Volume 10 Number 4 1982 Nucleic Acids Research

Construction and partial characterization of two recombinant cDNA clones for procollagen from chicken cartilage

Eero Vuorio*, Linda Sandell, Dean Kravis, Val C.Sheffield, Tuula Vuorio*, Albert Dorfman and William B.Upholt

Departments of Pediatrics and Biochemistry and the Joseph P.Kennedy, Jr. Mental Retardation Center, Hospital Box 413, The University of Chicago, Chicago, IL 60637, USA

Received 1 December 1981; Revised and Accepted 26 January 1982

ABSTRACT

Type II procollagen mRNA has been partially purified from embryonic chick sternal cartilage by guanidine hydrochloride extraction, sucrose gradient sedimentation and Sepharose 4B chromatography. Double stranded cDNA was synthesized using AMV reverse transcriptase and E. coli DNA polymerase I, tailed using terminal transferase and inserted into the Pst I site of pBR322. Two putative type II procollagen cDNA clones have been characterized. Both plasmids hybridize to 2 sternal RNA species, a major species of 5.3 kb and a minor species of 7 kb. These RNAs are present in total RNA from sterna and differentiated limb bud cultures but are not detected in RNA from stage 24 limb bud which has not yet differentiated to cartilage or in RNA from calvaria. The time of appearance of these RNAs during the differentiation of limb mesenchyme in culture parallels the appearance of translatable type II procollagen mRNA.

INTRODUCTION

Cartilage is characterized by an extensive extracellular matrix. Major components of this matrix include type II collagen and chondroitin sulfate proteoglycan. Collagen in extracellular matrix consists of triple helical molecules composed of 3 separate polypeptide chains. These chains are synthesized as polypeptides which are proteolytically cleaved subsequent to triple helix formation and secretion (1). At least 5 different collagen types (or 9 different collagen genes) have been identified (2). In the process of long bone development, at least 3 different collagen types are expressed (2,3). During cartilage differentiation, the core protein of cartilage proteoglycan and type II collagen are coordinately expressed and the synthesis of type I collagen is suppressed (4,5). Type II collagen is composed of 3 identical α1 (type II) chains and type I collagen is composed of 2 α1 (type I) chains and 1 α2 (type I)

0305-1048/82/1004-1175$2.00/0

chain. We have previously translated and identified mRNAs expressed in differentiating chick cartilage for the core protein of proteoglycan, type II procollagen and the α1 and α2 chains of type I procollagen (6). In order to extend these studies, we are currently preparing recombinant cDNA clones complementary to these mRNAs. Several other laboratories have reported recombinant cDNA clones corresponding to the mRNAs for the α1 and α2 chains of chick type I procollagen (7,8,9,10). Recent work using genomic clones for the α2 chain of type I procollagen suggests that collagen genes may be comprised of many (approximately 50 per gene) short coding sequences containing defined structural units of the triple helical region (11,12,13). This structural organization may have important implications concerning the evolution of collagen genes and the control of their expression.

MATERIALS AND METHODS

Purification of RNA

Total RNA was purified from 17-day embryonic chick sterna, stage 24 limb buds, differentiating high density limb bud cultures and 16-day embryonic calvaria by guanidine extraction as previously described (6). Large RNAs were prepared by applying 200 μg of total RNA (heated 2' at 65°C) to 36 ml 5-20% sucrose gradients containing 10 mM Hepes, pH 7.2. All RNA sedimenting 27S or faster was pooled and ethanol precipitated. Approximately 1 mg of this pooled RNA was heated for 2' at 65°C, applied to a 2.24 x 34 cm column of Sepharose 4B and fractionated into bound and unbound fractions as described by Frischauf *et al.* (14).

Synthesis of double-stranded cDNA

AMV reverse transcriptase (Life Sciences, Inc., lot G-479) was used to synthesize the first strand under conditions modified from Wahli *et al.* (15). The reaction contained 2.5 μM dNTPs and 800 U/ml reverse transcriptase. No Actinomycin D or pyrophosphate was used. The reactions were incubated at 44.5°C for 1 hr.

The second strand was synthesized as described by Wickens *et al.* (16) using *E. coli* DNA polymerase I (BRL) at 150 U/ml.

Routinely, 1-2 µg of double stranded cDNA was synthesized per 100 µg of RNA fractionated as described above. This partially purified RNA consists primarily of 27S rRNA.

The precipitated cDNA plus RNA was digested with 4000 U/ml of S1 nuclease (Sigma) at 37°C for 60 min. at a concentration of approximately 1.3 mg/ml of nucleic acid.

Homopolymer tailing

Terminal deoxynucleotidyl transferase (P-L, lot 730-11) was used to add homopolymer tails to the 3'-ends of cDNA. Conditions were essentially those described by Nelson and Brutlag (17). Reaction conditions were modified for each preparation of enzyme, dCTP and cDNA to obtain tails of approximately 15 dC residues per end. Representative conditions were 6 µg/ml of double stranded cDNA, 400 U/ml terminal transferase, 2.5 µM dCTP and 1 mM $CoCl_2$. Pst I cleaved pBR322 at 20 µg/ml was tailed as above with 2.5 µM dGTP and 40 U/ml terminal transferase.

Sucrose gradient sedimentation of tailed cDNA

Following phenol-chloroform extraction and passage over a Sephadex G-50 column, dC-tailed cDNA was sedimented through a 3-22% sucrose gradient (15). Fractions containing DNA with average size larger than 1000 base pairs were pooled, combined with tailed, phenol-chloroform extracted pBR322 and co-precipitated with ethanol.

Annealing reaction

Approximately 100 ng of tailed pBR322 and 20 ng of tailed double stranded cDNA precipitated as described above were dissolved in 100 µl of annealing buffer (100 mM NaCl, 10 mM Tris-HCl, pH 7.5, 1 mM EDTA), heated to 65°C for 1 hr. and allowed to cool slowly overnight to room temperature.

Transformation

Transformations were carried out in 50 mM $CaCl_2$ as described by Mandel and Higa (18). Approximately 250 transformants resistant to tetracycline were obtained per microgram pBR322 annealed with cDNA. Recombinant clones were tested for their resistance to ampicillin by replica plating on L broth agar plates containing 12.5 µg/ml ampicillin.

Colony hybridization

Colonies were grown on Whatman filter paper and amplified

by transfer to chloramphenicol plates after 7 hrs. on tetracycline plates. Filters were processed as described by Thayer (19). The filters were incubated for 16 hrs. at 65°C in 20 ml of prehybridization buffer (6 x SSC, 1 x Denhardt's solution) (19,20) and then hybridized for 20 hrs. at 65°C in 20 ml of the same buffer with 100 µg/ml sonicated denatured calf thymus DNA. ^{32}P-Labeled cDNA probe was prepared with reverse transcriptase using the partially purified type II procollagen mRNA fraction. After washing, the filters were exposed at -70°C with Kodak X-Omat film using intensifying screens.

RNA and DNA gel transfers

Total and fractionated RNAs were electrophoresed on 0.75% agarose gels after glyoxal denaturation (21). Transfer of RNA from gel to nitrocellulose filters was done as described by Thomas (22). In some experiments, diazobenzyloxymethyl (DBM) paper was prepared and used in RNA transfers as described by Alwine et al. (23). DNA transfers to nitrocellulose were done as described by Southern (24). Either whole plasmid or purified restriction fragments were labeled with ^{32}P-dCTP by nick translation (BRL nick translation kit) to a specific activity of about 10^7 cpm/µg. Prehybridization and hybridization of nitrocellulose filters and DBM papers were performed as described by Alwine et al. (23). Sonicated denatured calf thymus or salmon sperm DNA was added to a concentration of 250 µg/ml. Filters were hybridized at 42 or 52°C for 16-72 hrs. in 50% formamide, 5 x SSC and autoradiographed as described above.

Mapping of restriction endonuclease cleavage sites

Restriction endonucleases were obtained from BRL, New England Biolabs, or Biotech. Reaction conditions were as recommended by the manufacturers. Double digestions were done sequentially without removal of the first enzyme. Inserted sequences were purified subsequent to liberation with Pst I either by sedimentation for 16.5 hrs. at 2°C and 35 krpm on 3-22% sucrose gradients (0.1 M NaCl, 10 mM Tris HCl, pH 7.4, 1 mM EDTA), in a SW 50.1 rotor or by recovery from 8% polyacrylamide gels as described by Maxam and Gilbert (25). DNA samples were electrophoresed in agarose gels (0.7-1.5% agarose depending on the size of the fragments) containing 10 mM sodium phosphate

pH 7.0 and 0.5 μg/ml EtdBr or polyacrylamide gels (3-8%) containing 40 mM Tris, 20 mM sodium acetate, 2 mM EDTA, pH 8.3. Fine mapping was done by nick translating purified inserts and analysing fragments on 37 cm 8% polyacrylamide sequencing gels containing urea after further digestion with one or more enzymes. Phage λDNA, digested with Hind III or Eco RI, and pBR322 digested with Hinf I, Hha I, or Hae III were used as length markers.

RESULTS AND DISCUSSION

Partial purification of type II procollagen mRNA

Chick embryonic sterna were chosen as the tissue source for the isolation of type II procollagen mRNA due to the ease of RNA preparation and the relatively higher purity of chondrocytes in sterna compared with those from differentiating limb bud cultures. Total RNA was extracted with guanidine-HCl (6) to minimize degradation of the very large collagen mRNAs. Messenger RNAs for collagen and the core protein of cartilage chondroitin sulfate proteoglycan were separated from the bulk of smaller RNAs by sucrose gradient centrifugation. Frischauf et al. (14) have reported that the large rRNA and the mRNA for the α1 chain of type I procollagen from chicken selectively bind to Sepharose 4B in high salt and are eluted in lower salt. We have used a similar procedure to partially purify chicken type II procollagen mRNA. Figure 1 demonstrates cell-free translation of RNAs purified on Sepharose 4B. Column A shows the translation products obtained with total unfractionated sternal RNA. The major translation product, polypeptide d, has been previously identified as type II procollagen by immunoprecipitation and collagenase sensitivity (6). Polypeptides b and e have been identified as the α1 and α2 chains of type I procollagen. Polypeptides a and c are unidentified but are collagenase sensitive. These might correspond to other collagen types which have been reported to be present in cartilage at low levels (26). The mRNA for type II procollagen binds to Sepharose 4B under these conditions (Fig. 1C) and is virtually absent in the unbound fraction (Fig. 1B). Messenger RNA for core protein is found only in the unbound fraction, although the translation product is not visible in this exposure.

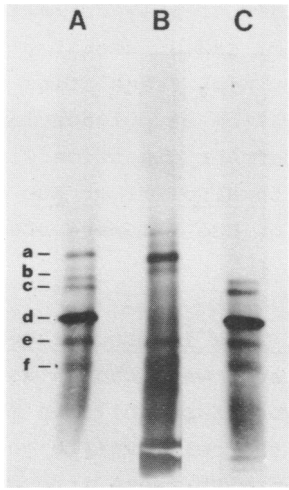

Figure 1. Cell-free translation of chick sternal RNA fractionated on Sepharose 4B. RNAs were translated in a wheat germ cell-free system using ^{35}S-methionine. Translation products were analyzed on 5% polyacrylamide gels containing sodium dodecyl sulfate as described previously (6). Lane A, total sternal RNA; lane B, RNA not bound to Sepharose 4B; lane C, RNA which binds to Sepharose 4B.

Construction and screening of cDNA clones

RNA enriched in type II procollagen sequences was transcribed by AMV reverse transcriptase using an oligo(dT) primer. The presence of the chicken large rRNA did not contribute to oligo(dT) dependent cDNA synthesis confirming results previously reported by King et al. (27). Second strand was synthesized using E. coli DNA polymerase I according to the method of Wickens et al. (16). Subsequent to S1 nuclease digestion to remove single stranded DNA and RNA, dC homopolymer tails were added using terminal deoxynucleotidyl transferase. The tailed fragments were fractionated on sucrose gradients to remove small fragments and annealed with pBR322 previously cleaved with Pst I and dG tailed. HB101 transformants were selected by tetracycline resistance and ampicillin sensitivity. Clones for further characterization were selected by colony hybridization using ^{32}P-labeled cDNA prepared from partially purified type II procollagen mRNA. Two clones, pCAR1 and pCAR2, containing inserts of 525 and 680 bp respectively, were selected for fur-

ther characterization. Pst I sites on both sides of the inserts remained intact.

Characterization of pCAR1 and pCAR2

In order to further characterize these 2 recombinant cDNA clones, the DNAs were labeled with ^{32}P-dCTP by nick-translation and hybridized to nitrocellulose or DBM blots of a number of different RNAs size fractionated on glyoxal gels. Lane B of Figure 2 shows the hybridization of pCAR2 to a blot of sternal RNA. The plasmid hybridizes to a major band at 5.3 kb. The probe also hybridizes to a minor RNA species of 7.2 kb which is not visible in this exposure but is visible in Figure 4. Plasmid pCAR1 shows an identical hybridization pattern. The smaller RNA (5.3 kb) is approximately the size expected to code for a single procollagen chain. pCAR2 also hybridizes to 5.3 and 7 kb RNAs in total RNA isolated from 8-day cultures of stage 24 chick limb mesenchyme which have differentiated into cartilage (Fig. 2, lane C and Fig. 4). However, pCAR2 does not hybridize

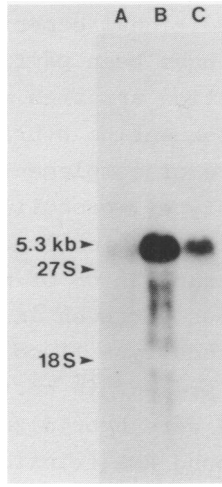

Figure 2. Hybridization of nick-translated pCAR2 to various RNAs transferred from a glyoxal gel to nitrocellulose. Total RNA (7.5 μg) from chick embryonic calvaria (A), embryonic sterna (B), or 8-day differentiating limb bud cultures (C) was treated with glyoxal, electrophoresed on a 0.75% agarose gel, transferred to nitrocellulose filters and hybridized with ^{32}P-labeled nick translated pCAR2. Chick rRNAs and bacteriophage λDNA digested with Eco RI were used as size markers.

significantly to gel transfers of calvarial RNA (Fig. 2, lane A), which contains little or no type II procollagen mRNA and large amounts of type I procollagen mRNAs.

Control experiments were done using a recombinant cDNA clone constructed in our laboratory containing sequences corresponding to prepro α1 (I) collagen (Upholt, Vuorio and Dorfman, unpublished) to eliminate the possibility that pCAR1 and pCAR2 might contain type I procollagen sequences. The α1 type I probe, unlike pCAR1 and pCAR2, hybridizes strongly to bands of 5.0 and 7.2 kb in total calvarial RNA similar to results reported by Rave, Crkvenjakow and Boedtker (28) and Adams et al. (29). The α1 type I probe hybridizes weakly to the same bands at 5.0 and 7.2 kb in total sternal RNA consistent with the results of cell-free translation of sternal RNA which shows low levels of mRNA for the precursors to the α1 chain of type I procollagen (band b in Fig. 1A). In comparative hybridizations to the same transfer, pCAR2 hybridizes to an RNA slightly larger than the α1(I) mRNA.

In order to further define the polypeptide for which this RNA might code, several additional experiments and limited amounts of DNA sequencing have been performed. Attempts to identify the mRNA by hybridization arrested translation (30) have been unsuccessful as, subsequent to hybrid arrest conditions (in the absence or presence of complementary plasmid), type II procollagen mRNA (also α1 type I procollagen mRNA) must be heated to 90°C to regain translational activity. Attempts to identify these clones by translation of mRNA selected by hybridization to plasmid have thus far been inconclusive.

Since Sepharose 4B binds type II collagen mRNA and does not bind most other mRNAs present in cartilage (Fig. 1), nick-translated pCAR1 and pCAR2 were hybridized to glyoxal gel transfers of the bound and unbound RNA fractions. Both plasmids (pCAR1 not shown) hybridize only with the RNA fraction which binds to Sepharose and not with unbound RNA (Figure 3).

To eliminate the possibility that the lack of hybridization of pCAR2 to the RNA fraction not binding to Sepharose 4B is due to degradation or poor transfer of this RNA to the filter paper, a recombinant cDNA plasmid containing sequences corre-

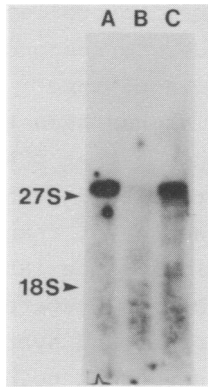

Figure 3. Hybridization of nick-translated pCAR2 to a glyoxal gel blot of sternal RNA fractionated on Sepharose 4B. Lane A, total sternal RNA; lane B, RNA not bound to Sepharose 4B; lane C, RNA which bound to Sepharose 4B.

sponding to α2 type I procollagen mRNA (Vuorio and Upholt, details of construction to be published separately) was also hybridized to these blots. The α2(I) mRNA is recovered in both the bound and unbound fractions from Sepharose 4B (see band e in Fig. 1). The α2(I) cDNA probe hybridizes clearly to bands of approximately 5 kb in both the bound and unbound fractions (data not shown).

Translation of bound and unbound RNA fractions is shown in Fig. 1. Translation products directed by the bound RNA fraction contain 6 abundant, clearly distinguishable high molecular weight polypeptides. Three of these (bands a, e, and f) are also present in translation products directed by unbound RNA (Fig. 1B) and 3 are unique to products of RNA which binds (Fig. 1C). Two of these 3 unique translation products have been previously identified. The major translation product (band d) with an apparent molecular weight of 160,000 is the precursor to the α1 type II procollagen chain and polypeptide b with an apparent molecular weight of 190,000 is the precursor to the α1 type I procollagen chain. The other major translation product unique to this RNA, band c, with an apparent molecular weight of 185,000 is unidentified but is collagenase sensitive. Since neither pCAR1 nor pCAR2 hybridizes significantly to calvarial

RNA (highly enriched in type I procollagen mRNAs), these plasmids do not contain sequences corresponding to the mRNA for the precursor of the α1 chain of type I procollagen. Assuming that the cDNA in the plasmids has been transcribed from a translatable RNA, these data suggest that pCAR1 and pCAR2 most likely correspond to sequences for type II procollagen, or, less likely (based on the relative abundance of translation products and the level of hybridization), for the 185,000 dalton unidentified polypeptide. If labeled cDNA is prepared from total sternal RNA using reverse transcriptase and hybridized to blots of sternal RNA, the major band observed is coincident with that identified by hybridization of nick-translated pCAR1 or pCAR2. Preliminary sequence data (see following section) indicate that pCAR1 contains base sequences coding for the carboxy-terminal propeptide region of a collagen chain and that these sequences are more closely related to pro-α1(I) mRNA sequences than to pro-α2(I) mRNA sequences.

As an additional means of characterizing these plasmids, the correlation between the time of appearance of the 5.3 and 7 kb RNAs during limb mesenchyme differentiation and the time of appearance of translatable type II procollagen mRNA was investigated. Total RNA was prepared from limb mesenchyme cultures on days 1, 2, 3, 5, and 6 of culture and studied by cell-free translation and by hybridization of nick-translated plasmid to RNA gel blots. Translation products obtained using these RNAs are shown in Figure 4A. The precursor to core protein (CP) is first visible on day 3 and is present in larger quantities on days 5 and 6. The precursor to type II procollagen (CII) is visible on days 5 and 6. The appearance of these products parallels the differentiation of the limb bud cultures into cartilage as measured by morphological characteristics, alcian blue staining and incorporation of $^{35}SO_4$ (31). Results obtained by hybridization of nick-translated pCAR1 or pCAR2 to blots prepared from agarose gel electrophoresis of the RNAs are shown in Figure 4B. Increasing amounts of hybridization to the 5.3 kb RNA parallels the cell-free translation of the type II procollagen mRNA. Identical results have been obtained using pCAR1.

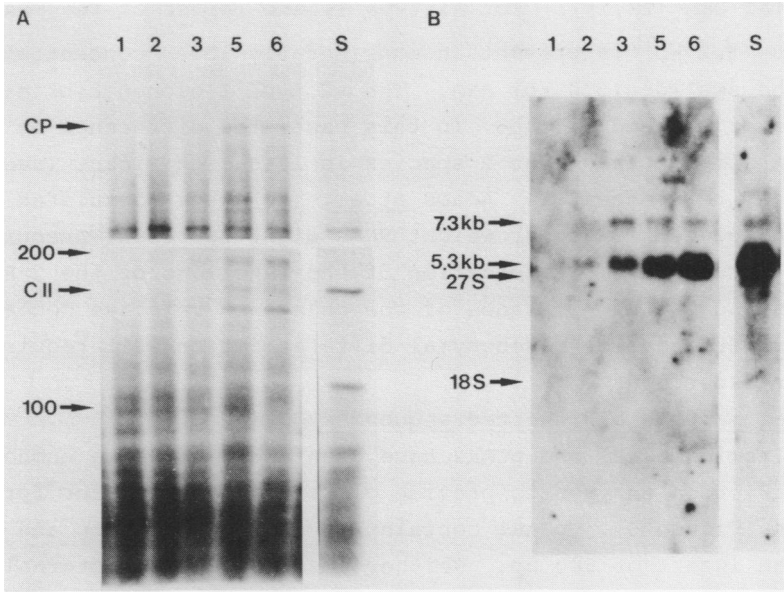

Figure 4. Characterization of total RNA prepared from differentiating limb bud cultures. Stage 24 chick limb mesenchymal cells were prepared and cultured at high density as previously described (6). Total RNA was prepared from a single series of differentiating plates after 1,2,3,5 or 6 days of culture. Total sternal RNA (S) is included for comparison. A, Sodium dodecyl sulfate-polyacrylamide gel electrophoresis of cell-free translation products. A longer exposure of the portion of the gel containing polypeptides larger than 200,000 daltons is shown so that both the proteoglycan core protein (CP) and type II collagen (CII) nascent translation products are visible. Arrows at 100 and 200 indicate the migration positions for polypeptides of 1×10^5 and 2×10^5 daltons respectively, as determined from molecular weight standards. B, Hybridization of ^{32}P-labeled pCAR2 to a glyoxal gel transfer of total RNA prepared from differentiating limb bud cultures and sterna.

These blots also show that the 5.3 kb and 7 kb RNAs apparently do not accumulate at the same rate during the differentiation of cultured limb mesenchyme. At day 3, the 5.3 and 7 kb RNAs are present at nearly equal concentrations. At days 5 and 6, the 7 kb RNA remains at approximately the same concentration or slightly decreases compared with day 3, while the 5.3 kb RNA increases dramatically.

Recombinant cDNA clones previously reported for α1 and α2 type I procollagen each hybridize to 2 RNA species in total

calvarial RNA (28,29). For α1 type I, the larger of the RNA species (7.2 kb) is present in considerably lower concentration than the smaller (5.0 kb) RNA. The α2 type I clones hybridize to RNAs of 5.0 and 5.7 kb. In this case, the difference in concentration between the 2 species is less significant (see Figure 2, reference 29). Adams et al. (29) have postulated that the higher molecular weight RNAs might represent precursors to the mature RNAs. Elucidation of the functions of the 2 RNA species and the significance of the change in relative concentrations during limb mesenchymal differentiation will require further experiments.

Physical mapping of inserted sequences

Plasmids pCAR1 and pCAR2 have been cleaved with a number of restriction enzymes to provide cleavage site maps for further studies (Figure 5). pCAR1 contains an insert of 525 bp and pCAR2 an insert of 680 bp. Neither insert contains internal cleavage sites for Pst I, Hind III, Eco RI, Sal I, Sma I or

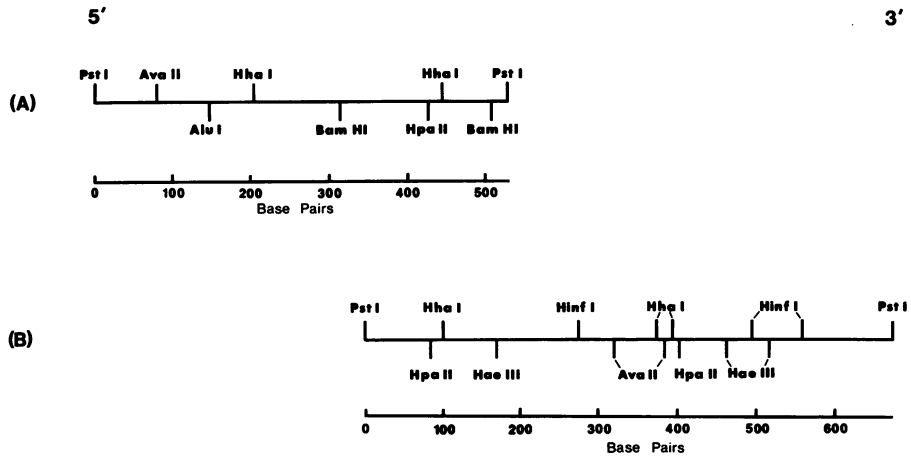

Figure 5. Restriction endonuclease cleavage maps of recombinant plasmids pCAR1 (A) and pCAR2 (B). The Pst I sites are those present in pBR322 at the insertion site. The inserts are oriented to indicate approximately the regions of overlapping sequence homology and to indicate their relative orientations with respect to the 5' and 3' ends of the mRNA to which they correspond. The restriction maps were determined by single and double digestions of both the intact plasmids and the purified inserts.

Bst EII. pCAR2 contains no Hae III or Hinf I sites, but does contain 2 sites for Bam HI and Hha I and one site for Ava II, Alu I and Hpa II. pCAR2 contains no Bam HI or Alu I sites, 3 sites for Hinf I, Hha I, and Hae III and 2 sites for Ava II and Hpa II.

Nick-translated probes prepared from either pCAR1 or pCAR2 hybridize to Southern transfers of both inserts, suggesting that these 2 inserts contain sequences in common. Since little discernible similarity exists in the restriction maps, this homology is most likely limited to overlapping ends. In order to determine the locations of the common sequences, several restriction fragments from each insert have been purified, nick-translated, and hybridized to Southern transfers of restriction digests of the other plasmid. One such experiment is shown in Fig. 6. This experiment shows first that a probe prepared from the left end of pCAR2 (330 bp Pst I-Ava II fragment) hybridizes to pCAR1 (Fig. 6A, lane 5) whereas the right end probe

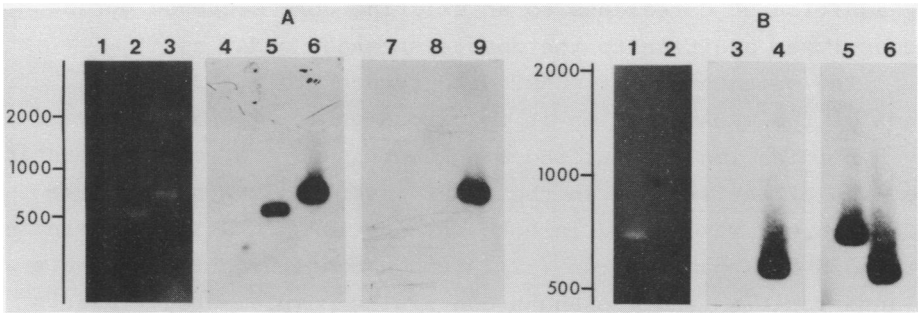

Figure 6. Determination of regions of homology between pCAR1 and pCAR2. Purified inserts from a pro-α1(I) recombinant cDNA clone (A, lanes 1, 4, and 7), pCAR1 (A, lanes 2, 5, and 8; B lanes 2, 4, and 6), and pCAR2 (A, lanes 3, 6, and 9; B, lanes 2, 4, and 6) were electrophoresed on agarose gels and transferred to nitrocellulose filter paper. A, lanes 1-3 and B, lanes 1 and 2, Ethidium Bromide stained gels. A, hybridization of nick translated pCAR2 probes. Lanes 4-6, 330 bp Pst I-Ava II fragment from the left side and lanes 7-9, 294 bp Pst I-Hpa II fragment from the right side. B, hybridization of nick translated pCAR1 probes. Lanes 3 and 4, 318 Pst I-Bam HI fragment from left side, and lanes 5 and 6, 439 bp Pst I-Ava II fragment from right side. Hybridizations were done at 52°C in 50% formamide, 5XSSC to eliminate cross hybridization with pro-α1(I) collagen cDNA sequences. A scale of molecular length in base pairs is at the left of each gel.

(274 bp Hpa II-Pst 1 fragment) does not hybridize (Fig. 6A, lane 8) thus limiting the region of homology in pCAR2 to the region extending 400 bp from the left side. Likewise, the right end probe of pCAR1 (439 bp Ava I Pst I fragment) hybridizes to pCAR2 (Fig. 6B, lane 5) and the left end probe (318 bp Bam HI-Pst I fragment) does not (Fig. 6B, lane 3), limiting the region of homology in pCAR1 to the 195 bp Bam fragment at the right end of pCAR1. (The very small Bam HI-Pst I fragment in pCAR1 has been shown by sequencing to consist entirely, excluding the cleavage sites, of dG:dC homopolymer sequence and has been shown not to be required for hybridization). Other experiments using a series of probes from pCAR2 and pCAR1 further limit the region of homology to the 170 bp Pst I-Hae III fragment of pCAR2. Preliminary sequence data obtained from the 84 bp Pst I-Hpa II fragment of pCAR2 show that this fragment contains sequences in common with the 112 bp Bam HI-Hpa II fragment of pCAR1 giving an overlap of approximately 150 bp. The Bam HI site is not present in pCAR2 and probably was created in pCAR1 by addition of C residues to an existing GGAT sequence during homopolymer addition to the double stranded cDNA insert.

The orientation of pCAR2 with respect to the 3' and 5' ends of the mRNA to which it is complementary was determined by hybridization of purified single stranded probes complementary to each of the strands in the pCAR2 insert to RNA blots. The pCAR2 insert was cloned in both orientations in M13mp7, a single stranded DNA phage vector (32). Phage containing inserts were plaque purified and 2 plaques containing complementary inserts were identified. The orientations of the inserts in the 2 plaques were determined by digestion of double stranded replicative form DNA with Hinf I which gives easily distinguishable fragment patterns dependent upon the orientation.

When attempts were made to prepare labeled probes from the single stranded phage DNAs with E. coli DNA polymerase I, inefficient extension of the universal primer of Anderson et al. (33) was observed. Similar primer extension problems have been reported by others using insertion into the Pst I site of M13mp7 (Dan Kuebbing, BRL, personal communication). This inefficient extension may be the result of the inability of

E. coli DNA polymerase I to read through the duplex generated by the reannealing of approximately 40 base pairs surrounding the Pst I sites. This duplex consists of 27 bp contributed by the inverted repeat present in the cloning site sequence plus approximately 15 G-C pairs contributed by the homopolymer tails of the insert. To circumvent this problem, we have purified insert restriction fragments (117 bp Pst I-Hinf I and 84 bp Pst I-Hpa II fragments from pCAR2) to be used as primers for the synthesis of strand specific probes. Probes were separated from M13mp7 template molecules by denaturation followed by agarose gel electrophoresis. Only the probe primed by the 117 bp Pst I-Hinf I fragment hybridized to mRNA (Fig. 7), giving the orientation of pCAR2 with respect to the 3' and 5' ends of the RNA as shown in Fig. 5.

The insert from pCAR1 has been cloned into a different vector, M13mp9, which has an asymetric cloning site substituted for the symmetric sequence in M13mp7 (J. Messing, personal communication). Preliminary sequence data have been obtained for the pCAR1 insert using the dideoxy technique (24) with the M13mp9 clones as templates with either the universal primer

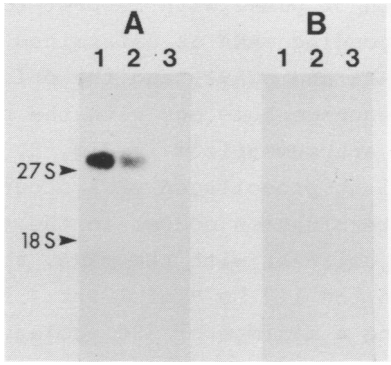

Figure 7. Determination of strand of pCAR2 complementary to mRNA. Probes prepared from M13mp7 recombinants were hybridized to glyoxal gel transfers of RNA prepared from (1) sterna, (2) 8-day differentiated limb bud cultures, and (3) stage 24 limb bud which has not differentiated into cartilage. A, probe prepared from M13mp7 recombinant 1 using the 117 bp Pst I-Hinf I fragment (right side) as primer. B, probe prepared from M13mp7 recombinant 2 using the 84 bp Pst I-Hpa II fragment (left side) as primer.

or pCAR1 restriction fragments as primers. Some additional sequence data were obtained using the Maxam-Gilbert technique (25) with purified restriction fragments from pCAR1. These data (to be published elsewhere) indicate that pCAR1 contains sequences coding for the carboxy-terminal propeptide of a procollegen chain. The sequences in pCAR1 begin at a position corresponding to base 298 (or amino acid 100) beyond the end of the triple helical region in the chicken pro-α1(I) collagen chain (positions are assigned in accordance with the results for pro-α1 (I) collagen mRNA published by Fuller and Boedtker, Ref. 35) and extending to base 795 (amino acid 265) which is 21 bp (7 amino acids) before the termination codon of chicken pro-α1(I) mRNA. Unfortunately, no amino acid sequence data are available for the carboxy-terminal region of type II procollagen. The preliminary sequence data, which cover approximately 90% of pCAR1 (324 bp) including 50-70 bp at either end show 75 and 66% homology with the published base sequences for pro-α1(I) and pro-α2(I) mRNAs respectively. This corresponds well with the degree of amino acid sequence homology reported for the triple helical regions of the bovine α1(I), α2(I) and α1(II) chains (36).

The orientation of pCAR2 with respect to the 3' and 5' ends of the corresponding mRNA as determined above, the apparent overlap between pCAR1 and pCAR2, and the orientation of pCAR1 as determined by sequence homology with the pro-α1(I) mRNA are all consistent and are summarized in Fig. 5. If the termination codon in the cartilage procollagen mRNA occurs at a position homologous to the termination codons in the type I procollagen mRNAs and pCAR2 is colinear with the mRNA, the poly A region would be limited to the 117 bp Hinf I-Pst I fragment at the right end of pCAR2 leaving a minimum of 380 nucleotides of noncoding sequence. This is somewhat longer than the noncoding sequences reported for the pro-α1(I) and pro-α2(I) mRNAs (133 and 300 nucleotides, respectively, Ref. 35).

In summary, 2 recombinant cDNA clones have been isolated which are complementary to an RNA of the size expected for a procollagen mRNA. This RNA has been found only in cartilagenous tissue, it copurifies with type II procollagen translational

activity on Sepharose 4B, and its time of appearance parallels the appearance of translatable mRNA for type II procollagen during the differentiation of chick limb mesenchyme in culture. Preliminary DNA sequence data indicate that pCAR1 is limited to the carboxy-terminal propeptide region of a procollagen chain and that these sequences share considerably more homology with chicken pro-α1(I) mRNA sequences than with pro-α2(I) mRNA sequences. These results strongly suggest that these clones are complementary to type II procollagen mRNA.

ACKNOWLEDGEMENTS

The expert technical assistance of Gloria Enriquez, James Mensch, and Ida Schaefer is appreciated. We gratefully acknowledge assistance from Yoshihiko Yamada in Maxam-Gilbert sequencing and gifts of pBR322 and HB101 from Tom Christianson and Murray Rabinowitz, and M13mp9 from D. Kuebbing, BRL. This work was supported by NIH grants HD-09402, HD-04583, GM-0781, and GM-7543.

*Current Address: Department of Medical Chemistry, University of Turku, Kiinamyllynkatu 10, 20520 Turku 52, Finland

REFERENCES

1. Bornstein, P. and Sage, H. (1980) Ann. Rev. Biochem. 49, 957-1003.
2. Fessler, J.H. and Fessler, L.I. (1978) Ann. Rev. Biochem. 47, 129-162.
3. von der Mark, K., von der Mark, H., and Gay, S. (1976) Dev. Biol. 53, 153-170.
4. Vertel, B.M. and Dorfman, A. (1979) Proc. Natl. Acad. Sci. U.S.A. 76, 1261-1264.
5. von der Mark, K. and von der Mark, H. (1977) J. Cell Biol. 73, 736-747.
6. Upholt, W.B., Vertel, B.M., and Dorfman, A. (1979) Proc. Natl. Acad. Sci. U.S.A. 76, 4847-4851.
7. Sobel, M.E., Yamamoto, T., Adams, S.L., DiLauro, R., Avvedimento, V.E., de Crombrugghe, B., and Pastan, I. (1978) Proc. Natl. Acad. Sci. U.S.A. 75, 5846-5850.
8. Lehrach, H., Frischauf, A.M., Hanahan, D., Wozney, J., Fuller, F., and Boedtker, H. (1979) Biochemistry 18, 3146-3152.
9. Yamamoto, T., Sobel, M.E., Adams, S.L., Avvedimento, V.E., DiLauro, R., Pastan, I., de Crombrugghe, B., Showalter, A., Pesciotta, D., Fietzek, P., and Olsen, B. (1980) J. Biol. Chem. 255, 2612-2615.
10. Lehrach, H., Frischauf, A.M., Hanahan, D., Wozney, J., Fuller, F., Crkvenjakov, R., Boedtker, H., and Doty, P. (1978) Proc. Natl. Acad. Sci. U.S.A. 75, 5417-5421.

11. Ohkubo, H., Vogeli, G., Mudryj, M., Avvedimento, V.E., Sullivan, M., Pastan, I., and de Crombrugghe, B. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 7059-7063.
12. Yamada, Y., Avvedimento, V.E., Mudryj, M., Ohkuba, H., Vogeli, G., Irani, M., Pastan, I., and de Crombrugghe, B. (1980) Cell 22, 887-892.
13. Wozney, J., Hanahan, D., Morimoto, R., Boedtker, H., and Doty, P. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 712-716.
14. Frischauf, A.M., Lehrach, H., Rosner, C., and Boedtker, H. (1978) Biochemistry 17, 3243-3249.
15. Wahli, W., Ryffel, G.U., Wyler, T., Jaggi, R.B., Weber, R., and Dawid, I.B. (1978) Dev. Biol. 67, 371-383.
16. Wickens, M.P., Buell, G.N., and Schimke, R.T. (1978) J. Biol. Chem. 253, 2483-2495.
17. Nelson, T. and Brutlag, D. (1979) Methods in Enzymology 68, 41-50.
18. Mandel, M. and Higa, A. (1970) J. Mol. Biol. 53, 159-162.
19. Thayer, R.E. (1979) Anal. Biochem. 98, 60-63.
20. Denhardt, D.T. (1966) Biochem. Biophys. Res. Commun. 23, 641-646.
21. McMaster, G.K. and Carmichael, G.C. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 4835-4838.
22. Thomas, P.S. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 5201-5205.
23. Alwine, J.C., Kemp, D.J., Parker, B.A., Reiser, J., Renart, J., Stark, G.R., and Wahl, G.M. (1979) Methods in Enzymology 68, 220-242.
24. Southern, E. (1979) Methods in Enzymology 68, 152-176.
25. Maxam, A.M. and Gilbert, W. (1980) Methods in Enzymology 65, 499-560.
26. Burgeson, R.E. and Hollister, D.W. (1979) Biochem. Biophys. Res. Commun. 87, 1124-1131.
27. King, C.R., Udell, D.S., and Deeley, R.G. (1979) J. Biol. Chem. 254, 6781-6786.
28. Rave, N., Crkvenjakov, R., and Boedtker, H. (1979) Nucleic Acids Res. 6, 3559-3567.
29. Adams, S.L., Alwine, J.C., de Crombrugghe, B., and Pastan, I. (1979) J. Biol. Chem. 254, 4935-4938.
30. Paterson, B.M., Roberts, B.E., and Kuff, E.L. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 4370-4374.
31. Levitt, D. and Dorfman, A. (1974) in Current Topics in Developmental Biology, Moscona, A. and Monroy, A., eds., Vol. 8, pp. 103-147, Academic Press, New York.
32. Messing, J., Crea, R., and Seeburg, P.H. (1981) Nucleic Acids Res. 9, 309-321.
33. Anderson, S., Gait, M.J., Mayol, L., and Young, I.G. (1980) Nucleic Acids Res. 8, 1731-1743.
34. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 5463-5467.
35. Fuller, F. and Boedtker, H. (1981) Biochemistry 20, 996-1006.
36. Mathews, M.B. (1980) in Biology of Collagen, Viidik, A. and Vuust, J., pp. 193-209, Academic Press, New York.