0022-202X/84/8203-0265$02.00/0
THE JOURNAL OF INVESTIGATIVE DERMATOLOGY, 82:265-269, 1984
Vol. 82, No. 3
Copyright © 1984 by The Williams & Wilkins Co.
Printed in U.S.A.

Isolation and Characterization of a Human Pro α_2(I) Collagen Gene Segment*

SHINGO TAJIMA, M.D.,† JENNY P. Y. TING, PH.D., SHELDON R. PINNELL, M.D., AND RUSSEL E. KAUFMAN, M.D.

Division of Dermatology and Division of Hematology and Medical Oncology, Department of Medicine, Duke University Medical Center, Durham, North Carolina, U.S.A.

Over 20 kilobase pairs of the human pro α_2(I) collagen gene have been isolated and characterized by restriction endonuclease mapping, cell-free translation of hybrid-selected RNA, and DNA sequence analysis. We have sequenced an exon and determined its length to be 108 base pairs (bp). This is consistent with the organization of chick and sheep collagen genes in that exons are multiples of 9 bp in length, frequently being 54 and 108 bp. The sequenced exon was bordered by a GT (guanine-thymine) at its 3′ end and an AT (adenine-thymine) at its 5′ end. This pattern has been found at all normal intron–exon junctions in eukaryotic cells. The amino acid sequence derived from DNA sequencing of this 108 bp exon revealed 88% homology compared to the amino acid sequence of bovine pro α_2(I). The bases encoded 12 Gly-X-Y triplets characteristic of the helical portion of collagen. A unique sequence Gly-Gly-Lys-Gly-Glu-Lys identified this fragment as α_2(I) collagen.

Collagen is a major component of connective tissue and at least 5 genetically distinct types of collagen have been identified [1]. Type I collagen is the most common of all types of collagen and comprises more than three-quarters of the collagen content of the body. Type I procollagen, its precursor, is composed of two separate polypeptide chains designated pro α_1(I) and pro α_2(I).

Recently it has been possible to isolate the entire chick [2, 3] and a portion of the sheep [4] and human [5] pro α_2(I) collagen genes. These genes have generated much interest because of their unique structural organization. First, they are the longest gene isolated to date, with 5 kilobase pairs (kb) of coding information distributed over 38 kb of DNA which includes almost 50 intervening sequences [3]. Secondly, coding blocks for the helical region range in size from 45 to 108 base pairs (bp) in multiples of 9 bp with 54 bp being the most common length. Carboxyterminal exons range in size from 189 to 444 bp [2]. While a cDNA representing the 3′ end of the human pro α_2(I) collagen message has been described [6,7], little is known about the general structural organization of the chromosomal gene. A portion of a human gene which may represent the human pro α_1(I) has recently been cloned into a cosmid [8]. This fragment has organizational structure similar to other vertebrate collagen genes.

In this paper we describe the isolation and partial characterization of an extensive region of the human pro α_2(I) collagen gene, including the precise location of several coding blocks.

MATERIALS AND METHODS

Materials

Recombinant bacteriophage DNA (SpC3) for sheep pro α_2(I) collagen was kindly provided by Dr. R. Crystal (N.I.H.). Nitrocellulose papers (BA 85) and DBM filters were from Schleicher and Schuell. Proline-free rabbit reticulocyte lysate was purchased from Bethesda Research Laboratories. Clostridial collagenase (form III) was from Advanced Biofactures. L-[2,3,4,5,^{3}H]proline and nucleotide triphosphates were from Amersham. The cDNA for chick pro α_1(I) (pCg54) and pro α_2(I) (pCg45) were obtained from Dr. H. Boedtker [9,10]. Restriction endonucleases and modifying enzymes were obtained from Bethesda Research Laboratories and used under conditions recommended by them.

Preparation and Screening of a Human Bacteriophage Library

High-molecular-weight DNA was isolated from human splenic tissue using the method of Blin and Stafford [11]. DNA was partially digested with Eco RI and size-fractionated on a 10–40% sucrose gradient. Fragments from 10 to 20 kb pairs were ligated to purified Eco RI arms of Charon 4A [12] and packaged into phage heads using in vitro packaging reagents [13].

The sheep DNA recombinant (SpC3) containing pro α_2(I) collagen gene sequences was digested with Eco RI and each purified Eco RI fragment was subcloned into plasmid pBR322 [14]. A plasmid recombinant containing the sheep 2.3 kb Eco RI collagen gene fragment was employed as a probe for screening the human bacteriophage library since it was known to have collagen coding sequences [4], and produced only 2 bands when hybridized to human genomic Eco RI-digested DNA. The fragment was labeled by "nick" translation [15,16]. Approximately 3×10^5 human bacteriophage recombinants were screened with the ^{32}P labeled 2.3 kb sheep collagen gene probe using the method of Benton and Davis [17]. A second screen of the human bacteriophage library was performed using the 5′ most Eco RI fragment of the first isolate to provide overlapping cloned regions.

Analysis of the Recombinant Clones

The isolated recombinant clones λHpro α_2(I)-1 and λHpro α_2(I)-2 were digested with Eco RI, and individual fragments purified by preparative agarose gel electrophoresis and electro-elution. Purified DNA fragments were subcloned into the plasmid pBR322. Restriction endonuclease mapping of each DNA fragment was determined by single or multiple enzyme digestion.

Coding sequences were identified by filter hybridization techniques using chick cDNA or sheep genomic DNA probes. Bacteriophage recombinant DNA fragments in agarose gels were transferred to nitrocellulose filters by a modified Southern technique [18,19], and hybridization of ^{32}P-labeled probe to filters containing DNA was essentially as described by Jeffreys and Flavell [20].

In order to identify moderately repetitive DNA sequences we hybridized ^{32}P-labeled human genomic DNA to λHpro α_2(I)-1 or λHpro α_2(I)-2 which had been digested with various enzymes as described above and transferred to nitrocellose membranes.

Hybridization Selected mRNA Translation

In order to verify the identity of the human collagen recombinant, we purified by hybrid selection the mRNA which had strong homology

Manuscript received June 27, 1983; accepted for publication October 24, 1983.

This work was supported by National Institutes of Health Research Grants AM17128 and AM28304. Dr. Tajima was supported in part by a Dermatology Foundation Fellowship. Dr. Kaufman was supported in part by the Searle Scholar Program. This is publication No. 134 from the Dermatology Research Laboratory at Duke University Medical Center.

* A portion of this work has been presented: Tajima S et al, Clin Res 30:611A, 1982.

† Present address: Department of Dermatology, Keio University School of Medicine, Shinjuku, Tokyo 160, Japan.

Abbreviations:
bp: base pair(s)
kb: kilobase(s)
SDS: sodium dodecyl sulfate

with λHpro α_2(I)-1. Total RNA was isolated from cultured human skin fibroblasts using 8 M guanidine hydrochloride [21,22]. Two hundred micrograms of total fibroblast RNA was hybridized with 10 μg of purified λH Pro α_2(I)-1 which had been covalently bound to a DBM filter. The hybridization reaction was performed in 50% formamide, 1 M NaCl, 0.01 M EDTA, 0.2% sodium dodecyl sulfate (SDS) and 50 μg/ml tRNA at 37°C for 24 h [23]. After incubation the filter was washed extensively in the hybridization buffer without RNA and all filter-bound RNA was eluted with 95% formamide with 3 washes. RNA was phenol extracted and precipitated with ethanol. The eluted RNA was used to direct the synthesis of polypeptides in a rabbit reticulocyte lysate cell-free translation system using [^{3}H]proline as a radioactive precursor [24].

Cell-free translation products were analyzed by 7.5% SDS polyacrylamide gel electrophoresis and subsequent fluorography. Bacterial collagenase digestion of cell-free translation products was performed as described previously [25].

DNA Sequence Analysis

The DNA sequence was determined by dideoxy termination of primed-chain synthesis [26]. Collagen coding sequences were first mapped to the 650 bp Eco RI-Pst I fragment of the 2.8^A kb fragment by filter hybridization. This fragment was purified by cloning into the appropriate sites of M13, mp9 [27]. Single-strand M13 recombinant phage DNA was then used as a template for synthesis of the complementary strand which was terminated at specific bases by dideoxynucleotide triphosphate incorporation. The sequence was verified by sequence analysis of the fragment after it was inserted in M-13 in the opposite direction.

RESULTS

Isolation of the Initial Recombinant

The 2.3 kb Eco RI fragment from the sheep collagen gene recombinant SPC-3 (R. Crystal, N.I.H.) was purified by subcloning into pBR322. The insert from this plasmid recombinant was labeled by nick translation and used as a molecular hybridization probe to screen a Charon 4A human bacteriophage library prepared from a patient who had no known collagen disorders. The initial screen produced a bacteriophage recombinant [λHpro α_2(I)-1] containing 10.8 kb of insert DNA and 4 Eco RI fragments [4.0 kb, 2.8 kb (doublet), and 1.2 kb]. Their relative order was determined by partial Eco RI digestion. The fragments were purified by subcloning into pBR322 and their detailed restriction endonuclease maps were determined from these recombinants (Fig 1). Two plasmid recombinants containing 2.8 kb Eco RI inserts were demonstrated to have markedly different restriction endonuclease maps, confirming that the 2.8 kb band was a doublet.

After ordering the fragments by partial restriction endonuclease mapping, the 2.8^A kb fragment was used to rescreen the library. An additional recombinant [λHpro α_2(I)-2] was isolated which extended the cloned DNA an additional 9.9 kb 5′ to the 2.8^A kb fragment (Fig 1). The 2.8^A kb fragment was shared by both recombinants. The total cloned DNA was 20.7 kb in length.

To confirm that the restriction endonuclease map of the cloned DNA was identical with genomic DNA, we used the cloned DNA Eco RI fragments as probes against endonuclease-digested human DNA in Southern blotting experiments. The restriction endonuclease map of the chromosomal gene corresponded to that derived for the cloned fragments with the exception of the 2.8^B kb probe. In this case only a smear was seen on the radioautograph, characteristic of patterns seen when moderately repetitive DNA is used as a probe [28].

To localize the region of the repetitive DNA, we digested the plasmid recombinant containing the 2.8^B kb Eco RI fragment with a number of restriction endonucleases and analyzed the size-fractionated fragments by Southern blotting using total genomic DNA labeled with P-32 as a probe. By hybridizing this probe to cloned DNA in 0.45 M NaCl at 68°C, we have previously demonstrated that only moderately to highly repetitive DNA fragments will produce a strong signal on the radioauto-

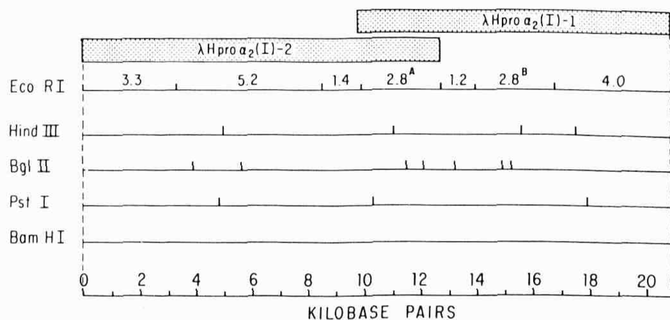

FIG 1. Restriction endonuclease map of λHpro α_2(I)-1 and λHpro α_2(I)-2. The restriction endonuclease sites for the overlapping recombinants are presented. Eco RI fragment *2.8^A* is common to both recombinants and contains the Eco RI-PsT I fragment which was sequenced. Fragment *2.8^B* contains moderately repetitive DNA which hybridizes to an Alu I sequence probe.

graph after an overnight exposure to radiographic film [29]. This experiment demonstrated strong hybridization to a 600 bp Bgl II-Hind III fragment. No significant signal was produced when the same probe was hybridized to the other Eco RI fragments derived from λHpro α_2(I)-1 or λHpro α_2(I)-2. We demonstrated that this repeated sequence belongs to the Alu I family of repeated DNA sequences. An Alu I sequence probe derived from a segment of DNA 4 kb 5′ to the human δ globin gene [29] hybridized strongly to the 2.8^B kb Eco RI fragment (data not shown).

Confirmation of the Identity of λHpro α_2(I)-1

Although the original screen of the human bacteriophage library was performed under conditions of high stringency, we could not be certain that the selected recombinant contained sequences coding for α_2(I) collagen. Myers et al [6] have recently reported the size of Eco RI fragments of genomic DNA which hybridize to their pro α_2(I) cDNA to include 4.0 kb and 2.8 kb Eco RI fragments. These are identical in size to fragments present in our recombinant λHpro α_2(I)-1.

In addition, we hybridized ^{32}P-labeled chick pro α_1(I) (pCG 54) and pro α_2(I) (pCG 45) cloned cDNAs to an Eco RI digest of λHpro α_2(I)-1. Only the pro α_2(1) cDNA, which contained sequences at the 3′ end of the α_2 mRNA produced a strong signal, hybridizing to the 4.0 and 2.8 kb fragments; only a faint signal was produced when the pCG 45 α_2(I) probe was hybridized to the purified 2.8^A kb fragment. The sheep 2.3 kb Eco RI fragment which had been used to screen the library and contained coding sequences from the middle of the sheep collagen gene hybridized only to the 2.8^A and 1.4 kb Eco RI fragments in λHpro α_2(I)-2. Thus the 5′ to 3′ orientation was determined and was confirmed by DNA sequence analysis.

To further establish the identity of this recombinant, we determined its ability to select by hybridization a unique species of mRNA. Ten micrograms of denatured λHpro α_2(I)-1 DNA was covalently bound to a DBM filter and hybridized to 200 μg of total fibroblast RNA. RNA remaining after medium stringency washes with 1 × SSC (0.15 NaCl, 0.015 M Na citrate, pH 7.5) at 65°C was eluted with 95% formamide, and used to direct the synthesis of proteins in a rabbit reticulocyte lysate system using [^{3}H]proline as the labeled amino acid. Analysis of the translation products of the hybrid-selected and total fibroblast RNA on an SDS-polyacrylamide gel demonstrated that the hybrid-selected RNA directed the synthesis only of a protein which comigrated with the α_2(I) procollagen band produced by total fibroblast RNA (Fig 2). No pro α_1(I) band was seen, whereas this was readily apparent when total fibroblast RNA was translated. Furthermore, treatment with collagenase destroyed the translation product representing pro α_2(I). Thus the bacteriophage recombinant λHpro α_2(I)-1 has homology

primarily to an RNA species which directs the synthesis of pro α_2(I) procollagen in the rabbit reticulocyte lysate system.

We attempted to characterize the RNA which was eluted from the λHpro α_2(I) DBM filter. Total fibroblast RNA and eluted RNA were fractionated on a methyl mercuric hydroxide agarose gel and transferred to a DBM filter. The 2.8^A kb Eco RI fragment was labeled and hybridized to the filter. The radioautograph of the filter demonstrated a consistent pattern of hybridization (Fig 3). A single intense band representing RNA of 5900 nucleotides was seen in the hybrid-selected RNA lane, a smear below this band representing degraded RNA. Two closely migrating bands were observed in the lane containing total fibroblast RNA while only a single band was seen in the lane containing poly A+ fibroblast RNA. These bands correspond to similar-sized RNA species identified by Myers et al [6]. We see an additional high-molecular-weight band in some preparations of hybrid-selected RNA which may represent a collagen precursor RNA species.

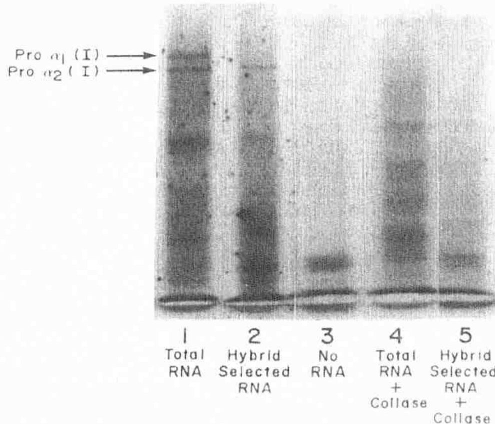

FIG 2. Fluorogram of SDS polyacrylamide gel electrophoresis of cell-free translation products synthesized by RNA hybridized to recombinant DNA clone λHPro α_2(I)-1. *Lane 1*, total RNA before hybridization; *lane 2*, RNA after hybridization to the recombinant λH Pro α_2(I)-1; *lane 3*, without RNA; *lane 4*, same as lane 1 except for digesting with bacterial collagenase (collase) prior to electrophoresis; *lane 5*, same as lane 2 except for digesting with bacterial collagenase prior to electrophoresis.

Sequence Determination of a Coding Block

We attempted to use the sheep 2.3 kb Eco RI fragment as a probe to identify coding sequences in the 2.8^A kb Eco RI fragment from which we hoped to obtain sequence information. We identified a 650 bp Eco RI-Pst I fragment at the 5′ end of the 2.8^A kb fragment which hybridized strongly to the sheep collagen probe. This fragment was purified, subcloned into M-13, mp9, and used as a template for dideoxy sequencing analysis. Eighty-three bases from the 5′ Eco RI site we identified a region with frequent G doublets. A continuous stretch of 108 bp was identified which contained no in-phase termination codons and encoded glycine at every third codon, including the first codon to produce the characteristic Gly-X-Y pattern (Fig 4). An unusual feature was the consecutive glycines at the 31st and 32nd codons in this coding block. No similar collagen-like sequence was identified immediately on either side of this coding block. However, consensus splice site signals were identified at the 5′ (AG) and 3′ (GT) ends of the presumed coding block. The consecutive sequence Gly-Gly-Lys-Gly-Glu-Lys is unique among known collagen sequences and precisely matches sequences 562–567 of bovine α_2(I) collagen (courtesy of Klaus Kuhn).

DISCUSSION

We have isolated and characterized a greater than 20 kb pair region of human DNA which has an organization typical of collagen genes. This DNA selectively hybridizes to RNA which directs the synthesis of pro α_2(I) collagen in a reticulocyte lysate translation system. We further demonstrated that this hybrid-selected RNA is almost 6000 bases in length, a value which agrees with determinations for pro α_2(I) mRNA reported by Myers et al [6].

Information gained from DNA sequencing strongly supports the identity of this gene fragment as pro α_2(I). The 108 bp exon is a characteristic of collagen genes and the repeating Gly-X-Y

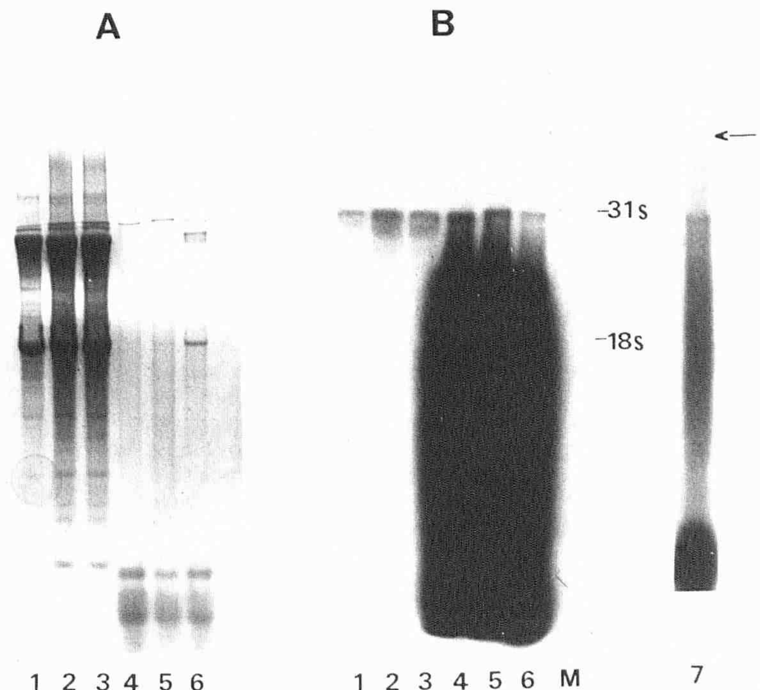

FIG 3. Analysis of hybrid-selected RNA. *Panel A* represents an ethidium bromide-stained 1.5% methyl mercuric hydroxide agarose gel containing RNA samples. Markers include 28s and 18s rRNA and LT-2 RNA, a 31s (6600 nucleotide) RNA virus mutant of VSV (30). *Lanes 1, 2, and 3* represent total fibroblast RNA. *Lanes 4, 5, and 6* represent hybrid selected fibroblast RNA. Lane 6 is contaminated with 18s and 28s rRNA which serves as internal size markers. A distinct band of RNA has been selected in lanes 4, 5, and 6. *Panel B* is a radioautograph of a filter to which the RNA from panel A has been transferred and hybridized with nick translated Eco RI fragment 2.8^A. *Lanes 1, 2, and 3* demonstrate hybridization to several bands including a 30s band. Single bands corresponding to the purified bands in panel A, *lanes 4, 5, and 6*, are seen in the corresponding lanes of panel B. Intense signal below these bands represents degraded hybrid selected collagen mRNA. *Lane 7* represents a radioautograph of poly A+ fibroblast RNA hybridized to the same probe. In many experiments, a very large RNA species is observed, but only in RNA derived from fibroblasts. This likely represents a precursor collagen RNA species.

```
GTGTTTGACTCAAG  GGT  GAA  CCT  GGT  GTG  GTT  GGT  GCT  GTG  |-Human
                GLY  GLU  PRO  GLY  VAL  VAL  GLY  ALA  VAL  |
                                                        PRO  |-Bovine α2(I)
                532  533  534  535  536  537  538  539  540  |

                GGC  ACT  GCT  GGT  CCA  TCT  GGT  CCT  AGT
                GLY  THR  ALA  GLY  PRO  SER  GLY  PRO  SER
                                         ALA
                541  542  543  544  545  546  547  548  549

                GGA  CTC  CCA  GGA  GAA  AGG  GGT  GCT  GCT
                GLY  LEU  PRO  GLY  GLU  ARG  GLY  ALA  ALA
                     ILE
                550  551  552  553  554  555  556  557  558

                GGC  ATA  CCT  GGA  GGC  AAG  GGA  GAA  AAG  GTACGTGTTGACCCTATTAC
                GLY  ILE  PRO  GLY  GLY  LYS  GLY  GLU  LYS
                     LEU
                559  560  561  562  563  564  565  566  567
```

FIG 4. DNA sequence of a coding block in fragment 2.8^A (see text). *Line 1* is the DNA sequence of a 142 bp segment from the 5′ end of fragment 2.8^A. *Line 2* represents the amino acid sequence derived from the DNA sequence. *Line 3* presents those amino acids in bovine α_2(I) collagen which are *different* from the human sequence. The bovine amino acids are numbered on the bottom line.

sequence which it encodes is a prerequisite of helical collagen sequences.

In order to precisely identify the sequence we systematically computer searched known collagen sequence information (courtesy of Klaus Kuhn). The unique run Gly-Gly-Lys-Gly-Glu-Lys was found only in an α_2(I) collagen chain and identified a region of close homology to the bovine α_2(I) collagen chain between residues 532–567 (see Fig 4). All polar residues were identical and other substitutions were conservative.

The sequenced fragment is located in the middle of our cloned DNA segment and likewise encodes a region near the middle of the procollagen chain. We can assume that much of the triple helical region is present in the cloned DNA fragments and that the restriction endonuclease map presented here may serve as a standard for comparing other normal and defective pro α_2(I) collagen genes. Furthermore since only single bands were seen on the Southern DNA blotting analysis of this patient's DNA using the cloned fragments as probes, both chromosomes are likely identical with regard to the restriction endonuclease sites reported. While it remains possible that this region is derived from a region which is a collagen pseudogene, further characterization and extensive DNA sequence analysis would be required to completely eliminate this possibility.

Two additional sources of information help establish the identity and relative location of our recombinant to the entire α_2(I) gene. Dalgleish et al have recently reported the isolation and characterization of a segment of DNA which includes sequences encoding the 3′ end of the human α_2(I) collagen gene [5]. We have been able to compare λHpro α_2(I)-1 and their recombinant (HpCl). Both recombinants contain 4 kb Eco RI fragments which cross-hybridize in Southern blotting analysis. Also, the restriction endonuclease map of our 4.0 kb fragment is identical to the published map of the 4.0 kb fragment in HpCl. Thus our recombinants appear to overlap and provide a restriction endonuclease map covering over 30,000 bp of DNA.

We also have compared our DNA sequence data with the recently published sequence of a human pro α_2(I) cDNA Hf-32 [7]. Our sequence data extends an additional 3 bases beyond the 5′ end of their cDNA (positions −1442 to −1340) and has the same amino acid designation. We note a single base difference G→A in the third position of codon 554. This change would not result in a change in the amino acid at that position and may represent a polymorphism. Furthermore we have identified a coding block 3′ to the 108 bp coding block. The sequence of this block is identical to the cDNA beginning at their position −1286. These coding blocks bracket a 54 bp coding block defined by the region from −1339 to −1287. Thus at least 3 coding blocks are present in the Eco RI fragment 2.8^A.

REFERENCES

1. Pinnell SR, Murad S: Disorders of collagen, The Metabolic Basis of Inherited Disease, 5th ed. Edited by JB Stanbury, JB Wyngaarden, DS Frederickson. New York, McGraw-Hill Book, 1982, pp 1425–1449
2. Wozney J, Hanahan D, Tate V, Boedtker H, Doty P: Structure of the pro α_2(I) collagen gene. Nature 294:129–135, 1981
3. Yamada Y, Avvedimento VE, Mudryi M, Ohkubo H, Vogeli G, Irani M, Pastan I, de Crombrugghe B: The collagen gene: evidence for its evolutionary assembly by amplification of a DNA segment containing on exon of 54 bp. Cell 22:887–892, 1980
4. Boyd CD, Tolstoshev P, Schafer MP, Trapnell BC, Coon HC, Kretschmer PJ, Nienhuis AW, Crystal RG: Isolation and characterization of a 15-kilobase genomic sequence coding for part of the Pro α_2 chain of sheep type I collagen. J Biol Chem 255:3212–3220, 1980
5. Dalgleish R, Trapnell BC, Crystal RG, Tolstoshev P: Copy number of a human type I α_2 collagen gene. J Biol Chem 257:13816–13822, 1982
6. Myers JC, Chu M-L, Faro SH, Clark WJ, Prockop DJ, Ramirez F: Cloning a cDNA for the pro-α_2 chain of human type I collagen. Proc Natl Acad Sci USA 78:3516–3520, 1981
7. Bernard MP, Myers JC, Chu M-L, Ramirez F, Eikenberry EF, Prockop DJ: Structure of a cDNA for the Pro α_2 chain of human type I procollagen. Comparison with chick cDNA for Pro α_2(I) identifies structurally conserved features of the protein and the gene. Biochemistry 22:1139–1145, 1983
8. Weiss EH, Cheah KSS, Grosveld FG, Dahl HHM, Solomon H, Flavell RA: Isolation and characterization of a human collagen α_1(I)-like gene from a cosmid library. Nucleic Acids Res 10:1981–1994, 1982
9. Lehrach H, Frischauf AM, Hanahon D, Wozney J, Fuller F, Erkuenjakow R, Boedtker H, Doty P: Construction and characterization of a 2.5 kilobase procollagen clone. Proc Natl Acad Sci USA 75:5417–5421, 1978
10. Fuller F, Boedtker H: Sequence determination and analysis of the 3′ region of chicken Pro-α_1(I) and Pro-α_2(I) collagen messenger ribonucleic acids including the carboxy-terminal propeptide sequences. Biochemistry 20:996–1006, 1981
11. Blin N, Stafford D: A general method of isolation of high molecular weight DNA from eukaryotes. Nucleic Acids Res 3:2303–2308, 1976
12. Blattner FR, Blechl AE, Denniston-Thompson K, Farber HE, Richards JE, Slightom JL, Tucker PW, Smithies O: Cloning human fetal γ globin and mouse α-type globin DNA: preparation and screening of shotgun collections. Science 202:1279–1284, 1978
13. Hohn B: In vitro packaging of lambda and cosmid DNA, Methods in Enzymology, vol 68. Edited by R Wu. New York, Academic Press, 1979, pp 299–309
14. Bolivar F, Rodriguez RL, Greene RJ, Betlach MC, Heynecker HL, Boyer HW: Construction and characterization of new cloning vehicles. Gene 2:95–113, 1977
15. Maniatis T, Jeffrey A, Kleid DG: Nucleotide sequence of the rightward operator of λ phage. Proc Natl Acad Sci USA 72:1184–1188, 1975
16. Benz EJ, Kretschmer PJ, Geist CE, Kantor JA, Turner PA, Nienhuis AW: Hemoglobin switching in sheep. J Biol Chem 254:6880–6888, 1979
17. Benton WD, Davis RW: Screening λgt recombinant clones by hybridization to single plaques in situ. Science 196:180–182, 1977
18. Southern EM: Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98:503–517, 1975
19. Tam JWO, Kaufman RE, Nienhuis AW: Analysis of globin gene structure in patients with β thalassemia by restriction endonuclease mapping. Hemoglobin 5:209–215, 1981
20. Jeffreys AJ, Flavell RA: A physical map of the DNA regions flanking the rabbit β-globin gene. Cell 12:429–439, 1977
21. Adams SL, Sobel ME, Howard BH, Olden K, Yamada KM, de Crombrugghe B, Pastan I: Levels of translatable mRNAs for cell surface protein, collagen precursors, and two membrane proteins are altered in Rous sarcoma virus-transformed chick embryo fibroblasts. Proc Natl Acad Sci USA 74:3399–3403, 1977
22. Tolstoshev P, Berg RA, Rennard SI, Bradley KH, Trapnell BC, Crystal RG: Procollagen production and procollagen messenger RNA levels and activity in human lung fibroblasts during periods of rapid and stationary growth. J Biol Chem 256:3135–3140, 1981
23. Adams SL, Alwine JC, de Crombrugghe B, Pastan I: Use of recombinant plasmids to characterize collagen RNAs in normal and transformed chick embryo fibroblasts. J Biol Chem 254:4935–4938, 1979
24. Tajima S, Pinnell SR: Regulation of collagen synthesis by ascorbic acid. Ascorbic acid increases type I procollagen mRNA. Biochem Biophys Res Commun 106:632–637, 1982

25. Freiberger H, Grove D, Sivarajah A, Pinnell SR: Procollagen I synthesis in human skin fibroblasts: effect of culture conditions on biosynthesis. J Invest Dermatol 75:425–430, 1980
26. Sanger F, Nicklen S, Coulson AR: DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467, 1977
27. Heidecker G, Messing J, Gronenborn B: A versatile primer for DNA sequencing in the M13mp2 cloning system. Gene 10:69–73, 1980
28. Kaufman RE, Kretschmer PJ, Adams JW, Coon HC, Anderson WF, Nienhuis AW: Cloning and characterization of DNA sequences surrounding the human γ-, δ-, and β-globin genes. Proc Natl Acad Sci USA 77:4229–4233, 1980
29. Adams JW, Kaufman RE, Kretschmer PJ, Harison M, Nienhuis AW: A family of long reiterated DNA sequences, one copy of which is next to the human beta globin gene. Nucleic Acids Res 8:6113–6128, 1980
30. Keene JD, Chien IM, Lazzarini RA: Vesicular stomatitis virus defective interfering particle containing a muted internal leader RNA gene. Proc Natl Acad Sci USA 78:2090–2094, 1981