Molecular cloning and carboxyl-propeptide analysis of human type III procollagen

Helen R.Loidl¹, Jane M.Brinker¹, Mary May², Taina Pihlajaniemi³, Scott Morrow², Joel Rosenbloom² and Jeanne C.Myers^{1,4*}

¹Connective Tissue Research Institute and ²School of Dental Medicine, University of Pennsylvania, Philadelphia, PA 19104, ³Dept. Biochemistry, Rutgers Medical School, Piscataway, NJ 08854, and ⁴Dept. Medicine and Human Genetics, University of Pennsylvania, Philadelphia, PA 19104, USA

Received 10 September 1984; Accepted 14 November 1984

ABSTRACT

Two human cDNA libraries prepared from normal fibroblast (GM3348) and rhabdomyosarcoma (CCL136) mRNAs were screened under cross hybridization conditions with a genomic fragment coding for exons 2 and 3 of the avian type III procollagen COOH-propeptide (Yamada, Y., Mudryj, M., Sullivan, M. and deCrombrughe, B. (1983) *J. Biol. Chem.* 258, 2758-2761). One cDNA clone containing a 1.12 kb insert was isolated from the CCL136 library and later used to identify a GM3348 derived clone with a 2.4 kb insert. Comparison of the human and avian type III C-terminal propeptides revealed much more divergence in the first 54 amino acids following the terminal cysteine of the triple helical region than is present in the $\alpha 1(I)$ and $\alpha 2(I)$ procollagen chains of these species. Analysis of poly (A⁺) RNA from normal human fibroblast and tumor cell lines showed that they differed greatly in the relative amounts of $\alpha 1(I)$, $\alpha 2(I)$, and $\alpha 1(III)$ mRNAs. Furthermore, as we previously reported for the $\alpha 1(I)$ and $\alpha 2(I)$ transcripts, multiple mRNAs also hybridize to the cloned $\alpha 1(III)$ DNA.

INTRODUCTION

Five major and at least five minor types of collagen comprise the most abundant protein family found in vertebrates (reviews 1,2; and ref. 3-7). In general, collagen is synthesized in the form of procollagen molecules composed of three polypeptide chains consisting of a central triple helical domain with the repeating Gly-X-Y sequence, and shorter carboxy and amino propeptides. The different chains are the products of probably more than fifteen genes which have recently begun to be isolated from several species (reviews 8,9; and ref. 10-17). Determination of the DNA sequences of the genes and their encoded amino acid sequences are fundamental to understanding the function of the various types of collagen and the factors controlling their synthesis. Previously, human cDNA clones coding for the $\alpha 1$ and $\alpha 2$ chains of the type I procollagen heterotrimer (pro $\alpha 1$)₂ (pro $\alpha 2$) were identified (18,19) and used to isolate overlapping genomic fragments (20,21). Analysis of these two genes, which are located on different chromosomes (22), showed that the coding region is frequently interrupted

by about 50 sometimes very large introns (20,21). Although the mRNA transcripts of these genes are normally found at a two to one ratio reflecting the type I collagen subunit composition, the levels can vary in disease states (23).

To study the coordinate and differential expression of several procollagens, we constructed two cDNA libraries containing sequences corresponding to different types of procollagen in addition to type I. Here we report the isolation and characterization of cDNA clones coding for about half of the human $\alpha 1(\text{III})$ procollagen chain. From the DNA sequence of these clones, part of the triple helical and 3' untranslated region and the entire C-terminal propeptide was derived. While the telopeptide and adjacent 5' part of the COOH-propeptide of avian (24,25) and human type III are more divergent than seen in the $\alpha 1(\text{I})$ and $\alpha 2(\text{I})$ chains of these species (26-28), the major part of the COOH-propeptide is conserved. Comparison of this domain in three human procollagen chains also shows significant homology in the regions where the cysteine residues and glycosylation attachment site are located. Together with elucidating the still undetermined role of type III, these probes will be valuable in defining mutations and regulatory abnormalities in genetic and acquired diseases of connective tissue.

MATERIALS AND METHODS

Enzymes and Isotopes

Restriction endonucleases were purchased from New England Biolabs, DNA polymerase I and proteinase K from Boehringer Mannheim, T4 DNA ligase and rabbit reticulocyte lysate from Bethesda Research Laboratories. All enzymes were used according to manufacturers suggestions. Formamide was obtained from Fluka, oligo dT cellulose and primers for DNA sequencing from Collaborative Research, and nitrocellulose filters from Schleicher & Schuell. Labeled isotopes were purchased from Amersham. Cell line GM3348 was obtained from the Human Genetic Mutant Cell Repository, Camden, New Jersey, and CCL136 and HT1080 from the American Type Culture Collection, Rockville, Maryland.

RNA Purification and cDNA Cloning

Poly (A^+) RNA was isolated by oligo dT cellulose chromatography and size fractionated on a sucrose-SDS gradient (19). First strand DNA synthesis was performed in a 150 μl reaction containing 5 μg of mRNA, 0.7 μg oligo dT 12-18 primer, 1.30mM dGTP, .53mM TTP, .32 mM dCTP, .27mM dATP and 55 units/ml of AMV reverse transcriptase purified by DEAE and CM sepharose

chromatography (29). Second strand DNA was synthesized with the Klenow fragment of DNA polymerase I after heating the reaction mix at 100° for 2 minutes (30). Following incubation for 1 hr. at 38°, the reaction was passed through a G50 sephadex column and the peak fractions were ethanol precipitated. The double-stranded DNA (100ng) was treated with 3500 units/ml of S₁ nuclease (Miles) at 25° for 30 minutes, and then size fractionated on a 5-20% neutral sucrose-SDS gradient. The pool containing the largest DNA was dC-tailed for 6 minutes at 11° with terminal transferase. 3ng of DNA was annealed with 5ng of pBR322 which had been cleaved with PstI and dG-tailed (New England Nuclear). The hybrid molecules were used to transform Escherichia coli strain MC1061 (31).

Screening of the Bacterial Colonies

The clones were replica plated onto nitrocellulose filters, grown for four hours and then the plasmid DNA was amplified by overnight incubation of the filters at 37° on Luria plates containing 12.5ug/ml tetracycline and 150 ug/ml chloremphenicol. The filters were processed (30), baked at 76° for 2 hours, and hybridized under low stringency (35% formamide, 6xSSPE, 34° for 40 hours) or high stringency (50% formamide, 5xSSPE, 40° for 20 hours). Final washing conditions were 1xSSC, 55° or .2xSSC, 65° respectively. The DNA probe used for isolating the initial human type III clone was 10ng/ml of the insert from a subcloned genomic fragment containing exons 2 and 3 coding for part of the COOH-propeptide of avian type III procollagen (24,25) provided by Y. Yamada. The human type III cDNA clone, E6, was isolated using 2ng/ml of the 5' PstI restriction fragment of RJ5 which was isolated by electroelution from an agarose gel. Inserts and plasmids were nick translated with ³²P nucleotides to a specific activity of 4-8x10⁸ cpm/ug.

DNA Sequencing

Plasmid DNA was restricted with PstI, or PstI and EcoRI and the fragments were isolated by electroelution after electrophoresis in 1% agarose gels. The DNA was extracted with phenol/chloroform and then cloned into the phage M13mp8,9 system (32). Restriction fragments derived from the cDNA inserts, now contained in the recombinant phage, were sequenced by the Sanger dideoxy method using a universal primer of 17 nucleotides.

Northern Blot Hybridization

Electrophoresis of poly (A⁺) RNA in .7% agarose, 2M formaldehyde gels, transfer to nitrocellulose and hybridization conditions were essentially as described (19,33).

RESULTS

Two sources of mRNA were investigated for the cloning of human type III procollagen sequences. One was a normal fibroblast cell line, GM3348, and the other a human rhabdomyosarcoma cell line, CCL136, that was reported to synthesize type III, and to a lesser extent types IV and V (34). Poly (A⁺) RNA was size fractionated and the 28-35S RNA analyzed by cell free translation for the presence of type III procollagen. The gradient fraction giving the highest yield of a protein that was sensitive to bacterial collagenase and which was slightly larger in size than an $\alpha 1(I)$ standard by polyacrylamide gel electrophoresis was used as template for AMV reverse transcriptase. Double stranded cDNA was prepared and inserted into the PstI site of the plasmid PBR322. Transformation of the recombinant molecules was carried out using E. coli strain MC1061 (31) which gave an efficiency of $0.5-1 \times 10^8$ colonies per ug of supercoiled pBR322. The initial screening of approximately 8000 colonies under cross hybridization

Figure 2. Hybridization of Human Procollagen cDNA Clones to Poly (A^+) RNAs.

Four tenths microgram of poly (A^+) RNA isolated from normal fibroblasts (GM3348) lane 1, a rhabdomyosarcoma (CCL136) lane 2, and a fibrosarcoma (HT1080) lane 3 was electrophoresed in a .7% agarose 2M formaldehyde gel at 30V for 20 hours (33). The RNA was transferred to a nitrocellulose filter at 4° for 16 hours, baked at 76° for two hours and hybridized to 5ng/ml of ^{32}P labelled DNA probe (specific activity $4\text{--}8 \times 10^8$ cpm/ μg). Autoradiography of the type III filter was about five times longer than for $\alpha 1(\text{I})$ and $\alpha 2(\text{I})$.

- Panel A Hybridization with the $\alpha 1(\text{III})$ cDNA clone, RJ5. The same results were obtained with the clone E6.
- Panel B Hybridization with the $\alpha 1(\text{I})$ cDNA clone, $\alpha 12$, (42) containing a 2.4 kb insert extending about 150 nucleotides 5' and 450 nucleotides 3' to Hf404 (28).
- Panel C Hybridization with the $\alpha 2(\text{I})$ cDNA clone, Hf32, containing a 2.2 kb insert (19).

conditions with avian type III DNA sequences (24,25) yielded eight positive signals from the GM3348 and one from the CCL136 libraries. Since two of these GM3348 clones were found to be coding for $\alpha 1(\text{I})$ and $\alpha 2(\text{I})$, the one CCL136 positive clone was next characterized. This cDNA clone, RJ5, contained a 1.12 kb insert that showed a different restriction pattern (Fig. 1) than that characteristic of the human $\alpha 1(\text{I})$ (18,28) or $\alpha 2(\text{I})$ (19,27) cDNA clones. When RJ5 was hybridized to Northern blots of GM3348 and CCL136 poly (A^+) RNAs, two mRNAs of about equal intensity and different from $\alpha 1(\text{I})$ and $\alpha 2(\text{I})$ were observed in each lane (Fig. 2).

In a subsequent screening, the 5' 216 bp PstI fragment of RJ5 was

```

892      GGA GAA AGT GGC CCT GCT GGC CCT GCT GGT GCT CCC GGT CCT GCT GGT TCC CGA GGT GCT CCT GGT CCT CAA GGC CCA CGT GGT GAC AAA      921
      GLY GLU SER GLY PRO ALA GLY PRO ALA GLY ALA PRO GLY PRO ALA GLY SER ARG GLY ALA PRO GLY PRO GLN GLY PRO ARG GLY ASP LYS

922      GGT GAA ACA GGT GAA CGT GGA GCT GCT GGC ATC AAA GGA CAT CGA GGA TTC CCT GGT AAT CCA GGT GCC CCA GGT TCT CCA GGC CCT GCT      951
      GLY GLU THR GLY GLU ARG GLY ALA ALA GLY ILE LYS GLY HIS ARG GLY PHE PRO GLY ASN PRO GLY ALA PRO GLY SER PRO GLY PRO ALA
      PRO

952      GGT CAG CAG GGT GCA ATC GGC AGT CCA GGA CCT GCA GGC CCC AGA GGA CCT GTT GGA CCC AGT GGA CCT CCT GGC AAA GAT GGA ACC AGT      981
      GLY GLN GLN GLY ALA ILE GLY SER PRO GLY PRO ALA GLY PRO ARG GLY PRO VAL GLY PRO SER GLY PRO PRO GLY LYS ASP GLY THR SER
      ALA THR

982      GGA CAT CCA GGT CCC ATT GGA CCA CCA GGG CCT CGA GGT AAC AGA GGT GAA AGA GGA TCT GAG GGC TCC CCA GGC CAC CCA GGG CAA CCA      1011
      GLY HIS PRO GLY PRO ILE GLY PRO PRO GLY PRO ARG GLY ASN ARG GLY GLU ARG GLY SER GLU GLY SER PRO GLY HIS PRO GLY GLN PRO
      SER

1012      GGC CCT CCT GGA CCT CCT GGT GCC CCT GGT CCT TGC      1023
      GLY PRO PRO GLY PRO PRO GLY ALA PRO GLY PRO CYS
      *

```

Figure 3. The DNA Sequence Encoding 132 Amino Acids of the Triple Helical Region of Human Type III Procollagen.

The protein sequence obtained from the DNA sequence of the type III clones, RJ5 and E6, was compared with the published human type III protein sequence derived by Edman degradation (35). Only four differences were found and they are shown on the bottom line. These occur at residues 924, 980, 981 and 983.

purified and hybridized under stringent conditions to the same set of filters and to about 4000 new colonies from a later transformation. Eight positive clones were identified and the largest clone, and the one extending most 5', E6, was later characterized by restriction mapping (Fig. 1) and Southern blot hybridization. E6 contains a 2.4 kb insert that overlaps RJ5 about 400 bases 3' and 900 bases 5', and therefore entirely encompasses the original isolate. The inserts of the other type III clones terminate 3' to the end of E6 as does one other positive from the original eight GM3348 clones and these were not examined further.

Comparison to Human Amino Acid Sequence Determined by Protein Sequencing

The 5' 216 bp PstI fragment of RJ5 contained entirely Gly-X-Y sequences as did 180 nucleotides of the 5' region of the adjacent 382 bp PstI fragment. Immediately at the 3' end of the 180 nucleotides were found two adjacent codons, TGC and TGT, that designate the cysteine residues responsible for formation of the disulfide bonds of type III collagen. According to the published amino acid sequences of type III, the triple helical region contains 1023 residues terminating at the 3' end with the first of the double cysteines (35). Therefore, the most 5' amino acid encoded by RJ5 is residue 892 (Fig. 3) and by comparison E6 extends to about 592.

Only a few differences were apparent when we compared the type III sequences derived from the clones with the reported human amino acid sequence determined by Edman degradation (34): at 924 a threonine versus

Figure 4. Comparison of the Human and Avian $\alpha 1(\text{III})$ Telo- and COOH-Propeptide Regions.

The middle two lines are the nucleotides and corresponding amino acids encoded by the human type III clones. The nucleotide differences found in the avian are shown on the top line, and where these designate a different amino acid is shown on the bottom line. Amino acid 1 is the most 3' of the double cysteines (the 5' cysteine is residue 1023 of the triple helical region shown in figure 2). The symbols indicate the deletion Δ in the avian sequence of three amino acids, insertion $\vee$ of one amino acid, the cysteine residues*, and the cleavage site $\uparrow$ between the telo- and COOH-propeptide, and the glycosylation attachment site which is underlined. The last 78 nucleotides are the beginning of the 3' untranslated region.

proline, at 980 a threonine versus alanine, at 981 a serine vs. threonine, and at 983 a histidine vs. serine. The remaining 128 amino acids were identical].

Analysis of the Telo- and Carboxy- Propeptides of $\alpha 1(\text{III})$, $\alpha 2(\text{I})$ and $\alpha 1(\text{I})$

Alignment of the human and chick type III (24,25) sequences beyond the collagenous region revealed considerable dissimilarity in the first 54 amino acids (Fig. 4). In the human, three amino acids (Gly Ile Gly) corresponding to residues 11-13 were inserted and one amino acid (Glu) between the two tryosines at positions 23 and 24 was deleted. Within the telopeptide (cleavage is presumed to occur between residues 25 and 26), 12

A

	55	PRO	ASP	GLY	SER	ARG	LYS	ASN	PRO	ALA	ARG	ASN	CYS	ARG	ASP	LEU	LYS	PHE	CYS	HIS	PRO	GLU	LEU	LYS	SER	GLY	GLU	106
$\alpha 1(\text{III})$																												
$\alpha 2(\text{I})$		GLU										THR											TRP	SER				TYR
$\alpha 1(\text{I})$		GLU		THR								THR	*					MET	*		SER	ASP	TRP					

	81	TYR	TRP	VAL	ASP	PRO	ASN	GLN	GLY	CYS	LYS	LEU	ASP	ALA	ILE	LYS	VAL	PHE	CYS	ASN	MET	GLU	THR	GLY	GLU	THR	CYS	106
$\alpha 1(\text{III})$																												
$\alpha 2(\text{I})$				ILE								THR	MET	GLU														
$\alpha 1(\text{I})$				ILE						*	ASN								*		ASP	PHE	PRO					*

B

	168	ALA	SER	GLN	ASN	ILE	THR	TYR	HIS	CYS	LYS	ASN	SER	ILE	ALA	TYR	MET	184
$\alpha 1(\text{III})$		GCT	TCC	CAG	AAC	ATC	ACA	TAT	CAC	TGC	AAA	AAT	AGC	ATT	GCA	TAC	ATG	ATG
$\alpha 2(\text{I})$		--C	--T	--C	--C	--C	--C	--C	--C	--G	--C	--C	--C	--C	--C	--C	--C	--C
$\alpha 1(\text{I})$		--C	--C	--C	--C	--C	--C	--C	--C	--G	--C	--C	--C	--G	--G	--C	--C	--C

* VAL

Figure 5. Conserved Regions in the COOH Propeptide of Human $\alpha 1(\text{III})$, $\alpha 2(\text{I})$ and $\alpha 1(\text{I})$.

Figure 5A shows residues 55-106 of the COOH propeptide of type III procollagen (top line) and the corresponding region of $\alpha 2(\text{I})$ (27) and $\alpha 1(\text{I})$ (28). Dashes indicate no difference from the type III amino acid sequence. Asterisks designate the cysteines found in $\alpha 1(\text{III})$.

Figure 5B shows residues 168-184 of the COOH-propeptide including the glycosylation attachment site (underlined). The two top lines show the type III amino acid and nucleotide sequence. The two bottom lines are the nucleotide sequences of the $\alpha 2(\text{I})$ (27) and $\alpha 1(\text{I})$ (28) cDNA clones. Dashes indicate no change from the type III DNA sequence. Only one change in the amino acid sequence is noted: Val at position 180 in the $\alpha 1(\text{I})$.

differences were observed between chick and human and one at the propeptidase cleavage site involved a charge change, Arg to Gly. In the region containing residues 26 to 54, 13 additional changes occurred including three charge differences, the most notable being a Lys in the human at residue 32 and a Glu in the chick. Therefore, a total of 25 changes in 54 amino acids (not counting the deletion and insertion) have been instituted. In the corresponding region of the chick (26) and human $\alpha 1(\text{I})$ (28) there are only two changes, and eight are present in the $\alpha 2(\text{I})$ not including the four amino acid insertion in the human (26,27). The fact that the human and avian type III genes have diverged more than $\alpha 1(\text{I})$ or $\alpha 2(\text{I})$ suggests that the latter two have evolved more recently or that changes in this region of $\alpha 1(\text{III})$ are tolerated.

However, when residues 55-106 were compared, very strong homology between the type III chick and human protein sequence was immediately evident (Fig. 4). Only three amino acids have diverged and none of these involved a difference in charge. Since it seemed important that the protein sequence in this area be conserved, we investigated the comparable region in the human $\alpha 1(\text{I})$ (28) and $\alpha 2(\text{I})$ (27). As can be seen in Fig. 5A, this segment of the three different procollagen COOH-propeptides also

shows striking homology with each other. Interestingly, and perhaps most significant is the presence in this domain of the first five of the eight (seven in $\alpha 2$) cysteine residues in $\alpha 1(\text{III})$, $\alpha 2(\text{I})$, and $\alpha 1(\text{I})$. These amino acids are involved in both inter- and intrachain disulfide bonds and thus the assembly of the procollagen molecules (9,36-38). Because of these functions there is apparently a selective pressure for maintaining a specific conformation.

Previously, other investigators have compared several different procollagen chains and found a conserved nucleotide and amino acid sequence surrounding positions 171-173 (Asn Ile Thr) that designate the glycosylation attachment site (25). In the DNA sequence coding for amino acids 170 to 179, there are thirty identical nucleotides in the avian $\alpha 1(\text{III})$ (25), and human $\alpha 2(\text{I})$ (27) and $\alpha 1(\text{I})$ (28). This homology is probably why we observed hybridization of the chick type III genomic subclone with both human $\alpha 1(\text{I})$ and $\alpha 2(\text{I})$ mRNAs even at stringent hybridization and washing conditions (not shown). However, since hybridization of the avian $\alpha 1(\text{III})$ DNA with the human $\alpha 1(\text{III})$ mRNA in the CCL136 cell line was very faint and inconsistent, we were not surprised to find that changes in these thirty nucleotides of the human type III had occurred. Although four out of thirty nucleotides have diverged, the amino acid sequence is identical (Fig. 3). This region contains the 6th cysteine residue. These results indicate that there does not appear to be conservation of the DNA sequence as long as the protein is relatively unchanged (Fig. 5B).

In the final third of the COOH-propeptide, inclusive of the last two cysteines, homology between human and avian $\alpha 1(\text{III})$ (Fig. 4) and human $\alpha 1(\text{I})$ and $\alpha 2(\text{I})$ is also apparent (25,27,28). Especially notable is the maintenance of the charged residues (Arg, Lys, Glu, Asp). The last fifteen amino acids of human and avian $\alpha 1(\text{III})$ are identical and more like the human $\alpha 1(\text{I})$ sequence (28) than the $\alpha 2(\text{I})$ (27).

Northern Blot Hybridization

Transcription of the human $\alpha 1(\text{I})$ and $\alpha 2(\text{I})$ procollagen genes results in multiple mRNAs in cultured fibroblasts (18-21). The difference in size (200-1000 nucleotides) is due to the use of different polyadenylation attachment sequences in the 3' untranslated region. To determine whether this feature also applied to the $\alpha 1(\text{III})$ gene, we hybridized the cloned $\alpha 1(\text{III})$ DNA to poly (A⁺) RNA from normal fibroblast and rhabdomyosarcoma cell lines. As shown in Fig. 2A, two mRNAs of about equal intensity are present. These are similar in size to the $\alpha 2(\text{I})$ mRNAs which range from

5500-6200 nucleotides (Fig. 2C, and ref. 19,20). Whether these different procollagen RNA transcripts are related to tissue specificity still remains to be determined. It is unlikely that their occurrence is fortuitous since we have recently identified a cDNA clone coding for another human procollagen chain which also hybridizes to multiple mRNAs. (Myers, J.C. et al, manuscript in preparation).

Hybridization of $\alpha 1(\text{III})$, $\alpha 1(\text{I})$ and $\alpha 2(\text{I})$ human procollagen cDNA clones to poly (A⁺) RNAs purified from normal and tumor cell lines also showed a significant difference in their representation (Fig. 2). Although all three procollagen genes are expressed in normal human fibroblasts, we have never detected $\alpha 2(\text{I})$ mRNA in the CCL136 tumor cell line (Fig. 2C) and only trace levels of $\alpha 1(\text{I})$ can be seen after long exposure (Fig. 2B). The amount of type III is comparable in both the normal and CCL136 cell lines (Fig. 2A), and in the latter vastly exceeds the small amount of $\alpha 1(\text{I})$. None of the three genes seem to be expressed in a human fibrosarcoma cell line, HT1080, which synthesizes type IV procollagen (Fig. 2 and ref. 39).

DISCUSSION

Analysis of the type III sequences provides a great deal of information about regions of the protein that have been maintained between species and the different procollagen chains. Comparison of the C-terminal propeptides of human and avian $\alpha 1(\text{III})$ shows strong homology after the variable 5' region, but in the three human procollagen chains two divergent and two conserved regions can be seen. The telopeptide and 5' part of the COOH-propeptide are divergent but the adjacent 3' region where the first five cysteines (four in $\alpha 2(\text{I})$) are located has been essentially maintained. The second variable region is followed by the conserved end of the COOH-propeptide which includes the other three cysteines and the glycosylation attachment site. The critical role of this domain in selection and association of the procollagen chains is probably responsible for the conservation of these regions (9,32). Support for this concept comes from our studies of an osteogenesis imperfecta patient whose DNA contains a four nucleotide frameshift deletion in the pro $\alpha 2(\text{I})$ collagen gene coding for the end of the carboxy-propeptide (38). Change of the last 33 amino acids including the position of the final cysteine prevents incorporation of pro $\alpha 2(\text{I})$ chains into the type I procollagen molecule and results exclusively in ($\alpha 1$)₃ homotrimer formation (38,40).

Whereas type I is the major constituent of bone and tendon (1,2,9), type III seems to predominate in tissues rich in smooth muscle cells (41).

The role of type III collagen has not been clearly defined but in another inherited disease of connective tissue, Ehlers Danlos type IV, rupture of viscera, uterus and large blood vessels is attributed to reduced amounts and/or structurally abnormal type III collagen (9,42). Several type I procollagen mutations resulting in skeletal abnormalities have recently been defined by the use of $\alpha 1(I)$ and $\alpha 2(I)$ DNA clones (37,38,43). Similar studies with these type III cDNA probes, and in the future genomic clones, will certainly elucidate the function and regulation of type III procollagen.

ACKNOWLEDGEMENTS

We are very grateful to Dr. Y. Yamada for providing the avian type III genomic subclones. The encouragement and helpful suggestions of Drs. N. Kefalides and E.J. Macarak are deeply appreciated. We also thank M. Mason for excellent typing of the manuscript and DNA sequences. These studies were supported by National Institutes of Health Research Grants AM33348, AM20553, DE-02623 and HL-29702.

*To whom correspondence should be addressed:

Connective Tissue Research Institute and University of Pennsylvania
3624 Market Street, Philadelphia, PA 19104

REFERENCES

1. Bornstein, P. and Sage, H. (1980) *Ann. Rev. Biochem.* **49**,957-1003.
2. Miller, E.J. and Gay, S. (1982) *Methods Enzymol.* **84**,3-32.
3. Burgeson, R.E., Hebda, P.A., Morris, N.P. & Hollister, D.W. (1982) *J. Biol. Chem.* **257**,7852-7856.
4. Jander, R., Rauterberg, J., and Glanville, R.W. (1983) *Eur. J. Biochem.* **133**,39-46.
5. Bentz, H., Morris, N.P., Murray, L.W., Sakai, L.Y. Hollister, D.W. and Burgeson, R.E. (1983) *Proc. Natl. Acad. Sci. USA* **80**,3168-3172.
6. Sage, H., Trueb, B. and Bornstein, P. (1983) *J. Biol. Chem.* **258**,13391-13401.
7. Gibson, G.J., Kielty, C.M., Garner, C., Schor, S.L., and Grant, M.E. (1983) *Biochem. J.* **211**,417-426.
8. Tate, V., Finer, M., Boedtker, H. and Doty, P. (1982) *Cold Spring Harbor Symp.* **47**,1039-1049.
9. Prockop, D.J. and Kivirikko, K.I. (1984) *New Engl. J. Med.* **311**, 376-386.
10. Lukens, L.N., Frischauf, A.M., Pawlowski, P.J., Brierley, G.T. and Lehrach, H. (1983) *Nucleic Acids Res.* **11**,6021-39.
11. Kurkinen, M., Barlow, D.P., Helfman, D.M., Williams, J.G. and Hogan, B.L.M. (1983) *Nucleic Acids. Res.* **11**,6199-6209.
12. Wang, S.Y., and Gudas, L.J. (1983) *Proc. Natl. Acad. Sci. USA* **80**,5880-5884.
13. Ninomija, Y., Showalter, A., vander Rest, M., Seidah, N.G., Chretien, M., and Olsen, B.R. (1984) *Biochemistry* **23**,617-624.
14. Ninomija, Y., and Olsen, B.R. (1984) *Proc. Natl. Acad. Sci. USA*

- 81,3014-3018.
15. Strom, C.M. and Upholt, W.B. (1984) *Nucleic Acids Res.* 12, 1025-1038.
16. Sandell, L.J., Prentice, H.L., Kravis, D. and Upholt, W.B. (1984) *J. Biol. Chem.* 259, 7826-7834.
17. Yamada, Y., L'fau, G., Mudryj, M., Obici, S., and deCrombrugge, B. (1984) *Nature* 310, 333-337.
18. Chu, M.L., Myers, J.C., Bernard, M.P., Ding, J.F., and Ramirez, F. (1982) *Nucleic Acids Res.* 10, 5925-5934.
19. Myers, J.C., Chu, M.L., Faro, S.H., Clark, W.J., Prockop, D.J., and Ramirez, F. (1981) *Proc. Natl. Acad. Sci. USA* 78, 3516-3520.
20. Myers, J.C., Dickson, L.A., de Wet, W.J., Bernard, M.P., Chu, M-L., DiLiberto, M., Pepe, G., Sangiorgi, F.O., and Ramirez, F. (1983) *J. Biol. Chem.* 258, 10128-10135.
21. Chu, M.-L., de Wet, W., Bernard M., Ding, J.F., Morabito, M., Myers, J.C., Williams, C., and Ramirez, F. (1984) *Nature* 310, 337-340.
22. Huerre, C., Junien, C., Weil, D., Chu, M-L, Morabito, M., VanCong, N., Myers, J.C., Foubert, C., Gross, M-S., Prockop, D.J., Boue, J., Kaplan, J.C., dela Chapelle, A., and Ramirez, F. (1982) *Proc. Natl. Acad. Sci. USA* 79, 6627-6630.
23. de Wet, W.J., Chu, M-L., and Prockop, D.J. (1983) *J. Biol. Chem.* 258, 14385-14389.
24. Yamada, Y., Mudryj, M., Sullivan, M. and de Crombrugge, B. (1983) *J. Biol. Chem.* 258, 2758-2761.
25. Yamada, Y., Kuhn, K., and de Crombrugge, B. (1983) *Nucleic Acids Res.* 11, 2733-2744.
26. Fuller, F. and Boedtker, H. (1981) *Biochemistry* 20, 996-1006.
27. Bernard, M.P., Myers, J.C., Chu, M.L., Ramirez, F., Eikenberry, E.F., and Prockop, D.J. (1983) *Biochemistry* 22, 1139-1145.
28. Bernard, M.P., Chu, M.L., Myers, J.C., Ramirez, F., Eikenberry, E., and Prockop, D.J. (1983) *Biochemistry* 22, 5213-23.
29. Myers, J.C., Ramirez, F., Kacian, D.L., Flood, M. and Spiegelman, S. (1980) *Anal. Biochem.* 101, 88-96.
30. Maniatis, T., Fritsch, E.F., Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, NY.
31. Casadaban, M.J. and Cohen, S.N. (1980) *J. Mol. Biol.* 138, 179-207.
32. Messing, J. (1983) *Methods Enzymol.* 101, 20-78.
33. Thomas, P.S. (1980) *Proc. Natl. Acad. Sci. USA* 77, 5201-5205.
34. Krieg, T., Timpl, R., Alitalo, K., Kurkinen, M. and Vaheri, A. (1979) *Febs. Lett.* 104, 405-410.
35. Seyer, J.M. and Kang, A.H. (1981) *Biochemistry* 20, 2621-2627.
36. Olsen, B. and Dickson, L. in *Chem. and Biol. Mineralized Connect. Tiss.* (ed. Veis. A.) Elsevier/North-Holland, NY (1981) 143-153.
37. Dickson, L.A., Pihlajaniemi, T., Deak, S.B., Pope, F.M., Nicholls, A., Prockop, D.J., and Myers, J.C. (1984) *Proc. Natl. Acad. Sci. USA*, 81, 4524-4528.
38. Pihlajaniemi, T., Dickson, L.A., Pope, F.M., Korhonen, V.R., Nicholls, A., Prockop, D.J., and Myers, J.C. (1984) *J. Biol. Chem.*, in press.
39. Pihlajaniemi, T., Myllyla, R., Alitalo, K., Vaheri, A. and Kivirikko, K.I. (1981) *Biochemistry* 20, 7409-7415.
40. Deak, S.B., Nicholls, A., Pope, F.M. and Prockop, D.J. (1983) *J. Biol. Chem.* 258, 15192-15197.
41. Mayne, R., Vail, M.S. and Miller, E.J. (1978) *Biochemistry* 17, 446-452.
42. Pyeritz, R.E., Stolle, C.A., Parfrey, N.A. and Myers, J.C. (1984) *Amer. J. Med. Genet.*, in press.
43. Chu, M.L., Williams, C.J., Pepe, G., Hirsch, J.L., Prockop, D.J. and Ramirez, F. (1983) *Nature* 304, 78-80.