Proc. Natl. Acad. Sci. USA
Vol. 75, No. 11, pp. 5417–5421, November 1978
Biochemistry

Construction and characterization of a 2.5-kilobase procollagen clone

(recombinant DNA/chicken embryo calvaria/double-stranded cDNA/restriction mapping/DNA sequence determination)

HANS LEHRACH*†, ANNA MARIA FRISCHAUF*†, DOUGLAS HANAHAN*, JOHN WOZNEY*, FORREST FULLER*, RADOMIR CRKVENJAKOV*‡, HELGA BOEDTKER*, AND PAUL DOTY*

* Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138

Contributed by Paul Doty, August 24, 1978

ABSTRACT **Recombinant bacterial plasmids have been constructed by inserting double-stranded chicken procollagen cDNA sequences linked to chemically synthesized decanucleotides containing *Hin*dIII sites into the *Hin*dIII site of pBR322. After transformation of *Escherichia coli* χ1776, colonies were selected by ampicillin resistance and recombinants containing procollagen sequences were identified by colony hybridization to ^{32}P-labeled procollagen cDNA. The inserts from three recombinant plasmids, pCg10, pCg13, and pCg45, were 1200, 2200, and 2550 base pairs long, respectively. Their sequence homology has been established by restriction mapping and crosshybridization of nick-translated plasmids to Southern blots of *Hpa* II fragments of the inserts. pCg45 has been positively identified as containing the pro α2 collagen sequence by partial determination of the DNA sequence of its ends: it has a short thymine-rich sequence at one end and a sequence coding for residues 478–499 in the chicken α2 chain at the other end.**

The analysis of the regulatory mechanisms that ensure the tightly controlled tissue and developmental stage-specific expression of the different collagen genes (1) will be essential for the understanding of developmental processes in eukaryotes. As an initial contribution to such an analysis we have isolated a highly purified mixture of the two type I procollagen mRNAs (2) and have prepared complementary DNA sequences to serve as hybridization probes for procollagen mRNA and gene sequences (3). Many experiments, however, require a far larger amount of highly purified procollagen sequences than can be obtained by mRNA isolation. To obviate this difficulty, we have used recombinant DNA technology to amplify procollagen cDNA sequences in *Escherichia coli*. Clones containing procollagen cDNA sequences were obtained using modifications of the method used by Goodman and his colleagues to clone rat proinsulin cDNA (4). We present here a description of three pro α2 clones characterized thus far.

MATERIALS AND METHODS

Construction of Recombinant Plasmids. Procollagen mRNAs, purified by binding embryonic chicken calvaria poly(A)-containing RNA to Sepharose 4B as described (3), served as template for the synthesis of the cDNA used for cloning. Double-stranded cDNA was synthesized and ligated to chemically synthesized decanucleotides containing *Hin*dIII sites (unpublished data). In brief, double-stranded cDNA was prepared by a procedure in which single-stranded cDNA is not isolated prior to second strand synthesis (5). It was converted into perfect duplex molecules with blunt ends by digestion with single-strand nuclease S1 from *Aspergillus oryzae* [purified from amylase (Sigma) as described by Vogt (6)] and then incubation with *E. coli* DNA polymerase I, a gift of Yvonne Chow, as described by Seeburg *et al.* (7). The product of the S1–DNA polymerase reaction was then ligated to *Hin*dIII linkers (a gift from Richard Scheller) that had been end labeled with ^{32}P by polynucleotide kinase as described by Heyneker *et al.* (8) with T4 DNA ligase prepared by the method of Panet *et al.* (9). After digestion with *Hin*dIII, the ligated double-stranded cDNA was separated from unligated linker by sedimentation on a 5–20% sucrose gradient for 8 hr at 50,000 rpm in a Beckman SW 56 rotor. Fractions of large, medium, and small size were pooled.

pBR322 (obtained from H. Boyer) was purified by CsCl and sucrose gradient centrifugation, and digested with *Hin*dIII. After phenol extraction and ethanol precipitation, the DNA was incubated with 1 unit of calf intestine alkaline phosphatase (Boehringer Mannheim, Indianapolis, IN) to remove the terminal phosphates and prevent self-ligation. The pBR322 DNA that had been cut by *Hin*dIII and treated with phosphatase was then ligated to each of the size fractions of double-stranded cDNA linked to *Hin*dIII half-sites (10).

Transformation and Identification of Recombinant Clones. Both the construction of chimeric plasmids just described and the transformation of *E. coli* χ1776 by these plasmids was performed in a P3 physical containment laboratory at Cold Spring Harbor, New York, in compliance with the National Institutes of Health guidelines for recombinant research.§

χ1776 was transformed by a transfection procedure described by Villa-Komaroff *et al.* (11) as follows: 200-ml cultures of χ1776 were grown to an optical density at 600 nm of 0.2/ml, centrifuged down, resuspended in 20 ml of 10 mM NaCl, repelleted, and resuspended in 10 ml of 70 mM $MnCl_2$/40 mM Na acetate, pH 5.6/30 mM $CaCl_2$. A 0.2-ml portion of the suspension was then added to sterile tubes containing 25 or 100 ng of ligated plasmid. Transformed colonies containing procollagen sequences were identified by colony hybridization by a modification of the procedure of Grunstein and Hogness (12) which makes it possible to screen plates containing several thousand colonies. The colonies were screened with ^{32}P-labeled procollagen cDNA that was over 80% pure (3). Strongly hybridizing colonies were selected and replated, single colonies were picked and grown in liquid culture, and the DNA was isolated by the procedure developed by Curtiss *et al.* (13). Cleared lysates were extracted with phenol, precipitated with ethanol, and freed from contaminating chromosomal DNA by banding in CsCl. A typical yield was 100 μg of recombinant plasmid DNA per liter of culture.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviation: bp, base pair.

† Present address: European Molecular Biology Laboratory, Postfach 10.2209, 6900 Heidelberg, Germany.

‡ Present address: Institute for Biological Research, 29 Novembra 142, Belgrade, Yugoslavia.

§ *Federal Register* (1976) **41**, 27902–27943.

Table 1. Size of *Hpa* II fragments of procollagen inserts (in base pairs)

pCg10	pCg13	pCg45
780*	950*	900*
300*	720	800*
	400*	720
115	115	115

* 5'-End-labeled fragments.

Restriction Mapping and Southern Gel Blotting. Inserts were released from recombinant plasmids by digestion of 10–180 μg of DNA with 100–400 units of *Hin*dIII prepared by the method of Old *et al.* (14) in 10 mM Tris·HCl, pH 7.6/5 mM $MgCl_2$/1 mM dithiothreitol/50 mM NaCl for 2 hr at 37°C. In some cases the digest was incubated for an additional 30 min with 1 unit of calf intestine alkaline phosphatase (Boehringer Mannheim, grade I). After alcohol precipitation, the inserts were separated from pBR322 by electrophoresis on 17-cm, 6% polyacrylamide slab gels and then extracted from the gel as described by Maxam and Gilbert (15). *Eco*RI, prepared by the method of Greene *et al.* (16), *Bam*HI and *Hae* III (New England Biolabs), *Hpa* II (Bethesda Research Laboratories), and *Hin*f (a gift of G. Sutcliffe) endonucleases were used to restrict the inserts. With *Hpa* II, *Bam*HI and *Eco*RI, the buffer was the same one used with *Hin*dIII. In *Hin*f digestions, the NaCl concentration was reduced to 20 mM; in *Hae* III digestions, 50 μg of bovine serum albumin (Sigma) was added per ml.

Inserts isolated as described above were restricted with *Hpa* II and the fragments were applied to one of two 1.6% agarose slab gels. After electrophoresis, the DNA was denatured and then transferred to nitrocellulose filters by minor modifications of Southern's procedure (17). pCg45 and pCg10 were nick translated (18) and then hybridized to separate nitrocellulose sheets for 16 hr at 65°C. After repeated washing, the nitrocellulose sheets were air-dried and autoradiographed.

DNA Sequence Determination. Inserts released as described above and treated with alkaline phosphatase were end-labeled with polynucleotide kinase. After digestion with *Hpa* II and *Eco*RI endonucleases, the resultant fragments were separated on an 8% polyacrylamide gel and the labeled ends were identified by autoradiography. They were then isolated from the gel and their sequences were determined as described by Maxam and Gilbert (15) with modifications made by Allan Maxam (personal communication).

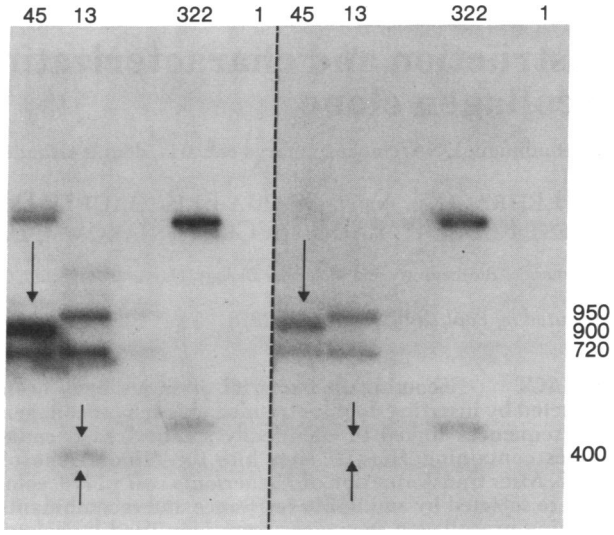

FIG. 2. Sequence homology of pCg10, pCg13, and pCg45 determined by Southern blot restriction mapping. (*Left*) Autoradiogram of nick-translated pCg45 plasmid hybridized to blots of *Hpa* II fragments of pCg45 insert (first lane) and of pCg13 insert (second lane) and to *Hin*f fragments of pBR322 (third lane). The nitrocellulose sheet was hybridized to 10^7 cpm of pCg45 in 8 ml. The fourth lane contains two *Hin*f fragments of the insert of pCg1, a procollagen clone that has no sequence homology to pCg45 (confirmed by the absence of any hybridization). (*Right*) Autoradiogram of nick-translated pCg10 plasmid (7×10^6 cpm) hybridized to blots containing the same fragments as on the *left*. Single arrow locates the 900- and 800-bp end fragments in the *Hpa* II digest of pCg45; double arrows locate the 400-bp fragment present in the *Hpa* II digest of pCg13.

RESULTS

Construction and Identification of cDNA Clones. Double-stranded procollagen cDNA, converted to perfect duplex molecules by treatment with S1 nuclease and *E. coli* DNA polymerase I, was ligated to chemically synthesized decanucleotides containing *Hin*dIII restriction sites and then digested

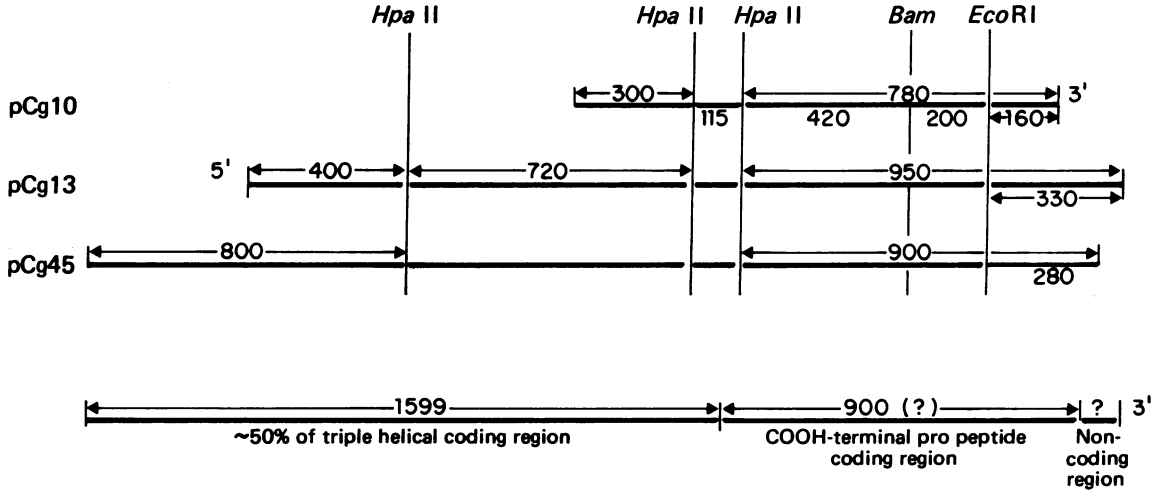

FIG. 1. Restriction map of pro α2 collagen clones. Relative orientation of fragments was determined by locating *Bam*HI or *Eco*RI site in the end fragment as indicated. Orientation of pro α2 collagen cDNA sequence relative to procollagen mRNA sequence was determined from primary sequence data given in Figs. 3 and 4.

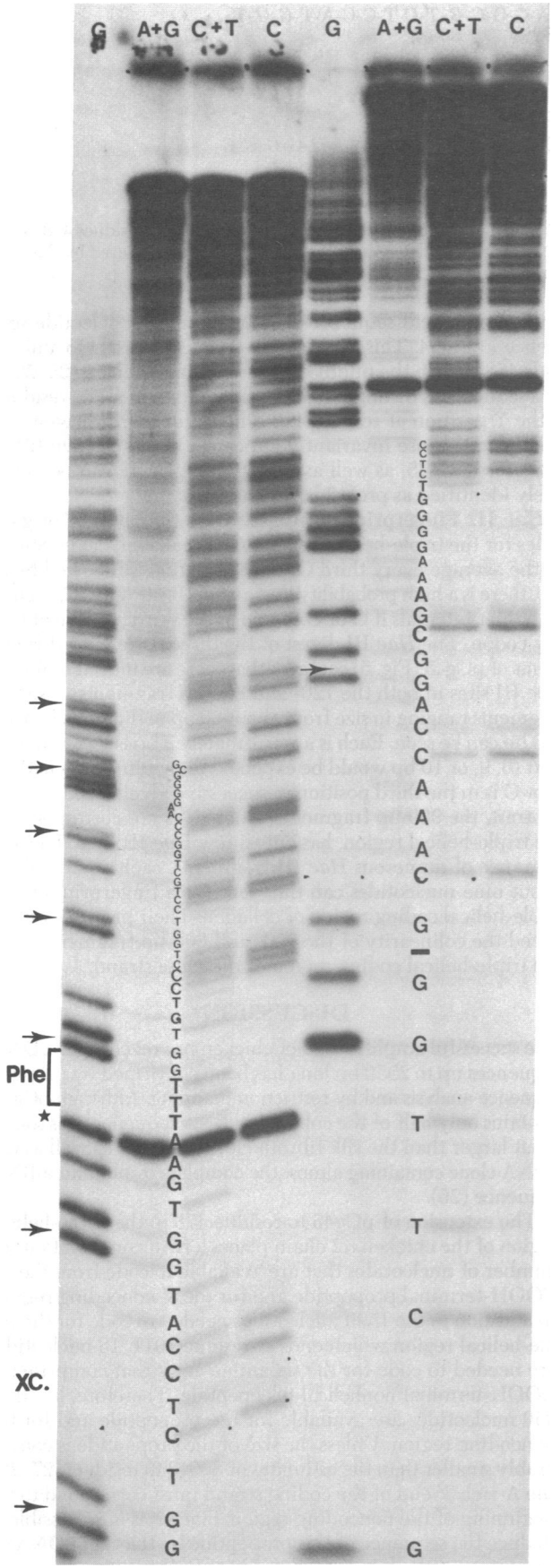

FIG. 3. Identification of pro α2 collagen coding end of pCg45 insert. Lanes are from left to right, in groups of four: *Hpa* II 800-bp end, first loading (long run); *Hpa* II 800-bp end, second loading (short run). Bar (▬) shows where *Hin*dIII linker joins cDNA. Arrows identify the G pairs corresponding to glycine codons. Star locates the position where a pair of adenines were unidentifiable because a nick in the DNA whose sequence was determined resulted in the appearance of an unreacted labeled fragment appearing as a heavily labeled band in all four lanes. The A doublet was identified in earlier sequence determinations of the *Hpa* II 800-bp end of pCg45 insert. XC, xylene cyanol.

with *Hin*dIII to create cohesive single-stranded ends. After fractionation on sucrose gradients, fractions of three sizes, which collectively covered a size range of 100–3000 base pairs (bp), were pooled. Each was inserted into pBR322 that had been cut at its single *Hin*dIII site and had had the 5′-terminal phosphates removed to prevent self-ligation. The resultant recombinant plasmids were then used to transform *E. coli* χ1776, and transformants containing procollagen cDNA sequences were identified by colony hybridization to ^{32}P-labeled procollagen cDNA that was over 80% pure (3). Although strongly hybridizing colonies were obtained from each transformation, we concentrated our efforts on characterizing inserts in clones resulting from transformation with the largest size fraction. The inserts from three of the largest clones, whose names were previously assigned by their location on nitrocellulose filters, are, in order of their size: pCg10, pCg13, and pCg45; their sizes are 1200, 2200, and 2550 bp, respectively.

Restriction Mapping and Southern Blotting of Procollagen Inserts. Restriction of the three inserts with *Hpa* II endonuclease resulted in three fragments from pCg10 and four fragments from pCg13 and pCg45. The sizes of the fragments are listed in Table 1. pCg45 and pCg13 share a 720-bp fragment, while all three have a 115-bp fragment in common. The end fragments of each were identified by labeling the 5′ ends with ^{32}P before restricting with *Hpa* II (Table 1).

pCg10, pCg13, and pCg45 were also restricted with *Eco*RI and *Bam*HI. All three were found to have a single *Eco*RI and *Bam*HI site near one end. Since the small *Bam*HI fragments were each 200 bp larger than the small *Eco*RI fragments, the two sites appeared to be separated by 200 bp. To confirm this, we digested all three plasmids with both *Eco*RI and *Bam*HI. Three fragments with identical mobility corresponding to an approximate size of 200 bp were produced (data not shown). A similar 200-bp *Bam*HI–*Eco*RI fragment has been reported to be located in double-stranded chicken calvaria procollagen cDNA, and a clone containing this sequence has been identified (19).

To determine the orientation of the ends of each of the inserts, we codigested them with *Hpa* II and either *Bam*HI or *Eco*RI to determine which *Hpa* II fragment contained the *Bam*HI or *Eco*RI site. As a result, a unique restriction map of the three inserts was deduced (Fig. 1).

To confirm the restriction maps of pCg10, pCg13, and pCg45, we electrophoresed *Hpa* II fragments of pCg45 and pCg13 inserts on two 1.6% agarose slab gels and blotted them onto nitrocellulose paper by the Southern method (17). One sheet was then hybridized to pCg45 while the other was hybridized to nick-translated pCg10. Fig. 2 shows the autoradiograms obtained after hybridization. pCg45 hybridized to each of its fragments as expected. It also hybridized to the 950-, 720-, and 400-bp fragments of the pCg13 insert and thus establishes the sequence homology of pCg45 and pCg13. Furthermore, it hybridized to the *Hin*f fragments of pBR322, as expected, since the intact recombinant plasmid was nick translated. The hybridization pattern of nick-translated pCg10 was similar but not identical to that of pCg45. In particular, pCg10 did not hybridize to the 800-bp fragment of pCg45 or

pG-C-T-T-G-G-*G-C*-A-A-A-C-C-A-*G-G-C*-G-A-A-A-G-G-*G-G-T*-C-T-C-C-A-T-*G-G-T*-

Gly - Lys - Hyp - Gly - Glu - Arg - Gly - Leu - His - Gly -
478

Hae III

G-A-A-T-T-T-*G-G-T*-G-T-C-C-C-T-*G-G-T*-C-C-T-G-C-T-*G-G-C*-C-C-A-A-G-G-*G-G-C*-

Glu - Phe - Gly - Val - Hyp - Gly - Pro - Ala - Gly - Pro - Arg - Gly -
488 498

FIG. 4. DNA sequence of the 5′ end of the sense strand of pCg45 and the corresponding amino acid sequence, identical to residues 478–499 of the chicken $\alpha 2$ collagen chain. The first nucleotides are the *Hin*dIII linker. Numbering of the amino acid residues follows that used by Fietzig and Kuhn (21), in which the first glycine of the triple helical region is numbered 1.

the 400-bp fragment of pCg13 as shown in Fig. 2, thereby confirming the restriction map in Fig. 1.

Primary Sequence of 5′ Ends of pCg45. Since pCg45 is 2550 bp long, one end of it was expected to contain part of the triple-helical coding region of the procollagen chain, and the sequences of both ends were therefore determined. The *Eco*RI end fragment had a very T-rich sequence following the *Hin*dIII site: $T_9GGT_8CT_3$. This could correspond either to the A-rich region commonly found at the 3′ end of the noncoding region or to the poly(A) end. Although the interspersion of bases other than A is more frequent than expected for a poly(A) end, silk fibroin mRNA has a poly(A) containing 17.5% of nucleotides other than A (20).

The nucleotide sequence of the 5′ end of the 800-bp *Hpa* II fragment is displayed in Fig. 3. The appearance of guanine pairs at regular intervals required by the $(Gly\text{-}X\text{-}Y)_n$ collagen sequence is immediately evident. This sequence corresponds to the amino acid sequence shown below the nucleotide sequence in Fig. 4. This in turn corresponds precisely to amino acids 478–499 in the middle of the chicken $\alpha 2$ chain (22, 23). This sequence is distinguished by the location of a Phe residue in the Y position at residue 489. Phe residues in collagen sequences otherwise invariantly appear in the X position (21). Therefore pCg45, as well as pCg13 and pCg10, can be definitely identified as pro $\alpha 2$ collagen clones.

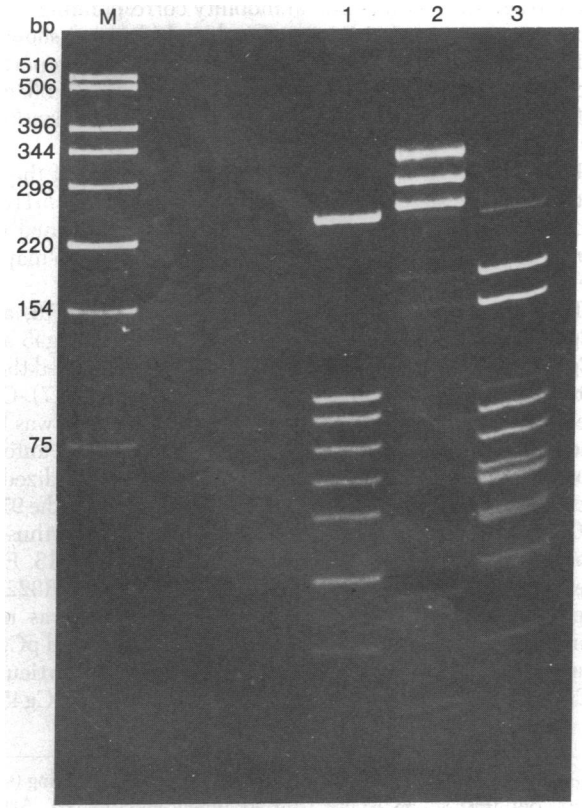

FIG. 5. Localization of frequently occurring GGCC sequences in pCg45 insert. *Hpa* II fragments of the pCg45 insert were digested with *Hae* III and electrophoresed on an 8% polyacrylamide slab gel. Lane M displays the *Hinf* fragments of pBR 322 used as size markers. *Hae* III digests of the *Hpa* II fragments are shown as follows: lane 1, digest of 720-bp *Hpa* II fragment; lane 2, digest of the 900-bp *Hpa* II fragment; and lane 3, digest of the 800-bp *Hpa* II fragment.

***Hae* III Fingerprint of pCg45.** More than half of pCg45 codes for the triple-helical region of the collagen chain. Since on the average every third Gly in this region is followed by a Pro, there is a high probability that a *Hae* III site, GGCC, occurs at regular intervals if G or C can be in the third position of the Gly codon. The *Hae* III digest of the three large *Hpa* II fragments of pCg45 (Fig. 5) confirms that there are indeed multiple *Hae* III sites in both the 720- and 800-bp fragments. A set of fragments ranging in size from about 30 bp to slightly less than 100 bp can be seen. Each is approximately 9 bp larger than the next (8, 9, or 10 bp would be expected depending on whether C or G is in the third position in successive glycine codons). In contrast, the 900-bp fragment, which does not correspond to the triple-helical region, has only a few *Hae* III sites. The appearance of numerous *Hae* III fragments, each separated by about nine nucleotides can thus serve as a fingerprint of the triple-helical coding region of collagen. Their presence established the colinearity of the 720- and 800-bp fragments with the triple-helical coding region of the sense strand.

DISCUSSION

The successful amplification of chicken pro $\alpha 2$ collagen cDNA sequences up to 2550 bp long has been confirmed by primary sequence analysis and by restriction mapping. Although pCg45 contains only half of the collagen mRNA sequence, it is somewhat larger than the silk fibroin clones (24, 25) as well as the cDNA clone containing almost the complete ovalbumin mRNA sequence (26).

The extension of pCg45 to residue 478 in the triple-helical region of the chicken $\alpha 2$ chain places a firm constraint on the number of nucleotides that are available to code from the $\alpha 2$ COOH-terminal propeptide and for the 3′ noncoding region. In addition to the 1599 nucleotides needed to code for the triple-helical region which ends at residue 1011, 18 nucleotides are needed to code for the six amino acids that comprise the COOH-terminal nonhelical teleopeptide. Therefore, less than 950 nucleotides are available for the propeptide and for the noncoding region. Unless the size of the propeptide is considerably smaller than the estimates of 300–340 residues (27, 28), the A-rich 3′ end of the coding strand must correspond to the beginning of the noncoding region. Since it is now possible to deduce the sequence of this propeptide by the rapid DNA sequencing methods, we will be able to find out its exact size as well as its sequence, which has not been determined so far.

The multiple *Hae* III sites found in the triple-helical coding region of the sense strand indicate there is no stringent selection

against C and/or G in the third position of the Gly codon, even when Gly is followed by Pro. Half of the Gly codons in the 66 nucleotides shown in Fig. 4 end in C. The very stable regions of secondary structure these GC regions are likely to generate (29) may explain why procollagen mRNAs are difficult to translate *in vitro*.

The availability of a sequence complementary to half of pro $\alpha 2$ collagen mRNA in unlimited amounts makes possible the study of the expression of type I collagen genes, as well as the screening for the pro $\alpha 2$ structural gene, and the study of primary transcripts of this gene in both tissues that make collagen and in tissues that do not.

While we have received the help of a large number of people, our deepest gratitude must be extended to James D. Watson and Joe Sambrook for letting us use the Cold Spring Harbor P3 Laboratory. For gifts of enzymes we are indebted to James Beard, Yvonne Chow, and Greg Sutcliffe. We thank Richard Scheller for his gift of the chemically synthesized decanucleotides. For helpful discussions, advice, and making available unpublished methods, we are indebted to Gray Crouse, Allan Maxam, Gary Buell, Marvin Wickens, and Lydia Villa-Komaroff. Finally we thank Tricia Bredbury for her help in RNA preparations and Doris Boger for her patience in typing this manuscript. This research was supported by the National Institutes of Health Grant HD-01229 and a grant from the Muscular Dystrophy Association, Inc.

1. Miller, E. J. (1976) *Mol. Cellular Biochem.* **13,** 165–192.
2. Boedtker, H., Frischauf, A. M. & Lehrach, H. (1976) *Biochemistry* **15,** 4765–4770.
3. Frischauf, A. M., Lehrach, H., Rosner, C. & Boedtker, H. (1978) *Biochemistry* **17,** 3243–3249.
4. Ullrich, A., Shine, J., Chirgwin, J., Pictet, R., Tischer, E., Rutter, W. J. & Goodman, H. M. (1977) *Science* **196,** 1313–1319.
5. Wickens, M. P., Buell, G. N. & Schimke, R. T. (1978) *J. Biol. Chem.* **253,** 2483–2495.
6. Vogt, V. M. (1973) *Eur. J. Biochem.* **33,** 192–200.
7. Seeburg, P. H., Shine, J., Martial, J. A., Baxter, J. P. & Goodman, H. M. (1977) *Nature (London)* **270,** 486–494.
8. Heyneker, H. L., Shine, J., Goodman, H. M., Boyer, H. W., Rosenberg, J., Dickerson, R. E., Narang, S. A., Itakura, K., Lin, S. & Riggs, A. D. (1976) *Nature (London)* **263,** 748–752.
9. Panet, A., van de Sande, J. H., Loewen, P. C., Khorana, H. G., Raae, A. J., Lillehaug, J. R. & Kleppe, J. (1977) *Biochemistry* **12,** 5045–5050.
10. Dugaiczyk, A., Boyer, H. W. & Goodman, H. M. (1975) *J. Mol. Biol.* **96,** 171–184.
11. Villa-Komaroff, L., Efstratiadis, A., Broome, S., Lomedico, P., Tizard, F., Naber, S. P., Chick, W. L. & Gilbert, W. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 3727–3731.
12. Grunstein, M. & Hogness, D. S. (1975) *Proc. Natl. Acad. Sci. USA* **72,** 3961–3965.
13. Curtiss, R., III, Inoue, M., Hsu, J. C., Alexander, L. & Rock, L. (1977) in *Molecular Cloning of Recombinant DNA*, eds. Scott, W. A. & Werner, R. (Academic, New York), pp. 90–114.
14. Old, R., Murray, K. & Roizes, G. (1975) *J. Mol. Biol.* **92,** 331–339.
15. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 560–564.
16. Greene, P. J., Betlach, M. C., Boyer, H. W. & Goodman, H. M. (1974) in *Methods in Molecular Biology*, ed. Wickner, R. B. (Dekker, New York), Vol. 7, p. 87.
17. Southern, E. M. (1975) *J. Mol. Biol.* **98,** 503–517.
18. Rigby, P. W. J., Dieckman, M., Rhodes, C. & Berg, P. (1977) *J. Mol. Biol.* **113,** 237–251.
19. Sobel, M. E., Yamamoto, T., Adams, S. L., Dilauro, R., de Crombrugghe, B. & Pastan, I. (1978) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **37,** 1408 (abstr. 767).
20. Lizardi, P. M., Williamson, R. & Brown, D. D. (1975) *Cell* **4,** 188–205.
21. Fietzek, P. P. & Kuhn, K. (1976) *Int. Rev. Connect. Tissues Res.* **7,** 1–60.
22. Dixit, S. N., Seyer, J. M. & Kang, A. H. (1977) *Eur. J. Biochem.* **73,** 213–221.
23. Dixit, S. N., Seyer, J. M. & Kang, A. H. (1977) *Eur. J. Biochem.* **81,** 599–607.
24. Morrow, J. F., Wozney, J. M. & Efstratiadis, A. (1977) in *Recombinant Molecules: Impact on Science and Society*, eds. Beers, R. F., Jr. & Bassett, E. G. (Raven, New York), pp. 407–417.
25. Morrow, J. F., Chang, N. T., Wozney, J. M., Richards, A. C. & Efstratiadis, A. (1977) in *Molecular Cloning of Recombinant DNA*, eds. Scott, W. A. & Wener, R. (Academic, New York), 161–171.
26. McReynolds, L., O'Malley, B. W., Nisbet, A. D., Fothergill, J. E., Givol, D., Fields, S., Robertson, M. & Brownlee, G. G. (1978) *Nature (London)* **273,** 723–728.
27. Fessler, L. I., Morris, N. P. & Fessler, J. H. (1975) *Proc. Natl. Acad. Sci. USA* **72,** 4905–4909.
28. Monson, J. M., Click, E. M. & Bornstein, P. (1975) *Biochemistry* **14,** 4088–4092.
29. Bachra, B. N. (1976) *J. Mol. Evol.* **8,** 155–173.