

**Amendment history:**

- [Correction](#) (April 1990)

## Frameshift mutation near the 3' end of the COL1A1 gene of type I collagen predicts an elongated Pro alpha 1(I) chain and results in osteogenesis imperfecta type I.

M C Willing, ... , D H Cohn, P H Byers

*J Clin Invest.* 1990;**85**(1):282-290. <https://doi.org/10.1172/JCI114424>.

**Research Article**

Osteogenesis imperfecta (OI) is a heterogeneous disorder of type I collagen of which OI type I, an autosomal dominant condition, is the mildest and most common form. Affected individuals have blue sclerae, normal stature, bone fragility without significant deformity and osteopenia. Fibroblasts from most affected individuals produce about half the expected amount of structurally normal type I collagen as a result of decreased synthesis of one of its constituent chains, pro alpha 1(I), but the nature of the mutations which result in OI type I are unknown. We describe a three generation family with OI type I in which all affected members have one normal COL1A1 allele and another from which the intragenic Eco RI restriction site near the 3' end of the gene is missing. Amplification by polymerase chain reaction and sequence determination of the normal allele and of the mutant allele in the domain that normally contains the Eco RI site demonstrated a 5-bp deletion from the mutant allele. The deletion changes the translational reading-frame beginning at the Eco RI site and predicts the synthesis of a pro alpha 1(I) chain that extends 84 amino acids beyond the [...]

**Find the latest version:**

<https://jci.me/114424/pdf>



# Frameshift Mutation Near the 3' End of the COL1A1 Gene of Type I Collagen Predicts an Elongated Pro $\alpha$ 1(I) Chain and Results in Osteogenesis Imperfecta Type I

Marcia C. Willing,\* Daniel H. Cohn,<sup>†</sup> and Peter H. Byers<sup>‡</sup>

Department of \*Pediatrics, \*\*Division of Medical Genetics, Department of Medicine, <sup>‡</sup>Department of Pathology, and <sup>††</sup>Center for Inherited Diseases, University of Washington, Seattle, Washington 98195; and <sup>§</sup>Division of Medical Genetics, Cedars Sinai Medical Center, Los Angeles, California 90048

## Abstract

Osteogenesis imperfecta (OI) is a heterogeneous disorder of type I collagen of which OI type I, an autosomal dominant condition, is the mildest and most common form. Affected individuals have blue sclerae, normal stature, bone fragility without significant deformity and osteopenia. Fibroblasts from most affected individuals produce about half the expected amount of structurally normal type I collagen as a result of decreased synthesis of one of its constituent chains, pro $\alpha$ 1(I), but the nature of the mutations which result in OI type I are unknown. We describe a three generation family with OI type I in which all affected members have one normal COL1A1 allele and another from which the intragenic Eco RI restriction site near the 3' end of the gene is missing. Amplification by polymerase chain reaction and sequence determination of the normal allele and of the mutant allele in the domain that normally contains the Eco RI site demonstrated a 5-bp deletion from the mutant allele. The deletion changes the translational reading-frame beginning at the Eco RI site and predicts the synthesis of a pro $\alpha$ 1(I) chain that extends 84 amino acids beyond the normal termination. Although the mutant pro $\alpha$ 1(I) chain is synthesized in an in vitro translation system, we are unable to detect its presence in intact cells, suggesting that it is unstable and rapidly destroyed in one of the cell's degradative pathways. Our analysis of individuals with OI type I from 20 families indicates that this is a unique mutation and suggests that the phenotype can result from multiple mechanisms that decrease the synthesis of normal type I procollagen molecules, including those that alter protein stability. (*J. Clin. Invest.* 1990; 85:282-290.) elongated pro $\alpha$ 1(I) • frameshift • OI type I

## Introduction

Osteogenesis imperfecta (OI)<sup>1</sup> is a clinically, biochemically and genetically heterogeneous disorder (1). The mildest and most

Presented in part at the Second International Conference on Molecular Biology and Pathology of Matrix, Philadelphia, June 1988 and the Annual Meeting of the American Society of Human Genetics, New Orleans, October 1988 and published in abstract form in 1988. (*Am. J. Hum. Genet.* 43:A207.)

Address reprint requests to Dr. Willing, Division of Medical Genetics, Department of Medicine, RG-25, University of Washington, Seattle, WA 98195.

Received for publication 19 April 1989 and in revised form 17 August 1989.

1. Abbreviations used in this paper: OI, osteogenesis imperfecta; PCR, polymerase chain reaction; RER, rough endoplasmic reticulum.

*J. Clin. Invest.*

© The American Society for Clinical Investigation, Inc.  
0021-9738/90/01/0282/09 \$2.00  
Volume 85, January 1990, 282-290

common form, OI type I, is characterized by normal or near normal stature, bone fragility without significant deformity, osteopenia, blue sclerae, and inheritance in an autosomal dominant fashion (1). Adult onset hearing loss and/or dentinogenesis imperfecta are seen in some kindreds. Family studies have demonstrated linkage of the OI type I phenotype to restriction fragment length polymorphisms in COL1A1, the gene encoding the pro $\alpha$ 1(I) chain of type I collagen, in the majority of cases, and to sites in COL1A2, the gene encoding the pro $\alpha$ 2(I) chain of type I collagen, in the minority (2-4). Little is known about the nature of the mutations that result in OI type I, except that dermal fibroblasts from affected individuals produce and efficiently secrete about half the expected amount of structurally normal type I collagen (4, 5). The abnormality in type I collagen production is reflected in altered ratios of type I to type III collagen in skin (4), and represents the most characteristic biochemical correlate of the phenotype. Barsh et al. (6) demonstrated that the disturbance in type I collagen production is due to decreased synthesis of pro $\alpha$ 1(I) chains in the patients they studied. Dermal fibroblasts from these individuals synthesized pro $\alpha$ 1(I) and pro $\alpha$ 2(I) chains in a 1:1 ratio instead of the expected 2:1 ratio. Additional analyses by Rowe et al. (7) and Genovese and Rowe (8) and by our laboratory (Willing, M. C., and P. H. Byers, unpublished observations) indicate that the decreased pro $\alpha$ 1(I) production reflects a reduction in steady-state COL1A1 mRNA levels. These data, taken together suggest that OI type I frequently results from mutations that affect the apparent expression of COL1A1, but the mechanisms by which expression is altered are not clear.

In this paper we describe a three-generation family with OI type I in which all affected members have a 5 bp deletion that disrupts the 3' intragenic Eco RI site in one COL1A1 allele. The deletion creates a shift in the translational reading-frame of the protein beginning at the Eco RI site, which is located 30 bp from the normal protein termination site, and predicts a protein that extends beyond the normal termination point by 84 amino acid residues. Although the mutant pro $\alpha$ 1(I) chain is synthesized in an in vitro translation system, we are unable to detect its presence in intact cells, suggesting that it is unstable, rapidly degraded, and thus unable to participate in heterotrimer assembly. These findings indicate that mutations that alter protein stability, in addition to those that interfere with nuclear processing (8) can result in decreased production of type I collagen and produce the OI type I phenotype.

## Methods

*Clinical summary.* The proband (II-2) (see pedigree in Fig. 2) is a 33-yr-old Caucasian male with normal stature, mild scoliosis, blue sclerae, normal hearing and dentition, and a history of multiple fractures that have healed without deformity. He has a son with a similar

phenotype and an unaffected daughter. Within the family there is clinical heterogeneity. Two individuals (I-2 and II-4) have had repeated fractures, bony deformity in the lower extremities, and severe osteopenia, which lead to above the knee amputations for both. Neither has upper extremity deformity and both have blue sclerae. Individual II-4 has a son who appears more mildly affected.

**Cell culture and electrophoretic analysis of labeled proteins.** Skin biopsies were taken from four affected (I-2, II-2, II-4, and III-1) and from one unaffected family member (II-3), after receiving appropriate consent. The biopsies were explanted and dermal fibroblasts were grown in Dulbecco-Vogt modified Eagle's medium (DME) that contained 10% fetal calf serum, 100 U/ml of penicillin, and 100 µg/ml of streptomycin. Labeling proteins with [2,3,4,5-<sup>3</sup>H]proline (Amersham Corp., Arlington Heights, IL; 101 Ci/mmol) and [<sup>35</sup>S]cysteine (Amersham Corp.; 600 Ci/mmol), harvesting of medium and cell layer proteins, SDS-PAGE, cleavage of proteins with cyanogen bromide (CNBr) in gels and two-dimensional analysis of CNBr peptides of type I collagen were all carried out as previously described (9, 10).

For short pulse experiments, dermal fibroblasts were plated in DME with 10% fetal calf serum at a density of 250,000 cells per 35-mm dish and allowed to attach overnight. An entire dish was used for each time point. The next day, the tissue culture medium was replaced with DME that contained 1.5% NuSerum (Collaborative Research, Inc., Bedford, MA) and ascorbic acid (50 µg/ml) and the cells were incubated for 2 h. After trypsinization, the cells from each dish were pelleted in microfuge tubes and resuspended in DME that contained 1.5% NuSerum and 100 µCi [<sup>3</sup>H]proline (Amersham Corp.). All labeling reactions were carried out at 37°C. At the end of each time period (5, 10, and 15 min), 1 ml of ice-cold DME was added to each tube, and the cells were rapidly pelleted and lysed in an inhibitor solution containing PBS, 11 mM EDTA, 10 mM PMSF, 45 mM Tris-HCl pH 7.4, 180 mM NaCl, and 4.5% SDS, and boiled for 2 min. Radiolabeled procollagens were analyzed by SDS-PAGE with 2 M urea under reducing conditions, and the separated proteins were identified following radioautofluorography with EN<sup>3</sup>HANCE (New England Nuclear, Boston, MA) as the fluor (9, 10).

**Restriction endonuclease analysis of genomic DNA.** Aliquots of genomic DNA prepared from cultured dermal fibroblasts or from lymphocytes were digested with Eco RI (New England Biolabs, Inc., Beverly, MA) according to manufacturer's specifications, the fragments were separated by gel electrophoresis in 0.8% agarose and transferred to nitrocellulose (11). The filter was hybridized separately with COL1A1 cDNA Hf677 fragments (12) and with a 5-kb genomic fragment from COL1A1 clone CG103 (13). Hybridization probes were labeled with α-[<sup>32</sup>P]dCTP (New England Nuclear, Inc., Boston, MA) by random primer extension (14).

**cDNA synthesis.** Dermal fibroblasts were grown to near confluence in DME with 10% FCS and 100 U/ml of penicillin and 100 µg/ml of streptomycin. 1 d before RNA isolation, cells were incubated in DME with 20% FCS and 50 µg/ml of ascorbic acid. Total cellular, nuclear, and cytoplasmic RNA were isolated as previously described (15, 16). Before cDNA synthesis, the RNA was resuspended in 40 mM Tris pH 7.8, 10 mM NaCl, 6 mM MgCl<sub>2</sub> and incubated with RNase-free DNase (Promega-Biotek, Madison, WI; 2.5 U/ml) at 37°C for 30 min. 10 µg of RNA was precipitated with the appropriate 3' synthetic oligonucleotide primer (Operon Technologies, Inc., San Pablo, CA), and resuspended in an RNase-free solution of 160 mM KCl, 10 mM Tris-HCl pH 8.0, 1 mM EDTA with 12.5 U of RNasin (Promega-Biotek; 25 U/µl). The RNA was denatured at 94°C for 5 min, followed by a 5-min incubation at 37°C to allow for primer annealing. The cDNA synthesis was carried out at 45°C for 30 min in a final reaction cocktail containing 10 mM MgCl<sub>2</sub>, 55 mM Tris pH 8.3, 80 µM deoxynucleotide triphosphates with 37.5 U of RNasin, and 40 U of AMV reverse transcriptase (Life Sciences, Inc., St. Petersburg, FL) (17). Before addition of the reverse transcriptase, the sample was split into two tubes and only one tube received the enzyme. The remaining tube served as a control for the polymerase chain reaction amplification (PCR, see below).

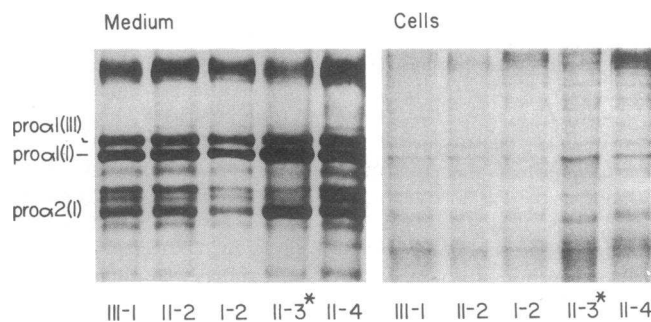
**Polymerase chain reaction amplification of genomic DNA and synthesized cDNA.** Aliquots of fibroblast DNA and synthesized cDNA served as templates for the PCR amplification (18). Sequence for the oligonucleotide primers was derived from sequencing appropriate portions of COL1A1 cDNA Hf677 and from published sequence information (12). One µg each of the appropriate HPLC-purified oligonucleotide primers was used per reaction, and conditions were those specified in the Perkin Elmer Cetus DNA amplification kit (Perkin Elmer Cetus, Norwalk, CT), except that a final concentration of 400 µM of each deoxynucleotide triphosphate was used instead of 200 µM. For amplification of genomic DNA, both primers were complementary to sequences in exon 52; their sequences are A: 5' CTGGGG-CAAGACAGTGATTGAATAC 3', B: 5' CAGTTTGGGTTGCTTGTCTGTTTCCG 3'. For amplification of cDNA, primer B and a new upstream primer (C), complementary to sequences in exon 51 were used; the sequence of C is: 5' CAACAGCCGCTTCACCTACAGCGTC 3' (see Fig. 8). After the PCR, aliquots were separated on a 6% polyacrylamide gel to assess the reaction products.

**Cloning of amplified genomic DNA and DNA sequence determination.** Amplified genomic DNA was separated by agarose gel electrophoresis in 2% low-melting-point agarose, and the 213-bp fragment was resected and purified (19). The isolated material was cleaved with Cla I and Msp I (New England BioLabs, Inc.) according to manufacturer's specifications, to yield a 131-bp fragment. 24 ng of cleaved DNA was ligated into the Acc I site of 100 ng of M13mp9, single-stranded DNA was isolated (20), and DNA sequence determined by the dideoxy chain-termination method (21). Confirmation of published sequence was determined after subcloning a fragment of COL1A1 cDNA Hf677 into M13mp9, and a fragment of genomic clone CG103 into M13mp8 (12, 22).

**Cell-free translation of mRNA.** 10 µg of total cellular RNA was used for in vitro translation in a methionine minus rabbit reticulocyte lysate (Promega) (23). Reaction conditions were as specified in the Promega kit except that 500 µCi of [<sup>3</sup>H]proline per sample was used as the radioactive label. All reactions were carried out at 26°C for 90 min. Each sample was subsequently divided into three equal aliquots and the newly synthesized proteins were precipitated with an equal volume of saturated ammonium sulfate at 4°C. Pelleted proteins from two aliquots were each resuspended in 80 mM CaCl<sub>2</sub>, 8 mM Tris-HCl, pH 7.4, 160 mM *N*-ethylmaleimide, and 1.6 mM PMSF. Proteins in one of the two samples were digested with 1.5 U of bacterial collagenase (24) (*Clostridium histolyticum*; Advance Biofactures Corp., Lynbrook, NY) for 4 h at 37°C before analysis of the synthesized proteins by SDS-PAGE. Proteins in the third aliquot were separated by SDS-PAGE in 5% acrylamide gels and subsequently digested in the gels with CNBr (9). The peptides were separated in a 12.5% acrylamide gel.

## Results

**Dermal fibroblasts from affected individuals synthesize decreased amounts of type I collagen.** We analyzed [<sup>3</sup>H]proline-labeled procollagens synthesized by dermal fibroblasts from four affected and from one normal family member by SDS-PAGE (Fig. 1). Cells from the unaffected control synthesized and secreted normal amounts of proα1(I), with the expected 2:1 ratio of proα1(I):proα2(I). There was a 3:1 ratio of proα1(I):proα1(III). In contrast, cells from affected family members secreted about half the expected amount of type I collagen. The secreted type I procollagen was normal and contained proα1(I) and proα2(I) chains in the expected 2:1 ratio. The ratio of proα1(I):proα1(III) was about 1:1. There was no evidence of excessive posttranslational lysyl hydroxylation or hydroxylysyl glycosylation (overmodification) or retention of aberrant material within the cell (Fig. 1 B), as is typically observed in more severe forms of OI (10, 25–28).

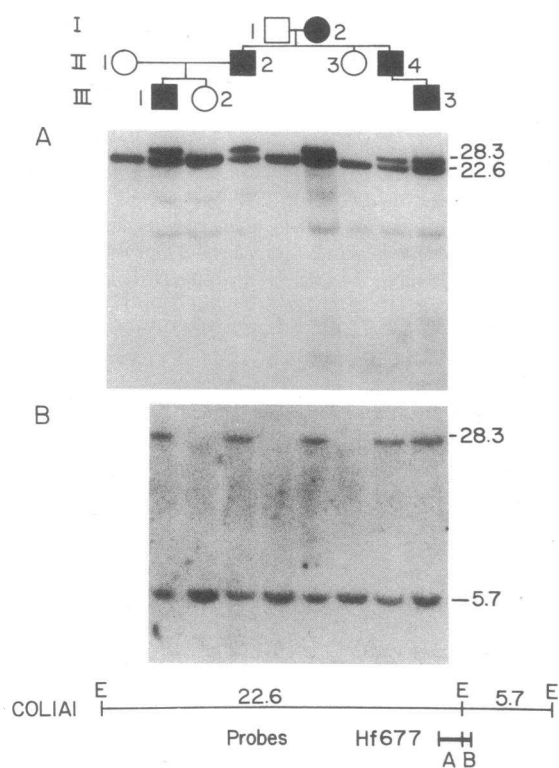


**Figure 1.** [<sup>3</sup>H]Proline-labeled pro $\alpha$  chains synthesized by dermal fibroblasts from affected and unaffected (\*) family members, separated by SDS-PAGE under reducing conditions. C, control; I-2, II-2, II-4, III-1, affected family members; II-4, normal family member. Dermal fibroblasts from affected individuals secrete less type I procollagen than the unaffected family member and have a pro $\alpha$ 1(I):pro $\alpha$ 1(III) ratio of 1:1, compared with the typical 3:1 ratio seen in normal cells.

**Identification of a disrupted intragenic Eco RI site in one COL1A1 allele in affected family members.** The COL1A1 gene, except for the portion that encodes the carboxyl-terminal 12 amino acids and the 3' untranslated region, is contained within a 22.6-kb Eco RI fragment, which is useful for assessing gene size and structural integrity (13). We analyzed Eco RI digests of genomic DNA from nine available family members (Fig. 2). In the digests from the four clinically normal individuals, the probe, a 1.4-kb fragment of COL1A1 cDNA Hf677 (labeled probe A in Fig. 2), recognized only a single fragment of 22.6 kb, while fragments of 22.6 kb and ~28.3 kb were recognized by the probe in the digests from the five affected individuals. This indicated that all of the family members with OI had one normal COL1A1 allele and another from which one of the intragenic Eco RI sites was missing. To determine which restriction site was altered, the same Southern blot was stripped of probe A and rehybridized separately with a 5-kb genomic fragment from the 5' end of the gene (data not shown), and with a 300-bp fragment of cDNA Hf677 that extends 3' from the intragenic carboxy-terminal Eco RI restriction site (Fig. 2 B). The latter probe (labeled probe B in Fig. 2) identified only the 28.3-kb fragment in the affected individuals, confirming that the 3' Eco RI site was missing in their DNA. All family members had a 5.7-kb fragment recognized by probe B.

**Identification of the mutation in one COL1A1 allele.** To sequence across the site of mutation in the normal and mutant alleles, we used the polymerase chain reaction to amplify a 213-bp fragment of genomic DNA in the domain containing the relevant Eco RI restriction site (Fig. 3) from an affected individual (III-1) and from a normal family member (II-3). To confirm that we had successfully amplified the appropriate fragment, we digested the DNA from both individuals with Eco RI (Fig. 3). As expected, the amplified material from the normal family member yielded only fragments of 118 and 95 bp, while material from the affected individual produced the 118- and 95-bp fragments as well as the full-length 213-bp fragment. The smaller fragments are derived from the normal allele, while the uncleaved fragment is the product of the mutant allele.

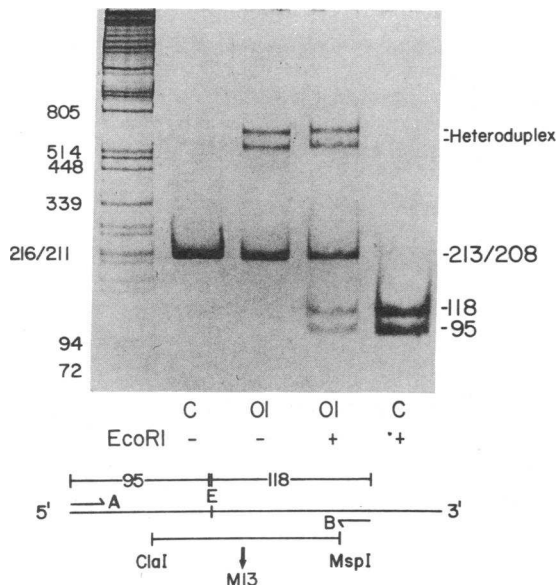
In addition to the full-length amplified material and the restriction products, we observed two more slowly migrating



**Figure 2.** Restriction endonuclease analysis of genomic DNA from available family members. (A) Eco RI digests hybridized with a 1.4-kb Eco RI fragment of COL1A1 cDNA Hf677 (indicated in the figure as probe A). (B) Identification of the disrupted Eco RI site. The same blot was rehybridized with the 300 nucleotide fragment of cDNA Hf677, extending 3' from the Eco RI site within the cDNA (indicated as probe B in the figure). The pedigree indicates affected (darkened figures) and unaffected (open figures) family members. DNA from II-1 has been omitted in B.

bands present only in amplified DNA from the affected individuals (only data from III-1 are shown). We used both restriction endonuclease mapping and denaturation-renaturation studies to show that these bands represent heteroduplexes formed between the products of the mutant and normal alleles. For the restriction mapping, DNA was purified from both bands of slower mobility and cleaved with different restriction endonucleases. In each case, the fragments generated were identical to those seen with cleavage of DNA isolated from the 213-bp fragment (data not shown), except that there was no cleavage with Eco RI. We reproduced the more slowly migrating bands by denaturing DNA isolated from the 213-bp fragment (a mixture of DNA from the normal and mutant alleles) and allowing these fragments to renature before electrophoresis. Finally, denaturation and renaturation of DNA isolated from the slowly migrating bands yielded the 213 bp band (Cohn, D. H., unpublished observations). The reason for the anomalous mobilities of the heteroduplexes is unknown, but presumably it reflects some secondary structure created by the hybridization of sense and anti-sense strands of different alleles, one of which has a mutation.

Taking advantage of natural restriction sites within the amplified 213-bp fragment, we cloned a 131-bp Cla I-Msp I fragment from III-1 into M13 and determined the nucleotide sequence across the region of the Eco RI site in the mutant and

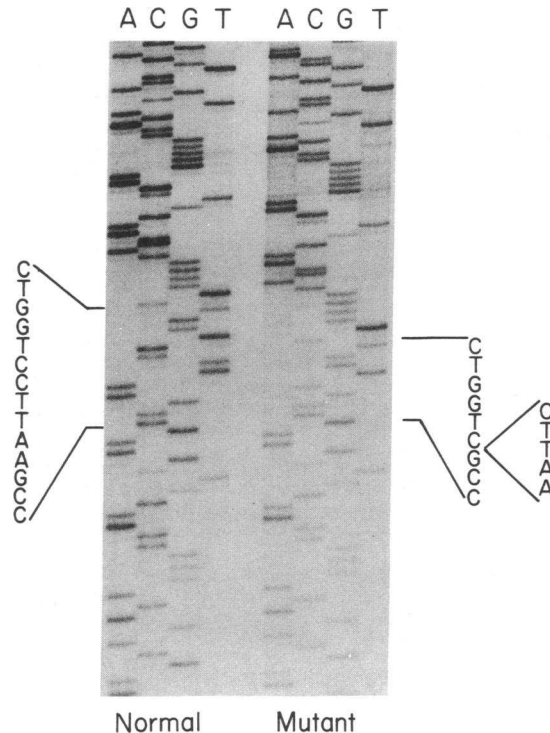


**Figure 3.** Strategy used to isolate, clone, and sequence the normal and mutant alleles in the domain containing the Eco RI site. Using the polymerase chain reaction, we amplified a 213-bp fragment of genomic DNA from II-3 (C) and a similar fragment from III-I (OI). Amplified material from each individual was digested with Eco RI. Positions of molecular size markers from bacteriophage DNA digested with Pst I are indicated down the left side of the figure. We cloned the normal and mutant ClaI-MspI fragments from III-I into M13 and determined the nucleotide sequence of both alleles. The horizontal arrows labeled A and B indicate the locations of the oligonucleotide primers used in the amplification reaction, while E marks the location of the Eco RI restriction site within the amplified fragment.

normal alleles (Fig. 4). The sequence from both alleles was identical, except that the mutant allele had a 5-bp deletion at the Eco RI site, creating a shift in the translational reading-frame of the protein beginning at the Eco RI site. The 213-bp band seen in Fig. 3, therefore, represents a mixture of 213- and 208-bp fragments derived from the normal and mutant alleles, respectively, and the heteroduplexes are formed by the annealing of sense and anti-sense strands of the normal and deletion allele. We have observed this same phenomenon with two other COL1A1 deletions (Wallis, G., P. H. Byers, C. Pruchno, and P. H. Byers, unpublished observations).

The nucleotide sequence of the mutant allele predicts a protein that extends 84 amino acids beyond the normal termination signal (Fig. 5). The nucleotide sequence corresponding to the extended pro $\alpha$ 1(I) chain was determined by sequencing appropriate portions of COL1A1 cDNA Hf677 (12), and genomic clone CG103 (13, 22). The new termination signal lies within the first canonical polyadenylation signal in the COL1A1 gene and immediately follows a 22-bp stretch of adenosine nucleotides. Of the 84 amino acid residues in the predicted pro $\alpha$ 1(I) extension, 14 are lysine residues and 10 of these lysines are contained within the last 19 amino acid residues, making the carboxyl-terminal portion of the protein highly charged.

**The search for the predicted protein.** We used short-pulse experiments and in vitro translation to identify the protein product of the mutant allele. For short pulse studies, fibro-



**Figure 4.** Nucleotide sequence around the Eco RI site in the normal allele (left) and in the same domain of the mutant allele (right) from III-I. Sequence was determined using the dideoxy chain-termination method. The mutant allele has a 5-bp deletion that disrupts the Eco RI site (far right vertical sequence).

blasts from one unaffected and four affected individuals were labeled for periods of 1–20 min with [ $^3$ H]proline. Cells from the normal individual synthesized pro $\alpha$ 1(I) and pro $\alpha$ 2(I) chains in the expected 2:1 ratio. In contrast, fibroblasts from the OI individuals synthesized less pro $\alpha$ 1(I) than the normal cells at each time point, and in fact, exhibited closer to a 1:2 synthetic ratio of pro $\alpha$ 1(I):pro $\alpha$ 2(I) (Fig. 6; only data from III-1 and II-4 are shown). The alteration in synthetic ratio was also observed in labeling periods of < 5 min (data not shown). In addition, cells from affected individuals appeared to produce larger amounts of pro $\alpha$ 1(III) than the normal cells.

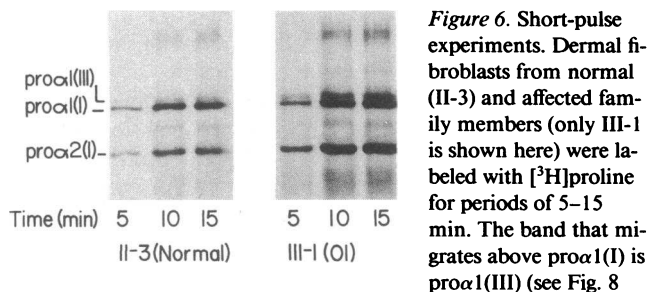
To identify the material migrating in the position of pro $\alpha$ 1(III), we digested [ $^3$ H]proline-labeled pro $\alpha$  chains separated in the first dimension by SDS-PAGE, with cyanogen bromide and separated the resultant peptides in a second dimension gel (Fig. 8 C). Only the normal type III collagen peptide pattern was observed, whether this was carried out on procollagens synthesized during pulses of 5–15 min, or on procollagens produced in 20-h labelings (data not shown). Because the predicted pro $\alpha$ 1(I) chain contains two new cysteine residues we also performed CNBr peptide mapping on [ $^{35}$ S]-cysteine-labeled procollagens. Although we expected a [ $^{35}$ S]-cysteine-labeled fragment from the aberrant pro $\alpha$ 1(I) chain that would migrate slightly faster than  $\alpha$ 1(I)CB3, we were unable to detect such a peptide (data not shown). The short-pulse data argue that the protein product of the mutant allele is either not synthesized or is rapidly degraded. Alternatively, the mutation while having its primary effect on translation termination may have a secondary effect on mRNA metabolism

5' Normal	1	ATC GAT GTG GCC CCC TTG GAC GTT GGT GCC CCA GAC CAG <u>GAA TTC</u> GGC TTC GAC GTT GGC CCT GTC TGC TTC CTG	75
		Ilu Asp Val Ala Pro Leu Asp Val Gly Ala Pro Asp Gln <u>Glu Phe Gly Phe Asp Val Gly Pro Val Cys Phe Leu</u>	
		Ilu Asp Val Ala Pro Leu Asp Val Gly Ala Pro Asp Gln <u>Arg Leu Arg Arg Trp Pro Cys Leu Leu Pro Val Asn</u>	
5' OI	1	ATC GAT GTG GCC CCC TTG GAC GTT GGT GCC CCA GAC CAG <u>CGG CTT CGA GGT TGG CCC TGT</u> CTG CTT CCT GTA AAC	75
5' Normal	76	TAA ACT CCC TCC ATC CCA ACC TGG CTC CCT CCC ACC GAA CCA ACT TTC CCC CGA ACC CGG AAA CAG ACA AGC AAC	150
		Ter *****	
		<u>Ser Leu His Pro Asn Leu Ala Pro Ser His Pro Thr Asn Phe Pro Pro Asn Pro Glu Thr Asp Lys Gln Pro Lys</u>	
5' OI	76	TCC CTC CAT CCC AAC CTG GCT CCC TCC CAC CCA ACC AAC TTT CCC CCC AAC CCG GAA ACA GAC AAG CAA CCC AAA	150
5' Normal	151	CCA AAC TGA ACC CCC CCA AAA GCC AAA AAA TGG GAG ACA ATT TCA CAT GGA CTT TGC AAA ATA TTT TTT TCC TTT	225
		*****	
		<u>Leu Asn Pro Pro Lys Ser Glu Lys Met Gly Asp Asn Phe Thr Trp Thr Leu Glu Asn Ilu Phe Phe Leu Cys Ile</u>	
5' OI	151	CTG AAC CCC CCC AAA AGC CAA AAA ATG GGA GAC AAT TTC ACA TGG ACT TTG GAA AAT ATT TTT TTC CTT TGC ATT	225
5' Normal	226	GCA TTC ATC TCT CAA ACT TAG TTT TTA TCT TTG ACC AAC CGA ACA TGA CCA AAA ACC AAA AGT GCA TTC AAC CTT	300
		*****	
		<u>His Leu Ser Asn Leu Val Phe Ilu Phe Asp Gln Pro Asn Met Thr Lys Asn Gln Lys Cys Ile Gln Pro Tyr Gln</u>	
5' OI	226	CAT CTC TCA AAC TTA GTT TTT ATC TTT GAC CAA CCG AAC ATG ACC AAA AAC CAA AAG TGC ATT CAA CCT TAC CAA	300
5' Normal	301	ACC AAA AAA AAA AAA AAA AAA AGA ATA AAT AAA TAA CTT TTT AAA AAA GGA AGC TT	359
		*****	
		<u>Lys Lys Lys Lys Lys Lys Asn Lys Ter*****</u>	
5' OI	301	AAA AAA AAA AAA AAA AAA AAG AAT AAA TAA ATA ACT TTT TAA AAA AGG AAG CTT	354

**Figure 5.** Nucleotide and protein sequences from the product of the normal allele and predicted sequence from the mutant allele. The 5-bp deletion (underlined in the top row of the figure) results in a frame-shift that predicts a protein which extends beyond the normal termination signal by 84 amino acid residues. The new reading-frame is indicated by the solid bar beginning at the Eco RI site. Nucleotide sequence was determined for the mutant and normal allele for nucleotides labeled 1–131 in the figure. Sequence data for the predicted protein was determined by sequencing a subclone of COL1A1 cDNA Hf677 that started at the Eco RI site, and extended to the end of the cDNA (nucleotides 40–303 in the figure). The remaining nucleotide sequence was determined by sequencing a subclone of COL1A1 genomic clone CG103 (nucleotides 40–359). Our nucleotide sequence data differ from the published sequence in several places. Nucleotide 96 is deoxycytosine in both the normal and OI alleles and in Hf677 from our sequence determination, instead of deoxythymidine (nucleotide 57, Fig. 4, clone d, reference 12). Immediately after residue 135 we found a run of three deoxyadenosine nucleotides instead of the two reported in Hf677 (nucleotides 97–99, Fig. 4, clone d, reference 12). In addition we found a polymorphism at nucleotide 166; the sequence in cDNA Hf677 is deoxythymidine (12), which we confirmed, while genomic clone CG103 has deoxycytosine at this position. Finally, we found a stretch of 22 deoxyadenosine nucleotides before the new termination signal (nucleotides 304–325), while the published sequence indicates a 19 base-pair stretch (nucleotides 225–243 in Fig. 6, reference 22).

and/or transport of the mRNA from the nucleus to the cytoplasm.

To investigate the stability of the mRNA from the mutant allele, we synthesized COL1A1 cDNAs from total, nuclear, and cytoplasmic fibroblast RNA from affected and normal family members, and used the polymerase chain reaction to amplify this material. We reasoned that if mRNA from both the normal and mutant alleles is present, we should detect heteroduplexes formed between the amplification products of the two alleles. To eliminate the possibility of contamination of RNA with genomic DNA that might amplify, we selected a 5' primer that is located in the exon immediately upstream of the one that contains the Eco RI site. This meant that any contaminating genomic DNA that amplified would contain



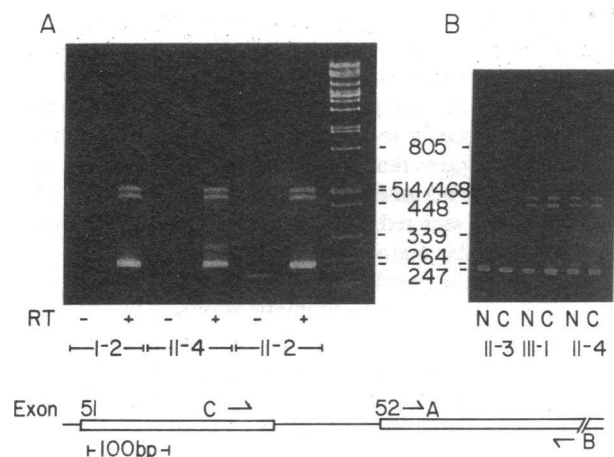
**Figure 6.** Short-pulse experiments. Dermal fibroblasts from normal (II-3) and affected family members (only III-1 is shown here) were labeled with [<sup>3</sup>H]proline for periods of 5–15 min. The band that migrates above proα1(I) is proα1(III) (see Fig. 8

C). Cyanogen bromide peptide maps of proα chains synthesized in short-pulse experiments by control and OI cells.

intronic sequence and thus be distinguished by size from amplified cDNA. Amplified cDNA produced separately from total cellular RNA from four affected individuals exhibited the characteristic heteroduplexes formed between the products of the normal and mutant alleles, while amplified cDNA from the normal family member produced only a single band of 269 bp (Fig. 7 A). Although these results confirm that mRNA is transcribed from the mutant COL1A1 allele in affected individuals, they do not allow us to determine if it is transported to the cytoplasm and available for translation. To this end, we amplified cDNAs derived from both nuclear and cytoplasmic RNA from two affected and one normal family member (Fig. 7 B). As expected, the amplified cDNAs synthesized from nuclear and cytoplasmic mRNA from the normal family member each yielded the expected 269-bp fragment. In addition, we observed the characteristic heteroduplexes in the amplified cDNAs derived from both nuclear and cytoplasmic mRNA from affected family members. These data argue that transport of the mutant COL1A1 mRNA from the nucleus to the cytoplasm and mRNA steady-state levels in the cytoplasm are probably not substantially different from those of the normal mRNA, but we cannot rule out the possibility of cross-contamination of RNAs from the two compartments during purification.

Although we were unable to detect the aberrant proα1(I) chain in cells, our cDNA amplification data suggested that the mRNA from the mutant COL1A1 allele was potentially available for translation. To investigate the ability of this mRNA to





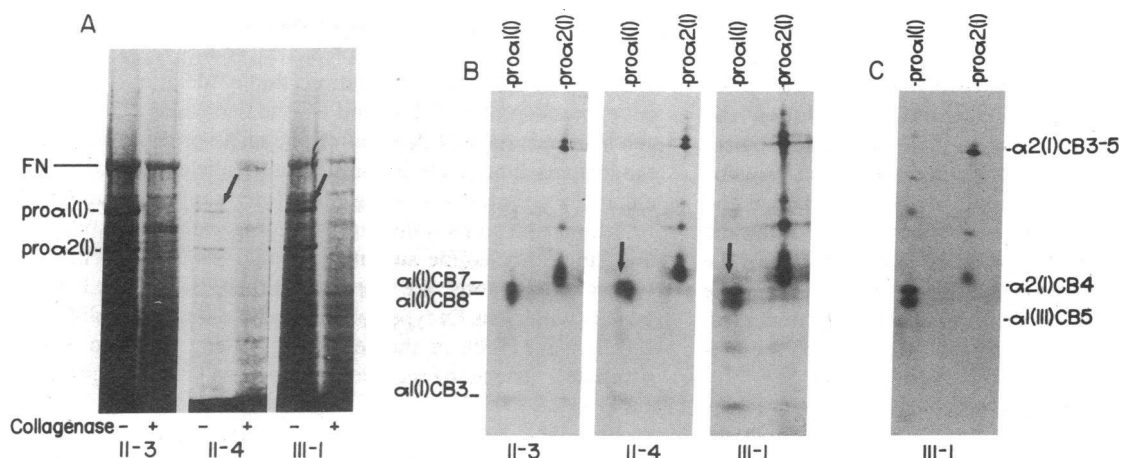
**Figure 7.** Polymerase chain amplification of OI and control cDNAs derived from total, nuclear, and cytoplasmic fibroblast RNA. (A) cDNAs synthesized from total cellular mRNA isolated from three different affected family members, I-2, II-4, and II-2, in the presence (+) or absence (-) of reverse transcriptase (RT). Molecular size markers, derived from digesting bacteriophage lambda DNA with Pst I, are indicated between parts A and B. Each individual has a 269/264-bp fragment corresponding to the normal and mutant alleles, and in addition each has the more slowly migrating heteroduplexes. (B) cDNAs synthesized from nuclear (N) and cytoplasmic (C) mRNA isolated from control (II-3) and OI (III-1 and II-4) dermal fibroblasts. The unaffected control has a single 269-bp fragment, while the two affected individuals each have a 269/264 nucleotide fragment and in addition, the heteroduplexes in the cDNAs derived from both nuclear and cytoplasmic mRNA. The diagram at the bottom illustrates the amplification strategy. The arrows C and B indicate the locations of the 5' and 3' oligonucleotide primers used for the cDNA amplifications, respectively. Arrows A and B represent the oligonucleotide primers used in the amplification of genomic DNA (Fig. 3).

direct translation *in vitro*, we translated mRNA in a cell-free rabbit reticulocyte lysate (Fig. 8 A). Aliquots of each *in vitro* translation reaction were treated with bacterial collagenase to aid in the identification of collagenous proteins. The mRNA from the unrelated controls (data not shown) and the unaffected family member directs the production of pro $\alpha$ 1(I) and pro $\alpha$ 2(I), with a synthetic ratio of roughly 2:1 (Fig. 8 A). The mRNA derived from the OI cells also directs the synthesis of pro $\alpha$ 1(I) and pro $\alpha$ 2(I), but the synthetic ratios are abnormal and resemble those seen in the short-pulse experiments. In addition, we observed a collagenous species migrating more slowly than the normal pro $\alpha$ 1(I) that was not observed in the short-pulse labelings. Cyanogen bromide peptide mapping of this chain indicated that it was a pro $\alpha$ 1(I) species (Fig. 8 B). These data confirm that the mRNA derived from the mutant COL1A1 allele can be translated, and suggest that our inability to detect the aberrant pro $\alpha$ 1(I) chain in cells reflects the presence of an unstable, rapidly degraded protein rather than a block in translation.

The decrease in type I procollagen production by cells from affected members in this family is correlated with remarkable alterations in collagen fibril morphology in their skin (29). Skin from the three affected individuals I-2, II-2 and II-4 have large numbers of composite fibril structures in dermis, similar to those seen in skin from individuals with several forms of Ehlers-Danlos syndrome (29).

## Discussion

Investigation into the biochemical and molecular basis of OI indicates that it is a heterogeneous disorder. Structural alterations and point mutations within the coding region of either COL1A1 or COL1A2 are known to produce the more severe



**Figure 8.** *In vitro* translation. (A) Pro $\alpha$  chains synthesized from total cellular mRNA isolated from control (II-3) and OI (II-4 and III-1) dermal fibroblasts. Digestion with bacterial collagenase (+) was used to identify collagenous proteins. The locations of the pro $\alpha$ 1(I) and pro $\alpha$ 2(I) chains and fibronectin (FN) are shown to the left of the figure. The control has a pro $\alpha$ 1(I)/pro $\alpha$ 2(I) synthetic ratio of 2:1, while the OI individuals have closer to a 1:2 ratio. Both affected individuals also have an additional band (shown by the arrow), which represents the aberrant pro $\alpha$ 1(I) chain (see below). (B) CNBr peptides of pro $\alpha$  chains synthesized *in vitro* from control (II-3) and OI (II-4 and III-1) mRNA. The pro $\alpha$  chains were separated in a first dimension gel under reducing conditions and cleaved in the gel with CNBr. After cleavage, the strip was placed at the top of the second dimension 12.5% polyacrylamide gel and the resultant peptides were separated in the second dimension under reducing conditions. The positions of the pro $\alpha$ 1(I) and pro $\alpha$ 2(I) chains are shown above the figure, while the positions of the peptide fragments derived from pro $\alpha$ 1(I) and pro $\alpha$ 2(I) are shown to the right and left of the figure, respectively. The arrow marks the position of the aberrant pro $\alpha$ 1(I) chain, seen only in the affected individuals. (C) Cyanogen bromide peptide map of pro $\alpha$  chains synthesized in short-pulse experiments by OI cells. There is no evidence of the aberrant pro $\alpha$ 1(I) chain as is seen in the peptide map of the pro $\alpha$  chains translated *in vitro*.

forms of OI (25). In most cases, dermal fibroblasts from affected individuals synthesize normal molecules and, in addition, structurally abnormal collagens that are overmodified, thermally unstable, and inefficiently secreted (25–28, 30–33). In contrast, fibroblasts from most patients with OI type I produce and efficiently secrete about half the expected amount of structurally normal type I collagen (4–6), and little is known about the nature of these mutations.

All affected members of this three generation family have a 5-bp deletion near the 3' end of one COL1A1 allele that produces a shift in the translational reading-frame and predicts a protein that extends beyond the normal termination signal by 84 amino acid residues. Although we are able to detect the aberrant pro $\alpha$ 1(I) chain in an *in vitro* translation assay, we find no evidence of it *in vivo* suggesting that if it is translated within the cell, it must have a very short half-life because even in the shortest pulse-label experiments (1 min), we cannot identify the aberrant protein.

The mRNA from the mutant allele appears to be processed and transported to the cytoplasm and thus potentially available for translation. What effect the translation of stretches of mRNA that normally remain untranslated has on mRNA and protein stability is unclear. Work on elongated globin variants (34–38) and phage and ribosomal RNAs (39, 40) suggests that these regions of mRNA may be essential for its stability or functional integrity, perhaps by forming loops that protect it from hydrolytic attack by 3' exonucleases. If mRNA secondary structure is disturbed in order for the elongated portion of the chain to be translated, the mRNA may become more susceptible to degradation. Alternatively, the unfolding process itself may impose a restraint upon elongation and alter the overall stoichiometry of translation. The elongated pro $\alpha$ 1(I) chain described in this paper is much longer than any of the aberrant globin proteins (41) and any secondary structure generated by the 3' untranslated region would be altered during translation of this long extension. If mRNA secondary structure plays a role here however, it may be to affect translation rather than mRNA stability, as the product of the mutant allele appears to have a stability comparable to that of the normal, judging from availability for reverse transcription and amplification.

After chain assembly, procollagen molecules follow the RER-Golgi-secretory vesicle route to the extracellular matrix (42). How the cell determines the destiny of aberrant proteins is unclear. From our data we cannot determine whether the abnormal pro $\alpha$ 1(I) is not synthesized within cells, or is rapidly destroyed in the lumen of the RER or elsewhere along the secretory pathway. The elongated protein may become trapped within the RER membrane and rapidly degraded, never entering the secretory pathway. Interestingly, the predicted protein sequence has a highly basic carboxyl-terminus similar to some nuclear targeting sequences (43), and of such density that it could act as a “stop-transfer” signal to block the final extrusion into the RER lumen (44).

The fate of free pro $\alpha$ 2(I) chains is also unknown, but we think it is unlikely that these normal chains will form heterotrimers with the elongated pro $\alpha$ 1(I) chains because the 84 amino acid extension would make interaction at the carboxyl-termini impossible. Accumulating evidence indicates that different pathways are used for handling normal versus aberrant protein chains, as well as partially or incorrectly assembled molecules (45–51). Thus, it is not unreasonable to think that different cell sorting mechanisms target free pro $\alpha$ 2(I) and abnormal pro $\alpha$ 1(I) chains for different degradative routes.

The persistence of structurally abnormal procollagen molecules inside the cell with variable secretion into the extracellular matrix is the biochemical hallmark of the more severe OI phenotypes, while decreased production of structurally normal type I collagen appears to be the hallmark of the OI type I phenotype. We suggest that having a “null allele” results in a milder phenotype because abnormal procollagen molecules are not generally secreted outside the cell and incorporated into the extracellular matrix.

The effect on the matrix of producing less type I procollagen than normal appears to be dramatic as judged by alteration in fibril biogenesis (29). We are uncertain why the effects are so striking but they may simply represent altered ratios of collagens and proteoglycan molecules in the extracellular environment. The production of other matrix components, including a variety of proteoglycan molecules, by these cells appears to be unaltered (Vogel, K., unpublished data and personal communication).

Although all affected members of the family described here have the same mutation, there is considerable intrafamilial clinical variability. We wondered if there might be subtle biochemical differences between family members with respect to the amount of abnormal pro $\alpha$ 1(I) produced or its intracellular fate, but no differences were observed. In reviewing the clinical findings, we noticed that the more severely affected family members (I-2 and II-4) had children with both mild and severe phenotypes, while the mildly affected individual (II-2) had an offspring with a mild phenotype. This suggests that there might be some other factor (not identified) segregating independently in this family that acts to modulate the final phenotype.

The mutation described in this paper is unique among mutations producing OI that we have examined. A similar mutation in the COL1A2 gene results in synthesis of a pro $\alpha$ 2(I) chain that is of normal size but has a different end (52). That pro $\alpha$  chain is not incorporated into molecules but is identified within cells (53).

Although the nature of mutations responsible for OI type I is largely unknown, structural alterations (5) and point mutations in (54, 55) or near the triple helical coding region (56–58) appear to be rare causes. Mutations that could alter the apparent expression of COL1A1 and lead to decreased type I collagen synthesis include deletion of one allele or structural rearrangements that create a nonfunctional allele, promoter and enhancer mutations, splicing mutations, small deletions and insertions that create a shift in the reading frame of the protein, and single nucleotide substitutions that result in premature termination codons. We suspect that there is etiologic heterogeneity within the OI type I phenotype and that many different mechanisms, both at the gene and protein level can act to ultimately reduce the secretion of type I collagen and produce OI type I.

## Acknowledgments

We wish to thank the family for their interest and help with the studies, Kathy Braun and Carole Rainer for their excellent technical assistance, Barbra Starman for advice on protein biochemistry, Dr. Art Grix for helping to obtain DNA samples from some family members, Dr. Francesco Ramirez for Hf677, Dr. Richard Gelinis for CG103, and Robert Underwood for his excellent photographic assistance.

This work was supported in part by grants from the National Institutes of Health (AR-21557 and GM-15253), a grant from the Arthritis Foundation and a Basic Science Research Grant from the University of Washington.



## References

1. Silience, D., A. Senn, and D. Danks. 1979. Genetic heterogeneity in osteogenesis imperfecta. *J. Med. Genet.* 16:101-116.
2. Sykes, B., P. Wadsworth, D. Ogilvie, J. Anderson, and N. Jones. 1986. Osteogenesis imperfecta is linked to both type I collagen structural genes. *Lancet.* ii:69-72.
3. Wallis, G., P. Beighton, and C. G. Mathew. 1986. Mutations linked to pro $\alpha$ 2(I) collagen gene are responsible for several cases of osteogenesis imperfecta type I. *J. Med. Genet.* 23:411-416.
4. Sykes, B., M. Francis, D. Phil, and R. Smith. 1977. Altered relation of two collagen types in osteogenesis imperfecta. *N. Engl. J. Med.* 296:1200-1203.
5. Byers, P. H., J. R. Shapiro, D. W. Rowe, K. E. David, and K. A. Holbrook. 1983. Abnormal  $\alpha$ 2-chain in type I collagen from a patient with a form of osteogenesis imperfecta. *J. Clin. Invest.* 71:689-697.
6. Barsh, G. S., K. E. David, and P. H. Byers. 1982. Type I osteogenesis imperfecta: a nonfunctional allele for pro $\alpha$ 1(I) chains of type I procollagen. *Proc. Natl. Acad. Sci. USA.* 79:3838-3842.
7. Rowe, D., J. Shapiro, M. Poirier, and S. Schlesinger. 1985. Diminished type I collagen synthesis and reduced alpha 1(I) collagen messenger RNA in cultured fibroblasts from patients with dominantly inherited (type I) osteogenesis imperfecta. *J. Clin. Invest.* 76:604-611.
8. Genovese, C., and D. Rowe. 1987. Analysis of cytoplasmic and nuclear messenger RNA in fibroblasts from patients with type I osteogenesis imperfecta. *Methods Enzymol.* 145:223-235.
9. Bonadio, J., K. A. Holbrook, R. E. Gelinas, J. Jacob, and P. H. Byers. 1985. Altered triple helical structure of type I procollagen is associated with prolonged survival in lethal perinatal osteogenesis imperfecta. *J. Biol. Chem.* 260:1734-1742.
10. Bonadio, J. F., and P. H. Byers. 1985. Subtle structural alterations in the chains of type I collagen produce osteogenesis imperfecta type II. *Nature (Lond.)* 316:363-366.
11. Southern, E. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98:503-517.
12. Chu, M.-L., J. C. Myers, M. P. Bernard, J.-F. Ding, and F. Ramirez. 1982. Cloning and characterization of fine overlapping cDNAs specific for the human pro $\alpha$ 1(I) collagen chain. *Nucleic Acids Res.* 10:5925-5934.
13. Barsh, G. S., C. L. Roush, and R. E. Gelinas. 1984. DNA and chromatin structure of the human  $\alpha$ 1(I) collagen gene. *J. Biol. Chem.* 259:14906-14913.
14. Feinberg, A., and B. Vogelstein. 1984. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 137:266-267.
15. Chomczynski, P., and N. Sacchi. 1987. Single-step method for RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* 162:156-159.
16. Greenberg, M. E. 1987. Preparation and Analysis of RNA. In *Current Protocols in Molecular Biology*. F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl, editors. John Wiley and Sons, New York. Vol I. 4.10.2.
17. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. In *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. 213-216.
18. Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science (Wash. DC)* 239:487-491.
19. Benson, S. A. 1984. A rapid procedure for isolation of DNA fragments from agarose gels. *Biotechniques.* 2:66-67.
20. Messing, J., R. Crea, and P. H. Seeburg. 1984. A system for shotgun DNA sequencing. *Nucleic Acids Res.* 9:309-321.
21. Sanger, F., S. Nicklen, and A. R. Coulson. 1979. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74:5463-5467.
22. Chu, M.-L., W. de Wet, M. Bernard, and F. Ramirez. 1985. Fine structural analysis of the human pro $\alpha$ 1(I) collagen gene. *J. Biol. Chem.* 260:2315-2320.
23. Bateman, J. F., S. Lamande, D. Chan, and W. G. Cole. 1987. Peptide analysis of collagen produced from cDNA by transcription and translation *in vitro*. *Biochem. J.* 245:393-398.
24. Peterkofsky, B., and R. Diegelmann. 1971. Use of a mixture of proteinase-free collagenases for the specific assay of radioactive collagen in the presence of other proteins. *Biochemistry.* 10:988-993.
25. Byers, P. H. 1989. Disorders of Collagen Biosynthesis and Structure. In *The Metabolic Basis of Inherited Diseases*. C. R. Scriver, A. L. Beaudet, W. S. Sly, and D. Valle, editors. McGraw-Hill Book Co., New York. 2805-2842.
26. Wenstrup, R. J., P. Tsipouras, and P. H. Byers. 1986. Osteogenesis imperfecta type IV: biochemical confirmation of genetic linkage to the pro $\alpha$ 2(I) gene of type I collagen. *J. Clin. Invest.* 78:1449-1455.
27. Wenstrup, R. J., A. S. Hunter, and P. H. Byers. 1986. Osteogenesis imperfecta type IV: evidence of abnormal triple helical structure of type I collagen. *Hum. Genet.* 74:47-54.
28. Wenstrup, R. J., D. Cohn, T. Cohen, and P. H. Byers. 1988. Arginine for glycine substitution in the triple-helical domain of the products of one  $\alpha$ 2(I) collagen allele (COL1A2) produces the osteogenesis imperfecta type IV phenotype. *J. Biol. Chem.* 263:7734-7740.
29. Holbrook, K. A., and P. H. Byers. 1986. Diseases of the extracellular matrix: structural alterations of collagen fibrils in skin. In *Connective Tissue Disease: The Molecular Pathology of the Extracellular Matrix*. J. Uitto and A. Perejda, editors. Marcel Dekker, Inc., New York. 101-140.
30. Barsh, G. S., C. L. Roush, J. Bonadio, P. H. Byers, and R. E. Gelinas. 1985. Intron-mediated recombination may cause a deletion in an  $\alpha$ 1(I) type I collagen chain in a lethal form of osteogenesis. *Proc. Natl. Acad. Sci. USA.* 82:2870-2874.
31. Bateman, J. F., D. Chan, I. D. Walker, J. G. Rogers, and W. G. Cole. 1987. Lethal perinatal osteogenesis imperfecta due to substitution of arginine for glycine at residue 391 of the  $\alpha$ 1(I) chain of type I collagen. *J. Biol. Chem.* 262:7021-7027.
32. Steinmann, B., V. H. Rao, A. Vogel, P. Bruckner, R. Gitzelmann, and P. H. Byers. 1984. Cysteine in the triple-helical domain of one allelic product of the  $\alpha$ 1(I) gene of type I collagen produces a lethal form of osteogenesis imperfecta. *J. Biol. Chem.* 259:11129-11138.
33. Willing, M. C., D. Cohn, B. Starman, K. A. Holbrook, C. Greenberg, and P. H. Byers. 1988. Heterozygosity for a large deletion in the  $\alpha$ 2(I) collagen gene has a dramatic effect on type I collagen secretion and produces perinatal lethal osteogenesis imperfecta. *J. Biol. Chem.* 263:8398-8404.
34. Clegg, J. B., D. J. Weatherall, and P. F. Milner. 1971. Hemoglobin Constant Spring, a chain termination mutant? *Nature (Lond.)* 234:337-340.
35. Clegg, J. B., D. J. Weatherall, I. Contopolow-Griva, K. Caroutsos, P. Pougouras, and H. Tsevrenis. 1974. Haemoglobin Icara, a new chain-termination mutant which causes  $\alpha$  thalassemia. *Nature (Lond.)* 251:245-247.
36. Derry, S., W. G. Wood, M. Pippard, J. B. Clegg, D. J. Weatherall, S. N. Wickramasinghe, J. Darley, S. Fucharoen, and P. Wasi. 1984. Hematologic and biosynthetic studies in homozygous hemoglobin Constant Spring. *J. Clin. Invest.* 73:1673-1682.
37. Laux, B. E., D. Dennis, and H. B. White III. 1973. Human  $\alpha$ -chain globin messenger: prediction of a nucleotide sequence. *Biochem. Biophys. Res. Commun.* 54:894-898.
38. White, H. B., B. E. Laux, and D. Dennis. 1972. Messenger RNA structure: compatibility of hairpin loops with protein sequence. *Science (Wash. DC)* 175:1264-1266.
39. Jou, W. Min, G. Haegeman, M. Ysebaert, and W. Fiers. 1972. Nucleotide sequence of the gene coding for the bacteriophage MS2 coat protein. *Nature (Lond.)* 237:82-88.
40. Fellner, P., C. Ehresmann, P. Stiegler, and J.-P. Ebel. 1972. Partial nucleotide sequence of 16S ribosomal RNA from *E. coli*. *Nat. New Biol.* 239:1-5.
41. Kazazian, H. H., Jr., C. E. Dowling, P. C. Waber, S. Huang, and W. H. Y. Lo. 1986. The spectrum of  $\beta$  thalassemia genes in China and southeast Asia. *Blood.* 68:964-966.

42. Byers, P. H., and J. Bonadio. 1985. The molecular basis of clinical heterogeneity in osteogenesis imperfecta: mutations in type I collagen genes have different effects on collagen processing. *In* Metabolic and Genetic Disease in Pediatrics. J. Lloyd and C. R. Scriver, editors. Butterworths, London. 56–91.
43. Kalderon, D., B. L. Roberts, W. D. Richardson, and A. E. Smith. 1984. A short amino acid sequence able to specify nuclear location. *Cell*. 39:499–509.
44. Sifers, R. N., S. Brashears-Macatee, V. J. Kidd, H. Muensch, and S. L. C. Woo. 1988. A frameshift mutations results in a truncated  $\alpha 1$ -antitrypsin that is retained within the rough endoplasmic reticulum. *J. Biol. Chem.* 263:7330–7335.
45. Valle, G., J. Besley, A. R. Williamson, T. R. Mosmann, and A. Colman. 1983. Post-translational fate of variant MOFC 315 chains in *Xenopus* oocytes and mouse myeloma cells. *Eur. J. Biochem.* 132:131–138.
46. Verbanac, K. M., and E. C. Heath. 1986. Biosynthesis, processing and secretion of M and Z variant human  $\alpha_1$ -antitrypsin. *J. Biol. Chem.* 261:9979–9989.
47. Kreis, T. E., and H. F. Lodish. 1986. Oligomerization is essential for transport of vesicular stomatitis viral glycoprotein to the cell surface. *Cell*. 46:929–937.
48. Copeland, C. S., R. W. Doms, E. M. Bolzau, R. G. Webster, and A. Helenius. 1986. Assembly of influenza hemagglutinin trimers and its role in intracellular transport. *J. Cell Biol.* 103:1179–1191.
49. Gething, M.-J., K. McCammon, and J. Sambrook. 1986. Expression of wild-type and mutant forms of influenza hemagglutinin: the role of folding in intracellular transport. *Cell*. 46:939–950.
50. Lippincott-Schwartz, J., J. S. Bonifacino, L. C. Yuan, and R. D. Klausner. 1988. Degradation from the endoplasmic reticulum: disposing of newly synthesized proteins. *Cell*. 54:209–220.
51. Berg, R. A., M. L. Schwartz, L. H. Rome, and R. G. Crystal. 1984. Lysosomal function in the degradation of defective collagen in cultured lung fibroblasts. *Biochemistry*. 23:2134–2138.
52. Deak, S. B., A. Nichols, F. M. Pope, and D. J. Prockop. 1983. The molecular defect in a nonlethal variant of osteogenesis imperfecta. *J. Biol. Chem.* 258:15192–15197.
53. Pihlajaniemi, T., L. A. Dickson, F. M. Pope, V. R. Korhonen, A. Nichols, D. Prockop, and J. C. Myers. 1984. Osteogenesis imperfecta: cloning of a pro $\alpha 2$ (I) collagen gene with a frameshift mutation. *J. Biol. Chem.* 259:12941–12944.
54. de Vries, W. N., and W. J. de Wet. 1986. The molecular defect in an autosomal dominant form of osteogenesis imperfecta. *J. Biol. Chem.* 261:9056–9064.
55. Starman, B. J., D. Eyre, H. Charbonneau, M. Harrylock, M. A. Weis, L. Weiss, J. M. Graham, Jr., and P. H. Byers. 1989. Osteogenesis imperfecta. The position of substitution for glycine by cysteine in the triple helical domain of the pro $\alpha 1$ (I) chains of type I collagen determines the clinical phenotype. *J. Clin. Invest.* 84:1206–1214.
56. Nichols, A. C., and F. M. Pope. 1984. An abnormal collagen  $\alpha$  chain containing cysteine in autosomal dominant osteogenesis imperfecta. *Br. Med. J.* 288:112–113.
57. Steinmann, B., A. C. Nicholls, and F. M. Pope. 1986. Clinical variability of osteogenesis imperfecta reflecting molecular heterogeneity: cysteine substitutions in the  $\alpha 1$ (I) collagen chain producing lethal and mild forms. *J. Biol. Chem.* 261:8958–8964.
58. Cohn, D. H., S. Apone, D. Eyres, B. Starman, P. Andreassen, H. Charbonneau, A. Nicholls, F. M. Pope, and P. H. Byers. 1988. Substitution of cysteine for glycine within the carboxyl terminal telopeptide of the  $\alpha 1$  chain of type I collagen produces mild osteogenesis imperfecta. *J. Biol. Chem.* 263:14605–14607.