JCI The Journal of Clinical Investigation

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.

B J Starman, … , J M Graham Jr, P H Byers

J Clin Invest. 1989;84(4):1206-1214. https://doi.org/10.1172/JCI114286.

Research Article

Skin fibroblasts grown from three individuals with osteogenesis imperfecta (OI) each synthesized a population of normal type I collagen molecules and additional molecules that had one or two alpha 1(I) chains that contained a cysteine residue within the triple-helical domain, a region from which cysteine normally is excluded. The patients had very different phenotypes. One patient with OI type I had a population of alpha 1(I) chains in which glycine at position 94 of the triple helix was substituted by cysteine; a patient with OI type III had a population of alpha 1(I) chains in which glycine at position 526 of the triple helix was substituted by cysteine; and the third patient, with OI type II, had a cysteine for glycine substitution at position 718 of the alpha 1(I) chain. From all three patients, molecules that contained two mutant chains formed interchain, intramolecular disulfide bonds, and although less stable to thermal denaturation than normal molecules, they were more stable than molecules that contained only a single mutant chain. These findings indicate that substitutions for glycine within the triple-helical domain of the alpha 1(I) chain are not invariably lethal and that their phenotypic effect largely depends on the nature of the substituting residue and its location in the chain.

Osteogenesis Imperfecta

The Position of Substitution for Glycine by Cysteine in the Triple Helical Domain of the Proα1(I) Chains of Type I Collagen Determines the Clinical Phenotype

Barbra J. Starman,* David Eyre,‡§ Harry Charbonneau,§ Maria Harrylock,§ Mary Ann Weis,‡ Lester Weiss,‖ John M. Graham, Jr.,¶ and Peter H. Byers* **‡‡

*Departments of *Pathology, ‡Orthopedics, §Biochemistry, and **Medicine, and ‡‡the Center for Inherited Disease, University of Washington, Seattle, Washington 98195; ‖Henry Ford Hospital, Detroit, Michigan 48202; and ¶Child Development Center, Dartmouth University, Hanover, New Hampshire 03755*

Abstract

Skin fibroblasts grown from three individuals with osteogenesis imperfecta (OI) each synthesized a population of normal type I collagen molecules and additional molecules that had one or two α1(I) chains that contained a cysteine residue within the triple-helical domain, a region from which cysteine normally is excluded. The patients had very different phenotypes. One patient with OI type I had a population of α1(I) chains in which glycine at position 94 of the triple helix was substituted by cysteine; a patient with OI type III had a population of α1(I) chains in which glycine at position 526 of the triple helix was substituted by cysteine; and the third patient, with OI type II, had a cysteine for glycine substitution at position 718 of the α1(I) chain. From all three patients, molecules that contained two mutant chains formed interchain, intramolecular disulfide bonds, and although less stable to thermal denaturation than normal molecules, they were more stable than molecules that contained only a single mutant chain. These findings indicate that substitutions for glycine within the triple-helical domain of the α1(I) chain are not invariably lethal and that their phenotypic effect largely depends on the nature of the substituting residue and its location in the chain.

Introduction

Osteogenesis imperfecta (OI)[1] is a remarkably heterogeneous disorder with phenotypes that range from lethal in the perinatal period to mild, virtually asymptomatic conditions (1, 2). A variety of molecular defects have been found to cause OI, including gene rearrangements (3–6), point mutations (7–10), genomic deletions or splicing mutations that result in deletion of single exon sequences from the mature mRNA and thus the protein (11, 12), mutations that alter the reading frame in the gene for either chain of type I procollagen (13),[2] and mutations that alter the amount of type I procollagen synthesized (14–16).

The relationship between the nature of the mutation and the clinical phenotype is only beginning to be explored. It appears that large rearrangements are likely to be lethal if they are expressed, and that the effect of small deletions may be determined by their chain location and the domain in the chain affected. We now present evidence that suggests that substitutions for glycine in the triple-helical domain of proα1(I) are not always lethal and that the phenotypic effects of such substitutions depend primarily on the nature of the substituting residue and its location in the chain.

Methods

Clinical summary

Family 1. The proband was a 21-yr-old woman of normal height who was active in a variety of sports and had had several previous fractures. On examination she was noted to have light blue sclerae but no dentinogenesis imperfecta or bone deformity. Her mother and all three siblings also had light blue sclerae and had had occasional fractures. None had bone deformity, hearing loss, short stature, or dentinogenesis imperfecta. Dermal fibroblasts were grown from the proband and her mother with appropriate consent.

Family 2. The proband was the second child of healthy nonconsanguinous parents from a town of ~ 10,000 inhabitants in Iraq. The pregnancy was uncomplicated; the child weighed 2.67 kg and was 16.5 in. long at birth. Apgar scores were 8 and 9. The x rays at birth were remarkable for wormian bones in the skull, gracile bones of the upper extremities, short, broad, and bowed femurs, and thin bowed tibias (Fig. 1). At 6 wk the child was bright and alert. There was a triangular appearance to the face and the sclerae were blue. There had been at least one long bone fracture since birth. Both the prior pregnancy and a subsequent pregnancy resulted in normal infants. At 3 yr of age her height was 28¼ in. and her weight was 7 kg, both below the third percentile. She had blue sclerae, opalescent dentin, and small, broken teeth. Her speech was normal and she was at least average in intelligence. She crawled up stairs and on and off furniture but did not stand without support. She had bilateral tibial, femoral, and humeral bowing. X rays confirmed the multiple bowing and demonstrated generalized osteopenia. Dermal fibroblasts were grown from skin biopsies of both parents and of the affected infant when the child was 6 wk old; all biopsies were taken with appropriate consent.

Family 3. The proband was the second child of a nonconsanguinous couple whose first child was normal. Ultrasound performed early in the third trimester identified a fetus with extremely short and bowed limbs and minimal calvarial mineralization. Elective delivery at 34 wk was performed and the infant died upon delivery. The radiologic picture was diagnostic of OI type II, the perinatal lethal form of OI.

Preparation of procollagen and collagen

Dermal biopsies were explanted and fibroblasts grown under standard conditions in DME (Grand Island Biologicals, Grand Island, NY) as previously described (17). To label collagenous proteins, dermal fibroblasts were incubated with 2,3,4,5-[^{3}H]proline (100 Ci/mM) or [^{35}S]-

Dr. Graham's present address is Ahmanson Pediatric Center, Medical Genetics and Birth Defects Center, Cedars-Sinai Medical Center, UCLA School of Medicine, Los Angeles, CA 90048.

Address correspondence to Dr. Peter H. Byers, Department of Pathology SM-30, University of Washington, Seattle, WA 98195.

Received for publication 11 April 1989 and in revised form 15 June 1989.

1. *Abbreviations used in this paper:* OI, osteogenesis imperfecta.
2. Willing, M., D. H. Cohn, and P. H. Byers. 1990. *J. Clin. Invest.* In press.

J. Clin. Invest.

0021-9738/89/10/1206/09 $2.00
Volume 84, October 1989, 1206–1214

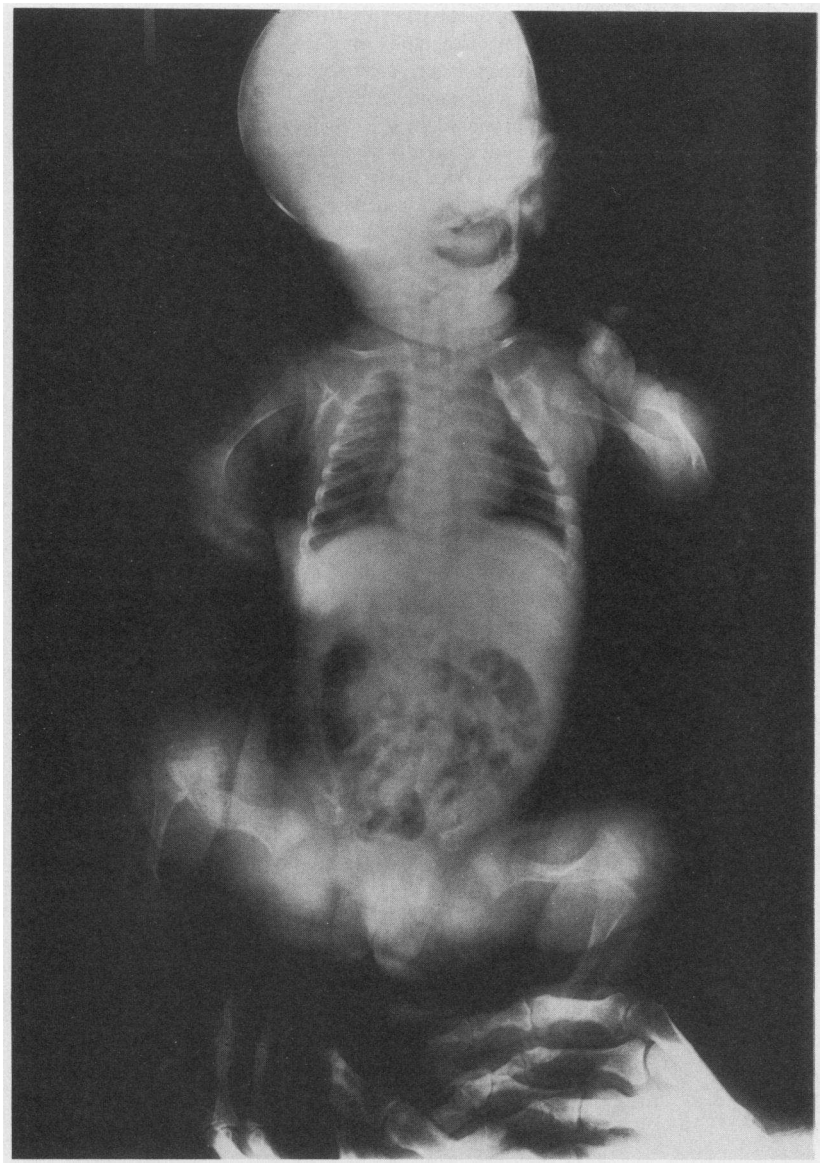


Figure 1. Radiograph of the proband in family 2. There are multiple wormian bones in the skull, the ribs are thin without evidence of previous fracture, the femurs are short and bowed, and the tibias are bowed.

cysteine (1,010 Ci/mM; both from Amersham Corp., Arlington Heights, IL) under conditions previously described (17, 18). Medium and cell-layer proteins were harvested separately in the presence of protease inhibitors (17, 18) and concentrated by ethanol precipitation (30% vol/vol) after addition of carrier collagen (25 μg/ml; Sigma Chemical Co., St. Louis, MO). Digestion of procollagen with pepsin (50 μg/ml; Boehringer Mannheim Biochemicals, Inc., Indianapolis, IN) was carried out for 2 h at 18°C in 0.5 M acetic acid adjusted to pH 2.0 with HCl. The reaction was stopped by the addition of pepstatin and lyophylization.

Characterization of collagen by SDS-PAGE

Proα chains were separated under reducing conditions in 5% acrylamide slab gels containing SDS and 2 M urea (17, 19); collagens were separated in the same gels under nonreducing conditions. In some experiments proteins separated in one dimension were cleaved in the gel with cyanogen bromide and the resultant peptides were analyzed by electrophoresis in a second dimension gel (18).

Determination of thermal stability

Thermal stability of pepsin-digested procollagens was determined as previously described (17, 18, 20).

Preparation of sample for determination of partial amino acid sequence

To prepare collagen chains for microsequence analysis, 10 100-mm dishes (Corning Glass Works, Corning, NY) were grown to confluence and proteins were labeled with [^{3}H]proline the first day and with [^{35}S]-cysteine the next day using conditions outlined above except that no carrier collagen was added to the harvested medium. The medium was collected, dialyzed into 30 mM Tris-HCl, 20 mM NaCl, 2 M urea, pH 7.6, at 4°C, applied to a DEAE-cellulose column (DE-52, 12 × 150 mm; Whatman Laboratory Products Inc., Clifton, NJ) in the same buffer, and eluted with a 300-ml linear gradient from 0 to 100 mM NaCl (21). Type I procollagen was dialyzed into 10 mM NH_4HCO_3 and lyophylized.

The methods for isolating the fragment of the triple-helical domain of the α1(I) chain that contained the cysteine residue differed slightly with each mutation.

In the first family the type I procollagen was cleaved with pepsin, the cysteine residues in the α chains were modified with 4-vinyl pyridine (22), and the chains were separated by reverse-phase HPLC (23, 24). The α1(I) chains were collected, dried, and then cleaved with trypsin and the ^{35}S-containing α1(I)CB5 peptide was recovered by a second HPLC step. This peptide was digested with trypsin (sequencing

grade; Boehringer Mannheim Biochemicals, Inc.) in 0.1 M NH_4HCO_3 and the resulting ^{35}S-containing fragment was again isolated by reverse-phase HPLC. The peptide was then dried and subjected to Edman degradation using a gas-phase automated microsequencer (model 470A; Applied Biosystems Inc., Foster City, CA). A portion of the output at each cycle was analyzed on line for PTH-amino acids (model 120A online PTH analyzer; Applied Biosystems Inc.) and the remainder was collected to determine the content of radioactivity by scintillation counting.

In the second family the cysteines in the isolated proα chains of type I procollagen were reduced and carboxymethylated with iodoacetate after reduction for 4 h in 8 M urea with DTT. Sodium iodoacetate was added to a final concentration of 20 mM and after 15 min a fresh aliquot of DTT was added and the sample was dialyzed into 0.1 M acetic acid before lyophylization. The dried sample was digested with cyanogen bromide for 24 h and the resulting peptides were separated by reverse-phase HPLC (see below for details). The tubes that contained α1(I)CB3 were pooled and dried. Purified α1(I)CB3 was cleaved at the single Asn-Gly bond in the peptide with hydroxylamine (25). The sample was dissolved in water and warmed to 40°C, an equal volume of cold 2 M NH_2OH, pH 10.5, was added, and the mixture was incubated at 35°C for 90 min. The sample was then frozen before separation of the fragments. The new peptides were resolved by HPLC (see below) and the ^{35}S-containing fragment was identified by liquid scintillation counting. The labeled fragment was dried and subjected to automated peptide sequence analysis. For the first eight cleavages 80% of the sample was committed to identification of PTH-amino acids and the remainder was monitored by liquid scintillation counting; the entire sample was counted for products of cleavages 9–29.

After purification of type I procollagen synthesized by the cells from the patient in family 3, the molecules were cleaved with pepsin and the cysteine residues in the α chains were reduced with DTT and alkylated with 4-vinyl pyridine. The chains were cleaved with cyanogen bromide and α1(I)CB7 was separated by reverse-phase HPLC. The peptide was then cleaved with trypsin and the fragment that contained ^{35}S was isolated by reverse-phase HPLC and then subjected to sequence determination as above.

Separation of peptides by HPLC

Whole α chains and peptides obtained from type I procollagen after cleavage with cyanogen bromide and cleavage of α1(I)CB3 with hydroxylamine were dissolved in 1% trifluoracetic acid and separated by reverse-phase chromatography on C4, C8, or C18 macroporous columns. The starting solvent, A, consisted of 5% (vol/vol) acetonitrile in 0.1% (vol/vol) trifluoroacetic acid. Solvent B consisted of 3:1 (vol/vol) acetonitrile/*n*-propanol in 0.1% trifluoroacetic acid. Whole α1(I) chains were separated on a Brownlee Aquapore RP-300 column (C8; 25 cm × 4.6 mm; Applied Biosystems Inc.) or a 214TP54 column (C4; 25 cm × 4.6 mm; Vydac, Hesperia, CA) by a gradient of 15–30% B in 50 min. Cyanogen bromide peptides were separated by a linear gradient of 5–15% B in 40 min (α1[I]CB5, family 1), or a complex gradient of 8–30% B in 60 min (α1[I]CB3, family 2; α1[I]CB7, family 3) on the Brownlee RP-300 column. Hydroxylamine peptides were separated by a linear 0–25% B gradient over a period of 40 min. Tryptic peptides were resolved on the same column using a linear gradient of 0–30% B in 60 min. All column effluents (1 ml/min) were monitored for absorbance at 220 nm, and for ^3H and ^{35}S radioactivity by liquid scintillation counting.

Results

Cysteine in the α1(I) chain of type I collagen. Dermal fibroblasts cultured from two of the three probands (probands from families 2 and 3) synthesized normal type I procollagen and type I procollagen whose chains had delayed electrophoretic mobility when examined by SDS-PAGE under reducing conditions (Fig. 2 *A*). The delay in mobility was abolished by incubation of cells in α,α′-dipyridyl to block prolyl and lysyl hydroxylation (not shown). In the third proband, the individual from family 1, we did not identify any alteration in mobility or in the efficiency of secretion of type I procollagen. When the amino- and carboxyl-terminal propeptides of type I procollagen were removed by cleavage with pepsin and the products were examined by SDS-PAGE in the absence of reduction of disulfide bonds, there was in each cell strain a band that migrated between the α1(I) chains and the trimer of α1(III) that was not synthesized by control cells or by cells from unaffected family members. The electrophoretic mobilities of the new band differed slightly with each cell strain. In all three cases the new chain migrated with chains in the region of α1(I) under reducing conditions (not shown). In addition, in the

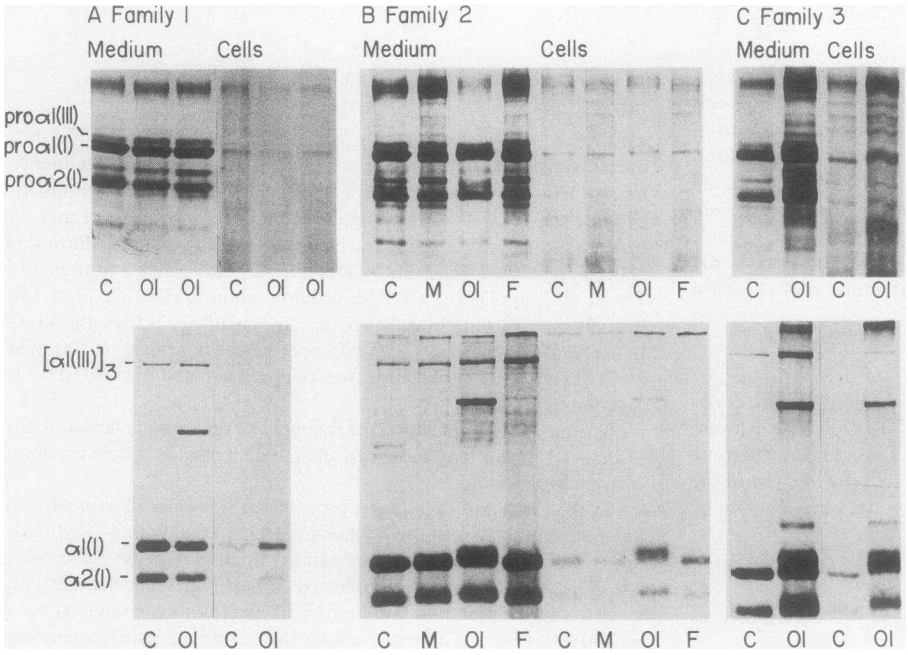


Figure 2. Proα and α chains synthesized by cells from family 1 (*A*), family 2 (*B*), and the proband in family 3 (*C*). In each group proα chains separated under reducing conditions are in the top panel and collagen chains separated under nonreducing conditions are in the bottom panel. In the bottom panel of each a new band is visible that migrates between α1(I) and $(\alpha 1[III])_3$; the mobility of the band differs in each patient. In *A* (*top panel*) the two OI panels are from the affected mother and one daughter (the proband); the control (*C*) is a nonfamily control. In the bottom panel only the proteins synthesized by the daughter's cells are shown. In *B* and *C*, *OI*, affected; *F*, unaffected father; *M*, unaffected mother; *C*, control.

samples from families 2 and 3 there was a band that migrated just slower than the normal α1(I) (Fig. 2 *B*). The amount of the abnormal band that remained within the cells differed among the three cell strains; there was virtually none in the cells from the proband in family 1 and increasing amounts in the probands from families 2 and 3, respectively.

Localization of the new cysteine residues in α1(I). To characterize the abnormal bands synthesized by cells from the probands, we separated the pepsin-treated [^{3}H]proline-labeled collagenous proteins in the first dimension under nonreducing conditions, cleaved the proteins in the gel with cyanogen bromide, and examined the peptides in a second dimension gel under nonreducing and reducing conditions. When the peptides were separated under reducing conditions, in each case the band that migrated between α1(I) chains and the $(\alpha 1[III])_3$ molecules gave rise to the peptides of α1(I) (Fig. 3). Under nonreducing conditions the expected peptides from the new chain synthesized by cells from the proband in family 1 were all present but the partial fragment α1(I)CB5-8 was missing (residues 86–401; Fig. 3 *A*), which suggested that the cysteine was in the domain of α1(I)CB5.

When the peptides from the α1(I) dimer synthesized by the cells grown from the proband in family 2 were separated under nonreducing conditions all the peptides except α1(I)CB3 were present and an additional band of about the size of a dimer of α1(I)CB3 was evident (Fig. 3 *B, left panel*). When the peptides were separated under reducing conditions in the second dimension, all the peptides of α1(I), including α1(I)CB3, were seen (Fig. 3 *B, right panel*). This indicated that the abnormal molecules synthesized by the proband contained a cysteine residue in the domain of α1(I)CB3 (triple-helical residues 402–551), a region from which cysteine is normally excluded. The electrophoretic mobility of α1(I)CB7, α1(I)CB8, and α1(I)CB6 from the dimer were normal; that of α1(I)CB3 was delayed. In contrast, the electrophoretic mobilities of α1(I)CB8 and α1(I)CB3 derived from the band that migrated slightly slower than α1(I) were delayed, while those of α1(I)CB7 and α1(I)CB6 were normal. These findings suggested that in this cell strain overmodification extended over a longer domain in molecules that contained a single mutant chain than in molecules that contained two mutant chains joined by a disulfide bond.

To better localize the new cysteine residue in proband 2, we separated [^{35}S]cysteine-labeled α chains in the first dimension by SDS-PAGE, cleaved them in the gel with hydroxylamine, and separated the resulting peptides in the second dimension. The major cysteine-containing band migrated more slowly than α1(I)CB7 (271 residues) and more rapidly than the partial cleavage fragment α1(I)CB7-6 (463 residues; not shown). This placed the new cysteine residue in both the α1(I)CB3 fragment (residues 402–551) and the large hydroxylamine fragment (residues 523–980) and located it between residues 523 and 551.

When the peptides from the α1(I) dimer synthesized by cells from the proband in family 3 were examined under nonreducing conditions all the peptides except for α1(I)CB7 were identified (Fig. 3 *C*). The peptides from the slowly migrating α1(I) chain all migrated more slowly than those from the normal chain except for α1(I)CB6, the carboxyl-terminal peptide of the triple-helical domain. These results indicated that overmodification of the chains in abnormal molecules began in α1(I)CB7 and extended to the amino-terminal end of the molecule and that the new cysteine was situated between residues 552 and 822.

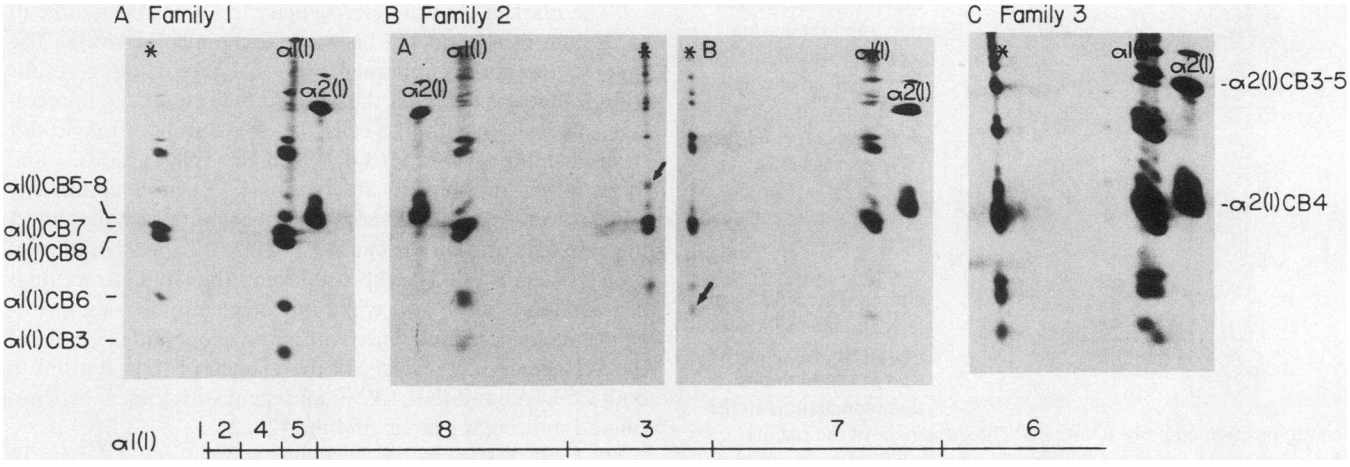


Figure 3. Cyanogen bromide peptides from the normal and abnormal chains and the α1(I) dimer synthesized by cells from the probands in families 1 (*A*), 2 (*B*), and 3 (*C*). [^{3}H]Proline-labeled procollagen secreted by cells from the proband were digested with pepsin and the α chains were separated in the first dimension under nonreducing conditions and cleaved in the gel with cyanogen bromide; the peptides were then separated in the second dimension under nonreducing conditions. For the right panel in *B* the peptides were separated under reducing conditions in the second dimension. The position of the new band in the first dimension gel is marked with an asterisk. *A,* Peptides from the dimer separated under nonreducing conditions show no evidence of the partial cleavage fragment α1(I)CB5-8. This locates the cysteine in α1(I)CB5. *B,* Under nonreducing conditions (*left panel*) α1(I)CB3 from the dimer (*) is not seen but a new band is present above α1(I)CB5-8 (*arrow*). Under reducing conditions α1(I)CB3 is present (*arrow*), although the electrophoretic mobility is slower than that of the normal α1(I)CB3. These findings indicate that the new cysteine is in α1(I)CB3. *C,* Under nonreducing conditions there is a new fragment from the α1(I) dimer that migrates near the position of α2(I)CB3-5 and all the peptides with the exception of α1(I)CB6 are overmodified. This indicates that the cysteine is located amino-terminal to α1(I)CB6 in the domain of α1(I)CB7. A diagrammatic representation of the cyanogen bromide peptides of the triple-helical domain of the α1(I) chain is located below.

Identification of substituted residues. We then labeled procollagens with [^{3}H]proline and [^{35}S]cysteine, purified type I procollagen by DEAE-cellulose chromatography, isolated by HPLC the cysteine-containing fragments created by chemical and/or enzymatic cleavage, and determined the radioactive and, when possible, the chemical sequence of the fragments. In this manner we determined that in family 1 the cysteine replaced glycine at position 94 of the triple helix of the α1(I) chain, in proband 2 the cysteine replaced glycine at position 526 of the triple helix of the α1(I) chain, and in proband 3 the cysteine replaced glycine at position 718 of the triple helix of the α1(I) chain (Fig. 4). In the collagens synthesized by the probands in families 1 and 3 the site of the cysteine for glycine substitutions was revealed by coincident ^{35}S radioactivity and the PTH derivative of pyridinylethyl cysteine together with a major running sequence for the tryptic peptide. In the peptide sequences around the mutant site in families 2 and 3 the proline that preceeded the cysteine appeared to be fully hydroxylated considering the PTH amino acid yield.

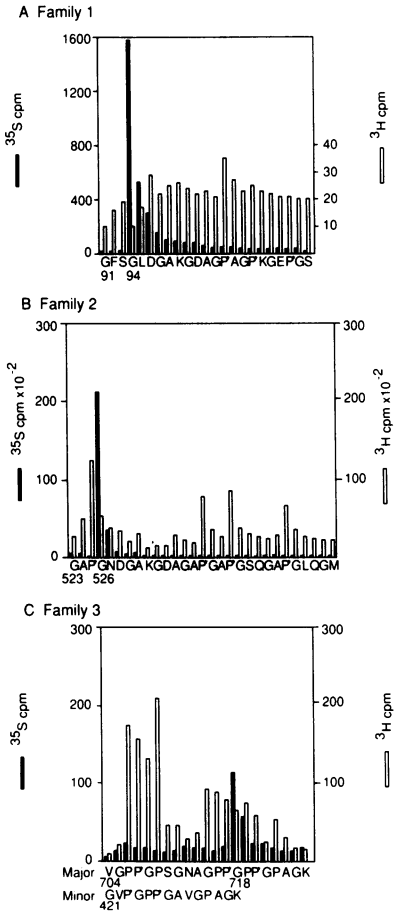


Figure 4. Profile of radiolabeled amino acids after automated Edman degradation sequence determination of the cysteine-containing fragments from the probands in families 1 (*A*), 2 (*B*), and 3 (*C*). The sequences below each panel are of the normal α1(I) chains; position 1 is the first glycine of the triple-helical domain. *A*, The radioactivity is not corrected for cleavage efficiency. Each residue was identified as the PTH derivative, except that the glycine at position 94 in the normal sequence was identified as a modified cysteine. *B*, The counts in the peaks are corrected for 92% cleavage efficiency at each cycle. For the first eight turns the PTH-derived amino acids were isolated and, although the levels were low, alanine in the second position, aspartic acid in the sixth position, and alanine in the eighth position could be identified. The placement of the radiolabeled proline residues and the directly identified residues confirmed the identity of the fragment and the isolation of [^{35}S]cysteine at the fourth cycle and indicated that there had been a cysteine for glycine substitution at residue 526 of the triple-helical domain of the α1(I) chain. *C*, Two tryptic peptides were identified for their full length during sequence analysis. The major sequence contained the residues in the triple helix of α1(I) from 704 to 725 and the minor sequence was derived from 421 to 434 of the triple helix of α1(I). All residues, including the modified cysteine at 718, were identified. Each residue in all three panels is identified by the single letter code: *A*, alanine; *D*, aspartic acid; *E*, glutamic acid; *G*, glycine; *K*, lysine; *L*, leucine; *M*, methionine; *N*, asparginine; *P*, proline; *Q*, glutamine; *R*, arginine; *S*, serine; *V*, valine; *P**, hydroxyproline.

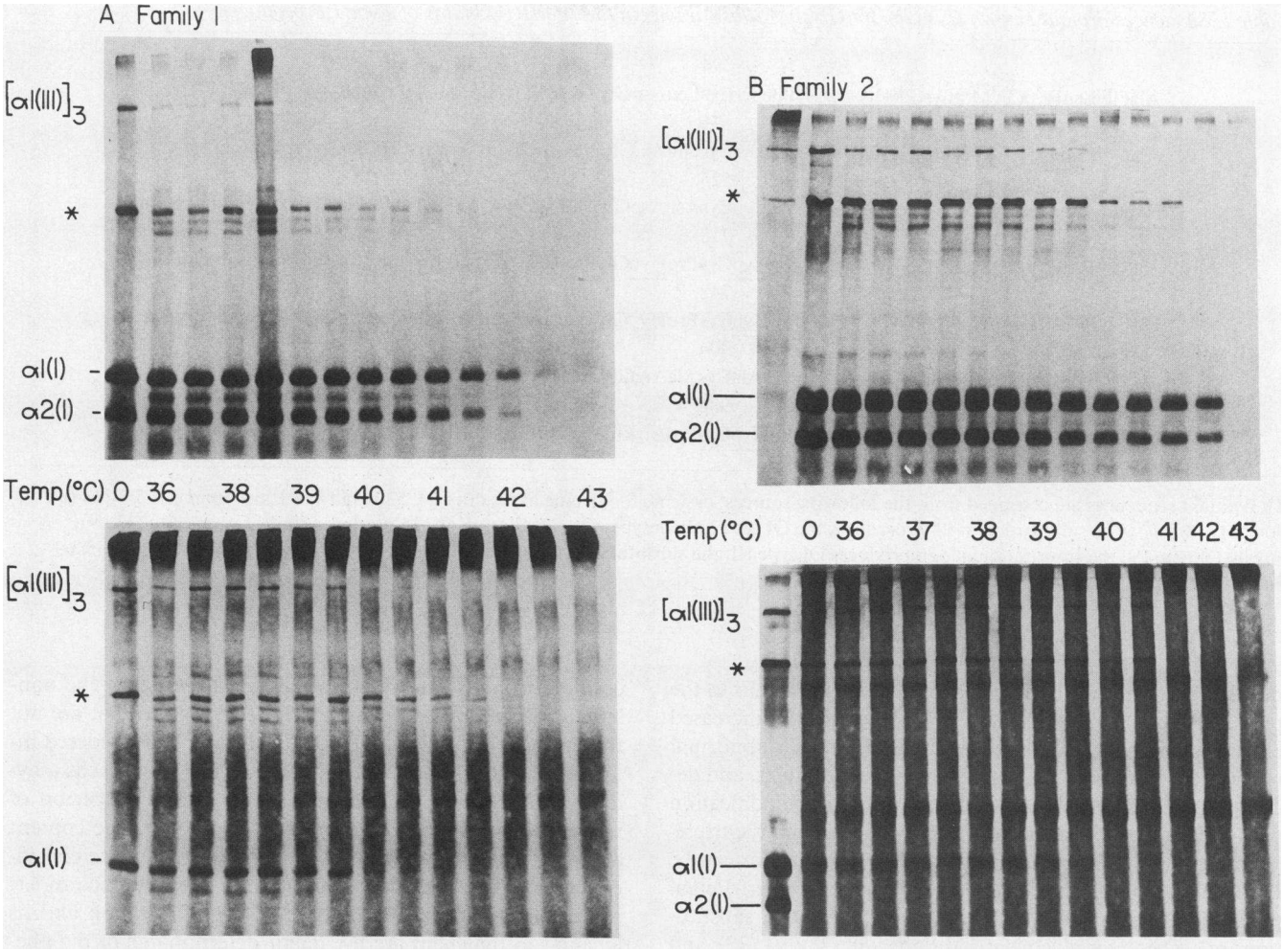
Abnormal molecules have a low thermal stability. In all three cell strains in which there is a cysteine for glycine substitution, the type I procollagen molecules that contained two abnormal chains (as evidenced by the presence of interchain disulfide bonds) melted at lower temperatures than control molecules, but at higher temperatures than those molecules that contained only a single abnormal chain (Fig. 5). In collagens synthesized by cells from proband 1, molecules that contained the disulfide bonded α1(I) dimer melted between 41 and 42°C, and the abnormal molecules that contained a single abnormal chain melted between 39 and 40°C; in those from proband 2, the dimer melted between 40 and 41°C, and the molecules that contained a single abnormal chain melted at ~ 39.5°C; in those from proband 3, the dimer melted between 37 and 38°C and the molecules that contained a single abnormal chain melted at ~ 37°C. In each instance the single mutant chain (cysteine labeled) in the abnormal molecules was susceptible to proteolysis at a lower temperature than the normal chain (proline labeled) that had become overmodified.

Discussion

We have identified three separate families in which the presence of a cysteine residue within the triple-helical domain of the proα1(I) chain of type I procollagen results in an OI phenotype. The clinical result of the mutations is strikingly different: glycine to cysteine at residue 94 of the triple helix results in the mild OI type I phenotype; glycine to cysteine at residue 526 results in the moderate OI type III phenotype; and glycine to cysteine at residue 718 results in the perinatal lethal OI type II phenotype.

The marked clinical heterogeneity in OI has been difficult to explain clearly on the basis of biochemical findings. The mildest dominantly inherited form, OI type I, often results from mutations that alter the amount of normal type I procollagen secreted by cultured cells (14–16). Rarely, it may result from deletion of all of a COL1A1 allele (Willing, M. C., and P. H. Byers, unpublished observations), but more commonly appears to result from a variety of other mutations that affect the synthesis of proα1(I) chains (14–16), the incorporation of proα1(I) chains into type I procollagen,[2] the rapid intracellular degradation of abnormal type I procollagen molecules, or evidence of an altered sequence within the telopeptide domain of the α1(I) chain (26); linkage of the phenotype to mutations in both COL1A1 and COL1A2 is consistent with these biochemical and molecular genetic findings (27, 28).

In most instances the perinatal lethal form, OI type II, results from substitutions for single glycyl residues within the triple helical domain of α1(I) or α2(I). These mutations appear to interfere with the formation of stable triple helix, delay secretion of molecules that contain mutant chains, alter the extent of posttranslational modification of all chains in molecules that carry a defective chain (17), interfere with amino-terminal propeptide cleavage, and result in an abnormal structure of type I procollagen molecules that may affect fibrillogenesis (29). In rare instances the OI type II phenotype may result from mutations that result in deletion or insertion of coding material (3–6).



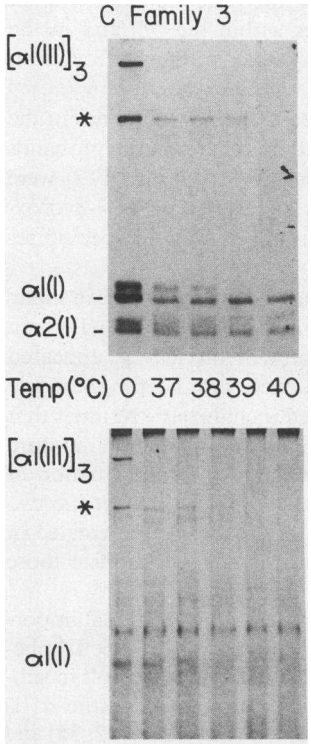


Figure 5. Thermal stability of the collagen molecules synthesized by cells from the probands in families 1 (*A*), 2 (*B*), and 3 (*C*). Proteins analyzed in each top panel were biosynthetically labeled with [^{3}H]-proline and those in the bottom were labeled with [^{35}S]-cysteine. In *B*, lane *1* of the top panel contains [^{35}S]-cysteine-labeled proteins and lane *1* of the bottom panel contains [^{3}H]proline-labeled proteins to confirm the position of the cysteine-labeled proteins.

Between these extremes lies a large group of individuals with OI, classified on clinical grounds as either OI type III (progressive deforming) or OI type IV (dominantly inherited with non-blue sclerae and often accompanied by dentinogenesis imperfecta) and ranging in severity from mild to moderately severe (1, 2). Relatively few mutations that produce these disorders have been identified. A 4-bp deletion near the 3′ end of the coding region of the proα2(I) chain that altered the last 12 residues of the chain and prevented its assembly into type I procollagen molecules produced an OI type III phenotype in the homozygous state but had little phenotypic effect in the heterozygotes (13, 30). Deletion of single exons encoding domains near the amino-terminal end of the triple helix of the proα2(I) chain (11) or near the center of the triple helix (31) have produced mild-to-moderate phenotypes, as has a substitution of arginine for glycine near the carboxyl-terminal end of the triple-helical domain of proα2(I) (32). This list, however, emphasizes the difficulty in providing a cohesive understanding of the genotype–phenotype relationship in the mild-to-moderate OI phenotypes.

We think that our findings in the three families studied here, in addition to the accumulation of mutations rapidly being identified, begin to help understand some aspects of the relationships between mutation and phenotype. It is striking that there is progressive improvement of phenotype as the same point mutation, substitution of cysteine for glycine,

Table I. Sequences around Known Cysteine for Glycine Substitutions in the Proα1(I) Chains of Type I Procollagen

94
AlaGlyLeuHypGlyMetLysGlyHisArgGlyPheSerCysLeuAspGlyAlaLysGlyAspAlaGlyProAlaGlyProLysGlyGluHyp
175
AlaGlyProHypGlyPheHypGlyAlaValGlyAlaLysCysGluAlaGlyProGlnGlyProArgGlySerGluGlyProGlnGlyValArg
526
HypGlyProAlaGlyProArgGlyAlaAsnGlyAlaHypCysAsnAspGlyAlaLysGlyAspAlaGlyAlaHypGlyAlaHypGlySerGln
718
ValGlyProHypGlyProSerGlyAsnAlaGlyProHypCysProHypGlyProAlaGlyLysGluGlyGlyLysGlyProArgGlyGluThr
748
ThrGlyProAlaGlyArgHypGlyGluValGlyProHypCysProHypGlyProAlaGlyGluLysGlySerHypGlyAlaAspGlyProAla
904
ThrGlyProAlaGlyProAlaGlyProValGlyProAlaCysAlaArgGlyProAlaGlyProGlnGlyProArgGlyAspLysGlyGluThr
988
AspGlyLeuAsnGlyLeuHypGlyProIleGlyProHypCysProArgGlyArgThrGlyAspAlaGlyProValGlyProHypGlyProHyp

The mutant sequences are compiled from the following sources: cysteine for glycine at residues 94, 526, and 718 (this report); 175;[3] 748 (8); 904 (34); 988 (7). The substitution at 94 procedures the OI type I phenotype; the substitution at 175 produces clinical variability (33) but largely OI type IV; the substitution at 526 produces OI type III; the substitutions at 718, 748, 904, and 988 all produce the OI type II lethal phenotype.

moves toward the amino-terminal end of the triple helix of the proα1(I) chain. This movement is accompanied by increased efficiency of secretion of molecules that contain abnormal chains, increased thermal stability of those molecules, and decreased extent along the molecule in which overmodification occurs. Because cysteine is normally excluded from the triple-helical domain of proα1(I), substitutions in the chain can be identified by the formation of an α1(I) chain dimer. Lethal point mutations have been identified that result in substitutions of cysteine for glycine at positions 988 (7), 904 (33), and 748 (8) in the triple-helical domain, and a milder form results from substitution for glycine at position 175 (34).[3] These families fit the generalization that mutations that result in substitution of cysteine for glycine are more severe in their phenotypic consequences when they occur near the carboxyl-terminal end of the triple-helical domain of proα1(I) than at the amino-terminal end.

This end-rule or phenotypic gradient for mutations along the proα1(I) chain probably depends on the nature of the substituting amino acid. We are aware, for example, that substitution of arginine for glycine at positions 391 (9) or 664 (10) results in a lethal OI phenotype and that substitution for glycine by arginine at position 154 results in the OI type III phenotype.[4]

Inspection of the sequence motifs around the lethal and nonlethal cysteine for glycine substitutions (Table I) does not clearly identify specific domains which, when altered, produce lethal phenotypes. For example, the substitutions at 718 and 748 both disrupt the hexapeptide GlyProHypGlyProHyp and are lethal. The substitution at 988 substitutes a glycine that succeeds a potentially hydroxylatable prolyl residue and the substitution is in a region bounded by additional hydroxyproline residues. In contrast, the substitution at 904 is in a region entirely devoid of hydroxyproline and would ordinarily be considered a relatively loose triple-helical domain. The nonlethal substitutions at 94 and 175 are in regions that are not rich in hydroxyproline and would probably be considered intermediate in stability. The substitution at 526 succeeds a hydroxylatable prolyl residue and is embedded in a region of moderate hydroxyproline content; if hydroxyproline content alone determined the phenotype, then this would be considered a candidate for a lethal mutation. Given these considerations, it appears that the nature of the substitution and its location in the chain are the major determinants of the phenotypic outcome of a point mutation. Because cysteine residues in the triple-helical domains of fibrillar collagens, in contrast to other substitutions, have the potential for covalent interactions with other molecules within the cell and in the matrix, cysteine substitutions may differ in their effects from those of other residues.

The prolyl residues that preceded the new cysteines in the abnormal α1(I) chains synthesized by cells from the probands in family 1 (position 93) and family 3 (position 717) were hydroxylated. These observations suggest that prolyl 4-hydroxylase may recognize general context as well as specific sequence.

The stability to thermal denaturation of the chains in molecules that contained one mutant and one normal α1(I) chain were not equivalent and the first proteolytic cleavage appeared in the abnormal chain. The sequences around the mutation sites in each chain we characterized include basic residues that would be cleaved by the trypsin in the incubation mixture (Table I). Our observations suggest that the conformation of the single mutant chain is altered more than that of the two normal chains and, further, that the disulfide bond created in molecules with two abnormal chains acts to stabilize those chains in a less available conformation.

Finally, the findings in family 2 have an additional important implication in that they confirm that the OI type III phenotype can result from dominant mutations that result in substitutions for glycine residues in the triple helix of the α1(I) chain of type I collagen. Until now, clinical genetic (2, 35) and

3. Hollister, D., personal communication.
4. Pruchno, C. J., B. J. Starman, and P. H. Byers, unpublished observations.

biochemical studies (13) have been interpreted to indicate that this form of OI is usually inherited in an autosomal recessive fashion. The findings in this family, additional sequence determination,[4] and our survey of individuals with moderately severe OI phenotypes (Wenstrup, R. J., M. C. Willing, B. J. Starman, and P. H. Byers, manuscript submitted for publication) all suggest that a significant proportion of individuals with the OI type III clinical phenotype result from dominant mutations which, in some families, are inherited.

Acknowledgments

We thank the families for their cooperation, Dr. Golder Wilson for supplying additional clinical information about family 2, Dr. Barbara McGillivray and Dr. Derek Applegarth for supplying clinical information and the cultured fibroblasts from the proband in family 3, Carole Rainer and Kathy Braun for assistance with cell culture and sample preparation, Stephen Apone for help with HPLC, Robert Underwood for photography and figure preparation, and David Hollister for providing sequence information before publication.

This research was supported in part by grants from the National Institutes of Health (AR-21557, GM-15253, AR-37318, and GM-15731) and a Clinical Research Grant (6-298) from the March of Dimes Birth Defects Foundation.

References

1. Byers, P. H. 1989. Disorders of collagen biosynthesis and structure. *In* Metabolic Basis of Inherited Disease. 6th edition, C. S. Scriver, A. L. Beaudet, W. S. Sly, and D. Valle, editors. 6th ed. McGraw-Hill Inc., New York. 2805–2842.

2. Sillence, D. O., A. Senn, and D. M. Danks. 1979. Genetic heterogeneity in osteogenesis imperfecta. *J. Med. Genet.* 16:101–116.

3. Chu, M.-L., V. Gargiulo, C. Williams, and F. Ramirez. 1985. Multiexon deletion in an osteogenesis imperfecta variant with increased type III collagen mRNA. *J. Biol. Chem.* 260:691–694.

4. 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 α1 type I collagen chain in a lethal form of osteogenesis imperfecta. *Proc. Natl. Acad. Sci. USA.* 82:2870–2874.

5. Byers, P. H., B. J. Starman, D. H. Cohn, and A. L. Horwitz. 1988. A novel mutation causes a perinatal lethal form of osteogenesis imperfecta: an insertion in one α1(I) collagen allele (COL1A1). *J. Biol. Chem.* 263:7855–7861.

6. Willing, M. C., D. H. Cohn, B. Starman, K. A. Holbrook, C. R. Greenberg, and P. H. Byers. 1988. Heterozygosity for a large deletion in the α2(I) collagen gene has a dramatic effect on type I collagen secretion and produces perinatal lethal osteogenesis imperfecta. *J. Biol. Chem.* 263:8398–8404.

7. Cohn, D. H., P. H. Byers, B. Steinmann, and R. E. Gelinas. 1986. Lethal osteogenesis imperfecta resulting from a single nucleotide change in one human proα1(I) collagen allele. *Proc. Natl. Acad. Sci. USA.* 83:6045–6047.

8. Vogel, B. C., R. R. Minor, M. Freund, and D. J. Prockop. 1987. A point mutation in a type I procollagen gene converts glycine 748 of the α1 chain to cysteine and destabilizes the triple helix in a lethal variant of osteogenesis imperfecta. *J. Biol. Chem.* 262:14737–14744.

9. 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 α1(I) chains of type I collagen. *J. Biol. Chem.* 262:7021–7027.

10. Bateman, J. F., S. R. Lamande, H.-H. M. Dahl, D. Chan, and W. G. Cole. 1988. Substitution of arginine for glycine 664 in the collagen α1(I) chain in lethal perinatal OI. *J. Biol. Chem.* 263:11627–11630.

11. Kuivaniemi, H., C. Sabol, G. Tromp, M. Sippola-Thiele, and D. J. Prockop. 1988. A 19-base pair deletion in the proα2(I) gene of type I procollagen that causes in-frame RNA splicing from exon 10 to exon 12 in a proband with atypical osteogenesis imperfecta and in his asymptomatic mother. *J. Biol. Chem.* 263:11407–11413.

12. Tromp, G., and D. J. Prockop. 1988. Single base mutation in the proα2(I) gene of type I procollagen that causes efficient splicing of RNA from exon 27 to exon 29 and synthesis of a shortened but in-frame proα2(I) chain. *Proc. Natl. Acad. Sci. USA.* 85:5254–5258.

13. Pihlajaniemi, T., L. A. Dickson, F. M. Pope, V. R. Korhonen, A. C. Nicholls, D. J. Prockop, and J. C. Myers. 1984. Osteogenesis imperfecta: cloning of a proα2(I) collagen gene with a frameshift mutation. *J. Biol. Chem.* 259:12941–12944.

14. Barsh G. S., K. E. David, and P. H. Byers. 1982. Type I osteogenesis imperfecta: a nonfunctional allele for proα1(I) chains of type I procollagen. *Proc. Natl. Acad. Sci. USA.* 79:3838–3842.

15. Rowe, D. W., J. R. 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.* 71:689–697.

16. Genovese, C., and D. W. Rowe. 1987. Analysis of cytoplasmic and nuclear messenger RNA in fibroblasts from patients with type I osteogenesis imperfecta. *Methods Enzymol.* 145:223–235.

17. Bonadio, J. F., and P. H. Byers. 1985. Subtle structural alterations in the chains of type I procollagen produce osteogenesis imperfecta type II. *Nature (Lond.).* 316:363–366.

18. Bonadio, J., K. A. Holbrook, R. E. Gelinas, J. Jacob, and P. H. Byers. 1985. Altered triple helical structure of type I procollagen in lethal perinatal osteogenesis imperfecta. *J. Biol. Chem.* 260:1734–1742.

19. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (Lond.).* 227:680–686.

20. Bruckner, P., and D. J. Prockop. 1981. Proteolytic enzymes as probes for the triple helical conformation of collagen. *Anal. Biochem.* 110:360–368.

21. Lichtenstein, J. R., P. H. Byers, B. D. Smith, and G. R. Martin. 1975. Identification of the collagenous proteins synthesized by cultured cells from human skin. *Biochemistry.* 14:1589–1594.

22. Andrews, P. C., and J. E. Dixon. 1987. A procedure for in situ alkylation of cysteine residues on glass fiber prior to protein microsequence analysis. *Anal. Biochem.* 161:524–528.

23. Eyre, D. R., and J. Wu. 1983. Collagen of fibrocartilage: a distinctive molecular phenotype in bovine meniscus. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* 158:265–270.

24. Eyre, D. R., F. D. Shapiro, and J. F. Aldridge. 1985. A heterozygous collagen defect in a variant of the Ehlers-Danlos syndrome type VII: evidence for a deleted amino-telopeptide domain in the proα2(I) chain. *J. Biol. Chem.* 260:11322–11329.

25. Bornstein, P. 1970. Structure of α1-CB8, a large cyanogen bromide produced fragment from the α1 chain of rat collagen: the nature of a hydroxylamine-sensitive bond and composition of tryptic peptides. *Biochemistry.* 9:2408–2421.

26. Cohn, D. H., S. Apone, D. R. Eyre, B. J. Starman, P. Andreassen, H. Charbonneau, A. C. Nicholls, F. M. Pope, and P. H. Byers. 1988. Substitution of cysteine for glycine within the carboxyl-terminal telopeptide of the α1(I) chain of type I collagen produces mild osteogenesis imperfecta. *J. Biol. Chem.* 263:14605–14607.

27. Sykes, B., D. Ogilvie, P. Wordsworth, J. Anderson, and N. Jones. 1986. Osteogenesis imperfecta is linked to both type I collagen structural genes. *Lancet.* ii:69–72.

28. Wallis, G., P. Beighton, C. Boyd, and C. G. Mathew. 1986. Mutations linked to proα1(I) collagen gene are responsible for several cases of osteogenesis imperfecta type I. *J. Med. Genet.* 23:411–416.

29. Vogel, B. E., R. Doelz, K. E. Kadler, Y. Hojima, J. Engel, and D. J. Prockop. 1988. A substitution of cysteine for glycine 748 of the

α1 chain produces a kink at this site in the procollagen I molecule and an altered *N*-proteinase cleavage site over 225 nm away. *J. Biol. Chem.* 263:19249–19255.

30. Nicholls, A. C., G. Osse, H. G. Schloon, H. G. Lenard, S. Deak, J. C. Myers, D. J. Prockop, W. R. F. Weigel, P. Fryer, and F. M. Pope. 1984. The clinical features of homozygous α2(I) deficient osteogenesis imperfecta. *J. Med. Genet.* 21:257–262.

31. Wenstrup, R. J., P. Tsipouras, and P. H. Byers. 1986. Osteogenesis imperfecta type IV. Biochemical confirmation of genetic linkage to the proα2(I) gene of type I collagen. *J. Clin. Invest.* 78:1449–1455.

32. Wenstrup, R. J., D. H. Cohn, T. Cohen, and P. H. Byers. 1988. Arginine for glycine substitution in the triple helical domain of the products of on α2(I) collagen allele (COL1A2) produces the osteogenesis imperfecta type IV phenotype. *J. Biol. Chem.* 263:7734–7740.

33. Constantinou, C. D., K. B. Nielsen, and D. J. Prockop. 1989. A lethal variant of osteogenesis imperfecta has a single base mutation that substitutes cysteine for glycine 904 of the α1(I) chain of type I procollagen. The asymptomatic mother has an unidentified mutation producing an overmodified and unstable type I procollagen. *J. Clin. Invest.* 83:574–584.

34. de Vries, W. N., and W. J. De Wet. 1986. The molecular defect in an autosomal dominant form of osteogenesis imperfecta: synthesis of type I procollagen containing cysteine in the triple-helical domain of pro-α1(I) chains. *J. Biol. Chem.* 261:9056–9064.

35. Sillence, D. O., K. K. Barlow, W. G. Cole, S. Dietrich, A. P. Garber, and D. L. Rimoin. 1986. Osteogenesis imperfecta type III: delineation of the phenotype with reference to genetic heterogeneity. *Am. J. Med. Genet.* 23:821–832.