in newborn animals correlated with regions of new bone growth. This is evident in newborn calvaria where only the areas close to the sutures stained strongly with X-gal. Cells in more mature areas of bone had low or no staining, suggesting that osteoblasts were losing transgene expression as they matured. This pattern of osteoblast staining was in sharp contrast with that observed in mice harboring specific segments of the mouse *pro α 1(I)* collagen promoter (Rossert et al., 1995) or a human *pro α 1(I)* construct (Liska et al., 1994) which were strongly expressed in all osteoblasts. Similarly, odontoblasts were X-gal negative in mice carrying the far-upstream *pro α 2(I)* collagen enhancer, whereas the same cells were strongly positive in newborn mice with the *pro α 1(I)* collagen gene promoter (Rossert et al., 1995).

Several different explanations are possible for the pattern of osteoblast expression in transgenic mice carrying the far-upstream *pro α 2(I)* collagen enhancer. One possibility is that the osteoblasts that form intramembranous bones and those that form endochondral bones respond to different genetic programs of development. Evidence consistent with the existence of a separate genetic program comes from the observation that a human genetic disease, cleidocranial dysplasia, shows malformation of membra-

nous bones such as the calvaria and clavicle, whereas long bones are largely unaffected (McKusick, 1988). If there were a separate genetic program for membranous and endochondral osteoblasts, it would exist in addition to a program that would direct expression of type I collagen genes in all osteoblasts, as evidenced by the activity of the *pro α 1(I)* collagen promoter in all osteoblasts of transgenic mice (Rossert et al., 1996).

A different hypothesis to account for the X-gal staining in early osteoblasts from membranous bones is that it could be due to transgene expression by the mesenchymal precursors of these cells (the half-life of the β -galactosidase enzyme in various cells of the embryo is unknown). In favor of this view, we have observed X-gal staining in some early chondrocytic cells that have just differentiated from mesenchymal precursors, but this staining is rapidly lost as the cells mature (data not shown). However, we found that the X-gal staining in osteoblasts from membranous bones is often stronger than that seen in surrounding fibroblasts, suggesting that new β -galactosidase enzyme is being synthesized in these cells. It is possible that osteoblasts or preosteoblasts at the edges of membranous bones are subjected to a boost in the expression of the type I collagen genes, which is reflected in strong X-gal staining in these cells in transgenic mice.

Although transgenic mice containing the far-upstream enhancer expressed the reporter gene only in cells in which the endogenous type I collagen genes are active, it is difficult to strictly determine from our results whether the far-upstream enhancer confers tissue specificity since the 350-bp proximal *pro α 2(I)* collagen promoter alone can direct a low level of tissue-specific expression of reporter genes in transgenic mice (Goldberg et al., 1992; Niederreither et al., 1992). However, a high level tissue-specific expression was also observed in transgenic embryos containing the far-upstream enhancer placed 5' of a 220-bp proximal *pro α 1(I)* collagen promoter. Our previous studies showed that in mice harboring the 220-bp proximal promoter linked to *lacZ*, no X-gal staining was detected (Rossert et al., 1996). When this promoter was linked to the gene for the highly sensitive luciferase enzyme, transgenic mice showed a low level of expression in extracts of the skin (Rossert et al., 1996). Hence, in mice containing the chimeric construct, in which the far-upstream enhancer of the *pro α 2(I)* collagen gene was placed 5' to the 220-bp proximal *pro α 1(I)* collagen promoter, it was likely that the tissue specificity was directed in large part by the upstream enhancer.

Examination of mice carrying *lacZ* transgenes driven by various promoter lengths of the mouse *pro α 1(I)* collagen gene strongly suggested that separate *cis*-acting elements in this gene controlled expression of reporter genes in different type I collagen-producing cells (Rossert et al., 1995, 1996). A similar conclusion was reached in studying the rat *pro α 1(I)* collagen promoter linked to a CAT reporter gene (Pavlin et al., 1992; Bogdanovic et al., 1994) and also a construction containing a segment of the human *pro α 1(I)* promoter and first intron linked to a growth hormone reporter gene (Liska et al., 1994). The mouse *pro α 2(I)* collagen gene may be regulated in a similar fashion with separate *cis*-acting elements driving expression at high levels in various cell types. Elements in the minimal promoter and

in the far-upstream enhancer seem to regulate expression of reporter genes mainly in fibroblasts and other mesenchymal cells, while elements needed for high levels of expression in odontoblasts, endochondral osteoblasts, and some tendon fibroblasts are lacking.

In summary, we have identified a potent enhancer element 15 to 17 kb upstream of the mouse *pro α 2(I)* collagen gene. We speculate that this far-upstream enhancer controls the expression of the transgene in mesenchymal cells such as fibroblasts and some osteoblasts but is not active in many osteoblasts or in odontoblasts. Work is in progress to better delineate the sequences around HS3 to HS5 that support the enhancing activity of this region and to determine the transcription factors which bind to this segment of DNA.

We wish to thank Silvio Garofalo and Dave Toman for the original work in identifying and mapping the cosmid clone of the *pro α 2(I)* collagen 5' flanking sequences and Drs. Farach-Carson and Platsoukas for generously donating two of the cell lines used in this study. We would also like to thank Pat McCauley for help in typing the manuscript.

The work reported in this paper was supported by National Institutes of Health grant number HL41264. Sequencing was supported by National Cancer Institute grant number CA-16672.

Received for publication 20 November 1995 and in revised form 28 May 1996.

References

Barsh, G.S., C.L. Roush, and R.E. Gelinis. 1984. DNA and chromatin structure of the human α 1(I) collagen gene. *J. Biol. Chem.* 259:14906-14913.

Beck, K.M., A.H. Seekamp, G.R. Askew, Z. Mei, C.M. Farrell, S. Wang, and L.N. Lukens. 1991. Association of a change in chromatin structure with a tissue-specific switch in transcription start sites in the α 2(I) collagen gene. *Nucleic Acids Res.* 19:4975-4982.

Bedalov, A., D.T. Breault, B.P. Sokolov, A.C. Lichtler, I. Bedalov, S.H. Clark, K. Mack, J.S. Khillan, C.O. Woody, B.E. Kream, et al. 1994. Regulation of the α 1(I) collagen promoter in vascular smooth muscle cells: comparison with other α 1(I) collagen-producing cells in transgenic animals and cultured cells. *J. Biol. Chem.* 269:4903-4909.

Bellard, M., G. Dretzen, A. Giangrande, and R. Phillippe. 1989. Nuclease digestion of transcriptionally active chromatin. *Methods Enzymol.* 170:317-346.

Bendeck, M.P., and B.L. Langille. 1991. Rapid accumulation of elastin and collagen in the aortas of sheep in the immediate perinatal period. *Circ. Res.* 69: 1165-1169.

Bogdanovic, Z., A. Bedalov, P.H. Krebsbach, D. Pavlin, C.O. Woody, S.H. Clark, H.F. Thomas, D.W. Rowe, B.E. Kream, and A.C. Lichtler. 1994. Upstream regulatory elements necessary for expression of the rat COL1A1 promoter in transgenic mice. *J. Bone Miner. Res.* 9:285-292.

Breindl, M., K. Harbers, and R. Jaenisch. 1984. Retrovirus-induced lethal mutation in collagen I gene of mice is associated with an altered chromatin structure. *Cell.* 38:9-16.

Chada, K., J. Magram, and F. Constantini. 1986. An embryonic pattern of expression of a human fetal globin gene in transgenic mice. *Nature (Lond.)* 319:685-689.

Cole, W.G. 1994. Collagen genes: mutations affecting collagen structure and expression. *Prog. Nucleic Acid Res. Mol. Biol.* 47:9-80.

D'Souza, R.N., K. Niederreither, and B. de Crombrughe. 1993. Osteoblast-specific expression of the α 2(I) collagen promoter in transgenic mice: correlation with the distribution of TGF- β 1. *J. Bone Miner. Res.* 8:1127-1136.

Gardner, E. 1971. Osteogenesis in the human embryo and fetus. In *The Biochemistry and Physiology of Bone*, 2nd ed. Vol II. G.H. Bourne, editor. Academic Press, New York, NY. pp. 77-118.

Goldberg, H., T. Helaakoski, L.A. Garrett, G. Karsenty, A. Pellegrino, G. Loz-

ano, S. Maity, and B. de Crombrughe. 1992. Tissue-specific expression of the mouse α 2(I) collagen promoter: Studies in transgenic mice and in tissue culture cells. *J. Biol. Chem.* 267:19622-19630.

Grosveld, F., M. Antoniou, M. Berry, E. De Boer, N. Dillon, J. Ellis, P. Fraser, O. Hanscombe, J. Hurst, A. Imam, et al. 1993. The regulation of human globin gene switching. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 339:183-191.

Hogan, B., R. Beddington, F. Constantini, and E. Lacy. 1986. *Manipulating the Mouse Embryo: A Laboratory Manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. 497 pp.

Kaltschmidt, C., M. Muller, G. Brem, and R. Renkawitz. 1994. DNase I hypersensitive sites far upstream of the rat tryptophan oxygenase gene direct developmentally regulated transcription in livers of transgenic mice. *Mech. Dev.* 45:203-210.

Khillan, J.S., A. Schmidt, P.A. Overbeck, B. de Crombrughe, and H. Westphal. 1986. Developmental and tissue-specific expression by the α 2 type I collagen promoter in transgenic mice. *Proc. Natl. Acad. Sci. USA.* 83:725-729.

Liau, G., D. Szapary, C. Setoyama, and B. de Crombrughe. 1986. Restriction enzyme digestions identify discrete domains in the chromatin around the promoter of the mouse α 2(I) collagen gene. *J. Biol. Chem.* 261:11362-11368.

Liska, D.J., M.J. Reed, E.H. Sage, and P. Bornstein. 1994. Cell-specific expression of α 1(I) collagen-hGH minigenes in transgenic mice. *J. Cell Biol.* 125: 695-704.

Mercer, E.H., G.W. Hoyle, R.P. Kapur, R.L. Brinster, and R.D. Palmiter. 1991. The dopamine β -hydroxylase gene promoter directs expression of *E. coli lacZ* to sympathetic neurons in adult transgenic mice. *Neuron.* 7:703-716.

McKusick, V.A. 1988. *Mendelian Inheritance in Man: Catalogs of Autosomal Dominant, Autosomal Recessive, and X-linked Phenotypes*, 8th ed. The Johns Hopkins University Press, Baltimore, MD. 153 pp.

McLeod, M.J. 1980. Differential staining of cartilage and bone in whole mouse fetuses by alcian blue and alizarin red. *Teratology.* 22:299-301.

Niederreither, K., R.N. D'Souza, and B. de Crombrughe. 1992. Minimal DNA sequences that control the lineage-specific expression of the *pro α 2(I)* collagen promoter in transgenic mice. *J. Cell Biol.* 119:1361-1370.

Pavlin, D., A.C. Lichtler, A. Bedalov, B.E. Kream, J.R. Harrison, H.F. Thomas, G.A. Gronowicz, S.H. Clark, C.O. Woody, and D.W. Rowe. 1992. Differential utilization of regulatory domains within the α 1(I) collagen promoter in osseous and fibroblastic cells. *J. Cell Biol.* 116:227-236.

Philajaniemi, T., and M. Rehn. 1995. Two new collagen subgroups: Membrane-associated collagens and types XV and XVIII. *Prog. Nucleic Acid Res. Mol. Biol.* 50:225-262.

Pogulis, R.J., and S.O. Freytag. 1993. Contribution of specific *cis*-acting elements to the activity of the mouse *pro α 2(I)* collagen enhancer. *J. Biol. Chem.* 268:2493-2499.

Rosert, J., H. Eberspaecher, and B. de Crombrughe. 1995. Separate *cis*-acting DNA elements of the mouse *pro- α 1(I)* collagen promoter direct expression of reporter genes to different type I collagen-producing cells in transgenic mice. *J. Cell Biol.* 129:1421-1432.

Rosert, J., S.S. Chen, H. Eberspaecher, C.N. Smith, and B. de Crombrughe. 1996. Identification of a minimal sequence of the mouse *pro α 1(I)* collagen promoter that confers high level osteoblast-specific expression in transgenic mice, and that binds an osteoblast-specific DNA binding protein. *Proc. Natl. Acad. Sci. USA.* 93:1027-1031.

Rossi, P., and B. de Crombrughe. 1987. Identification of a cell-specific transcriptional enhancer in the first intron of the mouse α 2(type I) collagen gene. *Proc. Natl. Acad. Sci. USA.* 84:5590-5594.

Rugh, R. 1990. *The Mouse: Its Reproduction and Development*. Oxford University Press, New York. 430 pp.

Sudo, H., H.A. Kodama, Y. Amagai, S. Yamamoto, and S. Kasai. 1983. In vitro differentiation and calcification in a new clonal osteogenic cell line derived from newborn mouse calvaria. *J. Cell Biol.* 96:191-198.

Townes, T.M., J.B. Lingrel, H.Y. Chen, R.L. Brinster, and R.D. Palmiter. 1985. Erythroid-specific expression of human β -globin genes in transgenic mice. *EMBO (Eur. Mol. Biol. Organ.) J.* 4:1715-1723.

Vuust, J., M.E. Sobel, and G.R. Martin. 1985. Regulation of type I collagen synthesis: Total *pro α 1(I)* and *pro α 2(I)* mRNAs are maintained in a 2:1 ratio under varying rates of collagen synthesis. *Eur. J. Biochem.* 151:449-453.

Wood, W.M., M.Y. Kao, D.F. Gordon, and E.C. Ridgway. 1989. Thyroid hormone regulates the mouse thyrotropin β -subunit gene promoter in transfected primary thyrotropes. *J. Biol. Chem.* 264:14840-14847.

Zhong, W., J. Mirkovitch, and J.E. Darnell, Jr. 1994. Tissue-specific regulation of mouse hepatocyte nuclear factor 4 expression. *Mol. Cell Biol.* 14:7276-7284.