Isolation and Partial Characterization of Genomic Clones Coding for a Human Pro-α1(II) Collagen Chain and Demonstration of Restriction Fragment Length Polymorphism at the 3′ End of the Gene

Anne Marie Nunez, Clair Francomano,[‡] Marian F. Young,[§] George R. Martin, and Yoshihiko Yamada*

Laboratory of Developmental Biology and Anomalies, National Institute of Dental Research, National Institutes of Health, Bethesda, Maryland 20205

Received February 4, 1985

ABSTRACT: We have isolated two clones containing 19 kilobases (kb) of the human gene coding for a pro-α1(II) collagen chain from human λ genomic DNA libraries. A 3′ clone, HC_2A, was selected by cross-hybridization with a cDNA clone containing sequences coding for the carboxy propeptide of chick type II procollagen. A second clone, HC_2B, was obtained by screening the library with the 5′ part of HC_2A. The sequence analysis of exon 3 corresponding to the C propeptide reveals the presence of stretches of conserved nucleotides between the human and the chick type II procollagen genes. On Northern blots, the human collagen clone hybridizes strongly to a 5.5-kb RNA for the rat type II procollagen chain. Finally, studies of genomic DNAs from normal individuals reveal the presence of a *Hin*dIII and a *Bam*HI polymorphic site at the 3′ end of the gene.

Cartilage contains chondrocytes surrounded by a homogeneous extracellular matrix composed primarily of collagen and proteoglycan. In addition to its cushioning function in joints, hyaline cartilage forms the skeletal primordia during embryogenesis and regulates the form and size of various bones. Alterations in its structures or synthesis undoubtably account for a variety of human chondrodystrophies. Indeed, altered collagen fibers have been noted in the tibial plate of certain patients with diastrophic dwarfism (Stanescu et al., 1982).

Type II collagen is the major collagen component of the cartilage matrix and is found only in cartilage and the vitreous humor of the eye (Von der Mark et al., 1976, 1979). Chondrocytes synthesize pro-α1(II) chains, assemble three such chains into a procollagen molecule in a process involving multiple posttranslation modifications, and transport the protein to the matrix where it is cleaved to collagen and deposited in small fibers.

We and others (Sandell et al., 1983; Strom & Upholt, 1984; Sandell et al., 1984; Young et al., 1984) have been studying the structure of the type II collagen gene. Previously, we isolated two cDNA clones covering approximately 80% of the C propeptide of the chick pro-α1(II) chain and a 4.5-kilobase (kb) genomic clone coding for the 3′ end of the pro-α1(II) gene (Young et al., 1984). The chick cDNA clone has been used as a probe to screen a human genomic library, and here we report the isolation and partial characterization of two genomic clones, HC_2A and HC_2B, containing approximately 19 kb of a human pro-α1(II) collagen gene and the existence of restriction polymorphic sites at the 3′ end of the gene.

Materials and Methods

Screening of a Human Library. HC_2A was obtained from a library of human genomic fragments in λ Charon 4A constructed by Lawn et al. (1978). It was screened by the method described by Maniatis et al. (1978). A 780 base pair (bp) chick pro-α1(II) collagen cDNA was isolated from plasmid pCs1 and labeled by nick translation (Maniatis et al., 1976) for use as a hybridization probe. Hybridization was performed at 46 °C for 36 h in 6× SSC, 10× Denhardt's solution, 0.1% sodium dodecyl sulfate (SDS), and 100 μg/ml salmon sperm DNA. DNA filters were washed twice with 1× SSC and 0.1% SDS at room temperature. HC_2B was obtained by the same method by screening a λ Charon 28 library constructed by Hieter et al. (1980) with as a probe the 3-kb *Eco*RI fragment of HC_2A.

DNA Preparation and Characterization. Phage and phage DNA were prepared as described by Maniatis et al. (1982). Restriction endonuclease cleavage site mapping was done by single and double digestions of HC_2A and HC_2B and by hybridization of nick-translated fragments to DNA gel blots. Two types of labeled probes were used to characterize the HC_2A clone. These are a chick pro-α1(II) as a 3′ probe and a chick pro-α1(II) genomic fragment as a 5′ probe. pCsI is a 780-bp cDNA that contains several untranslated nucleotides, the first three exons and the 45 most 3′-nucleotide residues of exon 4 of the mRNA coding for the C propeptide of chick pro-α1(II) collagen (Young et al., 1984). The 5′ probe is a 1.8-kb chick *Sma*I genomic fragment that contains exon 4 coding for the C telopeptide and part of the C propeptide and approximately 1.3 kb of adjacent 5′ sequences coding for the triple helical region of chick pro-α1(II) collagen (Sandell et al., 1983; Young et al., 1984).

DNA Sequencing. DNA fragments were isolated and labeled with [γ-^{32}P]ATP by polynucleotide kinase. Strand separation and DNA sequencing were carried out by the method described by Maxam & Gilbert (1980).

Isolation of RNA. Total RNA was isolated from a rat chondrosarcoma (Smith et al., 1975) by the guanidine method (Adams et al., 1977). Total RNA from human rhabdomyosarcoma CCL 136 (McAllister et al., 1969) and from human cell lines HT 1080 (Rasheed et al., 1974) and CRL 1502 (Hay et al., 1978) were isolated by centrifugation through a cesium

[‡]Present address: Johns Hopkins University School of Medicine, Baltimore, MD 21205.

[§]Present address: Mineralized Tissue Research Branch, National Institute of Dental Research, National Institutes of Health, Bethesda, MD 20205.

chloride cushion (Ullrich et al., 1977). Poly(A)-containing RNA was selected by oligo(dT)–cellulose chromatography (Aviv & Leder, 1972). We have not yet been successful in obtaining RNA from human hyaline cartilages.

Northern Hybridization. Poly(A)-containing RNA (4 μg/lane) was fractionated by electrophoresis on a 1% agarose gel in the presence of 2.2 M formaldehyde (Lehrach et al., 1977). The RNA in the gel was then transferred to nitrocellulose paper (Southern, 1975). A 4-kb *Eco*RI fragment containing sequence coding for the carboxy propeptide of HC_2A was used as probe. Human α1(I) probe Hf 677 (Chu et al., 1982) and human α1(III) (a gift of Dr. R. Crystal, unpublished data) and α1(IV) (S. Kato and Y. Yamada, unpublished data) probes were also hybridized to Northern blots. Hybridization was performed at 37 °C overnight in 50% formamide, 10× Denhardt's solution, 5× SSC, 0.1% SDS, 50 μg/mL salmon sperm DNA, and 200 μg/mL calf liver tRNA. After hybridization, the filters were washed 3 times with 0.1× SSC containing 0.1% SDS, at 37 °C.

Isolation of Genomic DNA from Nuclei of Leukocytes. Genomic DNA samples used in this study were collected from individuals with a variety of hemoglobinopathies whose β-globin genes were investigated at Johns Hopkins Hospital. A 15–20-mL sample of whole blood was collected in ethylenediaminetetraacetic acid (EDTA). The cells were pelleted, and the plasma was removed. Cells were then resuspended in an equal volume of 2× sucrose–Triton [320 mM sucrose, 10 mM tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl) (pH 7.6), 5 mM $MgCl_2$, and 1% Triton X-100]. The sample was brought to 50 mL with 1× sucrose–Triton, left to sit on ice for 1 h, and centrifuged at 4 °C at 1800 rpm for 20 min. The pelleted nuclei were resuspended in 75 mM NaCl and 24 mM EDTA (pH 8.0), homogenized, and lysed by adding sodium dodecyl sulfate (SDS) to a final concentration of 1% (w/v). Pronase was added and the solution incubated at 37 °C for 2 h. Subsequently, proteinase K was added and the solution incubated at 37 °C for 2 more h. Following three extractions with phenol and then chloroform, KCl was added to the aqueous phase to a concentration of 0.2 M, and 2.5 volumes of cold ethanol were layered over the aqueous phase. The DNA was then spun out from the interphase on the tip of a Pasteur pipet, washed with ethanol, and resuspended in 0.1× SSC (15 mM NaCl, 1.5 mM sodium citrate, pH 7.0). Aliquots were stored at 4 and −80 °C.

Restriction Endonuclease Analyses and "Southern" Blot Hybridization. The DNAs were digested with various restriction endonucleases under conditions recommended by the suppliers. The digested DNAs were then electrophoresed on 1% agarose gels, transferred to Schleicher & Schuell BA 85 nitrocellulose filters or Biodyne nylon membranes (Southern, 1975), and baked at 65 °C for 6 h. The filters were prehybridized at 37 °C for 12–24 h with a prehybridization solution containing 10 mM 4-(2-hydroxyethyl)-1-piperazinethanesulfonic acid (Hepes) buffer (pH 7.5), 50% formamide, 3× SSC, 0.1% SDS, salmon DNA (100 μg/mL), 0.2% bovine serum albumin, 0.2% Ficoll 400, 0.2% poly(vinylpyrrolidone), 0.05% sodium pyrophosphate, and yeast tRNA (200 μg/mL). The ^{32}P-labeled probe was heat-denatured, added to the filters in prehybridization solution, and allowed to hybridize for 24–48 h at 37 °C. Two fragments of HC_2A were used as probes (see top part of Figure 2), a 4-kb *Eco*RI fragment and a 2.7-kb *Sma*I fragment isolated from the 3-kb *Eco*RI fragment after subcloning in PUC-8 plasmid (Maniatis et al., 1982). The filters were then washed 3 times in 0.1× SSC and 0.1% SDS at 65 °C. Autoradiography was carried out with light-intensifying screens and two films, one to be developed in 24 h and the other in 2–7 days depending on the strength of the initial signal.

Results

Isolation of Clones. A λ Charon 4A phage library (Lawn et al., 1978) of human genomic fragments was screened with a cloned cDNA (pCs1) coding for most of the C propeptide of the chick pro-α1(II) chain (Young et al., 1984). About 3×10^5 plaques were screened, and two clones were isolated. Analysis of restriction enzyme cleavages showed that both clones contained the same human gene fragment. One of these clones, HC_2A, was taken for further study. The 5′, 3-kb *Eco*RI fragment of HC_2A was used as probe to screen a human λ charon 28 library (Hieter et al., 1980). Among 18 overlapping clones, HC_2B was selected on the basis of having the longest extension, 12 kb.

Restriction Analysis of HC_2A. To locate the position and orientation of the collagen gene sequences within the human genomic DNA insert, HC_2A DNA was digested with several restriction endonucleases, electrophoresed on an agarose gel, transferred to nitrocellulose paper, and hybridized to two different labeled probes. The first probe (3′ probe) is a 780-bp cDNA, called pCs1, containing 63 untranslated nucleotides, three exons, and the 45 most 3′-nucleotide residues of exon 4 of the mRNA coding for the C propeptide of the chick pro-α1(II) chain. HC_2A DNA is cleaved into six fragments by *Eco*RI (Figure 1). Of these six fragments, two represent the long and short arms of Charon 4A (20 and 10.8 kb, respectively). The remaining four fragments (4.5, 4.0, 3.0, and 0.9 kb) totaling 12.4 kb belong to the human-inserted DNA. The 3′ probe pCs1 hybridizes to a 4-kb *Eco*RI fragment, a 4.6-kb *Bam*HI fragment, 5.5- and 0.3-kb *Hind*III fragments, and two *Xba*I fragments of more than 20 kb (Figure 1B). The 5′ probe hybridizes to the 4-kb and also to a 3-kb *Eco*RI fragment, to a 4.6-kb *Bam*HI fragment, to a more than 20-kb *Xba*I fragment, and finally to a 10.5-kb *Hin*dIII fragment (Figure 1C). An additional, weak hybridization of the 5′ probe to a 5–6-kb *Hin*dIII fragment was also observed (Figure 1C), probably arising from a partial cleavage of the *Hin*dIII sites located in the center of the 4-kb *Eco*RI fragment (see Figure 2). The orientation of the human gene within the λ vector was determined as follows: pCs1 (3′ probe) hybridizes to the 4-kb *Eco*RI fragment (Figure 1B) while the 5′ probe hybridizes to the 4- and 3-kb *Eco*RI fragments (Figure 1C). This indicates that the 3-kb fragment is located 5′ to the 4-kb fragment. A *Hin*dIII fragment of 10.5 kb and a 4.6-kb *Bam*HI fragment (Figure 1C) containing the short arm of λ hybridize to the 5′ probe. Therefore, the short arm of the λ vector is located on the 5′ side of the inserted gene. HC_2B was studied with the same method using different fragments of HC_2A as probes. All these results and related similar experiments by double restriction have led to the development of the restriction maps shown in Figure 2. In summary, these two genomic clones contain approximately 19 kb of human collagen gene and 5.4 kb of 3′-flanking sequences. It should be noted that the two chick probes hybridize to four different fragments of the human DNA of the HC_2A clone that follow one another from the 3′ to 5′ direction (Figure 2). These included a 1.5-kb *Eco*RI fragment (fragment A) and a 0.32-kb *Hin*dIII fragment (fragment B) hybridizing to pCsI, then a 1.8-kb *Eco*RI *Hin*dIII fragment (fragment C), and finally a 3-kb *Eco*RI fragment (fragment D) hybridizing to the *Sma*I chick 5′ probe. This indicates that at least four regions of high homology exist at these levels between the chick pro-α1(II) chain gene and the human genomic fragment described here.

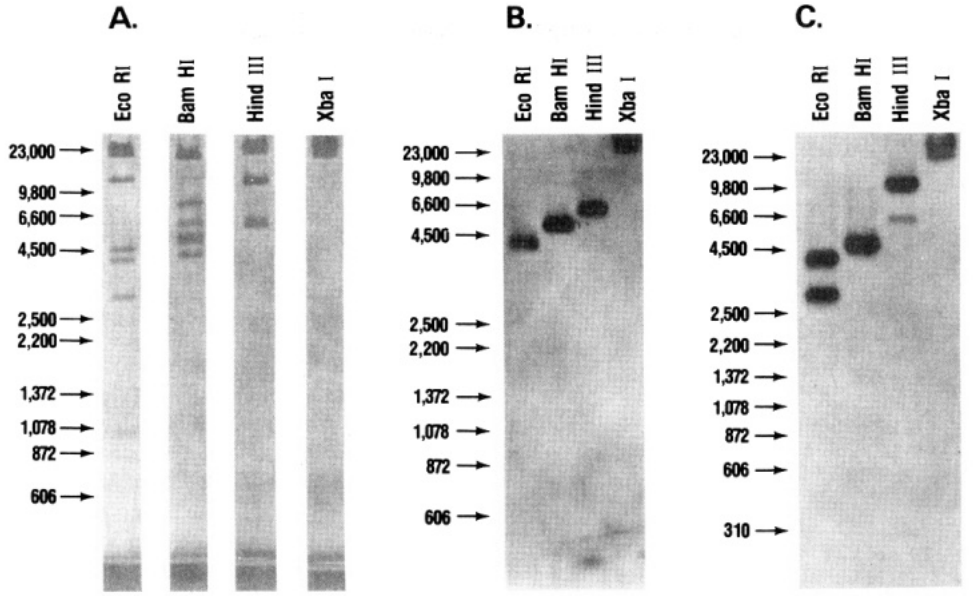

FIGURE 1: Cross hybridization of the human genomic clone HC_2A restriction fragments to the two chick pro-α1(II) collagen probes (3′ and 5′ probes, see text): (A) ethidium bromide stained agarose gel; (B) Southern hybridization with ^{32}P-labeled chick cDNA pCsl (3′ probe); (C) Southern hybridization with ^{32}P-labeled chick genomic 1.8-kb *Sma*I fragment (5′ probe).

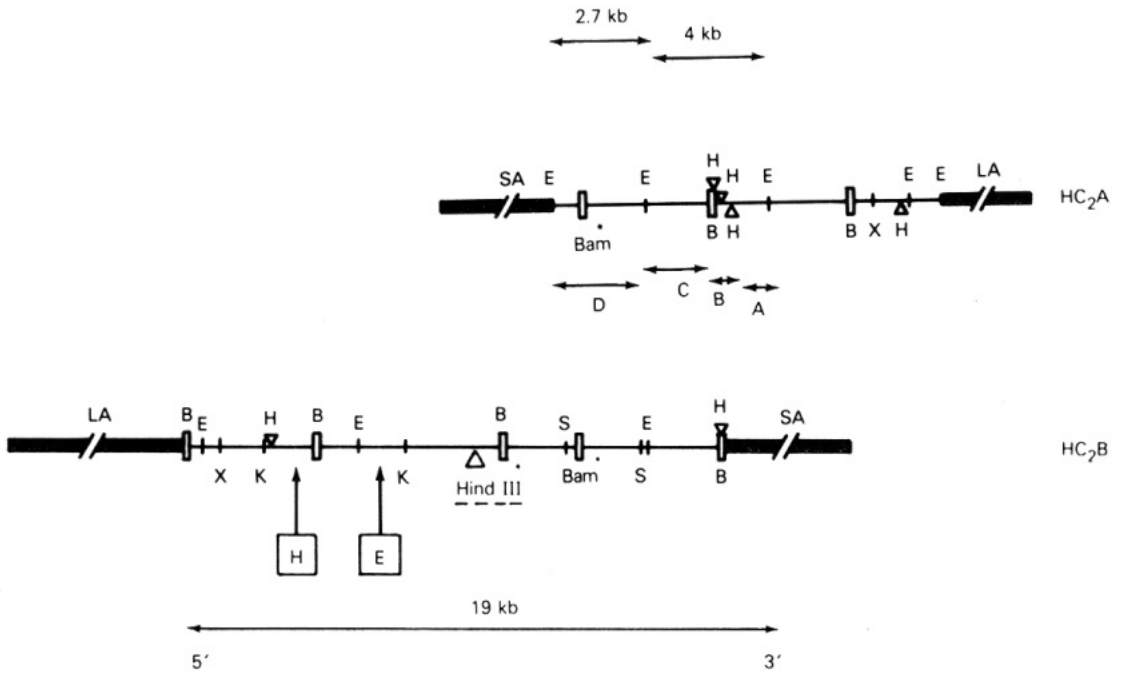

FIGURE 2: Restriction map of HC_2A and HC_2B DNAs. The heavy line represents the two arms of the λ vector. Restriction enzyme cleavage sites are indicated as follows: E, *Eco*RI; B, *Bam*HI; H, *Hin*dIII; X, *Xba*I; S, *Sma*I; K, *Kpn*I. The orientation of the clone is indicated. A and B represent two fragments hybridizing to pCsI and C and D two fragments hybridizing to the 5′ chick probe [*Sma*I 1.8-kb genomic fragment of chick pro-α1(II) clone]. Fragment B was sequenced. HindIII* (absent in HC_2B) and BamHI* are the two polymorphic sites. Boxed H and E indicate the sites absent in HC_2B and present in the α1(I)-like cosmid clone described by Weiss et al. (1982).

Sequence Analysis. The results discussed above imply that there is a high conservation of nucleotide sequence between the chick pro-α1(II) gene and the recombinant clone studied here. We have sequenced the 0.32-kb *Hin*dIII fragment that hybridizes to the chick clone pCs1 (Figure 3). The sequence analysis shows, by analogy with other collagen C propeptides, that the fragment contains the whole exon 3 of the C propeptide plus splicing signals and some intron sequences at its 3′ and 5′ ends. Several features indicate that this genomic clone codes for part of the human pro-α1(II) chain. The sequence of this region is not as highly conserved among different collagen species as is exon 2, the site of the carbohydrate attachment site (Yamada et al., 1983). First, this sequence is very different from the corresponding sequences of the human pro-α1(I), pro-α2(I) (Bernard et al., 1983ab), and pro-α1(III) chains (Loidl et al., 1984). It is also very different from the α1(IX) chain (Ninomiya et al., 1984). One particular feature is the absence of amino acid residue 141 due to a three-nucleotide deletion that is absent from the chick pro-αI(II) chain (Young et al., 1984) and the chick and human pro-α1(III) chain genes (Yamada et al., 1983; Loidl et al., 1984) but is present as a proline residue in the C propeptides of the human (Bernard et al., 1983a,b) and chick (Fuller & Boedtker, 1981) pro-α1(I) and pro-α2(I) chains. Second, in a comparison of the amino acid sequence of various C propeptides, the closest homology is found between this exon and

```
                                                                    T
(207)  gctttttgtgtctgtgcctgtctgagcccccatgggtgctgcctcttcccccctgca GGA GAC TAC
                                                                  Gly Asp Tyr

                        G       G                               C   A   A
(210)   TGG ATT GAC CCC AAC CAA GGC TGC ACC TTG GAC GCC ATG AAG GTT TTC TGC
(98)   Trp Ile Asp Pro Asn Gln Gly Cys Thr Leu Asp Ala Met Lys Val Phe Cys
                                                  Ile

                      A           C                 G  CC   C AGC  G   A C
(227)  AAC ATG GAG ACT GGC GAG ACT TGC GTC TAC CCC AAT CCA GCA AAC GTT CCC
(115)  Asn Met Glu Thr Gly Glu Thr Cys Val Tyr Pro Asn Pro Ala Asn Val Pro
                                                 Thr      Ser Ser Ile

       G                    C             CG   A   C        G     G
(244) AGG AAG AAC TGG TGG AGC AGC AAG AGC*AAG GAG AAG AAA CAC ATC TGG TTT
(132) Lys Lys Asn Trp Trp Ser Ser Lys Ser Lys Glu Lys Lys His Ile Trp Phe
      Arg                   Thr           Thr      Asp              Val

       C   G           C   C   T       C
(261) GCA GAA ACC ATC AAT GGT GGC TTC CAT gtgagtacctgggtgccctagatgatgagcaga
(149)  Gly Glu Thr Ile Asn Gly Gly Phe His
       Ala

       gatggctcctcaaactctttcttttctttctccctggaagc
```

FIGURE 3: Sequence analysis of fragment B of HC_2A (0.32-kb *Hin*dIII fragment). The nucleotide and amino acid sequences of the chicken pro-α1(II) collagen are taken from Young et al. (1984). The human and chick sequences were aligned to optimize homology. (First line) Nucleotide sequence in the chick where it differs from the human. (Second line) Human nucleotide sequence. (Third line) Human amino acid sequence. (Fourth line) Chick amino acid sequence where it differs from the human. (*) indicates the position of the additional amino acid residue found in the α1(I) and α2(I) genes.

the corresponding chick pro-α1(II) chain sequence rather than the exon sequences in the pro-α1(I), pro-α2(I), or pro-α(III) chains. The same result is found for nucleotide replacements (Table I). Third, the degree of divergence between HC_2A and human pro-α1(I) and pro-α2(I) exons 3 (Table I) is

Table I: Comparison of Amino Acid and Nucleotide Differences in Exon 3 between Human HC_2A Clone and Chick and Human Pro-α Chains

	amino acid replacement (% difference)	nucleotide changes (% difference)
chick pro-α1(II)	17	16
chick pro-α1(III)	33	28
chick pro-α1(I)	28	22
human pro-α1(I)	32	22
chick pro-α2(I)	40	31
human pro-α2(I)	43	31

consistent with the one observed between chick pro-α1(II) and chick pro-α2(I) and pro-α1(I) C propeptides (Young et al., 1984). Finally, we observe the existence of three long stretches of conserved nucleotides (14, 20, and 17) at the 5′ end of the exon when the human sequence is compared to the chick sequence (Figure 3). Such identical nucleotide stretches are not observed when the human sequence is compared to the corresponding regions of pro-α1(I), pro-α2(I), and pro-α1(III) chain genes.

Tissue-Specific Expression. Poly(A)-containing RNA preparations from rat chondrosarcoma (a source of type II procollagen mRNA), from human fibroblast cell line CRL 1502 (a source of types I and III procollagen mRNAs), from human cell line HT 1080 (a source of type IV procollagen mRNAs), and from human rhabdomyosarcoma cell line CCL136 (a source of types III, IV, and V procollagen mRNAs (Krieg et al., 1979) were blotted and hybridized to ^{32}P-labeled nick-translated probes. Four different probes were used: a probe for the human pro-α1(I) chain (Chu et al., 1982), the 4-kb *Eco*RI fragment of HC_2A described in this paper, a probe for the human pro-α1(III) chain (a gift from Dr. R. Crystal, NIH), and a probe for the human pro-α1(IV) chain (S. Kato and Y. Yamada, unpublished data). The data presented in Figure 4 clearly indicate that an mRNA slightly larger than 28 *S* from the chondrosarcoma hybridizes preferentially to the probe prepared from the HC_2A clone. This pattern was identical with the one observed with a rat pro-α1(II) chain cDNA as a probe (Kohno et al., 1984). However, the 4-kb *Eco*RI probe prepared from the human clone HC_2A (Figure 4A) also hybridizes with fibroblast mRNA, although less strongly, but not with mRNAs for the types III, IV, and V

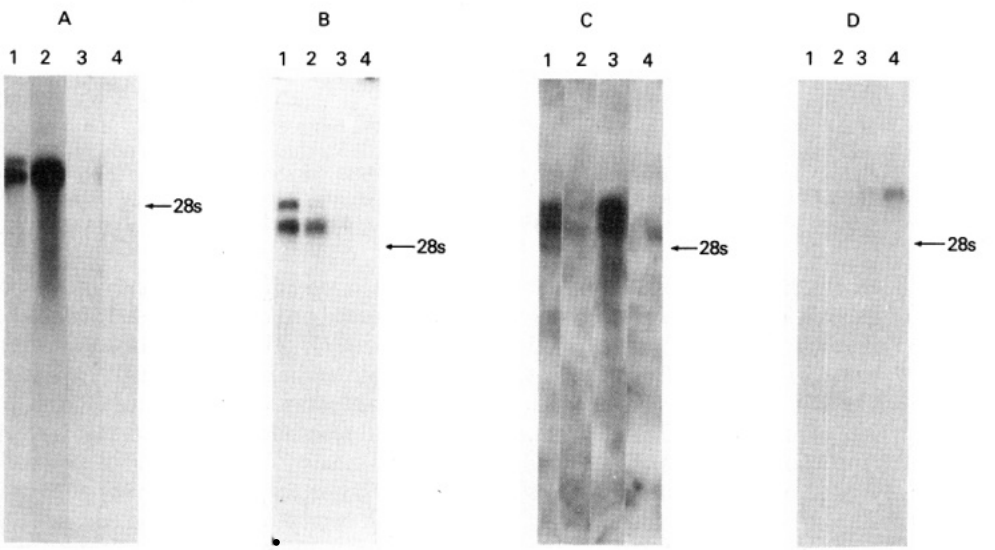

FIGURE 4: Northern hybridization. Poly(A)-containing mRNAs from human fibroblasts from cell line CRL1502 (lane 1), from rat chondrosarcoma (lane 2), from human rhabdomyosarcoma (lane 3), and from human cell line HT 1080 (lane 4) were separated on a formaldehyde agarose gel and transferred to nitrocellulose paper. The paper was hybridized with ^{32}P-labeled probes for the human pro-α1(I) chain (B), pro-α1(III) chain (C), and pro-α1(IV) chain (D) and with the 4-kb *Eco*RI fragment from HC_2A (A).

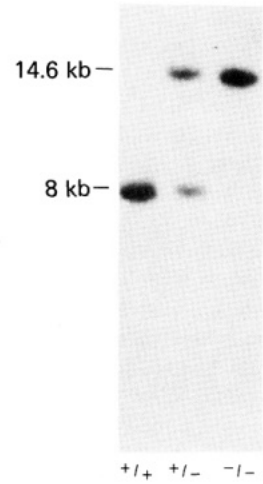

FIGURE 5: *Hin*dIII restriction fragment length polymorphisms. Human genomic DNAs from a control population were digested by *Hin*dIII, electrophoresed on 1% agarose gel, transferred to Biodyne nylon membranes, and hybridized to the 2.7-kb *Sma*I fragment of HC_2A clone. (+/+) indicates a homozygous individual for the presence of the restriction site, (+/−) a heterozygous individual, and (−/−) a homozygous individual for the absence of the site.

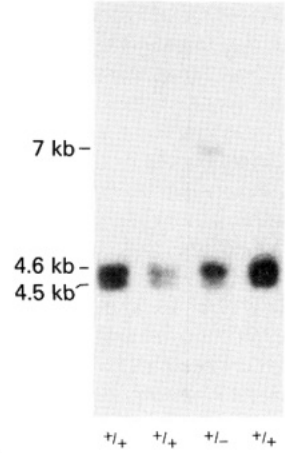

FIGURE 6: *Bam*HI restriction fragment length polymorphisms. Human genomic DNAs from a control population were digested by *Bam*HI, electrophoresed on 1% agarose gel, transferred to Biodyne nylon membranes, and hybridized to the 4-kb *Eco*RI fragment of HC_2A clone. (+/+) indicates a homozygous individual for the presence of the restriction site, (+/−) a heterozygous individual, and (−/−) a homozygous individual for the absence of the site.

procollagen chains. The human pro-α1(I) probe hybridizes with mRNA from the chondrosarcoma (Figure 4B). This is presumably due to a cross-hybridization between the α1(I) probe and the type II mRNA and to the presence of a pro-α1(I) message in the chondrosarcoma RNA preparation 1. Nonetheless, the ratio of the hybridization signals (chondrosarcoma RNA + HC_2A probe)/(fibroblast RNA + HC_2A probe) is several times higher than the ratio (chondrosarcoma RNA + α1 type I probe)/(fibroblast RNA + α1 type I probe). This quantitative difference observed despite the use of mRNAs and probes of different species strengthens our conclusion that the clones presented here indeed code for the human pro-α1(II) collagen gene. The human type III (Figure 4C) and type IV collagen probes (Figure 4D) hybridize very faintly to the rate type II message. Taken together with the sequence data, these results indicate that HC_2A contains sequences that encode for a portion of the human pro-α1(II) chain.

Study of Restriction Fragment Length Polymorphism among Normal Human Individuals. DNA from peripheral leukocytes of apparently normal Caucasian and black individuals was digested by restriction endonucleases *Hin*dIII, *Bam*HI, and *Eco*RI and hybridized to the 2.7- and 4-kb fragments from HC_2A shown in Figure 2. No *Eco*RI polymorphic site was detected. Two restriction polymorphic sites were illustrated as HindIII* and BamHI*. The *Hin*dIII polymorphism is illustrated in Figure 5. The 2.7-kb *Sma*I fragment was used as probe to characterize it. The HC_2B clone does not contain that site. This is a common polymorphism with an allelic frequency for the (+) allele of 0.47 in Caucasians (n = 36) and 0.37 in blacks (n = 33). In constrast, the *Bam*HI polymorphism, illustrated in Figure 6, with the 4-kb *Eco*RI fragment as probe is rare. Only one of 33 black controls was heterozygous for the marker. Two of her three children were similarly heterozygous, indicating that the marker is inherited in a Mendelian fashion (data not shown). All other black controls and all of the Caucasians studied were homozygous for the presence of this site.

Discussion

In summary, we have isolated two genomic clones, HC_2A and HC_2B, that contain approximately 19 kb of human DNA apparently coding for the pro-α1(II) chain of cartilage. This conclusion is based on several facts. First, there is a close homology in sequence between the HC_2A clone and a corresponding region of the chick pro-α1(II) chain gene. Second, the hybridization of two distinct regions of the chick pro-α1(II) chain to four different fragments of HC_2A indicates the presence of contiguous homologous segments. Third, the HC_2A clone hybridizes preferentially with a mRNA species from cartilage a little larger than 28 *S*. Vuorio et al. (1984) have also identified a mRNA of this size as the mRNA for human type II collagen.

Strom & Upholt (1984) have reported the isolation of human genomic clones containing 11 kb of DNA, which appear to code for a part of the pro-α1(II) chain. Their identification was based on similar lines of evidence, including sequence homologies with chick, cross-hybridization, and the presence of a more thermostable hybrid with chick pro-α1(II) DNA than with chick pro-α1(I) DNA. On the basis of similarities in our restriction maps, it would appear that their clones lie along the 3′ half of our clones.

Similarly, Weiss et al. (1982) described "an α1(I)-like" human clone, isolated from a cosmid library, which Cheah et al. (1984) and Stocker et al. (1984) found to be most like the human pro-α1(II) chain gene on the basis of similar types of evidence. Our restriction maps show considerable homology except at the 5′ end of the HC_2B clone. We note particularly, at this level, in our clone, the absence of an *Eco*RI and a *Hin*dIII site (their approximate locations are indicated on the HC_2B map, on Figure 2, by boxed E and H). This discrepancy could be due to the presence of rare polymorphic sites. In the case of the *Hin*dIII site H, its presence should be a very rare event since Dreisel et al. (1982) also looking for *Hin*dIII polymorphism using the whole α1(I) like clone as a probe were not able to detect it in their human control population (n = 34). The differences could also be due to possible rearrangements or other artifacts arising during cloning in the cosmid system.

DNAs from 33 Caucasian and 36 black control individuals were studied for the presence of restriction polymorphisms with a variety of restriction enzymes and the 2.7-kb *Sma*I and 4-kb *Eco*RI probes from HC_2A (indicated on top of maps, Figure 2). A common, *Hin*dIII polymorphism was found that has also been reported by Dreisel et al. in 1982 for the α1(I) like

gene believed to code for type II procollagen. A rarer, *Bam*HI polymorphism was also found. These polymorphisms should be of use in studying the familial association of genetic defects in type II collagen. Just as type I collagen defects occur in osteogenesis imperfecta and the Ehlers Danlos syndrome and type III collagen defects occur in the Ehlers Danlos type IV syndrome, defects in cartilage collagen should be observed in the chondrodystrophies.

The data relative to the pro-α1(II) chain gene concern mainly its 3′ end. In this region the general organization and the precise structure of chick and human pro-α1(II) chain genes seem to be very conserved and also resemble other genes coding for interstitial collagens. The portion of the gene coding for the chick pro-α1(II) C propeptide is divided into four exons (Sandell et al., 1984; Young et al., 1984) as in the chick pro-α2(I) and pro-α1(III) chain genes (Yamada et al., 1983). Information on human pro-α1(II) exon 4 (Strom & Upholt, 1984) and exon 3 (Figure 3) suggests the same overall structure. Strom & Upholt (1984) have reported the presence of an exon of 54 bp at the 5′ end of their human pro-α1(II) chain genomic clone. This length is characteristic of exons coding for helical regions in other interstitial collagen genes. The gene reported by Weiss et al. (1982) is about 35 kb long and nearly as long as the chick (Vogeli et al., 1980) and human (Bernard et al., 1983b) pro-α2(I) genes and less compact than the human pro-α1(I) gene (18 kb long) (Bernard et al., 1983b). Presumably, this will be mostly due to differences in intron structures although a further characterization of the gene is necessary to establish a better model.

References

Adams, S. L., Sobel, M. E., Howard, B. H., Olden, K., Yamada, K. M., de Crombrugghe, B., & Pastan, I. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 3399–3403.

Aviv, H., & Leder, P. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 1408–1412.

Bernard, M. P., Myers, J. C., Chu, M. L., Ramirez, F., Eikenberry, E. F., & Prockop, D. (1983a) *Biochemistry 22*, 1139–1143.

Bernard, M. P., Chu, M.-L., Myers, J. C., Ramirez, F., Eikenberry, E. F., & Prockop, D. (1983b) *Biochemistry 22*, 5213–5223.

Cheah, K. S. E., Griffin, F. R., Tilley, R., Grosveld, F. G., & Solomon, E. (1984) *J. Cell. Biochem. Suppl. 8B* (Abstr. 1568).

Chu, M. L., Myers, J. C., Bernard, P., Ding, J. F., & Ramirez, F. (1982) *Nucleic Acids Res. 19*, 5925–5934.

Cornfield, V. A., & Hay, R. J. (1978) *In Vitro 14*, 787–794.

Driesel, A. J., Schumacher, A. M., & Flavell, R. A. (1982) *Hum. Genet. 62*, 175–176.

Fuller, F., & Boedtker, H. (1981) *Biochemistry 20*, 2621–2627.

Hieter, P. A., Max, E. E., Seidman, J. G., Maizel, J. V., & Leder, P. (1980) *Cell (Cambridge, Mass.) 22*, 197–207.

Kohno, K., Martin, G. R., & Yamada, Y. (1984) *J. Biol. Chem. 259*, 13668–13673.

Krieg, T., Timpl, R., Alitalo, K., Kurkinen, M., & Vaheri, A. (1979) *FEBS Lett. 104*, 405–410.

Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G., & Maniatis, T. (1978) *Cell (Cambridge, Mass.) 15*, 1157–1175.

Lehrach, H., Diamond, D., Wozney, J. M., & Boedker, H. (1977) *Biochemistry 16*, 4743.

Loidl, H. R., Brinker, J. M., May, M., Pihlajaniemi, T., Morrow, S., Rosenbloom, J., & Myers, J. C. (1984) *Nucleic Acids Res. 12*, 9383–9394.

MacAllister, R. M., Melnyk, J., Finklestein, J. Z., Adams, E. C., & Gardner, M. B. (1969) *Cancer 24*, 520–526.

Maniatis, T., Kee, S. G., Efstratiadis, A., & Kafatos, F. C. (1976) *Cell (Cambridge, Mass.) 8*, 1630.

Maniatis, T., Hardison, R. C., Lacy, J., Lauer, J., O'Connor, C., Quon, D., Sim, G. K., & Efstratiadis, A. (1978) *Cell (Cambridge, Mass.) 15*, 687–701.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) in *Molecular Cloning*, pp 76–93, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Maxam, A. M., & Gilbert, W. (1980) *Methods Enzymol. 65*, 499–560.

Ninomiya, Y., & Olsen, B. R. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 3014–3018.

Rasheed, S., Nelson Rees, W. A., Toth, E. M., Arnstein, P., & Gardner, M. B. (1974) *Cancer 33*, 1027–1033.

Sandell, L. J., Yamada, Y., Dorfman, A., & Upholt, W. B. (1983) *J. Biol. Chem. 258*, 11617–11621.

Sandell, L. J., Prentice, A. L., Kravis, D., & Upholt, W. B. (1984) *J. Biol. Chem. 253*, 7826–7916.

Smith, B., Martin, G. R., Miller, E. J., Dorfman, A., & Swarm, R. (1975) *Arch. Biochem. Biophys. 166*, 181–186.

Southern, E. M. (1975) *J. Mol. Biol. 98*, 503–517.

Stanescu, R., Stanescu, V., & Maroteaux, P. (1982) *Collagen Relat. Res. 2*, 111–116.

Stocker, N., Cheah, K. S. E., Griffin, J. R., Grosveld, F. G., & Solomon, E. (1984) *J. Cell. Biochem. Suppl. 8B* (Abstr. 1573).

Strom, C. M., & Upholt, W. B. (1984) *Nucleic Acids Res. 12*, 1025–1038.

Ullrich, A. J., Shine, J., Chirgwin, R., Pictet, R., Tischer, E., Rutter, W. J., & Goodman, H. M. (1977) *Science (Washington, D.C.) 196*, 1313.

Vogeli, G., Avvedimento, E. V., Sullivan, M., Maizel, J. V., Lozano, G., Adams, S. L., Pastan, I., & de Crombrugghe, B. (1980) *Nucleic Acids Res. 8*, 1823–1837.

Von der Mark, H., Von der Mark, K., & Gay, S. (1976) *Dev. Biol. 48*, 237–249.

Von der Mark, K., Von der Mark, H., Timpl, R., & Trelstad, R. L. (1979) *Dev. Biol. 59*, 75–85.

Vuorio, E., Elima, K., Pulkkinen, J., & Viitanen, A. M. (1984) *FEBS Lett. 174*, 238–242.

Weiss, E. H., Cheah, K. S. E., Grosveld, F. K., Dahl, H., Solomon, E., & Flavell, R. A. (1982) *Nucleic Acids Res. 10*, 1981–1984.

Yamada, Y., Kühn, K., & de Crombrugghe, B. (1983) *Nucleic Acids Res. 11*, 2733–2744.

Young, M. F., Vogeli, G., Nunez, A. M., Fernandez, M. P., Sullivan, M., & Sobel, M. E. (1984) *Nucleic Acids Res. 12*, 4207–4228.