Expression of Two Nonallelic Type II Procollagen Genes during *Xenopus laevis* Embryogenesis Is Characterized by Stage-specific Production of Alternatively Spliced Transcripts

Ming-Wan Su,* Hiroaki R. Suzuki,‡ James J. Bieker,* Michael Solursh,‡ and Francesco Ramirez*

*Brookdale Center for Molecular Biology, Mt. Sinai School of Medicine, New York, New York 10029; and ‡The Department of Biology, University of Iowa, Iowa City, Iowa 52242

Abstract. The pattern of type II collagen expression during *Xenopus laevis* embryogenesis has been established after isolating specific cDNA and genomic clones. Evidence is presented suggesting that in *X. laevis* there are two transcriptionally active copies of the type II procollagen gene. Both genes are activated at the beginning of neurula stage and steady-state mRNA levels progressively increase thereafter. Initially, the transcripts are localized to notochord, somites, and the dorsal region of the lateral plate mesoderm. At later stages of development and parallel to increased mRNA accumulation, collagen expression becomes progressively more confined to chondrogenic regions of the tadpole. During the early period of mRNA accumulation, there is also a transient pattern of expression in localized sites that will later not undergo chondrogenesis, such as the floor plate in the ventral neural tube. At later times and coincident with the appearance of chondrogenic tissues in the developing embryo, expression of the procollagen genes is characterized by the production of an additional, alternatively spliced transcript. The alternatively spliced sequences encode the cysteine-rich globular domain in the NH_2-propeptide of the type II procollagen chain. Immunohistochemical analyses with a type II collagen monoclonal antibody documented the deposition of the protein in the extracellular matrix of the developing embryo. Type II collagen expression is therefore temporally regulated by tissue-specific transcription and splicing factors directing the synthesis of distinct molecular forms of the precursor protein in the developing *Xenopus* embryo.

T HE structural integrity of the vertebrate body is for the most part determined by the deposition of collagen molecules in the ECM of connective tissue. ECMs and implicitly collagens also play a dynamic role during development by regulating critical processes of cell adhesion, cell proliferation, and cell movement (for a review see Thiery, 1989). The role of ECM molecules during vertebrate development has been analyzed mostly in mouse and chicken embryos, particularly in reference to neural crest cell migration and limb formation (Lallier and Bronner-Fraser, 1990; Solursh, 1990). Amphibians, on the other hand, have provided a more suitable experimental system to study the function of ECM components during gastrulation and early morphogenesis (Keller and Winklbauer, 1990).

Unfortunately, very little is known about the identity and pattern of expression of collagen molecules during early stages of amphibian embryogenesis. At the time this work was undertaken, there had been just a single report showing that collagen synthesis, measured by hydroxyproline incorporation, commences in the *Xenopus* embryo at neurula stage and increases henceforth >500-fold (Green et al., 1968). Very recently, biochemical and immunohistochemical analyses have documented the presence of a maternal type VI-like collagen that is possibly involved in *Xenopus* gastrulation (Otte et al., 1990).

An understanding of how collagen expression is regulated during early stages of *Xenopus* development is undoubtedly a fruitful task, especially in view of recent evidence indicating that in this organism mesodermal specification is mediated by the same group of growth factors that modulate the synthesis of collagens and other ECM components in mammalian systems (Smith, 1989; Melton and Whitman, 1989). More generally, availability of collagen probes may provide sensitive cell lineage–specific molecular markers in induction studies (Gurdon, 1987). With these considerations in mind we searched for collagen genes that are regulated during *X. laevis* embryogenesis. As a result, we characterized the structure and developmental pattern of expression of two genes that correspond to the single-copy, cartilage-specific type II procollagen gene of higher vertebrates.

Materials and Methods

Culture Conditions

Female *Xenopus laevis* frogs, purchased from Xenopus I (Ann Arbor, MI),

were induced to ovulate with 100 U pregnant mare serum (Calbiochem-Behring Corp., La Jolla, CA), followed by a second injection of 1,000 U of human chorionic gonadotropin (Sigma Chemical Co., St. Louis, MO). Eggs were fertilized in vitro and embryos were cultured at 23°C according to the published protocols (Newport and Kirschner, 1982; Kimelman and Kirschner, 1987). Staging of embryos was according to the normal table of Nieuwkoop and Faber (1967).

Isolation and Sequencing of cDNA and Genomic Clones

Approximately 2×10^5 recombinants of a neurula stage (stage 17) cDNA library in the lambda-gt10 vector (Kintner and Melton, 1987) were screened with a full-size human type II procollagen cDNA (Su et al., 1989) at 37°C in 40% formamide, 6× SSC, 5× Denhart's solution, and 100 µg/ml of sheared salmon sperm DNA (Sambrook et al., 1989). Washing of the library filters was in 2× SSC at room temperature followed by several washes in decreasing salt conditions (down to 0.5× SSC) at 50°C. For the isolation of the type II procollagen genes, X. laevis genomic DNA, partially digested with Sau3AI and size fractionated on a sucrose gradient, was inserted into the BamHI restriction site of the lambda-DASH vector using a commercially available kit (Stratagene, La Jolla, CA) and following the manufacturer's recommendations. Screening of the genomic library and purification of positive recombinants were as for the standard protocols (Sambrook et al., 1989). Positive cDNA and genomic clones were subcloned in the pUC18 vector and sequenced using the dideoxynucleotide chain termination procedure on double stranded DNA (Zagursky et al., 1986). Sequencing of both DNA strands was achieved by generating progressively overlapping deletions with the exonuclease III/mung bean nuclease method (Sambrook et al., 1989) or by using synthetic oligonucleotide primers. Sequences were analyzed using the computer program of Mount and Conrad (1987).

RNA Purification, Northern Blot Analysis, and PCR Amplification

Total RNA was purified from unfertilized eggs and various embryonic stages by the guanidinium thiocyanate/cesium chloride method (Sambrook et al., 1989). Approximately 10 µg of RNA from each sample was electrophoresed in a 0.8% agarose gel in formamide/formaldehyde, transferred onto a Zeta-Probe nylon membrane (Bio-Rad Laboratories, Richmond, CA), and hybridized to cDNA uniformly labeled by random priming (Sambrook et al., 1989). The collagen probe was a 700-bp fragment that recognizes both 1(II) and 1(II)' sequences (general probe, see Fig. 2). The control probe was a cytoskeletal actin cDNA (Mohun et al., 1984). For polymerase chain reaction (PCR)[1] amplification of cDNA molecules, a commercial kit (Amersham Corp., Arlington Heights, IL) was used for the random primed synthesis of double stranded cDNA on 5 µg of template total RNA according to the manufacturer's recommendations. One-fourth of the reaction product was amplified with Taq DNA polymerase (Perkin Elmer-Cetus, Norwalk, CT) using collagen-specific oligonucleotide primers in a programmable heat block (Perkin Elmer-Cetus) following published conditions (Kawasaki and Wang, 1989). The same protocol was used for the PCR amplification of 1 µg of genomic DNA. Products of the PCR reaction, electrophoresed in a 2% NuSieve (FMC-BioProducts, Rockland, ME) agarose gel, were visualized by ethidium bromide staining or by autoradiography after gel electrophoresis of ^{32}P-end-labeled products separated in 7 M urea, 5% polyacrylamide (Sambrook et al., 1989). Products of the cDNA and genomic amplifications were subcloned into the SamI site of pUC18 and, in each case, five independent clones were sequenced as described above. Oligonucleotide primers were synthesized on an Applied Biosystems, Inc. (Foster City, CA) model 380A DNA synthesizer. Primer sequences are: exon 1, 5', ATGTTTAGCTTTGTGGATTCAAGG 3'; exon 6, 5' TACCCT-CAATACGTTCCTGGTTAC 3'; exon 49 reverse, 5' TTGCTGGCCTATCT-CAACCT 3'; exon 49 forward, 5' GGAACTTAGGTGTTTCGATTTGTA 3'; exon 50, 5' CAAGAAGAACTGGTGGG 3'; exon 51, 5' GGGTCTTGTAGT-GAAAT 3'.

In Situ Hybridization

In situ hybridization experiments were performed using albino Xenopus embryos according to the protocol previously described (Su et al., 1991). The

---

1. *Abbreviation used in this paper*: PCR, polymerase chain reaction.

700-bp long general probe or a 209-bp insert encoding the cystein-rich globular domain (exon 2 probe) were subcloned in the HindIII:EcoRI sites of the transcription vector pT7/T3-19 (Life Technologies Inc., Gaithersburg, MD). Uniformly ^{35}S-labeled sense and antisense RNAs were synthesized by transcribing 2 µg of linearized subclones with T3 or T7 RNA polymerase (Su et al., 1991). Both sense and antisense probes were used for all stages. The sense probe gave only background signal and is not illustrated.

Antibody Immunostaining

Whole mount staining of albino embryos was carried out as described in Dent et al. (1989) using a monoclonal antibody to chicken type II collagen (Linsenmayer and Hendrix, 1980). For binding visualization, secondary antibodies coupled to horseradish peroxidase were used, and the peroxidase reaction was done in the presence of diaminobenzidine. Antibody incubation was performed in Tris-buffered saline with 5% fetal calf serum without azide; washes contained 0.05% Tween-20. Stained embryos were embedded in paraffin and serial cross sections were examined to determine the cellular localization of the immunoreactivity. In some cases embryos were fixed in Bouin's fixative for 10 min at room temperature, embedded in paraffin, and sections processed for immunostaining.

Results

Identification of Type II Procollagen Clones

We screened a cDNA library derived from stage 17 embryonic mRNA with a human type II procollagen cDNA which, amongst all collagen probes, cross-hybridized best with X. laevis sequences. This led to the identification of several recombinants which, upon further characterization, were found to code for two polypeptides displaying nearly 96% sequence identity (Fig. 1). The conceptual amino acid translation of the cDNA sequences revealed that the polypeptides exhibit structural features of a fibril-forming procollagen molecule (Vuorio and de Crombrugghe, 1990). They consist of a long uninterrupted triple-helical domain flanked by two globular propeptides, and display a short leader sequence at the amino terminus (Fig. 2). Comparative analysis documented the highest level of sequence homology with the procollagen chain of human type II collagen (Table I). Characterization of selected portions of the corresponding X. laevis genes revealed an exon/intron organization nearly identical to the well-established structure of a prototypical fibrillar collagen gene (Fig. 3) (Vuorio and de Crombrugghe, 1990). Interestingly, both mRNAs display around the translation initiation codon an interrupted palindrome that can potentially form a stable hairpin loop structure (data not shown). Such a sequence is shared by all vertebrate fibrillar procollagen mRNAs with the exception of the type II (Vuorio and de Crombrugghe 1990).

Based on these data, the two frog genes were named α1(II) and α1(II'). In addition to the coding sequences, they differ slightly from each other in the restriction map and the relative organization of some exons (Fig. 3). Thus, they may represent allelic forms of the same gene or duplicated genes. To discriminate between these two possibilities, we used the PCR amplification technique on genomic DNA from tetraploid and homozygous diploid X. laevis, as well as from diploid Xenopus tropicalis (Kobel and Du Pasquier, 1986). Two distinct sets of primers were used. The first two oligonucleotides correspond to highly conserved sequences of exons 50 and 51. These exons in the α1(II) and α1(II)' genes are separated by a 180- and a 151-bp intron, respectively (Fig. 3 a). The other two primers flank a small 6-bp insertion in exon

```
a1(II)   MFSFVDSRTLVLFAATQVILLAVVRCQDEEDVLDTGSCVQHGQRYSDKDVWKPEPCQICVCDTGTVLCDDIICEESKDCPNAEIPFGECCPICPTEQSSTSSGQGVLKGQKGEPGDIKDV
a1(II)'                                   D                    E                                                    A
                                                     180
a1(II)   LGPRGPPGPQGPSGEQGSRGERGDKGEKGAPGPRGRDGEPGTPGNPGPVGPPGPP***GLGGNFAAQMTGGFDEKAGGAQMGVMQGPMGPMGPRGPPGPTGAPGPQGFQGNPG
a1(II)'  V              P       T               A       GPP                                                  S
             240                                                           300
a1(II)   EPGEPGAGGPMGPRGPPGPSGKPGDDGEAGKPGKSGERGPPGPQGARGFPGTPGLPGVKGHRGYPGLDGAKGEAGAAGAKGEGGATGEAGSPGPMGPRGLPGERGRPGSSGAAG
a1(II)'  S·                                                                      G                  S  S    A
                 360                                                   420
a1(II)   ARGNDGLPGPAGPPGPVGPAGAPGFPGAPGSKGEAGPTGARGPEGAQGPRGESGTPGSPGPAGASGNPGTDGIPGAKGSSGGPGIAGAPGFPGPRGPPGPQGATGPLGPKGETGD
a1(II)'                                                                           S                 AS
                     480                                                   540
a1(II)   PGVAGFKGEQGPKGEIGSAGPQGAPGPAGEEGKRGARGEPGAAGPNGPPGERGAPGNRGFPGQDGLAGPKGAPGERGVPGLGGPKGGNGDPGRPGEPGLPGARGLTGRPGDAGPQG
a1(II)'   I    H                                              L
                     600                                                   660
a1(II)   KVGPSGASGEDGRPGPPGPQGARGQPGVMGFPGPKGANGEPGKAGEKGLVGAPGLRGLPGKDGETGSQGPNGPAGPAGERGEQGPPGPSGFQGLPGPPGSPGEGGKPGDQGVPGEA
a1(II)'          A                                       G           A
                     720                                                   780
a1(II)   GAPGLVGPRGERGFPGERGSSGPQGLQGPRGLPGTPGTDGPKGASGPSGPNGAQGPPGLQGMPGERGAAGISGPKGDRGDTGEKGPEGASGKDGSRGLTGPIGPPGPAGPNGEKGES
a1(II)'    A                                    S           T                             P           L
                 800                                                   900
a1(II)   GPSGPPGIVGARGAPGDRGENGPPGPAGFAGPPGADGQSGLKGDQGESGQKGDAGAPGPQGPSGAPGPQGPTGVFGPKGARGAQGPAGATGFPGAAGRVGTPGPNGNPGPPGPPG
a1(II)'                             S  A                                 N
                     960                                                   1020
a1(II)   SAGKEGPKGVRGDAGPPGRAGDPGLQGAAGAPGEKGEPGEDGPSGPDGPPGPQGLSGQRGIVGLPGQRGERGFPGLPGPSGEPGKQGGPGSSGDRGPPGPVGPPGLTGPSGEPGREGN
a1(II)'            T       P                      S    N                                 A  T
                     1080                                                  1140
a1(II)   PGSDGPPGRDGATGIKGDRGETGPLGAPGAPGAPGAPGSVGPTGKQGDRGESGPQGPLGPSGPAGARGLAGPQGPRGDKGEAGEAGERGQKGHRGFTGLQGLPGPPGSAGDQGATG
a1(II)'              I    S   P           S           N                                           T
                         1200                                               1250
a1(II)   PAGPAGPRGPPGPVGPSGKDGSNGISGPIGPPGPRGRSGETGPSGPPGQPGPPGPPGPGPGIDMSAFAGLSQPEKGPDPMRYMRADQASNSL**PVDVEATLKSLNNQIENIRSPDG
a1(II)'  S  S                   P                A       S                                     S VLQTA   N  T    S
                         1320                                               1380
a1(II)   TKKNPARTCRDLKICHPEWKSGDYWIDPNQGCTVDAIKVFCDMETGETCVYPNPSKIPKKNWWSAKGKEKKHIWFGETINGGFQFSYGDDSSAPNTANIQMTFLRLLSTDASQNITYHCK
a1(II)'       L       E     E               N    S    S                                           L
                     1440
a1(II)   NSIAFMDEASGNLKKAVLLQGSNDVEIRAEGNSRFTYNALEDGCKKHTGKWSKTVIEYRTQKTSRLPIVDIAPMDIGGADQEFGVDIGPVCFL
a1(II)'
```

Figure 1. Comparison of the single-letter amino acid translation of *Xenopus laevis* type II procollagen cDNAs (termed in this figure al(II) and al(II)'). Asterisks indicate gaps inserted to give best alignment. The nucleotide sequence data, including non-coding segments, is available from EMBL/GenBank/DDBJ under accession numbers M63595 and M63596.

49 of αl(II)' (Fig. 3 *a*). Analysis of end-labeled PCR products separated by polyacrylamide gel electrophoresis suggested that a single type II procollagen gene is present in *X. tropicalis*, while there are two nonallelic copies in both tetraploid and homozygous diploid *X. laevis* (Fig. 3 *b*). These data are consistent with other findings suggesting that the tetraploid *X. laevis* contains duplicated forms of genes that are present in single copy in the diploid *X. tropicalis* and in other vertebrate species (Wahli and Dawid, 1980; Jeffreys et al., 1980; Stutz and Spohr, 1986; Fritz et al., 1989; Shuldiner et al., 1989).

Developmental Accumulation of Type II Procollagen mRNA

To establish the pattern of type II procollagen mRNA accumulation during *Xenopus* embryogenesis, RNA from different stages of development was analyzed by Northern blot hybridization. The probe used for this analysis was a cDNA sequence that covers the 3'-untranslated region of the mRNAs and which includes portions of the C-propeptide coding domain (Fig. 2). This probe is also termed general probe since it displays >95% nucleotide sequence identity between the αl(II) and αl(II)' mRNAs.
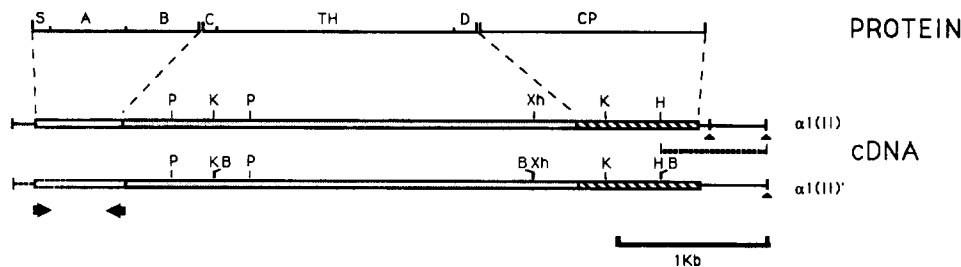
Figure 2. Composite of the restriction maps of the type II procollagen cDNAs shown in relation to the protein domains. Signal peptide (*S*), cysteine-rich globular region (*A*), and interrupted collagenous region (*B*) in the N-propeptide domain. NH₂-terminal (*C*) and COOH-terminal (*D*) telopeptides, and triple-helical region (*TH*) in the α-chain domain. CP signifies C-propeptide domain. In the cDNAs, letters indicate the following enzymes: *B*, BamHI; *H*, HindIII; *K*, KpnI; *P*, PstI; and *Xh*, XhoI. The dotted line denotes the cDNA fragment used for Northern and in situ hybridizations (general probe). Arrows and triangles indicate the positions of the exons 1 and 6 primers used in some of the PCR amplification experiments and the relative location of the polyadenylation sites, respectively. Note that the 5' sequence of αl(II)' is derived from analysis of a genomic clone (*broken line box*).

Table I. Amino Acid Homology between Frog and Human Collagen Propeptides*

X. laevis	N-propeptide		C-propeptide	
	α1 (II)	α1 (II)′	α1 (II)	α1 (II)′
	%			
Human α1 (II)	82	83	83	82
Human α1 (I)	53	52	73	73
Human α1 (III)	43	43	61	60
Human α2 (V)	50	51	56	55

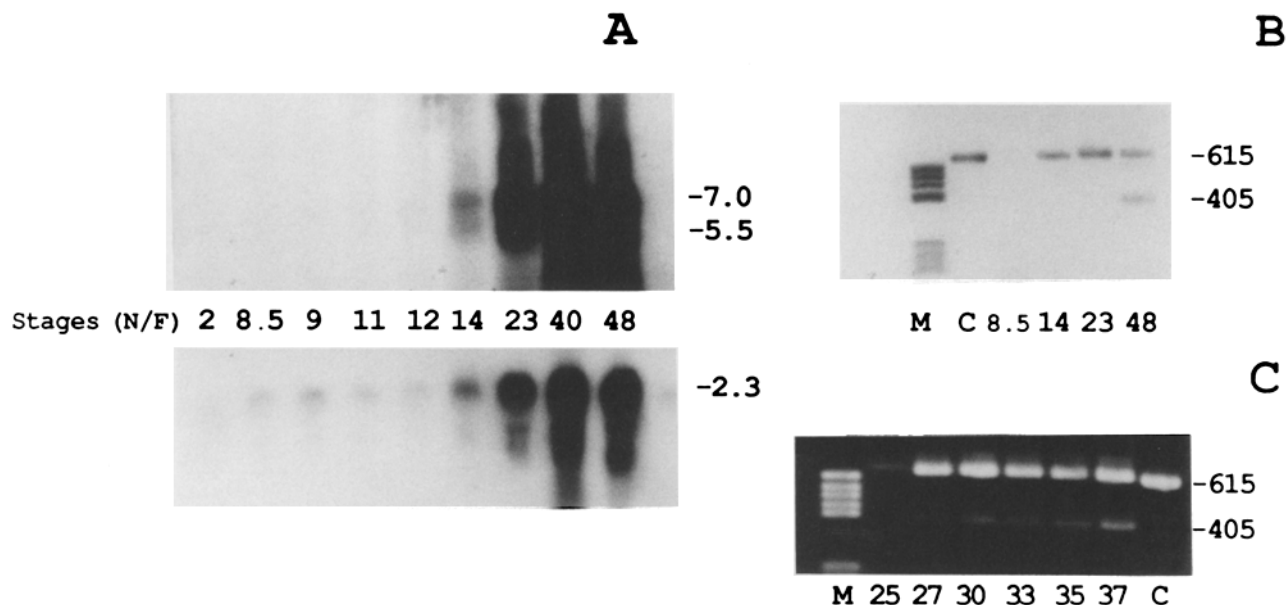* The triple-helical domains were not included in this analysis because of their highly repetitive Gly-X-Y triplet structure.

Procollagen transcripts were first detected shortly after the end of gastrulation in stage 14 embryos, and found to rapidly accumulate throughout the tadpole stage (Fig. 4 A). The resulting profile of steady-state type II procollagen mRNA accumulation closely resembles that of cytoskeletal actin at post-gastrulation stages (Fig. 5) (Mohun et al., 1984). These data confirm early biochemical results indicating that collagen synthesis in *Xenopus* begins at about neurula stage (Green et al., 1968).

Recent work on the human type II procollagen gene has revealed that distinct populations of chondrocytes produce transcripts in which the sequence of exon 2 is alternatively spliced (Ryan and Sandell, 1990). Exon 2 encodes the cysteine-rich globular region of the N-propeptide (Vuorio and de Crombrugghe, 1990). The N-propeptide of fibril-forming collagens has been postulated to play a regulatory role in fibrillogenesis and/or feedback inhibition of collagen biosynthesis (Wiestner et al., 1979; Paglia et al., 1981; Fleischmajer et al., 1981, 1983; Wu et al., 1986). Based on this observation, we tested whether type II procollagen expression during frog embryogenesis was associated with the generation of alternatively spliced products as well. To this end, reverse transcribed RNAs from various developmental stages were PCR amplified using primers specific for exons 1 and 6 sequences (Fig. 2). This identified an alternatively spliced, shorter product in the tailbud stage (Fig. 4 B), whose earliest appearance was timed at ∼8 h into this developmental stage (Figs. 4 C and 5). Incidentally, sequencing of PCR products confirmed that both type II procollagen genes are concomitantly transcribed at the neurula stage, and that their transcripts are alternatively spliced at the same time during the tailbud stage.
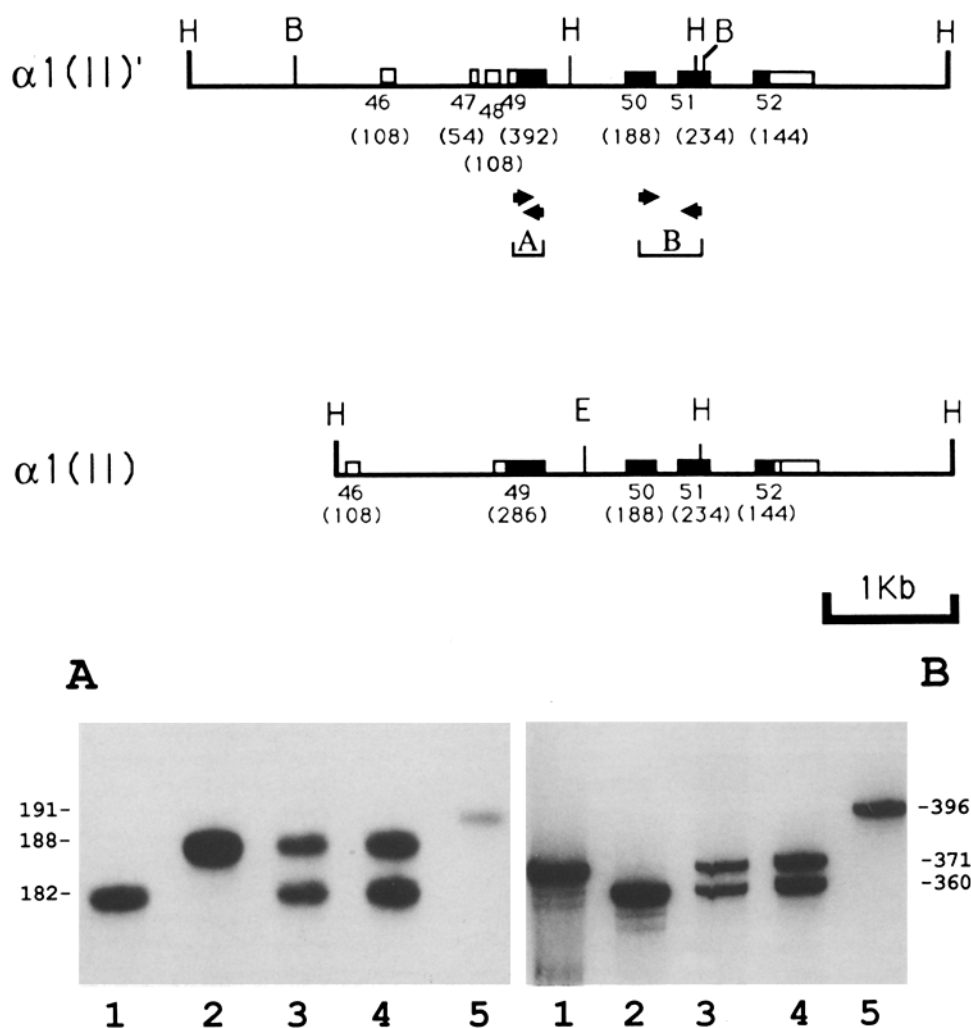
Figure 3. (Top) Partial restriction map and exon organization of the X. laevis type II procollagen genes. Exons (numbered according to the human gene) are denoted by black (C-propeptide), gray (triple helix), and white (3′ untranslated) rectangles. Numbers in parentheses signify size of exons in basepairs; note that the size of exon 52 refers only to its coding segment. B, BamHI; E, EcoRI; H, HindIII. Arrows indicate the relative positions of primers used in PCR amplification shown in A and B. In each autoradiogram, DNA samples are from α1(II) phage clone (1), α1(II)′ phage clone (2), tetraploid (3), and homozygous diploid (4) X. laevis, and X. tropicalis (5). Sizes of the bands are in basepairs.

A

B

C

Stages (N/F) 2 8.5 9 11 12 14 23 40 48

−7.0
−5.5

−2.3

M C 8.5 14 23 48

−615
−405

−615
−405

M 25 27 30 33 35 37 C

Figure 4. Pattern of accumulation and alternative splicing of type II procollagen mRNA. (*A*) Northern blot hybridization of total RNA (10 μg) to collagen and actin probes (*top and bottom*, respectively). Numbers indicate stages of *X. laevis* development according to the table of Nieuwkoop and Faber (1967) (N/F). Sizes of bands are expressed in kilobases. (*B*) PCR amplification of reverse transcribed RNA from the indicated (N/F) stages of development using exons 1 and 6 flanking primers (see Fig. 2). Tracks M and C contain DNA size markers and products from the PCR amplification of type II procollagen cDNA, respectively. Sizes of bands are in basepairs. A more precise determination of the time of appearance of the 405-bp band is shown in *C*.

### Tissue Distribution of Procollagen Transcripts and Protein Immunolocalization

The spatial distribution of type II procollagen mRNA was ascertained by in situ hybridization with two different antisense riboprobes. The first was the same general probe used in the Northern blot analysis, while the second con-

tained sequence encoding only the cystein-rich globular domain (exon 2 probe). As indicated below, the results observed with the two probes were essentially the same.

Based on the Northern blot analysis the earliest stage examined by in situ hybridization was stage 14 (Fig. 6). At this time of *Xenopus* development, the type II procollagen mRNA was detected in the notochord, sub-notochordal rod,
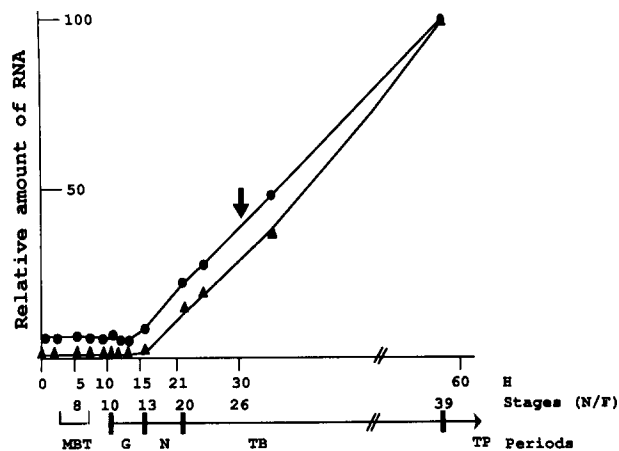
Figure 5. Overall profile of procollagen transcript accumulation plotted after quantification of the signals seen in Fig. 4 *A*. Values are normalized for embryo number. The established profile of cytoskeletal actin mRNA accumulation (*solid circles*) is shown for comparison. Its actual amount, relative to collagen mRNA (*solid triangles*), is 40 times greater at stage 14 and 5 times greater at stage 40. The arrow indicates the stage at which the alternatively spliced products first appear. Developmental periods and (N/F) stages are shown below the graph. *MBT*, midblastula transition; *G*, gastrula; *N*, neurula; *TB*, tailbud and *TP*, tadpole stages.
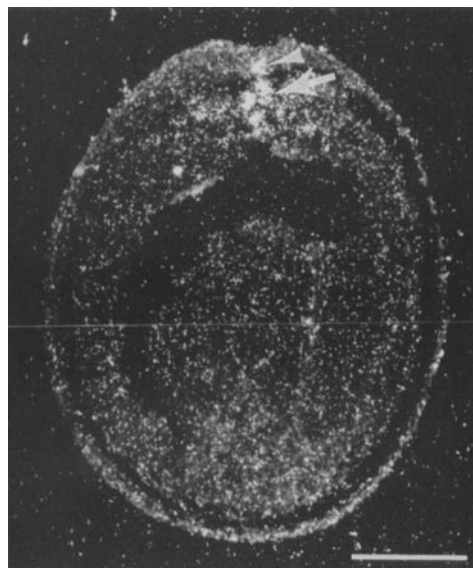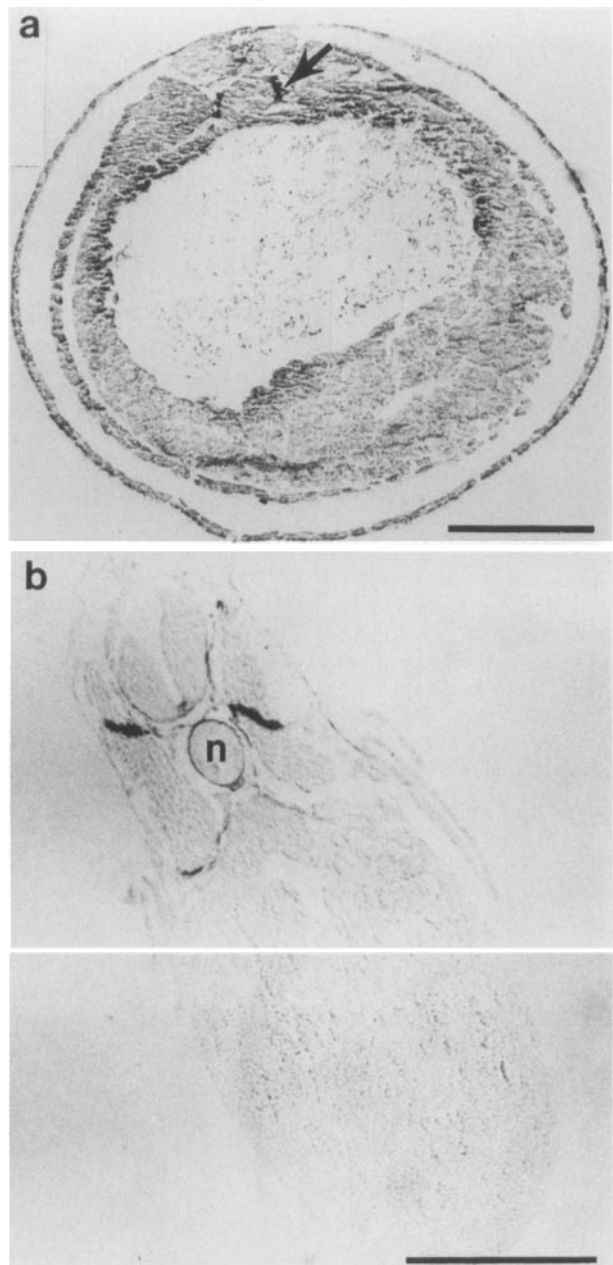

Figure 6. Dark field micrograph of a cross section through a stage 14 embryo probed for type II procollagen mRNA with an exon 2 sequence. Essentially identical results are obtained with the general probe (not shown). There is significant signal in the notochord (arrow), sub-notochordal rod, and peri-notochordal regions of the somites, as well as in the floor plate (arrowhead). Bar, 250 μm.
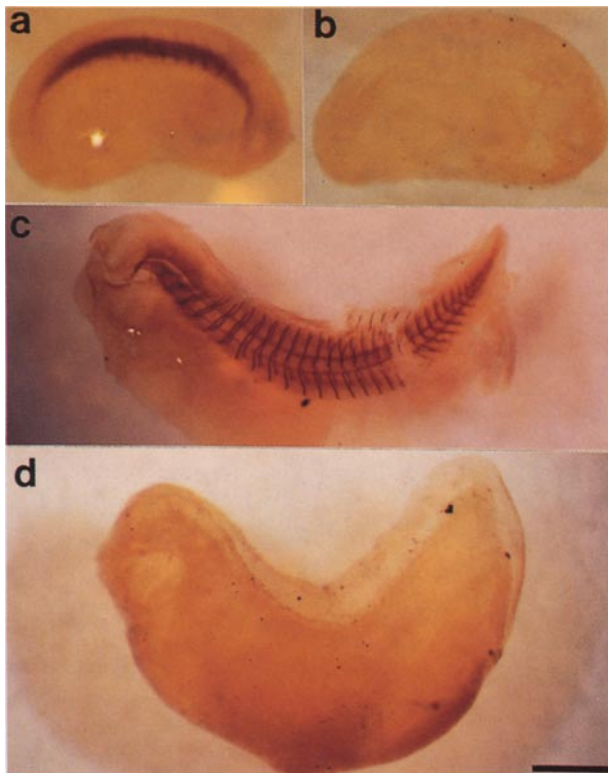
Figure 7. Immunohistochemical localization of type II collagen in stage 20 and 31 embryos. The embryos were processed by a whole mount procedure and bound antibody to type II collagen was localized with a peroxidase tag. At stage 20 (*a*), antigen can be observed in the peri-notochordal region which already appears segmentally organized. The embryo in *b* was treated identically to that in *a* except it was incubated with culture supernatant from the nonimmunoglobulin secreting parental myeloma cell line to serve as a negative control. Immunolocalization of type II collagen at stage 31 (*c*); *d* shows a control embryo. Note the immunoreactivity localized to the vertebrae. Bar, 500 μm.

the somites, and the floor plate. The protein must be translated shortly after this stage, since by stage 20 type II collagen can be detected by immunohistochemistry in this same region of the embryo. The perinotochordal localization of antigen can be observed in both whole mount (Fig. 7, *a* and *b*) and sectioned (Fig. 8 *a*) specimens. However, the accumulation of type II collagen in the ECM, as judged at the light microscopic level, does not exactly parallel the distribution of the mRNA. Antigen is largely restricted to the perinotochordal region at this stage. It is not known whether this difference between antigen and mRNA localization reflects limits in detection of the immunohistological method or differences in antigen accumulation.

In situ hybridization analysis of stage 21 (tailbud) embryos showed the developmental progression of type II procollagen expression. This is readily evident when developmentally younger posterior parts of the embryo are compared with developmentally more advanced anterior regions (Fig. 9). This comparison revealed a widespread but lower level of mRNA expression more posteriorly, followed by a more restricted but relatively elevated level of expression more anteriorly. For example, in the posterior region type II procollagen tran-

Figure 8. Immunohistochemical localization of type II collagen; (*a*) bright-field micrograph of a cross section through a stage 21 embryo that has been stained for type II collagen, demonstrating the localization of reaction product in the peri-notochordal region (*arrow*). *b* is a bright-field micrograph of a section through the same embryo shown in Fig. 7 *c*, without any additional staining in order to localize the immunoreactivity further. Note the presence of immunoreactivity in the basement membranes around the notochord (*n*) and on the ventral side of the neural tube, as well as in the floor plate, the sub-notochordal rod and in the developing vertebra. Bars, 250 μm.

scripts were detected at the highest level in the notochord. However, a significant signal was also seen throughout the somites and in the dorsal half of the lateral plate mesoderm. More anteriorly, transcripts were no longer detected in the lateral plate mesoderm but rather they were restricted to the
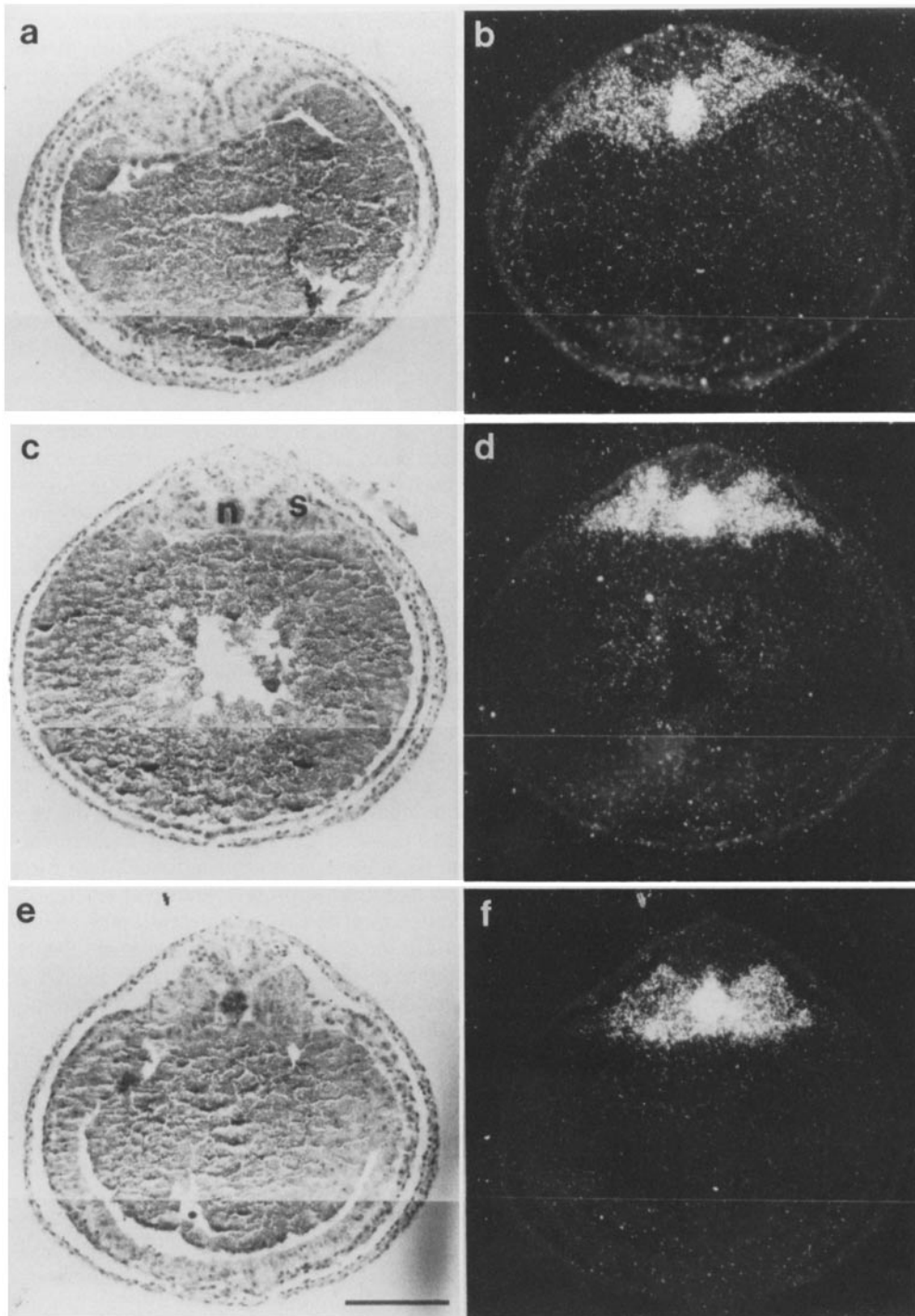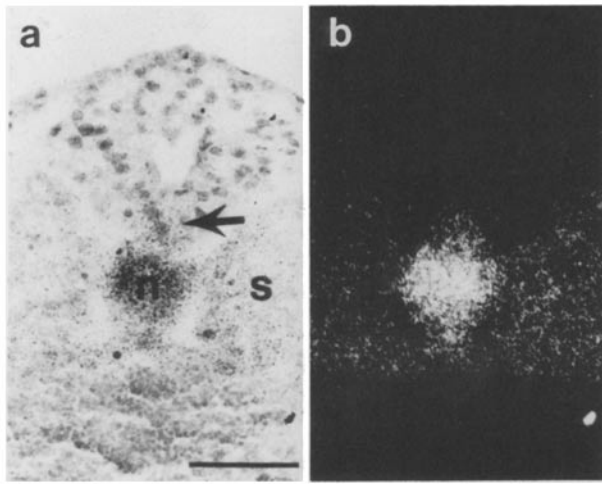
Figure 9. Bright-field (*a, c,* and *e*) and dark-field (*b, d,* and *f*) micrographs of cross sections at three different levels through a stage 21 embryo probed for type II procollagen mRNA with the general probe. *a* and *b* are cut at a posterior level showing a high level of mRNA in the notochord. In addition, mRNA is detected throughout the somites and in the more dorsal half of the lateral plate mesoderm. *c* and *d* are cut at the level of the midgut showing that the signal is no longer detected in the lateral plate mesoderm but is restricted to the notochord and somites. *e* and *f* are cut at a more anterior level of the same embryo in which the intense signal in the notochord and sclerotome can be noted. The level of label in the remainder of the somite is relatively weaker than in the peri-notochordal region (*n*, notochord; *s*, somite). Bar, 250 μm.

notochord and somites. In a more anterior section, the signal was relatively reduced in the dorsal part of the somite.

To illustrate the distribution of label in the perinotochordal region more clearly, a higher magnification micrograph is shown in Fig. 10. Signal can be detected in the sub-notochordal rod, the notochord, somites and floor plate. Additionally, antigen can be detected at these same locations by immuno-histochemistry. In both whole mount (Fig. 7, *c* and *d*) and sectioned (Fig. 8 *b*) preparations at stage 31, type II collagen is detected in the basement membrane around the notochord

and, on the ventral side of the neural tube besides the floor plate, the sub-notochordal rod and the developing vertebrae.

By the tadpole stage (stage 46), procollagen transcripts are largely restricted to chondrogenic regions of the embryo's head and trunk (Fig. 11). This distribution corresponds to the distribution of antigen observed by immunohistochemistry (Fig. 12, *a* and *b*). One noteworthy difference between the pattern of distribution of mRNA and antigen was seen in the ectoderm, where transcripts are detected but antigen does not accumulate.

Figure 10. Bright-field (*a*) and dark-field (*b*) micrographs of a section through the same embryo as that shown in Fig. 9 at a level between *c* and *e*, shown at higher magnification. Label can be clearly localized in the floor plate (*arrow*), notochord (*n*), and somites (*s*). Bar, 100 μm.

Having previously established that the shorter procollagen transcript appears during the tailbud stage, embryo sections from stage 46 (tadpole) were analyzed by in situ hybridization to a probe containing only the exon 2 sequence (Fig. 13). Comparative analysis of the hybridization data using the exon 2 and the general probe did not reveal any appreciable difference. These results suggest that, within the sensitivity of the method employed, the alternatively spliced products are co-expressed in the same tissues. It is also clear that both forms are expressed in cartilage. However, because of the limited resolution of the method we cannot exclude the possibility of small differences in coexpression at the cellular level.

Discussion

This report constitutes the first comprehensive analysis of collagen gene expression during the early stages of *Xenopus* development. Structural features indicated that the genes we isolated represent duplicated forms of the single copy gene that in higher vertebrates encodes the procollagen subunit of the homotrimeric type II collagen (Vuorio and de Crombrugghe, 1990). We also present evidence consistent with the idea that the type II procollagen gene duplicated probably as a result of *X. laevis* tetraploidization (Kobel and Du Pasquier, 1986). This is yet another example demonstrating that in *X. laevis* there are two distinct genetic loci for each one present in other vertebrate species (Wahli and Dawid, 1980; Jeffreys et al., 1980; Stutz and Spohr, 1986; Fritz et al., 1989; Shuldiner et al., 1989). Interestingly, also in this case both genes are transcriptionally active and regulated in an identical spatio/temporal manner during development.

The onset of transcription of the type II procollagen genes is at neurula stage. Based on in situ hybridization results, the type II procollagen mRNA is localized along the primary embryonic axis. At first glance, one might consider that its expression is associated with the developing vertebral column and the onset of chondrogenesis. In the case of the developing appendicular skeleton of the chick, type II procollagen mRNA is at first widely expressed throughout the limb mesenchyme and progressively becomes restricted to the differentiated cartilage elements (Swalla et al., 1988). Similarly, in the axial region in *Xenopus*, type II procollagen transcripts are detected more widely than are the presumptive vertebral chondrogenic precursors, including the lateral plate mesoderm, sub-notochordal rod, notochord, and the floor plate of the neural tube, in addition to all parts of the somites. With subsequent development, the mRNA becomes more spatially restricted in its expression, and upregulated in regions that will subsequently contribute to the vertebral column or to specific skeletal elements. However, the mRNA and even the protein can be detected at a number of additional sites that will not later undergo chondrogenesis.

Type II collagen expression is not unique to cartilage. Besides the notochord (Linsenmayer et al., 1973), neural retina (Newsome et al., 1976), and embryonic corneal epithelium (Linsenmayer et al., 1977) it has been detected in association with a number of other embryonic epithelial (Thorogood et al., 1986; Kosher and Solursh, 1989; Fitch et al., 1989), where its expression is developmentally regulated. In the case of the ectoderm, previous immunohistochemical studies demonstrated the presence of type II collagen is sub-ectodermal basement membranes (Kosher and Solursh, 1989; Fitch et al., 1989). However, the site of synthesis was not demonstrated. Our results show that in *Xenopus* type II procollagen mRNA is present in the ectodermal cells themselves. In addition, this study adds the floor plate in the ventral neural tube to the structures synthesizing and even accumulating type II collagen. The notochord and floor plate are known to play a central role in polarizing cell differentiation along the dorsal-ventral axis of the developing central nervous system (Hirano et al., 1991; Yamada et al., 1991). The recent observation that type II procollagen mRNA is also found in localized regions of the fetal mouse brain (Cheah et al., 1991) adds further support to the suggestion that this macromolecule may be involved in additional processes, besides chondrogenesis. The fact that its expression at such sites is transient is consistent with its having a role in the initiation of particular differentiation programs.

It is intriguing to note that coincident with the appearance of chondrogenic tissues in the developing embryo, expression of the type II procollagen genes is associated with a stage-specific change in pre-mRNA processing. Hence, temporal and spatial patterns of type II collagen expression seem to be regulated by specific transcription and splicing factors. As a result, two structurally distinct forms of the procollagen chain are conceivably synthesized and assembled at distinct developmental stages. We do not know the significance and function of these alternative gene products or how these precursor procollagens are sorted out and assembled by cells that synthesize both isoforms. Present evidence indicates that such a pattern may be broadly used during vertebrate embryogenesis (Ryan and Sandell, 1990). This in turn suggests that the temporal change in processing of type II procollagen transcripts might be a phenomenon responsive to tissue-specific signals likely to be connected with functional differences of the developing ECM.

More generally, the early localized accumulation of type II procollagen mRNA in the *Xenopus* embryo provides a
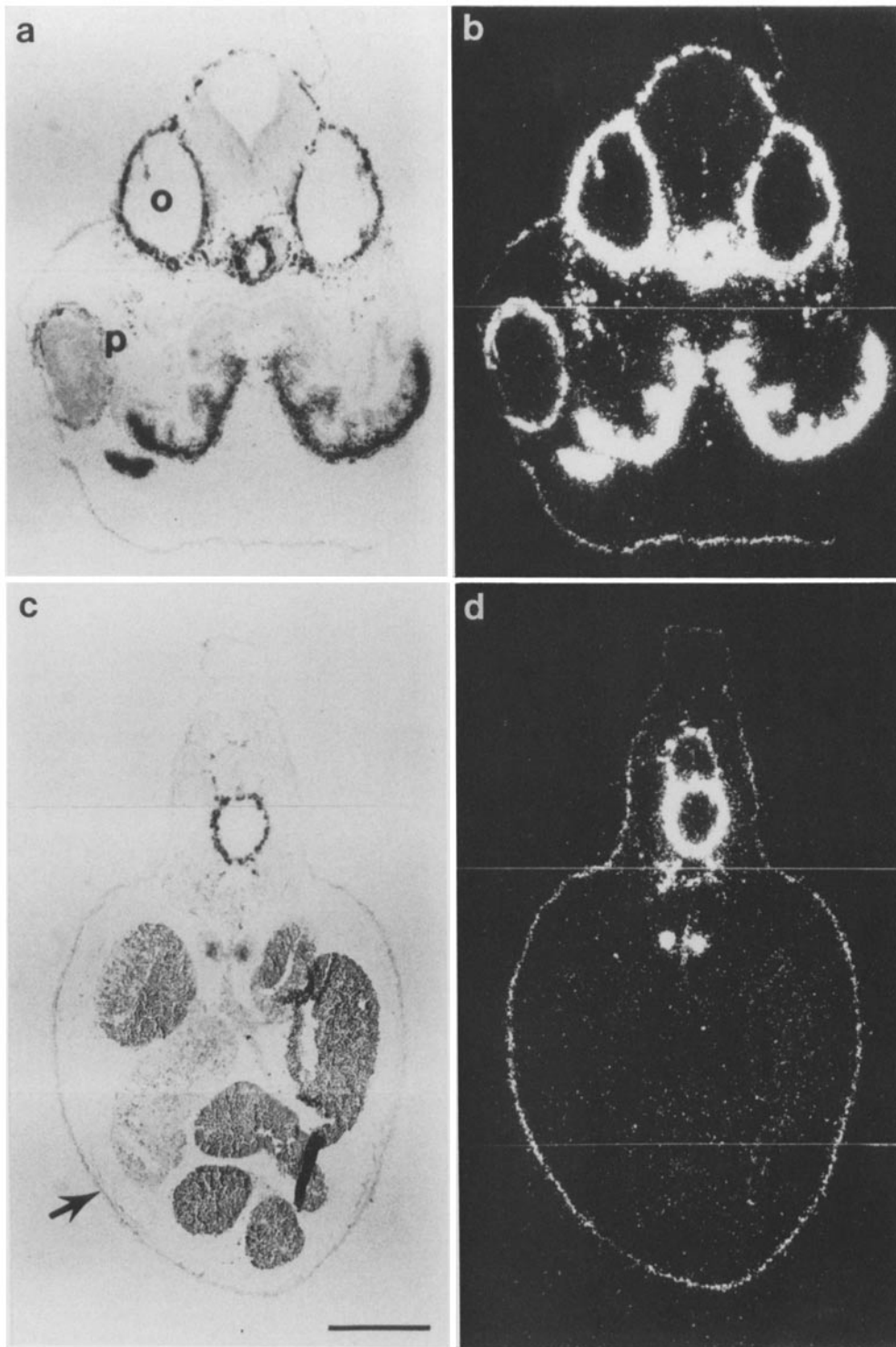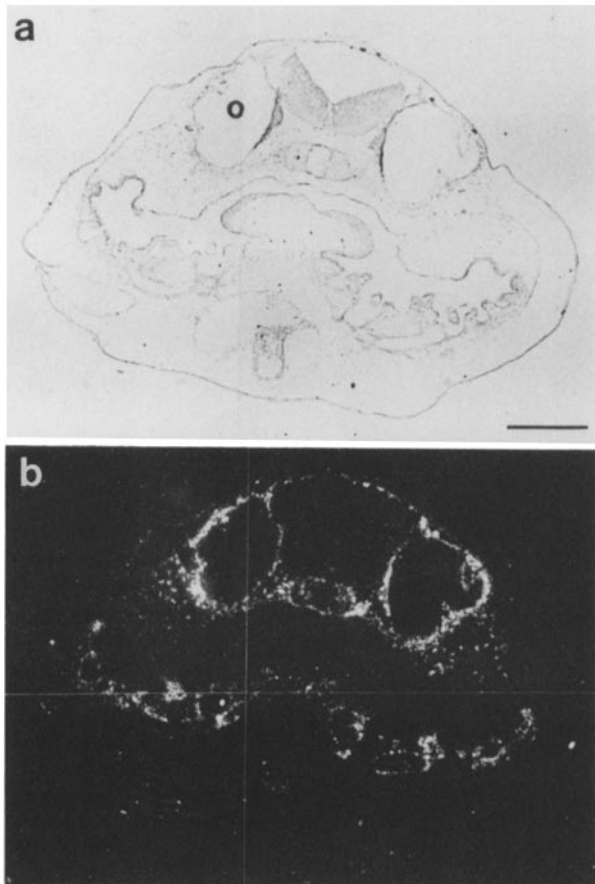
Figure 11. Bright-field (*a* and *c*) and dark-field (*b* and *d*) micrographs through the head (oblique cross section) and trunk (cross section) of a stage 46 embryo probed for type II procollagen mRNA with the general probe. In the head, mRNA is detected in the otocysts (*o*), chondrocranium, and periocular (*p*) tissue. In the section through the trunk, mRNA is detected in cells of the vertebra. In both the head and trunk regions, mRNA is detected in ectodermal cells (*arrow*). Bar, 250 μm.

potentially useful marker to study embryonic induction in tissues of mesodermal origin (Gurdon, 1987). To date, most of this work has been carried out using the muscle-specific actin gene (Mohun et al., 1984) and a keratin gene that is preferentially, but not exclusively, expressed in the notochord of the *Xenopus* embryo (LaFlamme et al., 1988). Experiments currently underway are analyzing the transcriptional features of the type II procollagen genes, including the expression of the alternatively spliced transcripts. These studies promise to elucidate some of the mechanisms and factors responsible for the regulation of these genes in embryonic tissues of mesodermal origin.
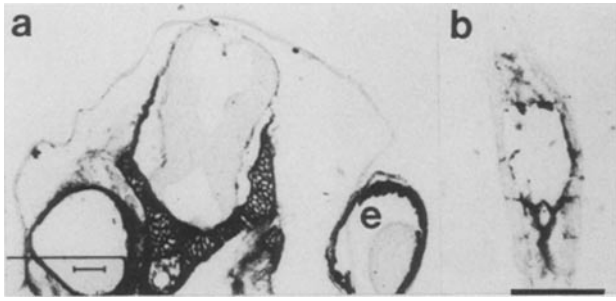
Figure 12. Immunolocalization of type II collagen at stage 46 in the head (*a*), as seen in frontal section, and the trunk (*b*), shown in cross section. Antigen is detected in the chondrocranium, eye (*e*), vertebra, and other regions expressing the mRNA, with the exception of the ectoderm. Bar, 250 µm.

Figure 13. Bright-field (*a*) and dark-field (*b*) micrographs of a cross section similar to that shown in Fig. 11 (*a* and *b*), but showing the pattern of hybridization with the exon 2 probe. The signal is generally weaker than that obtained with the general probe, but as far as the limit of resolution permits, is otherwise indistinguishable. (*o*, otocyst). Bar, 250 µm.

References

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.

Dent, J. A., A. G. Polson, and M. W. Klymkowsky. 1989. A whole-mount immunocytochemical analysis of the expression of intermediate filament protein vimentin in *Xenopus. Development (Camb.).* 105:61–74.

Fitch, J. M., A. Mentzer, R. Mayne, and T. F. Linsenmayer. 1989. Independent deposition of collagen types II and IX at epithelial-mesenchymal interfaces. *Development (Camb.).* 105:85–95.

Fleischmajer, R., R. Timpl, L. Tuderman, L. Raisher, M. Wiestner, J. S. Perlish, and P. N. Graves. 1981. Ultrastructural identification of extension aminopropeptides of type I and III collagens in human skin. *Proc. Natl. Acad. Sci. USA.* 78:7360–7364.

Fleischmajer, R., B. R. Olsen, R. Timpl, J. S. Perlish, and O. Lovelace. 1983. Collagen fibril formation during embryogenesis. *Proc. Natl. Acad. Sci. USA.* 80:3354–3358.

Fritz, A. F., K. W. Y. Cho, C. V. E. Wright, B. G. Jegalian, and E. M. De Robertis. 1989. Duplicated homeobox genes in *Xenopus. Dev. Biol.* 131:584–588.

Green, H. B., B. Goldberg, M. Schwartz, and D. D. Brown. 1968. The synthesis of collagen during the development of *Xenopus laevis. Dev. Biol.* 18:391–400.

Gurdon, J. B. 1987. Embryonic induction-molecular prospects. *Development (Camb.).* 99:285–306.

Hirano S., S. Fuse, and G. S. Sohal. 1991. The effect of the floor plate on pattern and polarity in the developing central nervous system. *Science (Wash. DC).* 251:310–313.

Jeffreys, A. J., V. Wilson, D. Wood, and J. P. Simons. 1980. Linkage of adult α- and β-globin genes in *X. laevis* and gene duplication by tetraploidization. *Cell.* 21:555–564.

Kawasaki, E. S., and A. M. Wang. 1989. Detection of gene expression. *In*: PCR Technology: Principles and Applications for DNA Amplification. H. A. Erlich, editor. Stockton Press, New York. 89–98.

Keller, R., and R. Winklbauer. 1990. The role of the extracellular matrix in amphibian gastrulation. *Sem. Dev. Biol.* 1:25–33.

Kimelman, D., and M. Kirschner. 1987. Synergistic induction of mesoderm by FGF and TGF-β and the identification of an mRNA coding for FGF in the early *Xenopus* embryo. *Cell.* 51:869–877.

Kintner, C. R., and D. A. Melton. 1987. Expression of *Xenopus* N-CAM RNA in ectoderm is an early response to neural induction. *Development (Camb.).* 99:311–325.

Kobel, H. R., and L. Du Pasquier. 1986. Genetics of polyploid *Xenopus. Trends Genet.* 2:310–315.

Kosher, R. A., and M. Solursh. 1989. Widespread distribution of type II collagen during embryonic chick development. *Dev. Biol.* 131:558–566.

LaFlamme, S. E., M. Jamrich, K. Richter, T. D. Sargent, and I. B. Dawid. 1988. *Xenopus* endo B is a keratin preferentially expressed in the embryonic notochord. *Genes & Dev.* 2:853–862.

Lallier, T., and M. Bronner-Fraser. 1990. The role of the extracellular matrix in neural crest migration. *Sem. Dev. Biol.* 1:35–44.

Linsenmayer, T. F., and M. J. C. Hendrix. 1980 Monoclonal antibodies to connective tissue macromolecules: type II collagen. *Biochem. Biophys. Res. Commun.* 92:440–446.

Linsenmayer, T. F., R. L. Trelstad, and J. Gross. 1973. The collagen of chick embryonic notochord. *Biochem. Biophys. Res. Commun.* 53:39–45.

Linsenmayer, T. F., G. N. Smith, and E. D. Hay. 1977. Synthesis of two collagen types of embryonic chick corneal epithelium *in vitro. Proc. Natl. Acad. Sci. USA.* 74:39–43.

Melton, D. A., and M. Whitman. 1989. Growth factors in early embryogenesis. *Annu. Rev. Cell Biol.* 5:93–118.

Mohun, T. J., S. Brennan, N. Dathan, S. Fairman, and J. B. Gurdon. 1984. Cell type-specific activation of actin genes in the early amphibian embryo. *Nature (Lond.).* 311:716–721.

Mount, D. W., and B. Conrad. 1987. Improved programs for DNA and protein sequence analysis on the IBM personal computer and other standard computer systems. *Nucleic Acids Res.* 14:11–16.

Newport, J., and M. Kirschner. 1982. A major developmental transition in early *Xenopus* embryos: II. Control of the onset of transcription. *Cell.* 30:675–686.

Newsome, D. A., T. F. Linsenmayer, and R. L. Trelstad. 1976. Vitreous body collagen. Evidence for a dual origin from the neural retina and hyalocytes. *J. Cell Biol.* 71:59–67.

Niewkoop, P. D., and J. Faber. 1967. Normal Table of *Xenopus laevis* (Daudin). 2nd Ed. Elsevier-North Holland, Amsterdam.

Otte, A. P., R. Debjani, M. Siemerink, C. H. Koster, F. Hochstenbach, A. Timmermans, and A. J. Durston. 1990. Characterization of a maternal type VI collagen in *Xenopus* embryos suggests a role for collagen in gastrulation. *J. Cell Biol.* 111:271-278.

Paglia, L. M., M. Wiestner, M. Duchene, L. A. Oullette, D. Horlein, G. R. Martin, and P. K. Muller. 1981. Effects of procollagen peptides on the translation of type II collagen messenger ribonucleic acid and on collagen biosynthesis in chondrocytes. *Biochemistry.* 20:3523-3527.

Ryan, M. C., and L. J. Sandell. 1990. Differential expression of a cysteine-rich domain in the amino-terminal propeptide of type II (cartilage) procollagen by alternative splicing of mRNA. *J. Biol. Chem.* 265:10334-10339.

Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular Cloning. A Laboratory Manual. 2nd Ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 545 pp.

Shuldiner, A. R., S. Phillips, C. T. Roberts, D. LeRoith, and J. Roth. 1989. *Xenopus laevis* contains two nonallelic preproinsulin genes. *J. Biol. Chem.* 264:9428-9432.

Smith, J. C. 1989. Mesoderm induction and mesoderm inducing factors in early amphibian development. *Development (Camb.).* 105:665-677.

Solursh, M. 1990. The role of the extracellular matrix molecules in limb development. *Sem. Dev. Biol.* 1:45-53.

Stutz, F., and G. Spohr. 1986. Isolation and characterization of sarcomeric actin genes expressed in *Xenopus laevis* embryos. *J. Mol. Biol.* 187:349-361.

Su, M. W., B. Lee, F. Ramirez, M. Machado, and W. Horton. 1989. Nucleotide sequences of the full length cDNA coding for human type II collagen. *Nucleic Acids Res.* 17:9473.

Su, M. W., H. R. Suzuki, M. Solursh, and F. Ramirez. 1991. Progressively

restricted expression of a new homeobox-containing gene during *Xenopus laevis* embryogenesis. *Development (Camb.).* 111:1179-1187.

Swalla, B. J., W. B. Upholt, and M. Solursh. 1988. Analysis of type II collagen RNA localization in chick wing buds by *in situ* hybridization. *Dev. Biol.* 125:51-58.

Thiery, J. P. 1989. Cell adhesion in morphogenesis. *In* Cell to Cell Signals in Mammalian Development. NATO ASI series. S. W. De Leat, J. G. Bluemink, and C. L. Mummery, editors. Springer-Verlag, Berlin. 109-128.

Thorogood, P., J. Bee, K. von der Mark. 1986. Transient expression of collagen type II at epitheliomesenchymal interfaces during morphogenesis of the cartilagenous neurocranium. *Dev. Biol.* 116:497-509.

Vuorio, E., and B. D. de Crombrugghe. 1990. The family of collagen genes. *Annu. Rev. Biochem.* 59:837-872.

Wahli, W., and I. B. Dawid. 1990. Isolation of two closely related vitellogenin genes, including their flanking regions, from a *Xenopus laevis* gene library. *Proc. Natl. Acad. Sci. USA.* 77:1437-1441.

Wiestner, M., T. Krieg, D. Horlein, R. W. Glanville, R. Fietzek, and P. K. Muller. 1979. Inhibiting effect of procollagen peptides on collagen biosynthesis in fibroblast cultures. *J. Biol. Chem.* 254:7016-7023.

Wu, C. H., C. B. Donovan, and G. Y. Wu. 1986. Evidence for pretranslational regulation of collagen synthesis by procollagen propeptides. *J. Biol. Chem.* 261:10482-10484.

Yamada, T., M. Placzek, H. Tanaka, J. Dodd, and T. M. Jessell. 1991. Control of cell pattern in the developing nervous system: polarizing activity of the floor plate and notochord. *Cell.* 64:635-647.

Zagursky, R. J., M. L. Berman, K. Baumeister, and N. Lomax. 1986. Rapid and easy sequencing of large linear double stranded DNA and supercoiled plasmid DNA. *Gene Anal. Tech.* 2:232-238.