Absence of mutations in the promoter of the COL1A1 gene of type I collagen in patients with osteogenesis imperfecta type I

Marcia C Willing, Rebecca L Slayton, Sara H Pitts, Sachi P Deschenes

Abstract

Osteogenesis imperfecta type I results from decreased production of structurally normal type I collagen as a result of a COL1A1 “null” allele. Steady state amounts of COL1A1 mRNA are reduced in both the nucleus and cytoplasm of dermal fibroblasts from most affected subjects. Mutations involving key regulatory sequences in the COL1A1 promoter, such as the TATAAA and CCAAT boxes, could alter steady state levels of mRNA, and therefore lead to this phenotype. To determine the frequency of such mutations in OI type I cell strains, we used PCR amplified genomic DNA in conjunction with denaturing gradient gel electrophoresis (DGGE) and SSCP, to screen the 5’ untranslated domain, exon 1, and a small portion of intron 1 of the COL1A1 gene. In addition, direct sequence analysis was performed on an amplified genomic DNA fragment that included the TATAAA and CCAAT boxes. Forty unrelated probands with OI type I, in whom no causative mutation was known, were included in the study. No mutations were identified in either the TATAAA or CCAAT boxes in any of the affected people. In addition, there was little evidence of sequence diversity among any of the 40 subjects. These data suggest that mutations in the COL1A1 promoter do not play a significant role in the aetiology of OI type I.

(J Med Genet 1995;32:697–700)

Osteogenesis imperfecta type I is the mildest form of dominantly inherited brittle bone disease.1 Dermal fibroblasts from most affected subjects produce and secrete about half the normal amount of type I collagen.2 A growing body of evidence indicates that mutations that alter the expression of COL1A1, the gene that encodes the proα1(I) chain of type I collagen, are responsible for these synthetic alterations.3–8 Most characterised mutations lead to allele specific reduction in steady state amounts of COL1A1 mRNA in both the nucleus and cytoplasm of cell.9 Although promoter mutations are recognised as a mechanism for altering levels of mRNA in a number of genetic disorders, their potential for causing OI type I has not been investigated. In this paper we report our mutational analysis of the 5’ untranslated region and first exon of the COL1A1 gene in 40 unrelated probands with OI type I. Our data suggest that mutations in this portion of the gene are rare causes of OI type I.

Subjects and methods

INCLUSION CRITERIA

Forty unrelated probands who met the clinical criteria for OI type I,1 and whose dermal fibroblasts produced about half the usual amount of type I collagen,9 were included in the study. The causative mutation responsible for the phenotype was unknown in all. The COL1A1 gene from each affected person had been previously screened by DGGE, using overlapping primer pairs that amplified 400 to 550 bp fragments of genomic DNA. In addition, about one-sixth of the gene had been screened by SSCP at the time of this analysis.

PCR AND MUTATION SCREENING

Aliquots of genomic DNA isolated either from white blood cells or from cultured dermal fibroblasts of affected subjects served as template for amplification by PCR.10 The sequences of the oligonucleotides used as primers for PCR were derived from published sequence information11 (figure). Amplification conditions were as previously described.12 Four overlapping COL1A1 domains, referred to as fragments A, B, C, and D (figure), were screened for mutations by denaturing gradient gel electrophoresis (DGGE)13,14 or SSCP,15,16 or both. For DGGE, which was performed on each fragment, the following linear gradients of DNA denaturant (formamide and urea) were used: fragment A, 50–80%; fragment B, 35–65%; fragment C, 45–75%; fragment D, 55–80%.9 SSCP was performed on fragments A and D, as previously described,9 except that allele identification was by silver staining, using Typhoon Silver Sequence film (Promega, Madison, WI), rather than by autoradiography. Potential mutations were identified by direct sequence analysis of PCR amplified genomic DNA, using the dyeoxy chain termination method.17

SEQUENCE ANALYSIS OF THE COL1A1 PROMOTER

Aliquots of amplified material from fragments B and C were electrophoresed in either 1% LMP, or in 0.8% NuSieve (FMC Bioproducts, Rockland, ME), and the sole 486 bp and 489 bp bands, respectively, were excised from the gel. DNA fragments isolated in LMP were
purified using the Promega Magic PCR Prep DNA Purification System; fragments isolated in NuSieve were used as templates without further purification. Sequence was determined by the dideoxy chain termination method.\textsuperscript{17} Sequence ambiguities were clarified by analysis of both sense and antisense strands. For comparison, amplified genomic DNA from nine unrelated controls was also sequenced.

**Results**

**ANALYSIS OF THE 5' END OF THE COL1A1 GENE**

**BY DGGE AND SSCP**

We amplified four overlapping COL1A1 fragments, which included 554 bp of 5' untranslated sequence, exon 1, and 65 bp of intron 1, from genomic DNA of 40 unrelated probands with OI type I. Fragment A spans from -434 to -187 of the COL1A1 promoter (in the figure, the CAP site is labelled +1), and overlaps with fragment B which includes promoter sequences -434 to +52. Fragment C encompasses -204 to +286, and includes portions of the COL1A1 promoter, exon 1 (+120 to +221), and a small portion of the 5' end of intron 1, which flanks the exon. Fragments B and C both contain the CCYAAAT and TATAAA boxes. Fragment C overlaps with fragment D, which spans from 0 to +266, and includes the 5' CAP site, the sequence immediately upstream of the translation initiation site, exon 1, and 45 bp of intron 1. Each fragment was screened by DGGE for the presence of potential mutations; fragments A and D were also screened by SSCP.

Using this approach, we identified one person (referred to as proband 1) with an aberrant DGGE banding pattern in both fragments B and C. Direct sequence analysis of fragment C showed a G to A substitution located 9 bp downstream of the CCYAAAT box, and a C to T transition 14 bp upstream of the TATAAA box (figure). Both were inherited from his unaffected father, whose dermal fibroblasts produced normal amounts of type I procollagen (data not shown). Another affected person (proband 2) had a variant SSCP conformer in fragment D. Direct sequence analysis of PCR products showed a G to A substitution at +117, 3 bp upstream of the translation start site (figure). No additional sequence variants were identified in any of the other probands.

**DIRECT SEQUENCE ANALYSIS OF CCYAAAT AND TATAAA BOXES**

To ensure that we had not overlooked a potential mutation in our screening, we performed direct sequence analysis on either fragment B.
or C. Sequence was obtained from −150 to +1 (CAP site), which includes the CCAAT and TATAAA boxes. There were no mutations in either of these regulatory sequences in any of the 40 affected people. In addition, except for the two substitutions noted above in proband 1, there were no deviations from the published sequence in this region, in any of the affected people, or in the nine controls.

Discussion

In most OI type I cell strains, COL1A1 mRNA is derived primarily from one allele. Several different molecular mechanisms, which lead to allele specific decrease in COL1A1 mRNA, have been identified in OI type I cell strains, including small deletions and insertions in exons that change the translational reading frame, and create new termination codons. Although mutations in the COL1A1 promoter could also limit the amount of mRNA available from one COL1A1 allele, none of the probands in this study has evidence of mutations in the promoter sequence analysed. The region screened encompasses the known regulatory motifs, including the TATAAA and CCAAT boxes, and sequences that resemble viral enhancer elements, as well as exon 1 and a small portion of intron 1. Although intron 1 is known to contain enhancer elements, we did not analyse these domains.

Our experience with mutation screening in OI type I cell strains indicates that mutation detection is optimised by using a combination of DGGE and SSCP, since some mutations will be detected by one, but not by the other technique. The only sequence variants found in proband 1 were initially identified by DGGE. We also sequenced portions of the COL1A1 promoter, including the TATAAA and CCAAT boxes, and did not identify additional sequence variants. Although the direct sequencing in this region confirmed the reliability of our approach to mutation detection in this region of the COL1A1 gene, we cannot eliminate the possibility that mutations further upstream in the promoter could be responsible for OI type I.

There appears to be little variation from the published sequence in most of our probands. Although proband 1 had two sequence variants in the promoter, both were inherited from his unaffected father. In this family, the proband's mother has OI, but does not have the sequence variants. No family members were available to provide insight into the relevance of the substitution identified in proband 2. The substitution is located 3 bp upstream of the translation initiation site in an area that is thought to be important for ribosome recognition and binding. At the (−3) position, A is the most conserved nucleotide, followed by G, suggesting that the substitution in this person is unlikely to be of functional significance.

Mutations involving the TATAAA and CCAAT boxes, as well as other nucleotides in the promoter region, have been identified in β thalassaemia and other haemoglobinopathies, familial hypercholesterolaemia, haemophilia B, and X linked chronic granulomatous disease. Most are single nucleotide substitutions that disrupt DNA protein interactions thought to be important in transcription initiation. As a class, promoter mutations appear to be rare causes of genetic disorders. For conditions whose molecular basis has been characterised extensively (for example, familial hypercholesterolaemia, Duchenne muscular dystrophy, haemophilia B, cystic fibrosis), they appear to account for fewer than 2% of the causative mutations. For β thalassaemia, which is associated with a null phenotype, mutations in the β globin gene promoter comprise about 15% of the total number of identified mutations. This relatively large proportion of documented promoter mutations may be related to the small size of the β globin gene. With only three exons and two introns, the target for mutation is small, compared to large genes, such as dystrophin and the LDL receptor. Our own experience with the COL1A1 gene supports this idea. We have identified 12 different COL1A1 mutations that lead to a null allele in OI type I families, including four splicing mutations, two single nucleotide substitutions that create new termination codons, and six small deletions/insertions within exons (MC Willing, unpublished data). These data, in combination with the present study of the COL1A1 promoter, suggest that most of the mutations responsible for OI type I will be in the coding portion of the gene.

We thank Dr Peter Byers for providing many of the cell strains for analysis. We thank the probands and the referring physicians for their interest and participation. This work was supported by grant AR40334 from the National Institutes of Health, USPHS, and a Biomedical Science Grant from the Arthritis Foundation, by Basic Research Grant 0634 from the March of Dimes Birth Defects Foundation, and by the Roy J Carver Charitable Trust at the University of Iowa.

4 Barsh GS, David KE, Byers PH. Type I collagen im-
7 Willing MC, Pruchno CJ, Byers PH. Molecular hetero-
8 Willing MC, Pruchno CJ, Atkinson M, Byers PH. Os-
Willing, Slayton, Pitts, Descenes


12 Willing MC, Cohn DH, Byers PH. Frameshift mutation near the 3' end of the COL1A1 gene of type I collagen predicts an elongated prov(1) chain and results in osteogenesis imperfecta type I. J Clin Invest 1990;85:282-90.


20 Korak M. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell 1986;44:283-92.


34 Top B, Uitterlinden AG, van der Zee A. Absence of mutations in the promoter region of the low density lipoprotein receptor gene in a large number of familial hypercholesterolemia patients as revealed by denaturing gradient gel electrophoresis. Hum Genet 1992;89:561-5.

