Am. J. Hum. Genet. 59:799–809, 1996

Premature Chain Termination Is a Unifying Mechanism for COL1A1 Null Alleles in Osteogenesis Imperfecta Type I Cell Strains

Marcia C. Willing,[1,2] Sachi P. Deschenes,[1,2] Rebecca L. Slayton,[3] and Erik J. Roberts[1,2]

[1]Department of Pediatrics and Colleges of [2]Medicine and [3]Dentistry, University of Iowa, Iowa City

Summary

Nonsense and frameshift mutations, which predict premature termination of translation, often cause a dramatic reduction in the amount of transcript from the mutant allele (nonsense-mediated mRNA decay). In some genes, these mutations also influence RNA splicing and induce skipping of the exon that contains the nonsense codon. To begin to dissect how premature termination alters the metabolism of RNA from the COL1A1 gene, we studied nonsense and frameshift mutations distributed over exons 11–49 of the gene. These mutations were originally identified in 10 unrelated families with osteogenesis imperfecta (OI) type I. We observed marked reduction in steady-state amounts of mRNA from the mutant allele in both total cellular and nuclear RNA extracts of cells from affected individuals, suggesting that nonsense-mediated decay of COL1A1 RNA is a nuclear phenomenon. Position of the mutation within the gene did not influence this observation. None of the mutations induced skipping of either the exon containing the mutation or, for the frameshifts, the downstream exons with the new termination sites. Our data suggest that nonsense and frameshift mutations throughout most of the COL1A1 gene result in a null allele, which is associated with the predictable mild clinical phenotype, OI type I.

Introduction

Osteogenesis imperfecta (OI) type I is a mild form of dominantly inherited brittle-bone disease, which results from decreased production of type I collagen (Sykes et al. 1977; Barsh et al. 1982; Wenstrup et al. 1990). A growing body of evidence suggests that null mutations in COL1A1, the gene that encodes the proα1(I) chain of type I collagen, result in this phenotype (Willing et al. 1990, 1994; Stover et al. 1993; Redford-Badwal et al. 1996). Steady-state levels of mRNA from one COL1A1 allele are decreased in most OI type I cell strains (Rowe et al. 1985; Genovese and Rowe 1987; Willing et al. 1992). Although several molecular mechanisms previously have been shown to alter the amount of COL1A1 mRNA available for translation (Willing et al. 1990, 1994; Stover et al. 1993; Redford-Badwal et al. 1996), data presented in the present paper suggest that premature termination is the most common mechanism by which RNA metabolism is altered.

Mutations leading to premature termination are a unique class of mutations, because, although they are expected to have their effect at the translational level and lead to shortened, truncated polypeptides, it is recognized that their primary effect is on RNA stability, with possibly a secondary effect on RNA splicing (Losson and Lacroute 1979; Baserga and Benz 1988, 1992; Daar and Marquat 1988; Urlaub et al. 1989; Magnus et al. 1990; Lim et al. 1992; Mashima et al. 1992; Dietz et al. 1993; Fisher et al. 1993; Gibson et al. 1993; Belgrader and Maquat 1994; Das et al. 1994). In an effort to understand how premature termination alters the expression of the COL1A1 gene, we have studied nonsense and frameshift mutations distributed over exons 11–49 of the gene, and we have determined the effect of each mutation on steady-state levels of mRNA and on splice site selection. Our data indicate that mutations that predict premature termination reduce steady-state amounts of COL1A1 mRNA from the mutant allele in both total cellular and nuclear RNA extracts. We find no evidence for a polar effect of mutation location on RNA levels. Although exon skipping has been described in association with premature termination in some genes, we did not detect this phenomenon in our cell strains. Our data suggest that premature-termination mutations have a predictable and uniform effect on COL1A1 gene expression, which ultimately leads to decreased production of type I collagen and to the mild phenotype associated with OI type I.

Methods

Inclusion Criteria

Each affected individual met the clinical criteria for OI type I (Sillence et al. 1979). In addition, dermal fibroblasts from each produced approximately half the expected amount of structurally normal type I procolla-

Received March 18, 1996; accepted for publication July 9, 1996.

Address for correspondence and reprints: Dr. Marcia C. Willing, Department of Pediatrics, 2JCP, University of Iowa, Iowa City, IA 52242. E-mail: marcia-willing@uiowa.edu

0002-9297/96/5904-0009$02.00

gen (Wenstrup et al. 1990). There was no evidence of posttranslational overmodification or retained collagenous material within the cells of any of the probands, as is seen in lethal and deforming OI cell strains.

PCR Amplification and SSCP Analysis

Genomic DNA isolated either from lymphocytes or from cultured dermal fibroblasts, of affected individuals, unaffected family members, and of unrelated controls, served as template for PCR amplification (Saiki et al. 1988). The sequences of the oligonucleotide primers used for PCR were derived from published information (D'Alessio et al. 1988; Westerhausen et al. 1991) and are available on request from the authors. PCR products were screened for potential mutations, by SSCP (Orita et al. 1989; Poduslo et al. 1991). Details of the SSCP protocol were as described elsewhere (Willing et al. 1994), except that allele identification was by silver staining, using Typon Silver Sequence film (Promega), rather than by autoradiography.

Sequence Analysis of Mutations

In preparation for sequence determination of mutations, amplified genomic DNA from each proband was directly ligated into the PCR vector (version 2.3), provided in the Invitrogen TA Cloning Kit. Transformants were screened for the presence of insert, by PCR. Aliquots of amplified material from individual clones were electrophoresed in 0.8% NuSieve (FMC Bioproducts). Single bands were excised from the gel and melted; 10-μl aliquots were used as template for sequence determination by the dideoxy-chain termination method (Sanger et al. 1979). Multiple individual subclones were sequenced from each proband to ensure that two or more independent examples of each variant sequence were obtained. When possible, mutations were confirmed by restriction-endonuclease analysis of PCR-amplified genomic DNA (F1, *Pvu*II; F2–F4, *Sfi*I; F5, *Ava*I; F6, *Bsu*36I; F8, *Bsa*HI; and F10, *Tsp*45I).

Determination of Allele-Specific Steady-State Levels of COL1A1 mRNA

RNA isolation.—Dermal fibroblasts were incubated with ascorbic acid (50 μg/ml) for 4 d, with daily medium changes, prior to isolation of total cellular and nuclear RNA. Total cellular RNA was isolated from dermal fibroblasts by the guanidinium thiocyanate-phenol-chloroform method of Chomczynski and Sacchi (1987). In preparation for isolation of nuclear RNA, a second set of fibroblasts was pelleted and washed twice with PBS at 4°C. Nuclei were isolated after the cells were incubated with NP-40 lysis buffer (10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, and 0.5% NP-40). To eliminate potential contamination with cytoplasmic RNA, nuclei were washed and repelleted several times, with fresh lysis buffer (Greenberg 1987). All washings were done at 4°C. RNA was extracted from nuclei by the method of Chomczynski and Sacchi (1987).

Restriction-endonuclease analysis of PCR-amplified cDNA.—The Perkin-Elmer GeneAmp RNA PCR Kit (Perkin Elmer) was used to synthesize and amplify cDNA derived from RNA transcripts of both COL1A1 alleles. In each case, the primer that was used to prime the RT reaction also served as the 3′ PCR primer. To optimize detection of small amounts of mutant mRNA, 35 amplification cycles were performed for each sample. Mutations in families F1–F6, F8, and F10 either disrupted or created a restriction-endonuclease recognition site. After the PCR, aliquots of each amplified cDNA product were cleaved with the appropriate enzyme (F1, *Pvu*II or *Msp*A1I; F2–F4, *Sfi*I; F5, *Ava*I; F6, *Ava*I; F8, *Bsa*HI; and F10, *Tsp*45I); restriction products were separated in a 6% polyacrylamide gel and were identified by ethidium bromide staining. Analyses were repeated at least twice for each mutation. Sequences of the oligonucleotide primers used for reverse-transcriptase PCR (RT-PCR) reactions were derived from published sequence information and are available on request.

Allele-specific oligonucleotide hybridization (ASOH).—ASOH was used to assess steady-state levels of mRNA from normal and mutant COL1A1 alleles, in families F7 and F9. cDNA was prepared by RT-PCR, from total cellular and nuclear RNA extracts, as above. Amplified products were separated by electrophoresis in 1.6% agarose and were transferred to Zeta-Probe nylon membranes (Bio-Rad) by an alkaline downward-transfer protocol (Zhou et al. 1994). Nucleic acid was bound to the membranes by UV cross-linking, followed by baking at 80°C for 30 min. Membranes were prehybridized for 30 min at 65°C in 1 mM EDTA, 0.5 M Na_2HPO_4, and 7% SDS. Hybridization was performed at 65°C for 6 h with either normal or mutant [γ-^{32}P] dATP (3,000 Ci/mmol; Amersham) end-labeled oligonucleotides (F7 normal allele, 5′ GGCCCCCCTGGTGCTGACGG 3′; F7 mutant allele, 5′ GGCCCCCCT*T*GGTGCTGACGG 3′; F9 normal allele, 5′ GGTCCCCCTGGCCCTGCCGG 3′; and F9 mutant allele, 5′ GGTCCCCCCTGGCCCTGCCG 3′). After the hybridization, membranes were washed twice at room temperature in 1 mM EDTA, 40 mM Na_2HPO_4, and 5% SDS for 5 min each, then in 1 mM EDTA, 40 mM Na_2HPO_4, and 1% SDS at 68°C for 10 min. Radiographic signals were identified by autoradiography using Kodak XAR-5 X-ray film. Duplicate hybridizations were performed for all samples.

Primer extension at the polymorphic MnlI *site.*—Steady-state levels of allele-specific mRNA in total cellular and nuclear RNA extracts were assessed for *Mnl*I heterozygotes (F1 and F2) by primer extension (Daar and Maquat 1988). Methodological details are as described elsewhere (Willing et al. 1992). The oligonucleotide, 5′ CAGGTACACTTTAACAGAGGGT 3′, was used to prime the extension reaction.

PCR Amplification with F1 Cloned cDNAs

RT-PCR was performed with 1 μg of F1 fibroblast nuclear RNA as template; amplification primers were homologous to sequences in exons 10 and 14. The PCR product was ligated into the vector provided in the Invitrogen TA Cloning Kit. Individual colonies were screened for the presence of insert, by PCR; normal and mutant clones were distinguished by cleavage of the PCR product by *Msp*A1I (the mutation disrupts the restriction site). One normal and one mutant clone were cultured separately, overnight, in LB broth with 50 μg ampicillin/ml. DNA was prepared by using the QIAprep Spin Plasmid Miniprep Kit, and stock solutions (30 fmol/μl) were made for both DNA preparations. Serial 10-fold dilutions (3, 0.3, 0.03, and 0.003 fmol/μl) were made for both cDNAs, starting from the original stocks; 2-μl aliquots of each dilution served as template for PCR amplification. Each amplification reaction contained 400 μM dNTPs, 500 mM KCl, 100 mM Tris-HCl pH 8.3, 2.0 mM $MgCl_2$, 0.1% gelatin, 2 × Triton X-100, 10 mM DTT, 10% glycerol, 150 ng of each oligonucleotide primer, 0.5 units of *Taq* polymerase, and 12.5 μCi of [α-^{32}P]ATP (3,000 Ci/mmol; Amersham). Normal and mutant plasmid DNA at each dilution were amplified separately, for 20 cycles (the logarithmic phase of amplification for this fragment). An aliquot of each final PCR product was separated in 6% acrylamide and was identified by autoradiography. In a second set of PCR reactions, equal amounts of normal and mutant plasmid DNA at each dilution were amplified together for 20 cycles. A 5-μl aliquot of each PCR product was digested with *Msp*A1I, and products derived from the normal and mutant alleles were identified by autoradiography, following electrophoresis as described above. Individual bands were quantified by an Ambis β-scanner.

Identification of Splicing Products by RT-PCR

To screen for products of aberrant splicing (exon skipping) that result from premature termination, RT-PCR was used to synthesize and amplify the domains of COL1A1 cDNA surrounding each mutation site; for the frameshift mutations, the exons surrounding their respective downstream termination codons were also amplified. For the F1, F3, F7, F9, and F10 mutations, the new termination signals were near the original mutations and therefore were included within the original amplified cDNA fragments. The F2, F4, and F8 mutations led to termination in more distant exons and therefore required additional amplifications, to span these domains. The amplification strategy for each mutation is outlined in table 1. Primers were derived from published sequence information, as noted above. RT-PCR products were separated by electrophoresis in 6% polyacrylamide and were identified by ethidium bromide staining.

Sequence Analysis of cDNA Clones to Determine the Effect of the F7 Mutation on Splicing

To determine the effect that the insertion mutation in family F7 (a mutation located in the last codon of exon 36) has on splicing, sequence analysis was performed on cDNA clones. Total cellular and nuclear RNA served as templates for RT-PCR using the Perkin Elmer GeneAmp RNA PCR Kit. Amplified material was cloned as described above, and sequence was determined on 139 clones (76 and 63 derived from total and nuclear RNA, respectively), by the dideoxy-chain termination method (Sanger et al. 1979).

Results

Mutation Identification

All mutations were identified by SSCP analysis of PCR-amplified genomic DNA. Prior to analysis, however, we suspected the presence of a deletion in families F8 and F10, because of heteroduplex formation identified by PAGE. SSCP primer pairs were designed to amplify 150–250-bp fragments that included intron-exon junctions and coding sequence, in order to optimize identification of donor and acceptor splice-site mutations, as well as mutations located in exons. Affected individuals from each family had an aberrant banding pattern that was not observed either in unaffected relatives or in controls. Two families (F2 and F3) had the same variant SSCP pattern, whereas the proband in family F4 had a similar but distinct banding pattern for the same domain. Affected individuals in families F8 and F9 had anomalous but different SSCP banding patterns for the COL1A1 domain that spanned exon 37 and surrounding introns (data not shown).

Sequence Analysis of Mutations

We identified eight out-of-frame deletion/insertions and two nonsense mutations in unrelated OI type I fami-

Table 1

Analysis of Splice-Site Selection in RNA of Probands, by RT-PCR

Family	Mutation Location (Exon)	Termination Location (Exon)	RNA Domain(s) Amplified (Exons)
F1	11	11	10–14
F2	17	24	14–19, 21–31
F3	17	18	14–19
F4	17	24	14–19, 21–31
F5	19	19	18–21
F6	31	31	29–34
F7	36	37	34–40
F8	37	46	34–40, 44–50
F9	37	38	34–40
F10	49	50	46–51

NOTE.—Total cellular RNA from each proband served as template for RT-PCR.

lies (fig. 1). All mutations are found in exons and occur in the heterozygous state. Families F1, F2, and F4 have single-base-pair deletions, whereas families F3, F7, and F9 have single-base-pair insertions. Families F8 and F10 have 16-bp and 13-bp deletions, respectively, which are the largest ones that we have detected in OI type I cell strains. The remaining families (F5 and F6) have single-base-pair substitutions that change arginine codons (CGA) to termination codons (TGA), in exons 19 and 31, respectively.

Mutations are distributed throughout the gene, from exon 11 (F1) to exon 49 (F10) (fig. 1*B*). As suggested by SSCP analysis, families F2, F3, and F4 have mutations in the same domain, as do families F8 and F9. The mutations in families F2, F3, and F4 cluster together in or immediately after six consecutive cytosine nucleotides (GGC CCC CCT), in exon 17. Families F2 and F3 have a single cytosine deletion and insertion, respectively, in this group of nucleotides, whereas family F4 has a deletion of the thymidine that immediately follows the cluster of cytosines. The mutation in family F9 also involves the insertion of a single cytosine in a similar sequence (GGT CCC CCT), in exon 37.

Each deletion/insertion leads to a shift in the translational reading frame and predicts premature termination downstream of the original mutation site, whereas the substitutions in families F5 and F6 directly produce premature termination, in exons 19 and 31, respectively. For the former group, the locations of the new termination codons vary, from either six amino acid residues away in the same exon (F1) or in close proximity in the next exon (F3 and F7) to much greater distances from the original mutation site (e.g., 266 amino acids downstream, in exon 46, for the F8 mutation). If the mRNA product of the mutant allele were translated, each deletion/insertion would result in a truncated-protein species that is missing variable portions of the triple-helical domain and the carboxy-terminal telopeptide and propeptide extension. Only the F10 mutation, located in exon 49, with a predicted termination signal in exon 50, would result in a protein with an intact triple-helical domain. The mutations in families F5 and F6 would also be expected to produce shortened proα chains, but, because they do not involve translational frameshifts, there would be no anomalous protein beyond the mutation site.

Premature Termination and Its Effects on Steady-State Amounts of COL1A1 mRNA

To determine the effect of each mutation on steady-state levels of COL1A1 mRNA, we used either restriction-endonuclease cleavage of PCR-amplified cDNA or hybridization of allele-specific oligonucleotides to cDNA, prepared from total cellular and nuclear RNA extracts. The mutations identified in families F1–F6, F8, and F10 either disrupt or create restriction-endonuclease recognition sites. For all cell strains except F1, mutation-specific endonuclease cleavage of amplified cDNA derived from total cellular and nuclear RNA yields products primarily from the normal allele, although restriction fragments derived from both COL1A1 alleles are present in approximately equal amounts in genomic DNA (fig. 2). Small amounts of mRNA from the mutant alleles are visible after 35 cycles of amplification in RNA extracts from the F2, F5, and F6 probands, and, although we were unable to visualize restriction products from the mutant allele in F10 RNA, the presence of heteroduplexes suggests that at least small amounts of mutant mRNA must also be available to serve as template for RT-PCR in this cell strain.

ASOH was used to assess steady-state levels of COL1A1 mRNA for the F7 and F9 mutations, since they were not amenable to analysis by restriction-endonuclease cleavage. Although, for these probands, cDNA products are readily identified by PAGE (fig. 3, rows A), hybridization with the respective mutant oligonucleotides detects almost no product from either mutant allele (fig. 3, rows C). In contrast, there are strong signals when the normal oligonucleotides are used as hybridization probes (fig. 3, rows B). These data indicate that, as is true for the other COL1A1 mutations, there is little detectable mRNA from the mutant allele in cells from the F7 and F9 probands.

The F1 cell strain is the only one in which we identified large amounts of mRNA from the mutant allele in both total cellular and nuclear RNA extracts (fig. 2*B*). Additional data from genomic DNA studies argue that this mutation predisposes to preferential amplification of products from the mutant allele. When the F1 mutation was initially sequenced, all of our clones had the deletion, suggesting that either there was preferential amplification of one allele in the genomic DNA or the proband was homozygous for the mutation. We thought that the latter alternative was unlikely, in view of the mild phenotype. Since the mutation disrupts both *Pvu*II and *Msp*A1I restriction-endonuclease sites, we used *Pvu*II to cleave an aliquot of the original PCR product that was used for genomic cloning. This yielded only the uncleaved product of the mutant allele, confirming the allele-specific–amplification hypothesis (fig. 4). With selection of different PCR primers, we eventually amplified and identified both alleles in genomic DNA (figs. 1*A* and 2*A*). Since the same sequence is found in the mRNA, this might also contribute to its preferential amplification as well and might give us the results shown in figure 2*B*. We used multiple sets of primers and conditions to determine whether this would alter the amount of RT-PCR product derived from the mutant allele, but the same result was obtained each time.

To further address the issue of preferential amplification of the F1 mutant-gene product, we used fixed amounts of normal and mutant cDNA isolated from

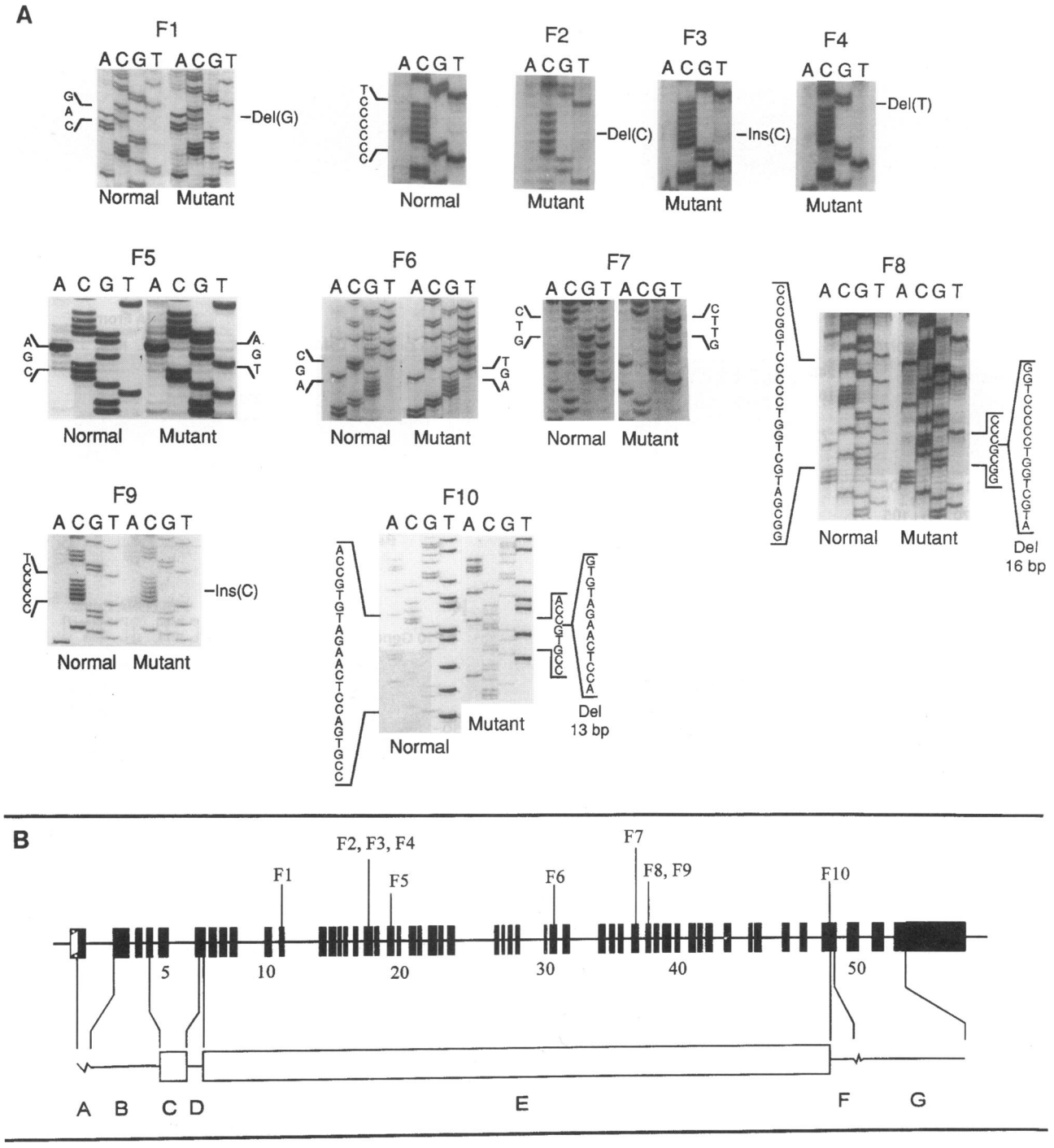

Figure 1 Sequence and location of COL1A1 mutations in OI type I families. *A*, Nucleotide sequence of normal and mutant alleles in cloned genomic DNA from OI probands in families F1–F10. Normal and variant sequences are shown for the region surrounding each mutation; deletions (Del) and insertions (Ins) are depicted. Nucleotide sequence from the sense strand is shown for F1–F5 and F7–F10, while the antisense strand is displayed for the F6 mutation. *B*, Location of mutations in the COL1A1 gene. Blackened boxes and horizontal lines symbolize exons and introns, respectively. Exon numbers appear below the blackened boxes. Corresponding protein domains are shown below the diagram: A = signal sequence; B = amino-terminal propeptide globular domain; C = amino-terminal propeptide triple-helical domain; D = amino-terminal telopeptide; E = triple helix; F = carboxyl-terminal telopeptide; and G = carboxyl-terminal propeptide.

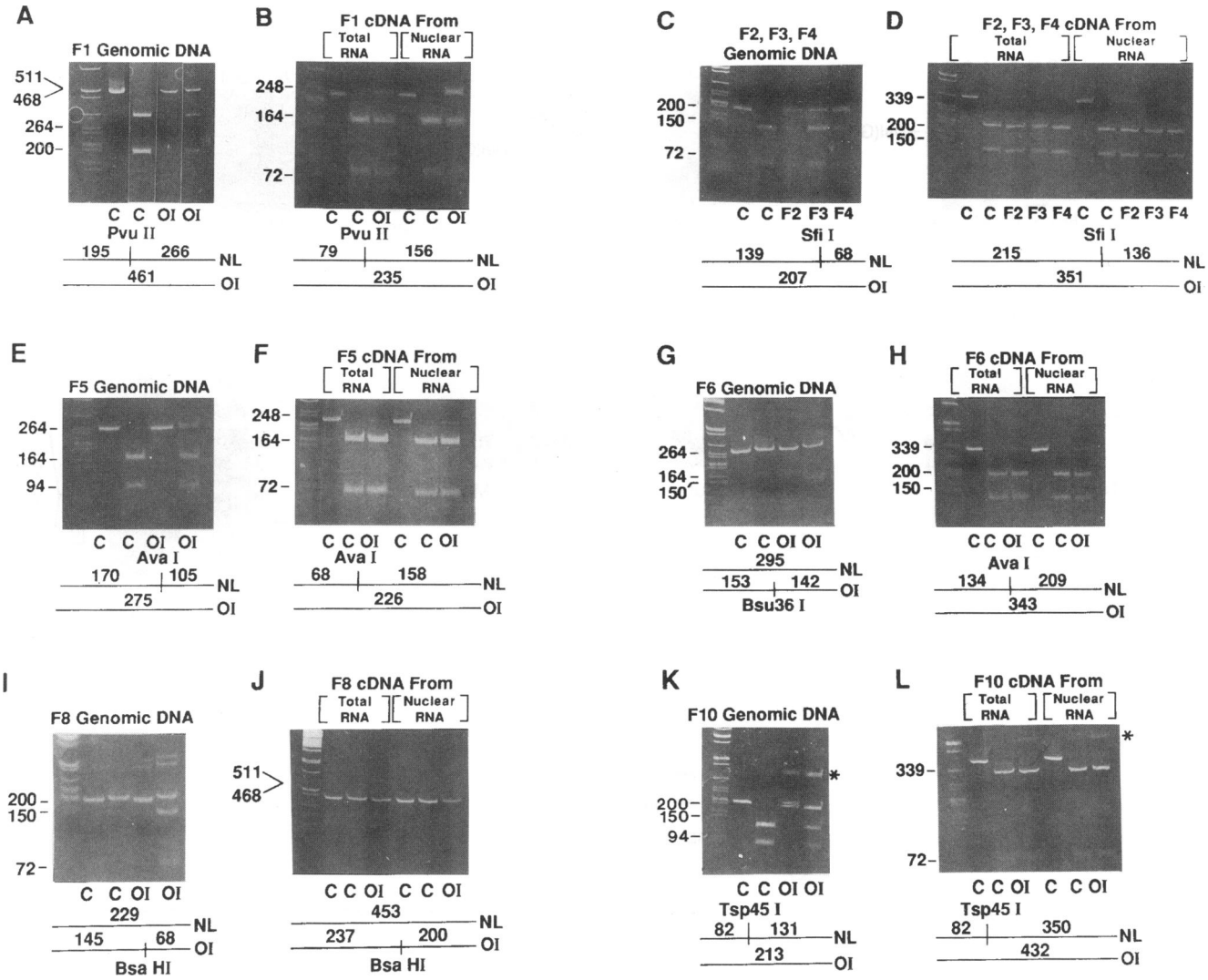

Figure 2 Mutation-specific restriction-endonuclease cleavage of PCR-amplified genomic DNA and cDNA from control (lanes C) and OI (lanes OI) probands in families F1–F6, F8, and F10. Mutation-specific cleavage of PCR-amplified genomic DNA from each proband demonstrates the presence of two COL1A1 alleles (*A, C, E, G, I,* and *K*). Cleavage of cDNA derived from total cellular and nuclear RNA extracts yields products primarily from the normal allele (*B, D, F, H, J,* and *L*). In *K* and *L*, an asterisk (*) identifies heteroduplexes.

individual clones, amplified them separately and together, and examined the end products first by ethidium bromide staining and then by radioactive labeling during the PCR (fig. 5). Allele-specific products were distinguished by *Msp*A1I cleavage, when normal and mutant cDNAs were competitively amplified together. Regardless of whether the mutant cDNA is amplified alone or in conjunction with identical amounts of normal cDNA, the end product of the mutant cDNA template is always more than that of the normal cDNA, at every dilution tested. The discrepancy is most dramatic when normal and mutant cDNAs are amplified together; quantification indicates that the amount of final PCR product with the normal cDNA is 12%–26% of that in the mutant, depending on the dilution. These data confirm our hypothesis of preferential amplification.

Verification of PCR-Based Results with Primer Extension

To confirm PCR-based assessments of RNA levels in OI cell strains, we used primer extension at the polymorphic *MnlI* site at the 3′ end of the COL1A1 gene, to quantify steady-state amounts of allele-specific mRNA in total cellular and nuclear RNA extracts. Probands F1 and F2 were heterozygous at the *MnlI* site and therefore were amenable to analysis. Steady-state amounts of mRNA from the mutant allele were reduced in total cellular and nuclear RNA extracts from both probands (fig. 6), verifying the RT-PCR results. These data support previous primer-extension studies on other OI type I cell strains, which identified the nucleus as the primary site of mutant RNA degradation (Willing et al. 1994).

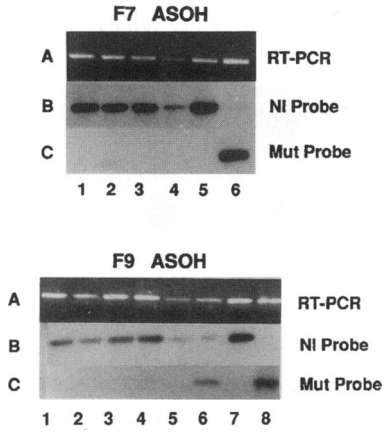

Figure 3 ASOH of PCR-amplified cDNA from F7 and F9 probands. Rows A, Total cellular and nuclear RNA from control and OI cells (F7 [*upper panel*] and F9 [*lower panel*]) served as template for RT-PCR (lanes 1–4 in both panels). Amplified genomic or cloned material was included to verify hybridization stringency (*upper panel,* lanes 5 and 6; and *lower panel,* lanes 5–8). Rows B, Hybridization of [γ^{32}P] end-labeled normal oligonucleotide probe to PCR products blotted onto nylon. Rows C, Hybridization of [γ^{32}P] end-labeled mutant oligonucleotide probe to PCR products blotted onto nylon. F7 and F9 ASOH lanes are as follows: lane 1, control total RNA; lane 2, control nuclear RNA; lane 3, OI total RNA; and lane 4, OI nuclear RNA. Hybridization stringency for the F7 ASOH was verified by including amplified DNA from normal (*upper panel,* lane 5) and mutant (*upper panel,* lane 6) cDNA clones. Hybridization stringency for the F9 ASOH was verified by including amplified genomic DNA from the control (*lower panel,* lane 5) and the F9 proband (*lower panel,* lane 6), as well as amplified DNA from normal (lane 7) and mutant (lane 8) genomic clones. Rows A demonstrate the presence of amplified material in all lanes. The normal probes hybridize to RT-PCR material in lanes 1–4 and to the hybridization controls. The mutant probe fails to hybridize to RT-PCR material from either proband (lanes 3 and 4 in both panels), demonstrating decreased amounts of mRNA from the mutant allele in total cellular and nuclear RNA extracts. The mutant probe hybridizes to appropriate hybridization controls, however (*upper panel,* lane 6; and *lower panel,* lanes 6 and 8).

Premature Termination and Its Effect on Splicing

To determine whether our mutations alter RNA splicing, we amplified cDNA fragments that included both the site of the original mutation and, for the frameshift mutations, the downstream exons that contain the new termination signals (table 1). For the F1, F3, F7, F9, and F10 mutations, premature termination occurs in close proximity to the original mutation, making it possible to amplify cDNA fragments that spanned 4–6 exons and included both sites. For the F2, F4, and F8 mutations, where termination occurs at more distant sites, additional amplification reactions were performed that included these downstream sites. Using this approach, we found no evidence of either shortened products or heteroduplexes that might suggest the presence of exon skipping. For the F2, F4, and F7 mutations, we identified, in addition to the expected product, at least one additional cDNA fragment, but it was also present in controls, suggesting that it is a nonspecific finding and is not a consequence of aberrant splicing. These data suggest that premature termination in these COL1A1 alleles does not predispose to significant exon skipping.

To further investigate the relationship between premature termination and splicing, we sequenced cDNA clones prepared from total cellular and nuclear RNA extracts from an affected member of family F7. Since the F7 mutation predicts premature termination at the 5′ end of exon 37 and has the potential to alter recognition of the 5′ splice site for this exon (the insertion mutation occurs in the last codon of exon 36), we thought that it would be the most likely, of the 10 mutations, to produce aberrant splicing. When 76 cDNA clones derived from total RNA were sequenced, we identified 51 clones with the normal sequence and 6 with the insertion. In all cases, exons 36–38 were intact and correctly spliced. The remaining 19 clones, of varying sizes, were sequenced to determine whether they represented products of alternative splicing; none were derived from the COL1A1 gene. Sequence results for the cDNA clones derived from nuclear RNA were identical, including the proportions of normal and mutant clones. Thus, although two different mechanisms (premature termination and alteration of donor splice-site recognition) might have predisposed to alternative splicing, we found, in our sequence analysis, no support for this phenomenon.

Discussion

The ability of premature translation termination to reduce steady-state levels of the corresponding mRNA (nonsense-mediated mRNA decay) is a phenomenon common to most eukaryotes, including yeast, plants, *Caenorhabditis elegans, Drosophila,* humans, and other mammals (Losson and Lacroute 1979; Baserga and Benz 1988, 1992; Daar and Marquat 1988; Urlaub et al. 1989; Magnus et

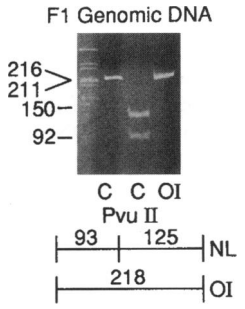

Figure 4 Preferential amplification of the mutant COL1A1 allele in F1 genomic DNA. Amplified genomic DNA from the control (lanes C) and F1 proband (lane OI), cleaved with *Pvu*II: lane 1, control DNA before cleavage; lane 2, control DNA cleaved; and lane 3, OI DNA cleaved. Control DNA is cleaved into two fragments, demonstrating the presence of only the normal COL1A1 alleles, whereas cleavage of amplified DNA from the F1 proband yields only the 218-bp product of the mutant allele.

al. 1990; Belgrader and Maquat 1994; Belgrader et al. 1994; Zhang et al. 1994). Several studies have demonstrated that transcription efficiency is not altered by the presence of nonsense codons, suggesting that the effect is primarily on RNA stability (Losson and Lacroute 1979; Urlaub et al. 1989; Baserga and Benz 1992). Recent analyses of frameshift and nonsense mutations in the human triosephosphate isomerase (TPI) gene indicate that nuclear mRNAs harboring nonsense codons are targeted for destruction after they are fully processed, suggesting that the mRNA is lost during or immediately before transport across the nuclear membrane (Belgrader et al. 1994). Although our analyses did not distinguish between unprocessed (heterologous RNA [hnRNA]) and mature mRNA, they support the same general conclusion.

The results of the F1 mutation deserve additional comment, since, at first glance, they appear to be at odds with data on the other COL1A1 mutations. Our initial genomic DNA data (fig. 4) suggested that this mutation predisposes to preferential amplification of products

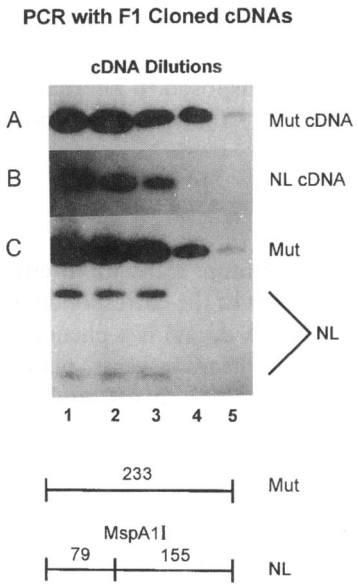

Figure 5 Amplification of mutant and normal cloned F1 cDNA. Rows A and B, Standardized amounts of 10-fold serial dilutions (lanes 1–5) of normal and mutant cloned cDNAs amplified separately for 20 cycles at 94°C (1.5 min), 60°C (30 s), and 72°C (30 s). Row C, Standardized amounts of normal and mutant cDNAs amplified together, then cleaved with *Msp*A1I to identify allele-specific PCR products. Products of the mutant allele remain uncleaved: lane 1, 60 fmol; lane 2, 6 fmol; lane 3, 0.6 fmol; lane 4, 0.06 fmol; and lane 5, 0.006 fmol. When amplified separately, the final PCR product of the normal cDNA is less than that of the mutant cDNA, at each concentration. Quantification with an Ambis beta-scanner indicates that the PCR products of the normal cDNA represent 6%–54% of the mutant, for starting concentrations of 0.006 fmol and 60 fmol, respectively. When fixed amounts of normal and mutant cDNA are competitively amplified together, the discrepancy is more obvious; the PCR product of the normal cDNA never represents >26% of that of the mutant cDNA. Normal and mutant cDNAs were isolated from individuals clones prepared from F1 fibroblast nuclear RNA.

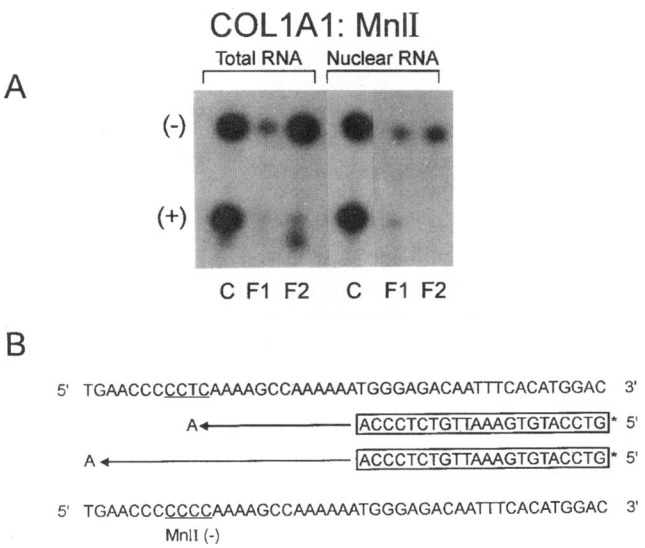

Figure 6 Primer extension at the polymorphic COL1A1 *Mnl*I site. *A*, Total cellular and nuclear mRNA from control cells (lane C), which direct the synthesis of approximately equal quantities of cDNA derived from COL1A1 *Mnl*I (+) and (−) alleles. In contrast, total cellular and nuclear mRNA from OI individuals in families F1 and F2 (lanes F1 and F2) direct the synthesis of predominantly one cDNA species. *B*, Schematic of the primer-extension reaction. The oligonucleotide primer end-labeled (denoted by an asterisk [*]) with ^{32}[P]γ ATP is enclosed by boxes, and the *Mnl*I site is underlined. Terminations (leftmost A's) are reached 15 nucleotides upstream of the primer in the *Mnl*I (+) allele and 24 nucleotides away in the *Mnl*I (−) allele.

from the mutant allele. Precedence for this phenomenon has already been described elsewhere (Eldadah et al. 1995). Preferential amplification of one allele could be related to a secondary-structure change that results from the mutation. In the F1 cell strain, the G deletion is located in the middle of the 54-bp exon 11 and falls within a short palindrome (CAGCTG). We considered the possibility that deletion of the G disrupts a short hairpin structure, making PCR amplification of the mutant allele kinetically more favorable. Our amplification data derived by use of fixed amounts of normal and mutant cloned cDNA support this hypothesis. Finally, to corroborate PCR-based analyses, we used primer extension at the *Mnl*I site, a site distant to that of the mutation, and showed that steady-state levels of mRNA from the mutant allele are reduced in both total cellular and nuclear RNA extracts. These data, taken together, argue that premature termination promotes RNA degradation in the F1 cell strain as well.

Our findings of reduced steady-state levels of mutant COL1A1 mRNA in the nucleus of the cell appear to contradict studies by Stover et al. (1993) and Redford-Badwal et al. (1996), which demonstrate nuclear retention of aberrant mRNAs in OI type I cell strains. In their most recent report, null alleles are generated by exclusion of RNA transcripts from the cytoplasm, thus eliminating their translation into protein and their incor-

poration into heterotrimers (Redford-Badwal et al. 1996). Included in their study is a cell strain, referred to as "patient 263," which we had characterized previously by primer extension (the F7 cell strain in Willing et al. 1994) and had reported to have markedly reduced steady-state levels of mutant mRNA in both total cellular and nuclear RNA extracts. We have not looked for mutant transcripts in the cytoplasm of this particular cell strain or in any of the others described here, and thus we are not certain whether the small amount of mutant RNA detected in some cell strains is preferentially retained in the nucleus.

Instead, we have focused on nuclear events that may be important in nonsense-mediated mRNA decay. Using quantitative RT-PCR with nuclear RNA extracts, we find that levels of hnRNA from mutant and normal alleles are approximately equal, whereas the mutant transcript is significantly reduced in processed mRNA, in comparison with that in the normal allele (Slayton et al. 1995). In most of the 12 cell strains examined to date (including Redford-Badwal's patient 263), the mutant mRNA represents ~5%–10% of the normal mRNA, suggesting that the RNA product from the mutant allele is lost either during or immediately after splicing. In view of this finding, we wondered whether the ability of Redford-Badwal et al. to detect retained mRNA within the nucleus of OI type I cell strains reflects differences in the efficiency of amplification of RNA transcripts, which are present in relatively low copy number, in conjunction with the use of SSCP to identify these transcripts. Their PCR buffer is different from ours (we use the buffer in the Perkin Elmer GeneAmp RNA PCR Kit) and includes both β-mercaptoethanol and 1% polyoxyethylene-9-lauryl ether, which may be relevant as a means of explaining potential amplification differences. Regardless of when the mutant mRNA is lost, the effect on type I collagen production is still the same, however.

In some genes, position of the nonsense mutation influences mRNA stability, such that mutations located more 5′ in the gene decrease mRNA levels more dramatically than do those toward the 3′ end of the gene (Urlaub et al. 1989; Cheng et al. 1990; Mashima et al. 1992; Zhang et al. 1994). Mutations involving the last exon of a gene typically do not trigger mRNA destruction, suggesting that genes have boundaries beyond which nonsense mutations are not recognized. The position effect of the last exon also appears to be true in the COL1A1 gene. Whereas nonsense and frameshift mutations distributed throughout most of the gene uniformly lead to decreased steady-state levels of allele-specific mRNA, a 5-bp deletion in the last exon of the COL1A1 gene results in approximately equal amounts of mRNA from both COL1A1 alleles, in the nucleus and cytoplasm of cells from affected individuals (Willing et al. 1990). With the mutations presented here, we found no evidence for a linear relationship between mutation location and steady-state levels of mutant mRNA. These data corroborate previous studies of nonsense and frameshift mutations in exons 22, 46, 47, and 49 of the COL1A1 gene (Willing et al. 1994) and suggest that premature termination has a uniform effect on mRNA stability throughout the gene.

We also examined the effect of premature termination on splice-site selection. In many genes, premature termination influences RNA processing and leads to skipping of the exon containing the new termination codon (Dietz et al. 1993; Fisher et al. 1993; Gibson et al. 1993; Belgrader et al. 1994; Das et al. 1994). Location of the mutation within the exon, its effect on RNA secondary structure, and the order of intron removal for the gene may influence this phenomenon, since not all nonsense mutations within the same gene facilitate exon skipping (Belgrader et al. 1994). Although the mutations described here are distributed throughout the COL1A1 gene, at varying distances from 5′ and 3′ splice sites, we found no evidence for exon skipping in any of these COL1A1 alleles. Although it is possible that RT-PCR did not detect infrequent alternative splicing events either at the mutation sites or at more distant sites, we think that it is more likely that our ability to detect exon skipping was biased by studying only individuals with the mild OI type I phenotype. Exon skipping in the COL1A1 gene often leads to lethal or severe deforming OI (Byers 1993). For a structural macromolecule such as type I collagen, where mutations have a dominant-negative effect, restoration of an open reading frame to create a more functional protein would not have an ameliorating effect on the clinical phenotype, as has been suggested for recessive disorders. It should be noted that, in the fibrillin gene, which encodes another extracellular matrix protein, Dietz et al. (1993) observed exon skipping as a result of a nonsense mutation in a patient with severe manifestations of Marfan syndrome.

Several models have been proposed to explain the phenomenon of nonsense-mediated mRNA decay (reviewed in Urlaub et al. 1989), but none adequately explains all of the available experimental data. The translational translocation model suggests that transport of mature mRNA out of the nucleus is coupled with the splicing and translation processes, at the nuclear membrane, whereas the nuclear-scanning-of-translational-frames model proposes a mechanism, within the nucleus, that scans RNA molecules for the presence of nonsense codons. It is possible that cells have more than one mechanism for recognizing and degrading nonsense-containing mRNAs. Any model proposed to explain nonsense-mediated mRNA decay in the COL1A1 gene must take into account that this gene encodes a secreted protein whose synthesis occurs on membrane-bound ribosomes of the rough endoplasmic reticulum (RER) (Kielty et al. 1993). Unless the RER were an integral part of the nuclear membrane, it would be hard to imag-

ine how the translational translocation model could explain recognition and degradation of mutant COL1A1 mRNA. Instead, we think that it is more likely that there is a nuclear surveillance system that scans the hnRNA for the presence of premature-termination codons. This surveillance mechanism potentially could occur at the time of RNA splicing, when exons are defined. Cis-acting sequences, as well as alterations in RNA secondary structure, might contribute to the recognition process, but trans-acting factors, analogous to the smg and UPF proteins identified in *C. elegans* and yeast, respectively (Leeds et al. 1992; Pulak and Anderson 1993), would also be required to mediate RNA degradation.

Mutations that lead to premature translation termination appear to be the most common molecular cause of OI type I. Of the 21 COL1A1 mutations that we have identified, 15 lead to premature termination, as a result of either translational frameshifts or single nucleotide substitutions (Willing et al. 1994; M.C.W., unpublished data). Another mutation (the 5-bp deletion in exon 52) extends termination to beyond the normal site (Willing et al. 1990). The remaining five mutations are splicing defects, and, of these, four disrupt 5′ donor splice-site recognition and lead to either cryptic splicing or intron retention within the mature mRNA. Both of these alternative splicing pathways indirectly lead to frameshifts and premature termination in downstream exons. A similar observation has been reported by Stover et al. (1993), who described a 5′ donor splice-site mutation, in intron 27, which leads to intron retention in the mature mRNA. Thus, in OI type I cell strains, almost all identified mutations that lead to null alleles do so as a result of premature termination. Premature termination seems to have a predictable effect on steady-state levels of COL1A1 mRNA and, ultimately, on type I collagen and, for this reason, may be an explanation for the relatively uniform clinical phenotype associated with OI type I, compared with the more variable nature of OI types III and IV.

Acknowledgments

We thank the families and referring physicians for their interest and participation. We are grateful to Dr. Peter Byers, both for making OI cell strains available for analysis and for his continued enthusiasm and support. This work was supported by grant AR 40334 from the National Institutes of Health, by a Biomedical Science Grant from the Arthritis Foundation, and by the Roy J. Carver Trust at the University of Iowa.

References

Barsh GS, David KE, Byers PH (1982) Type I osteogenesis imperfecta: a nonfunctional allele for proα1(I) chains of type I procollagen. Proc Natl Acad Sci USA 79:3838–3842

Baserga SJ, Benz EJ (1988) Nonsense mutations in the human β-globin gene affect mRNA metabolism. Proc Natl Acad Sci USA 85:2056–2060

——— (1992) β-Globin nonsense mutation: deficient accumulation of mRNA occurs despite normal cytoplasmic stability. Proc Natl Acad Sci USA 89:2935–2939

Belgrader P, Cheng J, Zhou X, Stephenson LS, Maquat LE (1994) Mammalian nonsense codons can be *cis* effectors of nuclear mRNA half-life. Mol Cell Biol 14:8219–8228

Belgrader P, Maquat LE (1994) Nonsense but not missense mutations can decrease the abundance of nuclear mRNA for the mouse major urinary protein, while both types of mutations can facilitate exon skipping. Mol Cell Biol 4: 6326–6336

Byers PH (1993) Osteogenesis imperfecta. In: Royce PM, Steinmann B (eds) Connective tissues and its heritable disorders. Wiley-Liss, New York, pp 317–351

Cheng J, Fogel-Petrovic M, Maquat LE (1990) Translation to near the distal end of the penultimate exon is required for normal levels of spliced triosephosphate isomerase. Mol Cell Biol 10:5215–5225

Chomczynski P, Sacchi N (1987) Single-step method for RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem 162:156–159

Daar IO, Marquat LE (1988) Premature translation mediates triosephosphate isomerase mRNA degradation. Mol Cell Biol 8:802–813

D'Alessio M, Bernard M, Pretorious P, de Wet W, Ramirez F (1988) Complete nucleotide sequence of the region encompassing the first twenty-five exons of the human proα1(I) collagen gene (COL1A1). Gene 67:105–115

Das S, Levinson B, Whitney S, Vulpe C, Packman S, Gitschier J (1994) Diverse mutations in patients with Menkes disease often lead to exon skipping. Am J Hum Genet 55:883–889

Dietz HC, Valle D, Francomano CA, Kendzior RJ Jr, Pyeritz RE, Cutting GR (1993) The skipping of constitutive exons in vivo induced by nonsense mutations. Science 259:680–682

Eldadah ZA, Grifo JA, Dietz HC (1995) Marfan syndrome as a paradigm for transcript-targeted preimplantation diagnosis of heterozygous mutations. Nat Med 1:798–803

Fisher CW, Fisher CR, Chuang JL, Lau KS, Chuang DT, Cox RP (1993) Occurrence of a 2-bp (AT) deletion allele and a nonsense (G-to-T) mutant allele at the E3 (DBT) locus of six patients with maple syrup urine disease: multiple-exon skipping as a secondary effect of the mutations. Am J Hum Genet 52:414–424

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

Gibson RA, Hajianpour A, Murer-Orlando M, Buchwald M, Mathew CG (1993) A nonsense mutation and exon skipping in the Fanconi anaemia group C gene. Hum Mol Genet 2: 797–799

Greenberg ME (1987) Preparation and analysis of RNA. In: Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (eds) Current protocols in molecular biology. Vol 1: Preparation and analysis of RNA. John Wiley & Sons, New York, p 4.10.2

Kielty CM, Hopkinson I, Grant ME (1993) The collagen family: structure, assembly, and organization in the extracellular

matrix. In: Royce PM, Steinmann B (eds) Connective tissues and its heritable disorders. Wiley-Liss, New York, pp 112–113

Leeds P, Wood JM, Lee B-S, Culbertson MR (1992) Gene products that promote mRNA turnover in *Saccharomyces cerevisiae*. Mol Cell Biol 12:2165–2177

Lim S-K, Sigmund CD, Gross KW, Maquat LE (1992) Nonsense codons in human β-globin mRNA result in the production of mRNA degradation products. Mol Cell Biol 12: 1149–1161

Losson R, Lacroute F (1979) Interference of nonsense mutations with eukaryotic messenger RNA stability. Proc Natl Acad Sci USA 76:5134–5137

Magnus T, Baserga SJ, Stolle C, Takeshita K, Benz EJ (1990) Metabolism of non-translatable globin mRNAs arising from premature translation termination codons. Ann NY Acad Sci 612:55–66

Mashima Y, Murakami A, Weleber RG, Kennaway NG, Clarke L, Shiono T, Inana G (1992) Nonsense-codon mutations of the ornithine aminotransferase gene with decreased levels of mutant mRNA in gyrate atrophy. Am J Hum Genet 51:81–91

Orita M, Iwahana H, Kanazawa H, Hayashi K, Sekiya T (1989) Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. Proc Natl Acad Sci USA 86:2766–2770

Poduslo SE, Dean M, Kolch U, O'Brien SJ (1991) Detecting high-resolution polymorphisms in human coding loci by combining PCR and single-strand conformation polymorphism (SSCP) analysis. Am J Hum Genet 49:106–111

Pulak R, Anderson P (1993) mRNA surveillance by the *Caenorhabditis elegans smg* genes. Genes Dev 7:1885–1897

Redford-Badwal DA, Stover ML, Valli M, McKinstry MB, Rowe DW (1996) Nuclear retention of COL1A1 messenger RNA identifies null alleles causing mild osteogenesis imperfecta. J Clin Invest 97:1035–1040

Rowe D, Shapiro J, Poirier M, Schlesinger S (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

Saiki RK, Gelfand DH, Stoffel S, Scharf R, Higuchi R, Horn GT, Mullis KB, et al (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239:487–491

Sanger, Nicklen FS, Coulson AR (1979) DNA sequencing with chain-termination inhibitors. Proc Natl Acad Sci USA 74: 5463–5467

Sillence D, Senn A, Danks D (1979) Genetic heterogeneity in osteogenesis imperfecta. J Med Genet 16:101–116

Slayton RL, Deschenes SP, Willing MC (1995) Premature termination mutations in the COL1A1 gene interfere with mRNA stability during RNA processing in osteogenesis imperfecta type I cell strains. Am J Hum Genet Suppl 57:A9

Stover ML, Primorac D, Liu SC, McKinstry MB, Rowe DW (1993) Defective splicing of mRNA from one COL1A1 allele of type I collagen in nondeforming (type I) osteogenesis imperfecta. J Clin Invest 92:1994–2002

Sykes B, Francis M, Phil D, Smith R (1977) Altered relation of two collagen types in osteogenesis imperfecta. N Engl J Med 296:1200–1203

Urlaub G, Mitchell PJ, Ciudad C, Chasin LA (1989) Nonsense mutations in the dihydrofolate reductase gene affect RNA processing. Mol Cell Biol 9:2868–2880

Wenstrup RJ, Willing MC, Starman BJ, Byers PH (1990) Distinct biochemical phenotypes predict clinical severity in nonlethal variants of osteogenesis imperfecta. Am J Hum Genet 46:975–982

Westerhausen AI, Constantinou C, Pack M, Pend M, Hanning C, Olsen A, Prockop DJ (1991) Completion of the last half of the structure of the human gene for the proα1(I) chain of type I procollagen (COL1A1). Matrix 11:375–379

Willing MC, Cohn DH, Byers PH (1990) Frameshift mutation near the 3′ end of the COL1A1 gene of type I collagen predicts an elongated proα1(I) chain and results in osteogenesis imperfecta type I. J Clin Invest 85:282–290

Willing MC, Deschenes SP, Scott DA, Byers PH, Slayton RL, Pitts SH, Arikat H, et al (1994) Osteogenesis imperfecta type I: molecular heterogeneity for COL1A1 null alleles of type I collagen. Am J Hum Genet 55:638–647

Willing MC, Pruchno CJ, Atkinson M, Byers PH (1992) Osteogenesis imperfecta type I is commonly due to a COL1A1 null allele of type I collagen. Am J Hum Genet 51:508–515

Zhang Z-X, Wakamatsu N, Mules EH, Thoma GH, Gravel RA (1994) Impact of premature stop codons on mRNA levels in infantile Sandhoff disease. Hum Mol Genet 3:139–145

Zhou MY, Xue D, Gomez-Sanchez EP, Gomez-Sanchez CE (1994) Improved downward capillary transfer for blotting of DNA and RNA. Biotechniques 16:58–59