Minimal DNA Sequences That Control the Cell Lineage-specific Expression of the proα2(I) Collagen Promoter in Transgenic Mice

Karen Niederreither,* Rena N. D'Souza,‡ and Benoit de Crombrugghe*

*Department of Molecular Genetics, M. D. Anderson Cancer Center, and ‡Department of Anatomical Sciences, Health Science Center, Dental Branch, University of Texas, Houston, Texas 77030

Abstract. The pattern of expression of the proα2(I) collagen gene is highly tissue specific in adult mice and shows its strongest expression in bones, tendons, and skin. Transgenic mice were generated harboring promoter fragments of the mouse proα2(I) collagen gene linked to the *Escherichia coli* β-galactosidase or firefly luciferase genes to examine the activity of these promoters during development. A region of the mouse proα2(I) collagen promoter between −2,000 and +54 exhibited a pattern of β-galactosidase activity during embryonic development that corresponded to the expression pattern of the endogenous proα2(I) collagen gene as determined by in situ hybridization. A similar pattern of activity was also observed with much smaller promoter fragments containing either 500 or 350 bp of upstream sequence relative to the start of transcription. Embryonic regions expressing high levels of β-galactosidase activity included the bulbus arteriosus, valves of the developing heart, sclerotomes, meninges, limb buds, connective tissue fascia between muscle fibers, osteoblasts in newly formed bones,

fibroblasts in tendons, periosteum, dermis, and peritoneal membranes. The pattern of β-galactosidase activity was similar and included within the extracellular immunohistochemical localization pattern of transforming growth factor-β1 (TGF-β1). The −315–−284 region of the proα2(I) collagen promoter was previously shown to mediate the stimulatory effects of TGF-β1 on the proα2(I) collagen promoter in DNA transfection experiments with cultured fibroblasts. A construct containing this sequence tandemly repeated 5' to a very short α2(I) collagen promoter (−40–+54) showed preferential activity in tail and skin of 4-wk-old transgenic mice. Except for low expression of the transgene in bone, this pattern mimics the expression of the endogenous proα2(I) collagen gene. We propose the hypothesis that the tissue-specific expression of the proα2(I) collagen gene during embryogenesis is controlled by both TGF-β1 and cell-specific transcription factors; one of these could interact directly or indirectly with either the −315–−284 or the −40–+54 segment.

TYPE I collagen is the most abundant extracellular matrix protein in vertebrates and is a member of a family of multiple collagen types. It is a heterotrimer composed of two α1(I) chains and one α2(I) chain, which are coordinately regulated. Type I collagen is first detected at day eight during murine embryonic development (Adamson and Ayers, 1979; Leivo et al., 1980) and is crucial for complete vertebrate development (Adamson, 1982; Reddi et al., 1984). Homozygous disruption of the proα1(I) collagen gene, caused by the integration of retroviral DNA into the first intron of the mouse gene, results in embryonic lethality at 12–13 d of gestation (Schnieke et al., 1983; Harbers et al., 1984).

Expression of both proα1(I) and proα2(I) collagen follows a distinct tissue-specific pattern. RNase protection assays show high levels of RNA encoding these polypeptides in bone, tendon, and skin, with much lower levels in other tissues (Slack et al., 1991; Goldberg et al., 1992). The sequences responsible for this tissue-specific regulation have not been identified. However, several regions have been

identified in the mouse proα2(I) and human, rat, and mouse proα1(I) collagen genes that regulate transcription (Karsenty et al., 1988; Rossi et al., 1988; Karsenty et al., 1991; Ritzenthaler et al., 1991). Within the 350-bp sequence upstream of the start of transcription of the mouse proα2(I) collagen gene, there are several functional *cis*-acting elements that stimulate proα2(I) gene transcription (Karsenty et al., 1988; Hatamochi et al., 1988; Maity et al., 1988; Rossi et al., 1988). A binding site for CCAAT transcription factor/nuclear factor-1 (CTF/NF-1) is present at −300. The −300 site mediates the effect of TGF-β1, which, when added to fibroblasts in culture, acts to stimulate the activity of a transfected α2(I) promoter-CAT chimeric gene. Mutations in the NF-1 site abolish this TGF-β1 inducibility (Rossi et al. 1988).

TGF-β1 is a potent inducer of Type I collagen expression and may play an important role as a developmental regulator. TGF-β1 and some of its related family members are expressed in restricted patterns during development (Akhurst et al., 1990) and several other family members such as activin and possibly bone morphogenetic factors may influence

mesoderm formation during early stages of *Xenopus* development (Thomsen et al., 1990; Koster et al., 1991).

Tissue culture transfection assays (Schmidt et al., 1986) and transgenic mice (Khillan et al., 1986) have been used to identify the regions of the mouse α2(I) collagen gene that confer tissue specificity. The α2(I) promoter from −2,000 to +54 directed CAT reporter gene expression in a pattern mimicking that of the endogenous Type I collagen genes in adult mice (Khillan et al., 1986). Sequences present in the first intron of the α2(I) collagen gene which had previously been reported to contain an enhancer element were not required to direct tissue-specific expression (Khillan et al., 1986; Goldberg et al., 1992). When the region from −350 to +54 of the proα2(I) collagen promoter was linked to CAT, transgenic mice displayed the same tissue-specific expression pattern, except that expression in the brain was also observed (Goldberg et al., 1992). Others have reported that the human proα1(I) collagen gene requires 2.3 kb of upstream promoter sequences to direct tissue-specific expression in transgenic mice (Slack et al., 1991). Furthermore, the 5' flanking region of the proα1(I) collagen gene between −3.5 and −2.3 kb has been implicated in high-level expression in osteoblasts (Pavlin et al., 1992).

In this paper, we further define the ability of the proα2(I) collagen gene promoter to direct tissue-specific expression in transgenic mice. We have used the β-galactosidase reporter gene to localize transgene expression during embryonic development. Through promoter deletions we show that 350 bp of promoter sequences linked to β-galactosidase were sufficient to confer a very high degree of tissue and cell specificity. This pattern mimicked the distribution of endogenous proα2(I) collagen during development as determined by in situ hybridization and was similar to the immunohistochemical localization of TGF-β1. A region of the α2(I) collagen promoter from −315 to −284 which mediates the effects of TGF-β1 was tandemly repeated 5' to a short α2(I) collagen promoter (−40−+54) linked to the firefly luciferase reporter gene. This chimeric gene showed a high degree of tissue-specific activity in transgenic animals.

Materials and Methods

Plasmid Constructions

Regions of the proα2(I) collagen promoter were inserted upstream of the translational start site of a *lacZ* expression vector (placF) (Mercer et al., 1991) which contains the mouse protamine gene sequences that supply an intron and polyadenylation signal (see Fig. 1). A 2,054-bp *Hind*III fragment (for −2,000 *lacZ*) from pR40 was subcloned by blunt end ligation into the *Sma*I site of placF. A 554-bp *Xba*I to *Hind*III fragment from PR40 (Rossi et al., 1988) was made blunt using the Klenow enzyme and deoxynucleotide triphosphates; after addition of *Xba*I linkers this fragment was ligated into the *Xba*I site of placF to create −500 *lacZ*. A 404-bp *Bgl*II to *Hind*III fragment, also from PR40 (Rossi et al., 1988) was similarly ligated into the *Xba*I site of placF to create −350 *lacZ*.

A segment consisting of four tandem copies of the proα2(I) collagen promoter sequence from −315 to −284 inserted upstream of the −40−+54 proα2(I) collagen DNA segment was introduced into the polylinker region of the firefly luciferase vector (pA3luc) to create pLAG1 (Goldberg et al., 1992). The pA3luc vector contains an SV-40 polyadenylation signal and splice sites 3' to the luciferase gene and a polyadenylation cassette 5' to the site in which the promoter region is inserted. In transfection studies, this 5' polyadenylation segment inhibits read-through transcription (Wood et al., 1989).

Generation of Transgenic Mice

Transgenic mice were produced by pronuclear injection of gene construct into B6D2F2 fertilized eggs (Hogan et al., 1986). The −2,000, −500, and −350 *lacZ* plasmids were digested with *Asp*718 and *Hind*III, and pLAG1 was cleaved with *Aat*II and *Sac*I to release the inserts from plasmid vector sequences (see Fig. 1). Founder mice carrying the introduced gene constructions were identified by Southern analysis of tail DNA (Southern, 1975).

β-galactosidase and Luciferase Expression

β-galactosidase activity was assayed in collagen-*lacZ* transgenic mouse lines by staining embryos with X-Gal using techniques previously described (Mercer et al., 1991). Embryos were stained for 1-2 d. No additional staining appeared after 2 d. After staining, embryos were paraffin embedded. 6-μm-thick sections were counterstained with eosin.

To examine the expression of the luciferase reporter gene in LAG1 transgenic mice, representative organs from 4-wk-old animals were dissected from surrounding tissue, placed on dry ice, homogenized in a buffer containing 100 mM Tris, pH 7.8, and 1 mM DTT, and microfuged. The supernatant was assayed for luciferase activity in the presence of excess D-luciferin (0.3 mg/ml), 15 mM MgSO₄, 5 mM ATP, and 1 mM DTT. Known quantities of luciferase containing fibroblast cell extract were homogenized along with tissues from a normal mouse to determine whether these tissues contained inhibitors of luciferase activity. No inhibitors of luciferase activity were present in any of the mouse tissues. Measurements were taken using a Monolight 2010 luminometer. Protein levels were measured using the Bradford assay with BSA as standard (Bradford, 1976). Results were expressed as luciferase light units per 1 mg of protein.

In situ Hybridization

Embryos isolated at 11.5-15.5 d postcoitum (E11.5 to E15.5) were fixed in 10% formalin for 24 h or more (depending on size). After dehydration with increasing concentrations of ethanol, embryos were embedded in paraffin and sectioned onto silane-treated slides at a thickness of 6 μm. A 200-bp *Pst*I-*Eco*RI fragment coding for the COOH-terminal region of the rat proα2(I) collagen gene (pα2R2, kindly provided by Dr. David Rowe, University of Connecticut, Farmington, CT) was directionally cloned into Bluescript + (Stratagene Inc., La Jolla, CA) to give pKα2R2. Plasmids were linearized downstream of the inserted fragment with *Pst*I and antisense riboprobes generated using T7 RNA polymerase. Similarly, plasmid pKα2R2 was digested with *Eco*RI, transcribed by T3 RNA polymerase to generate sense RNA. Probes were labeled with ³⁵αS-rUTP to a specific activity of 10⁵ cpm/μl. In situ hybridizations were performed as previously described (Wikinson et al., 1987) with the following modifications. Sections were postfixed in 4% formaldehyde, treated with 40 μg/ml proteinase K for 7.5 min, and treated with 0.1 M triethanolamine containing 0.25% acetic anhydride for 10 min. Slides were heated for 2 min at 98°C and cooled on ice before hybridization. The hybridization mix containing 50% formamide, 0.3 M NaCl, 1× Denhardt, 5 mM EDTA, 10 mM Tris, 10% dextran, 10 mM DTT, 10 mM NaHPO₄, 500 μg/ml tRNA, 200 μg/ml calf thymus DNA, and the denatured probe was added to each slide and further incubated at 48°C overnight. Sections were treated with RNAse at 37°C for 30 min using 20 μg/ml RNAse A and 1 μg/ml RNAse T₁. After washing, slides were dipped in a 50% mixture of NTB-2 (Eastman Kodak Co., Rochester, NY) and 0.6 M NH₄Ac, further exposed at 4°C for 3-4 d for autoradiography, and subsequently developed. All sections were counter-stained with hematoxylin and photographed under darkground illumination.

Results

Generation of Transgenic Mice

Transgenic mouse lines were generated using the mouse α2(I) collagen promoter–reporter gene constructions shown in Fig. 1. Founder mice and progeny carrying the various gene constructions were identified. The number of copies of the transgenes in each line were characterized by Southern blot analysis (Table I). Four out of eight lines carrying the construct containing 2,000 bp of the α2(I) collagen promoter
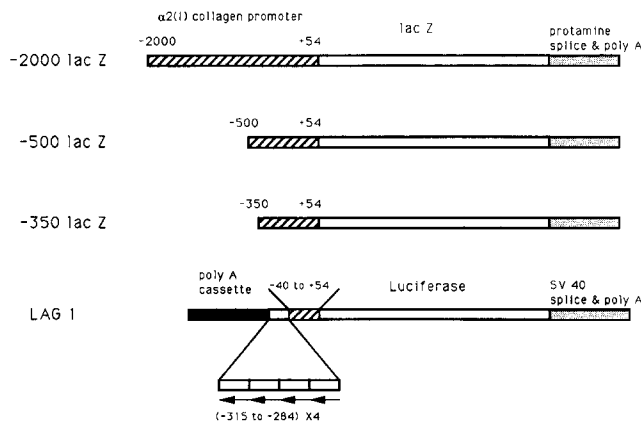
Figure 1. Representations of DNA constructs used to generate α2(I) collagen-*lacZ* and α2(I) collagen–luciferase transgenic mice. See Materials and Methods for details.

fused to the *lacZ* gene exhibited detectable β-galactoside activity as assayed by X-gal staining of embryos. Three out of seven lines carrying the construction containing 500 bp upstream of the start of transcription expressed detectable β-galactosidase activity. Three out of eight lines bearing the construct containing the −350-bp promoter, showed transgene expression. Variations were seen in the levels of transgene expression between lines with the same promoter construction and are likely influenced by the site of transgene integration. Overall, however, there was a remarkable similarity in the pattern of expression of β-galactosidase in these different strains.

In contrast to the *lacZ*-containing strains, which were examined at various stages during embryonic development, transgenic mouse lines created with the luciferase reporter gene were tested by measuring luciferase activity in individual tissue of 4-wk-old founder mice (see Materials and Methods for details). At this age mice show higher levels of proα1(I) collagen RNA than at later times (Slack et al., 1991). Four of the five founders that were positive for the luciferase gene by Southern hybridization expressed the chimeric gene containing the tandemly repeated −315-−284 sequence of the α2(I) collagen promoter. This higher percentage of expression for mice carrying the luciferase transgene than for mice with a β-galactosidase transgene can be accounted for by the much greater sensitivity of the assay for the luciferase enzyme.

α2(I) Collagen LacZ Transgenes Direct Tissue-specific Expression during Murine Development

The distinct transgene expression pattern directed by the

construction harboring a −2,000-bp promoter is shown in whole embryos in Fig. 2. At 9.5 d of embryonic development (E9.5; Fig. 2 *A*) transgene expression was first seen in the bulbus arteriosus, the most anterior segment of the tubular embryonic heart (Fig. 2 *A*). 1 d later at E10.5 (Fig. 2 *B*), a more pronounced X-gal staining appeared at the base of the limb buds and in the visceral arches. The upper ridge of the region of the head bordering the fourth ventricle and the upper region above the notochord also showed β-galactosidase expression. Staining in the telencephalic region was also seen in some lines. At E12.5 (Fig. 2 *C*), flaring and indentation of the future digits occurred. A distinct radiating pattern of X-gal staining appeared in the developing digits. Marked staining was also visible in regions of the limbs closer to the body. These regions represent areas of early mesenchymal tissue condensation in forming bones and joints. A diffuse stain at the back of the head corresponds to staining in the developing meninges. Segmented expression in the sclerotomes was visible in the tail of the embryo; slightly rostral to this, the developing vertebrae show more intense expression. Transgene expression was also seen in the nasal region and forehead. At E13.5 (Fig. 2 *D*) digit formation was evident with transgene expression further demarcating this event. There was transgene expression in a region surrounding the areas where the hair follicles of the face (vibrissae) appear. Other areas of transgene expression included the external portion of the developing ear (pinna), tail, skin, mesenchymal regions of the mandible, and both cranial and spinal meninges evident around the head and above the spinal cord. Fig. 2 *E* exemplifies the transgene expression that occurred in the tail and developing limbs of the −2,000 *lacZ* carrying mice at E15.5 (Fig. 2 *E*).

Additional transgenic lines were generated containing truncations of the −2,000-bp promoter to determine whether smaller segments would confer similar patterns of expression. Two lines carrying the −500 *lacZ* construct and two lines carrying the −350 *lacZ* construct showed β-galactosidase expression similar to mice carrying the −2,000-bp promoter. Fig. 2, *F* and *G* show the pattern of staining in E10.5 and 12.5-d embryos carrying the −350-bp promoter. In a few other lines (not shown) β-galactosidase expression was restricted to regions such as the mandibular region, meninges, and sclerotomes. In three lines (one harboring the −2,000 *lacZ* construct and two with the −350 *lacZ* gene) small areas of promiscuous β-galactosidase activity occurred in the brain. Otherwise, no inappropriate transgene expression was observed.

Histological Localization of α2(I) Collagen LacZ Transgenes

As with β-galactosidase staining in whole-mount embryos,

Table I. Copy Number of Transgenes in Mouse Lines in which There Was a Detectable Level of Expression of Either β-galactosidase or Luciferase (LAG1)

| −2000LacZ		−500LacZ		−350LacZ		LAG1	
Line	Copy number	Line	Copy number	Line	Copy number	Line	Copy number
7	6–8	13	7	13	5	7	3
16	1	23	1	37	8–10	9	8
30	1	26	1	40	40	17	30
38	5					24	20
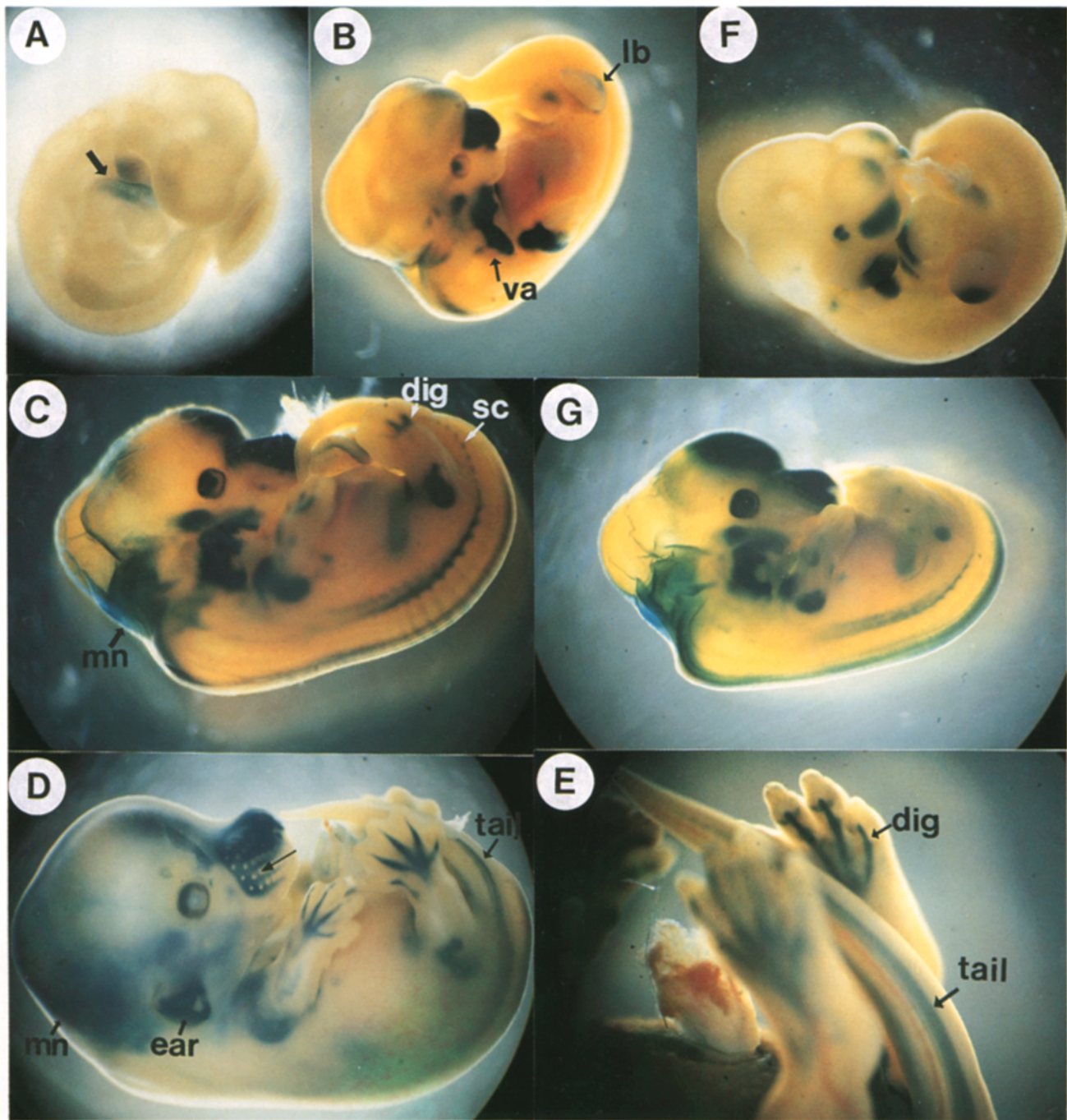
Figure 2. Expression of the −2,000 and −350 *lacZ* promotor constructs in whole mount embryos. Blue color indicates positive staining for β-galactosidase. *A–E* show expression for the −2,000 *lacZ* construction for line 7; *F* and *G* show comparable expression with the −350 *lacZ* construction line 37. (*A*) E9.5 embryo showing expression in the primitive heart (*arrow*). (*B*) E10.5 embryo showing expression in limb buds (*lb*) and visceral arches (*va*). (*C*) E12.5 embryo showing *lacZ* expression in developing digits (*dig*), meninges (*mn*), and sclerotomes (*sc*). (*D*) E13.5 −2,000 *lacZ* embryo showing expression in digits, connective tissue surrounding developing hair follicles in the nose, meninges (*mn*), ear, and tail. (*E*) E15.5 −2,000 *lacZ* embryo showing expression in tail and developing foot. *F* and *G* show E10.5 and E12.5 −350 *lacZ* embryos showing expression comparable with the −2,000 *lacZ* expression shown in *B* and *C*, respectively.

the transgenic embryos carrying the −2,000, −500, and −350 α2(I) collagen promoter-*lacZ* construct showed striking similarities in β-galactosidase staining when further examined histologically (see Table II). Staining clearly occurred in very discrete areas of tissues. The histological sections shown in Fig. 3 illustrate some examples of this ex-

quisite cell specificity. At E10.5, β-galactosidase expression was evident in the mesenchymal cells of the bulbus arteriosus (Fig. 3 *A*). A day later in development, at E11.5, β-galactosidase expression was found in the sclerotomes (Fig. 3 *B*). At this stage, staining in the heart was also restricted to cells of the endocardial cushion, a precursor to the developing
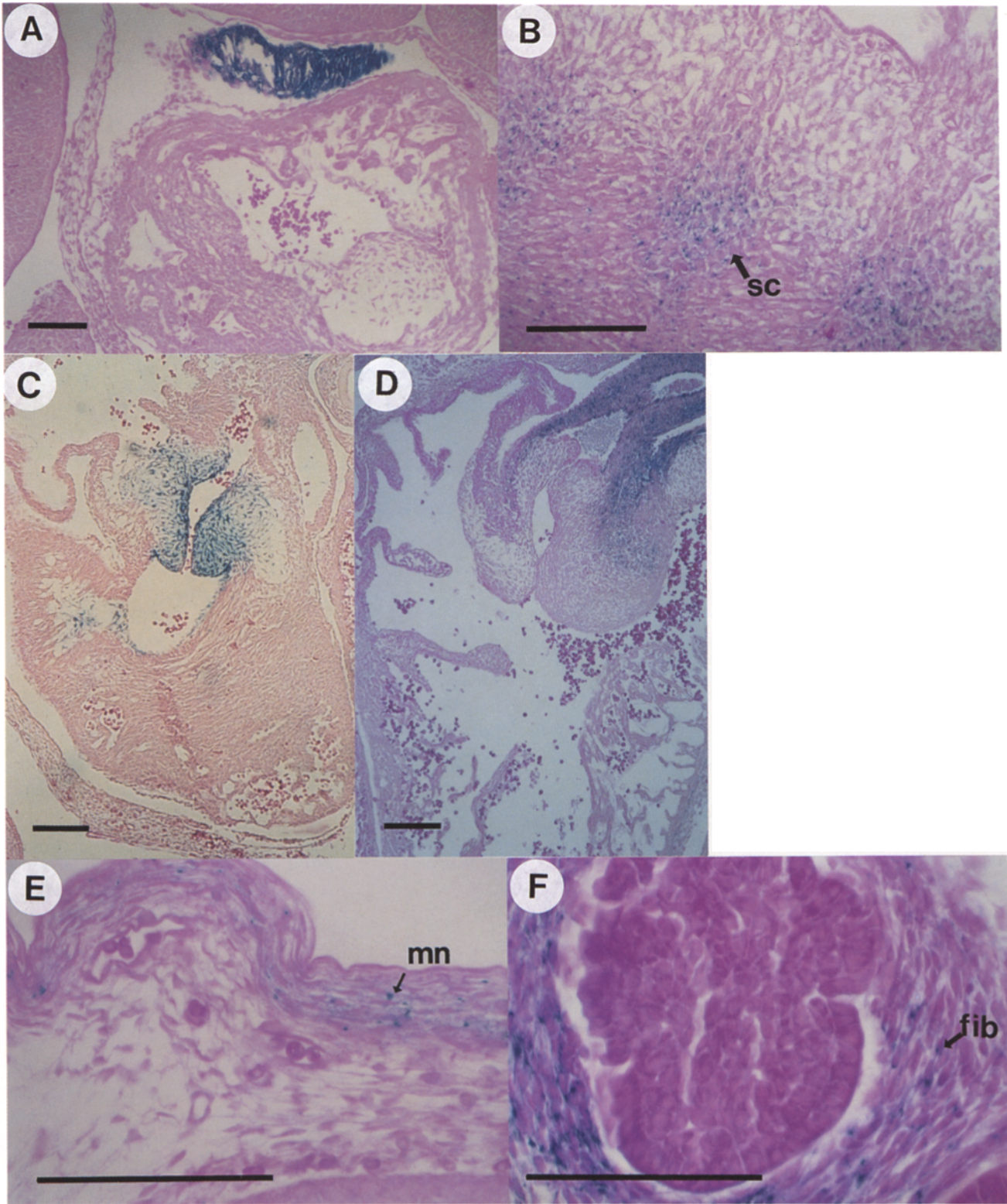
Figure 3. α2(I) collagen–*lacZ* fusion gene expression in sagittal sections of early embryos as indicated by blue stain. (*A*) Transgene expression at E10.5 day in bulbus arteriosus of developing heart (−2,000 *lacZ*, line 7). (*B*) E11.5 expression in sclerotomes (−500 *lacZ*, line 23). (*C*) E11.5 expression in heart valves (−500 *lacZ*, line 23). (*D*) Expression in the bulbus arteriosus at E12.5 (−350 *lacZ*, line 37). (*E*) *lacZ* expression in the meninges overlying the brain at E12.5 (−350 *lacZ*, line 37). (*F*) Transgene expression restricted to the fibroblastic connective tissue (*fib*) surrounding the submandibular gland at E13.5 days (−350 *lacZ*, line 37). Note the absence of expression in the central parenchymal glandular tissue. Bars, 100 μm.
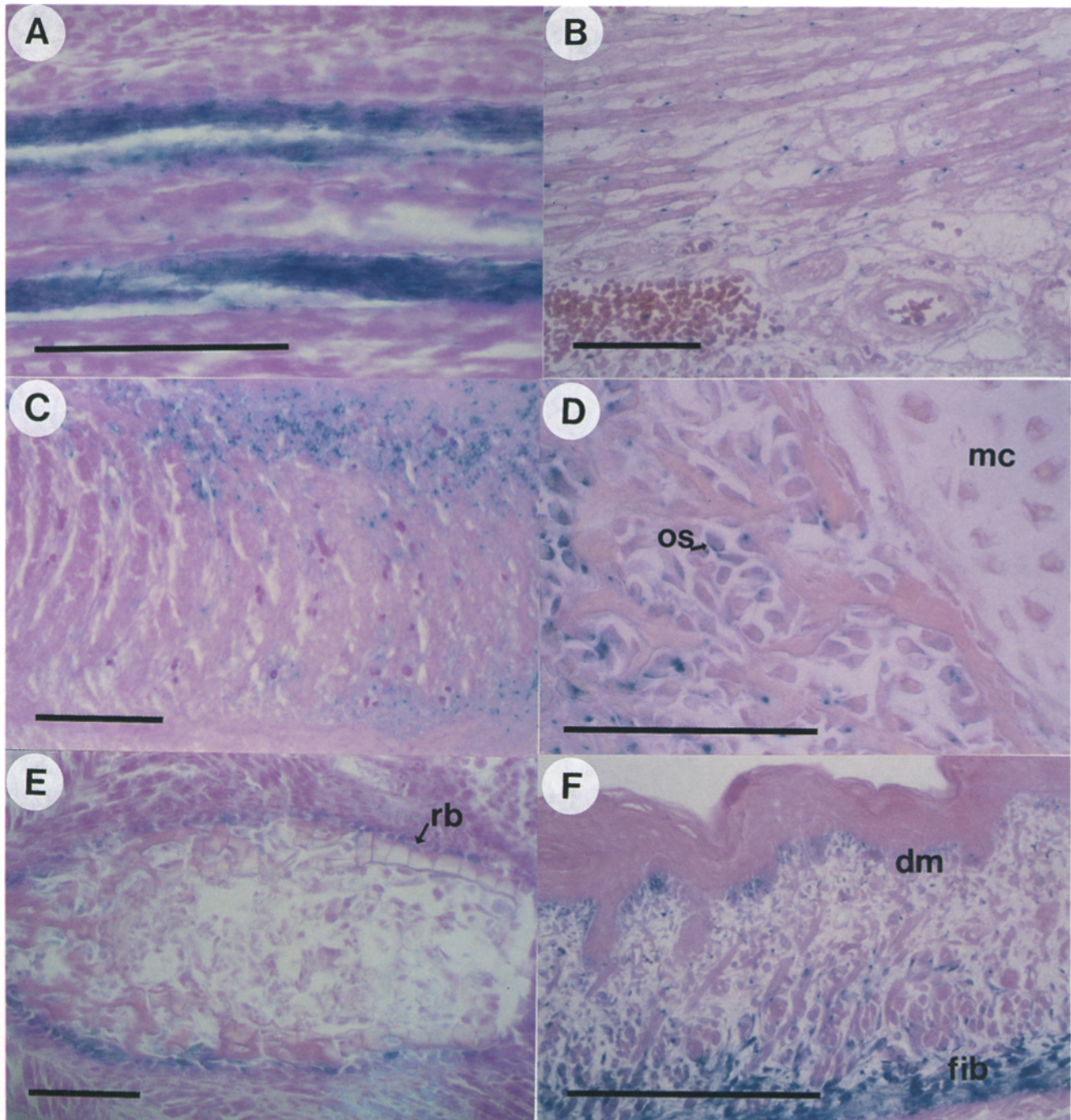
Figure 4. α2(I) collagen-*lacZ* expression later in development. (*A*) Transgene expression in tail tendon at E15.5 (−2,000 *lacZ*, line 7). *LacZ* expression in the extracellular region probably indicates diffusion of stain. (*B*) β-galactosidase expression in fibroblasts of the endomysium surrounding muscle fibers at E15.5 (−2,000 *lacZ* construction, line 7). (*C*) Transgene expression in connective tissue cells interspersed between muscle fibers of the tongue in an E14.5 embryo (−500 *lacZ*, line 23). (*D*) Transgene expression in osteoblasts (*os*) of the mandibular bone at E15.5 (−2,000 *lacZ*, line 30). Note the lack of α2(I) collagen-*lacZ* expression in adjacent chondrocytes in Meckels cartilage (*mc*). (*E*) E15.5 expression in the periosteum of the ribs (*rb*) (−2,000 *lacZ*, line 30). (*F*) *LacZ* expression in the dermis (*dm*) and underlying connective tissue (*fib*) in newborn mouse skin (−2,000 *lacZ*, line 30). Bars, 100 μm.

atrioventricular heart valves, as well as in the fibroblasts within the subendothelial layer of the endocardium but was absent in the myocardium itself (Fig. 3 *C*).

At E12.5, transgene expression in cells of the heart valves and bulbus arteriosus was apparent (Fig. 3 *D*). Expression in meningeal cells overlaying the brain (Fig. 3 *E*) and spinal cord was also seen. The specificity of the pattern of expres-

sion conferred by the α2(I) collagen promoter was also demonstrated by X-gal staining which is restricted to the fibroblast-rich matrix surrounding the submandibular gland, and is not visible within the central parenchymal tissue (Fig. 3 *F*).

At E15.5, when organogenesis is essentially complete, cell-specific transgene expression was again clearly appar-

Table II. Examples of Embryonic Structures Expressing α2(I) Collagen β-galactosidase Transgenes

| Age	Structure	−2,000LacZ	−500LacZ	−350LacZ
E11.5	Endocardium	+	+	+
	Heart valves	+	+	+
	Sclerotomes	+	+	+
E13.5	Meninges	+	+	+
	Capsule of submandibular gland	+	ND	+
E15.5	Fascia between muscles	+	+	+
E17.5	Dermis	+	+	+
	Osteoblasts	+	+	+
	Odontoblasts	ND	+	ND

ent. Fig. 4 A shows that the zone of densely arranged fibroblasts in tail tendon stained for β-galactosidase. Transgene expression was also found solely in the fibroblasts of the endomysium surrounding individual muscle fibers (Fig. 4 B). Fig. 4 C shows similar transgene expression in connective tissue cells interspersed between skeletal muscle fibers

of the tongue. At E15.5, ossification is beginning to occur in certain bones such as the mandibular bone which showed one of the earliest intramembranous ossification regions. *LacZ* expression was seen in osteoblasts adjacent to newly formed bone matrix, but no expression was evident in the chondrocytes of Meckels cartilage (Fig. 4 D). In the developing ribs, before the onset of endochondral ossification, transgene expression was found in the fibrous periosteum but not in the chondrocytes under the periosteum (Fig. 4 E).

Transgene expression in newborn mouse skin is shown in Fig. 4 F. β-galactosidase expression was found in the fibroblasts of the dermis and in cells of the subdermal connective tissue but was absent in the keratinized epithelial cells.

The discrete cell-specific activity observed with the −2,000-bp promoter was exactly reproduced with the −500- and −300-bp promoter throughout embryonic development. Both the −500- and −350-bp promoters were active in osteoblasts in newly formed bones (data not shown). Very recently, we have also observed that the −500-bp promoter was active in odontoblasts (Table II).

The very discrete, tissue- and cell-specific X-gal staining at different times of development is similar and is included in the patterns of extracellular TGF-β1 expression during mouse development (Heine et al., 1987; Flanders et al.,
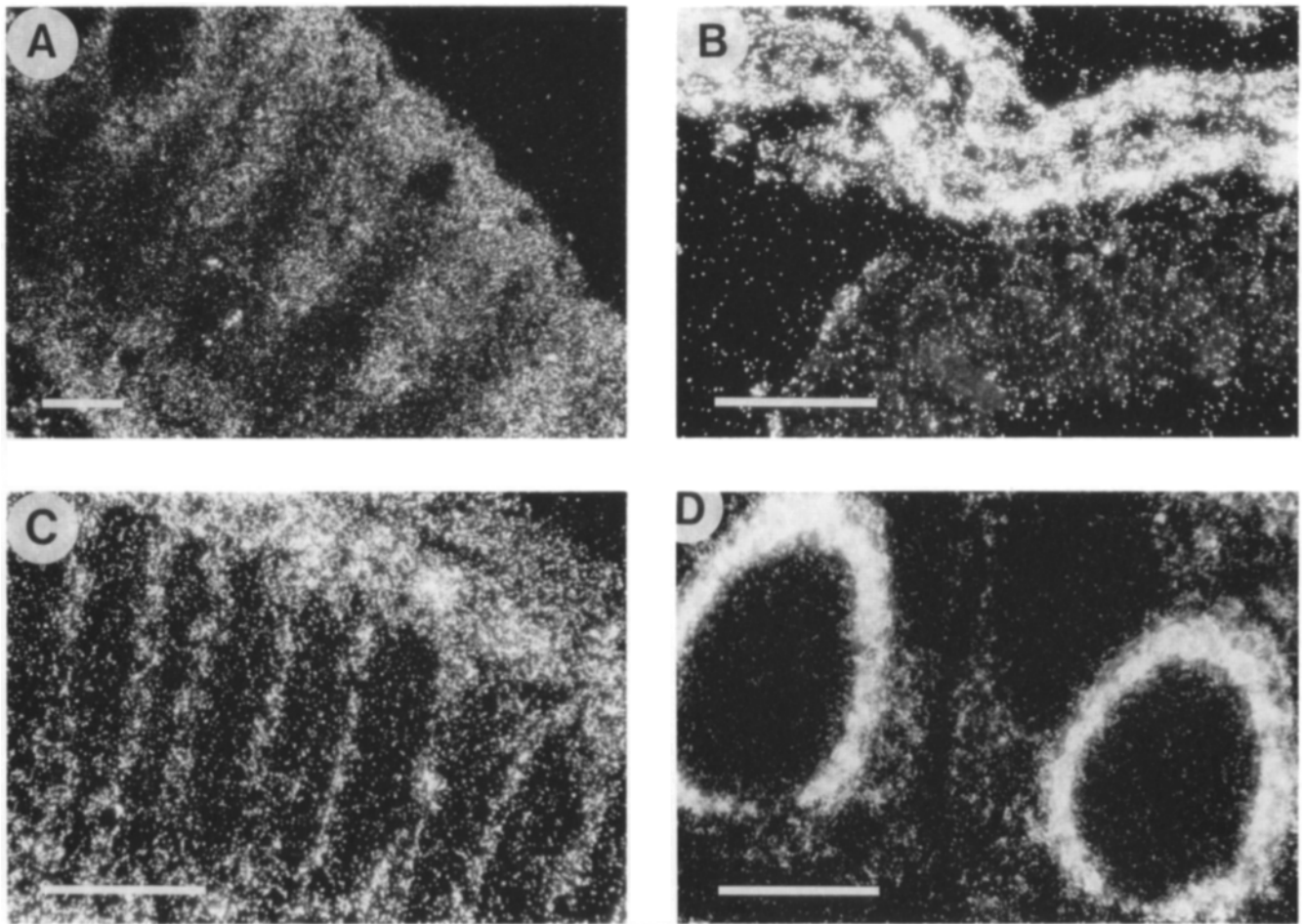
Figure 5. In situ hybridization for α2(I) collage RNA in the embryonic mouse. Sagittal sections were hybridized with antisense rat α2(I) collagen RNA and photographed under dark-field illumination. (*A*) E11 expression in sclerotomes. (*B*) E12.5 expression in meninges above brain. (*C*) E13.5 expression in connective tissue cells interspersed between muscle fibers of the tongue. (*D*) E14.5 expression in periosteum of ribs. Bars, 100 μm.

1989; Akhurst et al., 1990; D'Souza et al., 1990; data not shown). For instance, an identical pattern of immunoreactive TGF-β1 was found, in the developing cardiac cushion at E11.5 (see Fig. 3 *B* for transgene expression). A similar discrete localization of TGF-β1 was also found in sclerotomes at E11.5 (Heine et al., 1987; data not shown). TGF-β1 is present in the extracellular matrix surrounding meningeal fibroblasts in an identical pattern as was seen for β-galactosidase in Fig. 3 *E*. Other examples of this striking similarity between TGF-β1 localization and X-gal staining are in the connective tissue of fascia separating groups of myotubes (Fig. 4, *B* and *C*) and in the newly formed ossification centers (Fig. 4 *D*) (Heine et al., 1987; R. D'Souza, unpublished observation).

The Expression Pattern of the α2(I) Collagen-LacZ Transgenes Mimics That of the Endogenous proα2(I) Collagen Gene

Expression of the endogenous mouse α2(I) collagen gene was examined by in situ hybridization of embryos using an antisense rat proα2(I) collagen RNA probe. The pattern of embryonic proα2(I) collagen RNA was identical to the pattern of proα1(I) collagen RNA (Cheah et al., 1991 and data not shown). Expression of the β-galactosidase transgene was very similar to the pattern of expression of the endogenous gene, although in some areas, such as the peritoneal membranes, X-gal staining was not as uniform as the pattern that was seen by in situ hybridization. Fig. 5 shows representative examples that highlight the similarities between the tissue-specific expression of the endogenous proα2(I) collagen gene and that of the transgenic construction. At E11, the in situ hybridizations showed discrete signals of expression in the sclerotomal cells (Fig. 5 *A*), in a similar yet broader pattern than the X-gal staining shown in Fig. 3 *B*. Fig. 5 *B* shows high signal intensity in meningeal fibroblasts at E12.5 in a pattern that mimics transgene expression shown in Fig. 3 *E*. Similarly, hybridization in fibroblasts within the connective tissue surrounding the muscles of the tongue (Fig. 5 *C*) at E14.5 was very similar to the X-gal staining of Fig. 4 *C*. At this stage of development, periosteal cells surrounding the developing ribs show a strong hybridization signal (Fig. 5 *D*) that, again was similar to the staining in transgenic embryos (Fig. 4 *E*), and hybridization occurred in fully differentiated osteoblasts within the mandible (data not shown), like that seen for β-galactosidase staining (Fig. 4 *D*).

A Region of the α2(I) Collagen Promoter from −315 to −284 Tandemly repeated Upstream of a Minimal Promoter Can Direct Luciferase Expression in a Tissue-Specific Manner

Previous transient expression studies using cultured fibroblasts had indicated that a region within the α2(I) collagen promoter between −315 and −284 mediates the increased activity of a transfected α2(I) collagen promoter in response to TGF-β1 (Rossi et al., 1988). To examine the possible role of this region in directing the tissue-specific expression of the α2(I) collagen promoter, tandem copies of this sequence were inserted upstream of a very short α2(I) collagen promoter (−40−+54) linked to the highly sensitive firefly luciferase reporter gene (Fig. 1). The short promoter contains a TATA element to which the general transcription factor TFIID pre-

Table III. Luciferase Activity (Luciferase Light U/mg Protein) of LAG1 Transgenic Mice at 4 wk of Age

	LAG1-7	LAG1-9	LAG1-17	LAG1-24
	%	%	%	%
Tail	100	100	100	100
Skin	15	189	24	27
Calvaria	6.7	1.9	7.4	1.1
Liver	0.2	0.1	0.5	0
Muscle	0.2	7	4.0	0
Heart	0.3	0.5	4.0	0.1
Lung	0	0.2	7.7	0.2
Brain	0.9	5.1	9.4	11.3
Testes	27	1.2	17.5	2.2
Kidney	0	0.3	2.8	0.1
Thymus	0	0.2	1.2	0.2
Small Intestine	0.1	0.5	9.6	0

Luciferase values in the different tissues are represented as percent of the values in tail. Luciferase light U/mg protein in tail were 41,400 for LAG1-7, 781,000 for LAG1-9, 96,500 for LAG1-17, and 922,200 for LAG1-24.

sumably binds. In transient expression experiments of fibroblasts in culture, the −40−+54 segment itself is only 1% as active as the −350-bp promoter (Goldberg et al., 1992). Mice, positive for the transgene, were killed at 4 wk of age and assayed for luciferase activity in various tissues (Table III). In all four founders, transgene expression was highest in the tail. Expression in the skin, though more variable, was also elevated. However, the expression of the transgene did not mimic the high level of endogenous Type I collagen RNA found in calvaria. This may suggest that a promoter element necessary for high expression of the transgene in bone is absent from this construction. Tissues that have very low levels of endogenous α2(I) collagen RNA, such as liver, lung, and kidney (Goldberg et al., 1992), also showed very low levels of luciferase activity. Moderate to low levels of luciferase activities were found in brain in three founders and in several other tissues in some founders. These activities could be attributed to the integration site of the transgene since none were consistent across all lines. These results suggest that the −315−−284 sequence plays an important role in mediating the activity of the α2(I) collagen gene in skin and tail. The lack of high level expression of LAG1 in bone warrants further study.

Discussion

Expression of the Type I collagen genes follows a specific and discrete pattern. The present study was performed to examine in detail the tissue-specific activity of promoter segments of the gene for the mouse α2 chain of Type I collagen during embryonic development and to determine which minimal sequences were capable of conferring this discrete tissue-specific activity. The following three major conclusions can be drawn from this study:

First, although the activity of the −350-bp promoter was generally less than that of the −2,000-bp promoter, the pattern of expression of the −350-bp promoter-*lacZ* chimeric gene was the same as that of the −2,000-bp promoter. Both clearly mimicked the expression of the endogenous gene during embryonic development. Hence, the segment between −350 and +54 of the mouse α2(I) collagen promoter contained sufficient information to confer correct temporal

and spatial expression during embryonic development to a reporter gene. This expression was also remarkably cell type-specific, e.g., when the proα2(I) collagen gene was transcriptionally active in a tissue such as muscle, transgene expression was limited to the fibroblasts in the endomysium surrounding the muscle fibers. Since the *lacZ* promoter construction containing the −350−+54 segment was also active in osteoblasts, sufficient information existed in this segment to direct bone-specific transgene expression.

The spatial and temporal activity of the proα2(I) collagen promoter resembled closely the extracellular immuno-histochemical localization of TGF-β1 during embryonic development (Heine et al., 1987; Flanders et al., 1989). It is likely that regardless of the exact location of the cells where TGF-β1 is synthesized, its extracellular localization is critical for its function during embryonic development. Furthermore, several lines of evidence suggest that control of the amounts of active TGF-β1 in the extracellular matrix available for interaction with its signaling receptors is a highly regulated process (for review see Massague, 1990).

Second, when a segment of the α2(I) collagen promoter between −315 and −284 was tandemly repeated upstream of a very short α2(I) collagen promoter (−40−+54), itself linked to the luciferase gene, this chimeric gene displayed a high degree of tissue-specific expression in 4-wk-old founder mice. It should be noted, however, that this construct was only poorly expressed in bones, suggesting that an additional sequence may be needed to determine the correct activity of this promoter in osteoblasts. Since the −350 *lacZ* chimeric gene was active in osteoblasts, it is likely that such a putative bone-specific element would be located within this −350-bp segment. The promoter with the tandemly repeated −315−−284 sequence was much less active than the −2,000- and −350-bp promoters and we were unable to examine its activity in transgenic embryos when it was linked to a β-galactosidase reporter gene (data not shown). Our results, however, suggest an important role for either the −315−−284 sequence or the −40−+54 sequence or both in conferring tissue specificity. It is clear that other *cis*-acting elements exist within the −350-bp proximal promoter and probably outside of it which contribute to the overall strength of the promoter.

The observed correlation between the tissue-specific spatial and temporal activity of the α2(I) collagen promoter and the presence of immunoreactive extracellular TGF-β1 is not simply coincidental but suggests a role for TGF-β1 in the tissue- and cell-specific activity of the proα2(I) collagen gene during embryonic development. Several other observations support this hypothesis. TGF-β1 increases the steady-state levels of type I collagen RNAs in fibroblasts in culture (Raghow et al., 1987; Ignotz et al., 1987; Massague et al., 1990). When injected subcutaneously in newborn mice, TGF-β1 also increases the accumulation of type I collagen at the site of injection (Roberts et al., 1986). In addition, TGF-β1 was shown to stimulate the activity of a −2,000-bp mouse α2(I) collagen promoter chimeric genes in transient expression experiments of fibroblasts (Rosa et al., 1988). The same −315−−284 segment of the α2(I) collagen promoter that was used in a short promoter construction conserving a high degree of tissue specificity in transgenic mice, could mediate the effects of TGF-β1 on the activation of a transfected proα2(I) collagen promoter (Rossi et al., 1988).

TGF-β1 is a member of a family of polypeptides of which several have important roles in early embryonic decisions. Both activins (Thomsen et al., 1990) and TGF-β2 (Rosa et al., 1988) have been shown to induce the synthesis of mesodermic markers in early frog embryos. The temporal and spatial correlation between the activity of the α2(I) collagen promoter-*lacZ* transgene, the expression of endogenous Type I collagen genes, and the extracellular localization of TGF-β1 during embryonic development in the mouse (Heine et al., 1987; Flanders et al., 1989; Heine et al., 1990; R. D'Souza, unpublished results), together with the ability of TGF-β1 to activate the α2(I) collagen promoter in fibroblasts in culture, suggest that Type I collagen genes may constitute an important end point in the pathways controlled by TGF-β1. Type I collagen is a major structural protein during embryonic development. Along with Type I collagen proteins, other collagen proteins, metalloproteinases, and protease inhibitors are synthesized during fibroblast differentiation. The synthesis of these other components is also controlled by TGF-β1 in fibroblasts in culture (for review see Roberts et al., 1988).

Our working model is that the tissue-specific activity of the Type I collagen genes during embryonic development would be controlled both by TGF-β1 and by cell-specific transcription factors that include one which would bind, directly or indirectly, either to the −315−−284 sequence or to the −40−+54 segment. This factor still needs to be identified. Our experiments also raise the question whether TGF-β1 induces the differentiation of fibroblasts from more primitive mesenchymal cells during embryonic development. Since in many tissues the patterns of immuno-histochemical localization of TGF-β2 and TGF-β3 as well as the distribution of the corresponding transcripts overlap with those of TGF-β1 (Pelton et al., 1990; Pelton et al., 1991; Schmid et al., 1991), it is possible that these factors also act in stimulating fibroblast differentiation and Type I collagen expression.

It is difficult to estimate the overall level of expression of the *lacZ* transgenes driven by proα2(I) collagen promoters. Indeed, the observed activity of β-galactosidase is dependent on the rate of synthesis of *lacZ* mRNA but also on its stability and on the stability of the *E. coli* β-galactosidase itself. Furthermore since X-gal staining was performed on whole embryos estimation of the levels of transgene expression, especially in less accessible tissues of older embryos, may not be accurately represented. We believe, nonetheless, that expression of the transgenes that were tested was weaker than that of the endogenous proα2(I) collagen gene. Transfections of DNA into fibroblasts in culture using promoter constructions linked to either the CAT or the firefly luciferase gene have indicated that even the α2(I) collagen promoter constructions containing −2,000 bp were considerably weaker than similar α1(I) collagen promoters, long-terminal repeat viral promoters, or a SV-40 early promoter (Schmidt et al., 1986; Goldberg et al., 1992). Therefore, it is possible that the proximal 2,000 bp of the α2(I) collagen promoter, though tissue specific, lack appropriate enhancer regions necessary to direct high-level expression that are located elsewhere. A locus control region located ~40 kb 5′ from the β-globin gene, confers high-level tissue-specific transgene expression which is independent of the site of integration (Grosveld et al., 1987).

Our results in transgenic mice establish the very high

degree of tissue and cell specificity of Type I collagen promoters. TGF-β1 plays a major role in eliciting the excessive accumulation of extracellular matrix components, including Type I collagen, in certain animal models with fibrotic diseases (Border et al., 1990). Hence, it is possible that the same mechanisms which control the synthesis of fibroblast-specific extracellular matrix components during embryonic development could also be abnormally activated in fibrotic lesions.

References

Adamson, E. D. 1982. The effect of collagen on cell division, cellular differentiation and embryonic development. *In* Collagen in Health and Disease. J. B. Weiss and M. W. Jayson, editors. Churchill Livingstone, Edinburgh, Scotland. 218–243.

Adamson, E. D., and S. E. Ayers. 1979. The localization and synthesis of some collagen types in developing mouse embryos. *Cell.* 16:953–965.

Akhurst, R. J., S. A. Lehnert, D. Gatherer, and E. Duffie. 1989. The role of TGF-β1 in mouse development. *Ann. NY Acad. Sci.* 593:259–271.

Akhurst, R. J., S. A. Lehnert, A. Faissner, and E. Duffie. 1990. TGF beta in murine morphogenetic processes: the early embryo and cardiogenesis. *Development (Camb.)* 108:645–656.

Border, W., S. Okuda, L. Languino, M. Sporn, and E. Ruoslahti. 1990. Suppression of experimental glomerulonephritis by antiserum against transforming growth factor β1. *Nature (Lond.).* 346:371–374.

Bradford, M. M. 1976. A rapid and sensitive method for the quantitation of nanogram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72:248–254.

Cheah, K. S. E., E. T. Lau, P. K. C. Au, and P. P. L. Tam. 1991. Expression of the mouse α1(II) collagen gene is not restricted to cartilage during development. *Development (Camb.).* 111:945–953.

D'Souza, R. N., R.-P. Happonen, N. M. Ritter, and W. T. Butler. 1990. Temporal and spatial patterns of transforming growth factor β1 expression in developing rat molars. *Arch. Oral. Biol.* 35:957–965.

Flanders, K. C., N. L. Thompson, D. S. Cissel, E. Van Obberghen-Schilling, C. C. Baker, M. E. Kass, L. R. Ellingsworth, A. B. Roberts, and M. B. Sporn. 1989. Transforming growth factor-β1: histochemical localization with antibodies to different epitopes. *J. Cell Biol.* 108:653–660.

Goldberg, H., T. Helaakoski, L. A. Garrett, G. Karsenty, A. Pellegrino, G. Lozano, S. Maity, and B. de Crombrugghe. 1992. Tissue-specific expression of the mouse α2(I) collagen promoter: studies in transgenic mice and in tissue culture cells. *J. Biol. Chem.* 267:19622–19630.

Grosveld, F., G. B. Van Assendelft, D. R. Greaves, and G. Kollias. 1987. Position-independent, high-level expression of the human β-globin gene in transgenic mice. *Cell.* 51:975–985.

Harbers, K., M. Kuehn, H. Delius, and R. Jaenisch. 1984. Insertion of retrovirus into the first intron of the α1(I) collagen gene leads to embryonic lethal mutation in mice. *Proc. Natl. Acad. Sci. USA.* 81:1504–1508.

Hatamochi, A., P. T. Golumbek, E. Van Schaftingen, and B. de Crombrugghe. 1988. A CCAAT DNA binding factor consists of two different components that are both required for DNA binding. *J. Biol. Chem.* 263:5940–5947.

Heine, U. I., E. F. Munoz, K. C. Flanders, L. R. Ellingsworth, H.-Y. P. Lam, N. L. Thompson, A. B. Roberts, and M. B. Sporn. 1987. Role of transforming growth factor-β in the development of the mouse embryo. *J. Cell Biol.* 105:2861–2876.

Heine, U. I., E. F. Munoz, K. C. Flanders, A. B. Roberts, and M. B. Sporn. 1990. Colocalization of TGF-β1 and collagen I and III, fibronectin and glycosaminoglycans during lung branching morphogenesis. *Development (Camb.).* 109:29–36.

Hogan, B., F. Costantini, and E. Lacy. 1986. Manipulating the Mouse Embryo: A Laboratory Manual. Cold Spring Harbor, NY.

Ignotiz, R. A., T. Endo, and J. Massague. 1987. Regulation of fibronectin and Type I collagen mRNA levels by transforming growth factor-β. *J. Biol. Chem.* 262:6443–6446.

Karsenty, G., and B. de Crombrugghe. 1991. Conservation of binding sites for regulatory factors in the coordinately expressed α1(I) and α2(I) collagen

promoters. *Biochem. Biophys. Res. Commun.* 177:538–544.

Karsenty, G., P. Golumbek, and B. de Crombrugghe. 1988. Point mutations and small substitution mutations in three different upstream elements inhibit the activity of the mouse α2(I) collagen promoter. *J. Biol. Chem.* 263:13909–13915.

Khillan, J. S., A. Schmidt, P. Overbeek, B. de Crombrugghe, and H. Westphal. 1986. Developmental and tissue-specific expression directed by the α2(I) collagen promoter in transgenic mice. *Proc. Natl. Acad. Sci. USA.* 83:725–729.

Koster, M., S. Plessow, J. H. Clement, A. Lorenz, H. Tiedemenn, and W. Knochel. 1991. Bone morphogenetic protein 4, a member of the TGF-β family, in early embryos of Xenopus laevis: analysis of mesoderm inducing activity. *Mech. Dev.* 33:191–199.

Leivo, I., A. Vaheri, R. Timpl, and J. Wartiovaara. 1980. Appearance and distribution of collagens and lamin in the early mouse embryo. *Dev. Biol.* 76:100–114.

Maity, S. N., P. T. Golumbek, G. Karsenty, and B. de Crombrugghe. 1988. Selective activation of transcription by a novel CCAAT binding factor. *Science (Wash. DC).* 241:582–585.

Massague, J. 1990. The transforming growth factor-β family. *Annu. Rev. Cell Biol.* 6:597–641.

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 and other neurons in adult transgenic mice. *Neuron.* 7:703–716.

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 α1(I) collagen promoter in osseous and fibroblastic cells. *J. Cell Biol.* 116:227–236.

Pelton, R. W., M. E. Dickinson, H. L. Moses, and B. L. M. Hogan. 1990. In situ hybridization analysis of TGFβ3 RNA expression during mouse development: comparative studies with TGFβ1 and β2. *Development (Camb.).* 110:609–620.

Pelton, R. W., B. Saxena, M. Jones, H. L. Moses, and L. I. Gold. 1991. Immunohistochemical localization of TGFβ1, TGFβ2, and TGFβ3 in the mouse embryo: Expression patterns suggest multiple roles during embryonic development. *J. Cell Biol.* 115:1091–1105.

Raghow, R., A. E. Postlethwaite, J. Koski-Oja, H. L. Moses, and H. Kang. 1987. Transforming growth factor-β increases steady-state levels of type I procollagen and fibronectin messenger RNAs posttranscriptionally in cultured human dermal fibroblasts. *J. Clin. Invest.* 79:1285–1288.

Reddi, A. H. 1984. Extracellular matrix and development. *In* Extracellular Matrix Biochemistry. K. A. Piez and A. H. Reddi, editors. Elsevier Science Publishing Co., Inc., New York. 375–412.

Ritzenthaler, J. D., G. H. Goldstein, A. Fine, A. Lichtler, D. W. Rowe, and B. D. Smith. 1991. Transforming growth factor-β activation elements in the distal promoter regions of the rat α1 type I collagen gene. *Biochem. J.* 280:157–162.

Roberts, A. B., M. B. Sporn, R. K. Assoian, J. M. Smith, N. S. Roche, L. M. Wakefield, U. I. Heine, L. A. Liotta, V. Falanga, J. H. Kehrl, and A. S. Fauci. 1986. Transforming growth factor type-beta: rapid induction of fibrosis and angiogenesis in vivo and stimulation of collagen formation in vitro. *Proc. Natl. Acad. Sci. USA.* 83:4167–4171.

Roberts, A. B., K. C. Flanders, P. Kondaiah, N. I. Thompson, E. Van Obberghen-Schilling, L. Wakefield, P. Rossi, B. de Crombrugghe, J. L. Heine, and M. B. Sporn. 1988. Transforming growth factor β: biochemistry and roles in embryogenesis, tissue repair, remodeling and carcinogenesis. *Recent Prog. Horm. Res.* 44:157–197.

Rosa, F., A. B. Roberts, D. Danielpour, L. L. Dart, M. B. Sporn, and I. B. David. 1988. Mesoderm induction in amphibians: the role of TGF-β2-like factors. *Science (Wash. DC).* 239:783–785.

Rossi, P., G. Karsenty, A. B. Roberts, N. S. Roche, M. B. Sporn, and B. de Crombrugghe. 1988. A nuclear factor binding site mediates the transcriptional activation of type I collagen promoter by transforming growth factor-β. *Cell.* 52:405–414.

Schmid, P., D. Cox, G. Bilbe, R. Maier, and G. K. McMaster. 1991. Differential expression of TGFβ1, β2 and β3 genes during mouse embryogenesis. *Development (Camb.).* 111:117–130.

Schmidt, A., P. Rossi, and B. de Crombrugghe. 1986. Transcriptional control of the mouse α2(I) collagen gene: Functional deletion analysis of the promoter and evidence for cell specific expression. *Mol. Cell Biol.* 6:347–354.

Schnieke, A. H., K. Harbers, and R. Jaenisch. 1983. Embryonic lethal mutation in mice induced by retrovirus insertion into the α1(I) collagen gene. *Nature (Lond.).* 304:315–320.

Slack, J. L., D. J. Liska, and P. Bornstein. 1991. An upstream regulatory region mediates high-level, tissue specific expression of the human α1(I) collagen gene in transgenic mice. *Mol. Cell Biol.* 11:2066–2074.

Southern, E. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98:503–517.

Thomsen, G., T. Woolf, M. Whitman, S. Sokol, J. Vaughan, W. Vale, and D. A. Melton. 1990. Activins are expressed early in Xenopus embryogenesis and can induce axial mesoderm and anterior structures. *Cell.* 63:485–493.

Wilkinson, D. G., J. A. Bailes, and A. P. McMahon. 1987. Expression of the proto-oncogene int-1 is restricted to specific neural cells in the developing mouse embryo. *Cell.* 50:79–88.

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.