Complete Sequence of the 23-Kilobase Human COL9A3 Gene

DETECTION OF GLY-X-Y TRIPLET DELETIONS THAT REPRESENT NEUTRAL VARIANTS*

We report the complete sequence of the human COL9A3 gene that encodes the α3 chain of heterotrimeric type IX collagen, a member of the fibril-associated collagens with interrupted triple helices family of collagenous proteins. Nucleotide sequencing defined over 23,000 base pairs (bp) of the gene and about 3000 bp of the 5′-flanking sequences. The gene contains 32 exons. The domain and exon organization of the gene is almost identical to a related gene, the human COL9A2 gene. However, exon 2 of the COL9A3 gene codes for one -Gly-X-Y- triplet less than exon 2 of the COL9A2 gene. The difference is compensated by an insertion of 9 bp coding for an additional triplet in exon 4 of the COL9A3 gene. As a result, the number of -Gly-X-Y- repeats in the third collagenous domain remains the same in both genes and ensures the formation of an in-register triple helix. In the course of screening this gene for mutations, heterozygosity for separate 9-bp deletions within the COL1 domain were identified in two kindreds. In both instances, the deletions did not co-segregate with any disease phenotype, suggesting that they were neutral variants. In contrast, similar deletions in triple helical domain of type I collagen are lethal. To study whether α3(IX) chains with the deletion will participate in the formation of correctly folded heterotrimeric type IX collagen, we expressed mutant α3 chains together with normal α1 and α2 chains in insect cells. We show here that despite the deletion, mutant α3 chains were secreted as heterotrimeric, triple helical molecules consisting of three α chains in a 1:1:1 ratio. The results suggest that the next noncollagenous domain (NC2) is capable of correcting the alignment of the α chains, and this ensures the formation of an in-register triple helix.

Type IX collagen is a structural component of hyaline cartilage and vitreous of the eye. It is a heterotrimeric molecule composed of three genetically distinct polypeptide chains: α1, α2, and α3 (1). The protein is characterized by interruptions in the triple helix, and it consists of three collagenous domains (COL1, COL2, and COL3, numbered from the C terminus) that are joined by four small noncollagenous domains (NC1 to NC4) (2, 3). In addition to interrupted triple helices, type IX collagen is a fibril-associated collagen and thus belongs to the FACIT subgroup of collagens (4).

Type IX collagen is attached to the surface of type II collagen fibrils by lysine-derived covalent cross-links between the COL2 domain and telopeptides of type II collagen (5–8). Because the flexible NC3 domain enables the COL3 and NC4 domains to project out of the fibril surface, it has been suggested that the COL3 and NC4 domains may play a role in mediating interactions between collagens and noncollagenous components of hyaline cartilage (6, 9, 10). The NC3 domain of the α2(IX) chain also has an attachment site for a glycosaminoglycan side chain (11). Results from a recent study indicate that the NC1 domain of the three α chains contains all of the necessary information for chain selection and assembly (12). The COL1 domain may play a critical role in the anti-parallel binding to the fibril surface, although this has not been directly demonstrated. There are no known or proposed functions for the NC2 domain as yet.

Transgenic mice expressing α1(IX) cDNA with a large in-frame deletion of the sequences encoding a part of the COL3 domain, the entire NC3 domain, and part of the COL2 domain develop abnormalities in cartilage collagen fibril structure and a phenotype similar to human osteoarthritis and a mild chondrodysplasia (13). Degenerative joint disease was also seen in separate lines of transgenic mice with inactivation of the Col9a1 gene (14). All of these findings suggest that type IX collagen is not essential for cartilage development, but it is required for maintaining the normal structural integrity of cartilage.

Linkage to COL9A2 has been reported in two families with autosomal dominant multiple epiphyseal dysplasia (MED)1 (15–16). In one of these families, a splice site mutation leading to an in-frame 12-amino acid residue deletion in the COL3 domain has been identified. Multiple epiphyseal dysplasia comprises a genetically heterogeneous group of disorders characterized by shared clinical findings ranging from mild joint stiffness and pain in large joints to early onset osteoarthritis (17). Mutations in cartilage oligomeric matrix protein have also

---

* The work was supported in part by grants from the Academy of Finland (to L. A.-K.) and the Arthritis Foundation (to M. L. W.) and National Institutes of Health Grants AR50481 and EY09908 (to R. M.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

**To whom correspondence should be addressed. Tel.: 358-8-5375756; Fax: 358-8-5375811; E-mail: Leena.Ala-Kokko@oulu.fi.

1 The abbreviations used are: CSGE, conformation sensitive gel electrophoresis; MED, multiple epiphyseal dysplasia; PCR, polymerase chain reaction; bp, base pairs; kb, kilobase pairs; nt, nucleotide(s); PAGE, polyacrylamide gel electrophoresis.
been shown to cause MED (18–20), and linkage studies support the existence of additional MED loci (21). The COL9A1 and COL9A3 genes are logical candidates for such loci. The complete cDNA sequences for the α3(IX) chain are currently available for chick (22, 23) and human (24). Here we report the complete genomic organization and sequences of the human COL9A3 gene. Also, we report that two unrelated families have different 9-bp deletions in the same region of the COL1 domain that are neutral variants of the gene. These are the first examples of deletions within the triple helical domain of a collagens protein that are neutral variants.

**EXPERIMENTAL PROCEDURES**

**Isolation of Phage and P1 Clones for the Human COL9A3 Gene**—A probe (p1061) was prepared by reverse transcription-PCR for screening of phage genomic libraries using primers H10 and H2 as described (24). A second probe was prepared using a sense primer H41 (5′-AAA TCA GGC TCT CGA AGC TC, nt 2032–2051) with an antisense primer H42 (5′-TCT TTA CAC AAA TGC TAT GC, nt 2355–2374) to amplify a 342-bp PCR product, p424, that originates in NC1 and extends into the 3′-untranslated region of the human α3(IX) cDNA (24). The probes p342 and p1061 were 32P-Pradilabeled by nick-translation and used to screen replicate filters from a human EMBL3 PS6/T7 genomic library (CLONTECH Laboratories, Palo Alto, CA) as described previously (24). Two unique clones, gR2B21 and gR5B51, were isolated and sequenced (24).

For PCR screening of genomic P1 libraries for the human COL9A3 gene, the primer pairs were designed on the basis of published sequences for the human α3(IX) cDNA (24) and genomic organization of the mouse Col9a2 gene (25). To amplify the 5′-end sequences of the gene, the primer pair C93-F4 (5′-CAG GAA AGC GGG GGA AAC CAG, nt 200–220 from the start of translation in the human cDNA) and C93-R5 (5′-GTC CAT CTC GTC CAG TCA GAC, nt 277–257) was used. The primer pair C93-F32 (5′-CCT GCC AAG GAG CCG TGT TAG G, nt 200–220 from the start of translation in the human cDNA) and C93-RUTR (5′-CTT TTA CAC AAA TGC TAT GC, nt 2249–2227) was used for the amplification of the 3′-end sequences of the COL9A3 gene. The primers C93-F4 and C93-R5 corresponded to the sequences in exons 4 and 5, and the primers C93-F32 and C93-RUTR corresponded to the sequences in exons 32 of the mouse Col9a2 gene. PCRs were performed in a 40-μl reaction volume using 50–100 ng of genomic DNA, 0.25 μM each primer, 200 μM each dNTP, 1.5 mM MgCl2, and 1 unit of Taq polymerase (AmpliTaq, Perkin-Elmer). Thermal cycling conditions were 1 min at 94 °C, 1 min at 60 °C, and 1 min at 72 °C for 30 cycles. The primer pairs amplified single bands of about 700 bp (C93-F4 and C93-R5) and 250 bp (C93–32F and C93–RUTR), and were used for screening a human P1 library (Genome Systems, Inc.). The screening resulted in identification of three positive P1 clones: P1-C93A, P1-C93B, and P1-C93C (Genome Systems control numbers 12269, 12270, and 12271 and clone addresses DMPC-HFF#1-270-C3, DMPC-HFF#1-753-B10, and DMPC-HFF#1-1082-B5, respectively).

**Characterization of P1 Clones**—To isolate DNA, the P1 clones were cultivated overnight in 3 ml of LB and 25 μg/ml kanamycin, and 2.5 ml of the overnight culture was grown in 75 ml of LB for 1.5 h. After addition of isopropl-1-thio-β-β-galactopyranoside to a final concentration of 0.5 mm, incubation was continued for 5 h followed by centrifugation at 10,000 × g for 10 min in 10-ml aliquots and P1 DNA isolation with a standard plasmid isolation protocol (26). Sequencing of the P1 clone was performed by cycle sequencing (Cycle Sequencing Kit, Amersham Pharmacia Biotech). Sequencing primers were designed on the basis of the published cDNA sequences for the human α3(IX) chain (24) and the genomic organization of the mouse Col9a2 gene (25). Intronic sequences between exons 10 and 11, 12 and 13, 26 and 27, and 31 and 32 were amplified using Expand long template PCR system (Roche Molecular Biochemicals). The PCR products were purified using an agarose gel extraction protocol (QIAEX II gel extraction kit, Qiagen), followed by cloning into pUC 18 vector (SURE clone ligation kit, Amersham Pharmacia Biotech), and sequencing (T7 sequencing kit, Amersham Pharmacia Biotech). Sequencing reactions were analyzed on a 6% polyacrylamide gel. Heteroduplex Analysis—Human control and patient DNA was extracted from whole blood using standard methods and used for PCR amplifications. PCR primers were designed from the intronic sequences to amplify separately each exon of the COL9A3 gene. The product sizes varied from 200 to 400 bp and contained at least 80 bp of 5′ and 3′ intronic sequences. Genomic DNA was amplified in a 40-μl volume with thermal cycling of 45 s at 94 °C, 45 s at 60–62 °C, and 1 min at 72 °C for 30 cycles followed by a final extension at 72 °C for 10 min. Heteroduplexes were generated by denaturing the samples at 95 °C for 5 min and reannealing for 30 min at 68 °C. The concentration and quality of PCR products was estimated analyzing 5 μl of each reaction in 1.5% agarose gel. CSGE was used to scan the PCR products for mutations that generated heteroduplexes (27). A CSGE gel consisted of 10% polyacrylamide, 9.9:1 ratio of acrylamide to 1,4-bisacryloylpyperazine (Fluka), 10% ethylene glycol, 15% formamide (Life Technologies, Inc.), 0.1% ammonium persulfate, and 0.07% TEMED in 0.5× TTE (44 mM Tris, 14.5 mM Taurine, 0.1 mM EDTA buffer, pH 9.0) buffer. Gel electrophoresis was performed with a standard DNA sequencing apparatus (Life Technologies, Inc.) using 0.5× TTE as the electrode buffer. Prior to electrophoresis, 3–15 μl or 25–75 ng of sample was mixed with loading buffer (10× stock solution of 30% glycerol containing 0.25% of both xylene cyanol FF and bromphenol blue). The gel was pre-electrophoresed at 45 W for 15 min, and the samples were electrophoresed at 45 W for 5 h at room temperature. After electrophoresis, the gel was stained with ethidium bromide (1 μg/ml), destained with water, and photographed. Samples containing heteroduplexes were analyzed by direct PCR product sequencing (T7 Sequenase PCR product sequencing kit, United States Biochemical). Some PCR products were purified from agarose gel, and 60 ng of purified product was cloned into pUC18 vector and sequenced. Several clones were sequenced to obtain sequences for both alleles.

**Expression and Analysis of Recombinant Type IX Collagen in Insect Cells**—For amplification of the α3(IX) chain containing the Gly-Pro-Pro deletion, specific primers were designed on the basis of the published cDNA sequences (24). Two oligonucleotides, R9A3DEL (5′-GTT TTA CAC AAA TGC TAT GC, nt 250–750) and F9A3DEL (5′-GTT TTA CAC AAA TGC TAT GC, nt 200–250), both containing a generated CapI cleavage site, were designed to exon 30. R9A3DEL was used for PCR amplification with oligonucleotide M29B (5′-CCC GGC GCC GTC TAG ACT CGG CCA CGG) that corresponded to the 5′-noncoding region and F9A3DEL with oligonucleotide MH30 (5′-TGGC GGC GCC GTC TAG ACT CGG CCA CGG) that corresponded to the 3′-untranslated region of the human α3(IX) cDNA.
GGC GTC CTT GTC TCT AGA TTC CTC ACG) that corresponded to the 3'-noncoding region of the α3(IX) cDNA. A DNA template for PCR amplification was α3(IX) cDNA transcribed from total RNA extracted from human fetal cartilage. PCR was performed in a 40-μl volume with thermal cycling of 45 s at 94 °C, 45 s at 60 °C, and 1 min at 72 °C for 30 cycles followed by a final extension at 72 °C for 10 min. The primer pairs amplified 1729-bp 5'-end sequences (M29B and R9A3DEL) and 404-bp 3'-end sequences (F9A3DEL and MH30) of the α3(IX) cDNA.

The PCR products were digested with CspI, purified using QIAEX II gel extraction kit (Qiagen) followed by ligation into pVL1392 vector, and sequenced using cDNA specific primers (ABI PRISM model 377 sequencer, Perkin-Elmer; ABI PRISM dye terminator cycle sequencing ready reaction with AmpliTaq DNA polymerase, FS, Perkin-Elmer).

The deleted α3(IX) cDNA was transfected into Spodoptera frugiperda (Sf9, Invitrogen) insect cells using BaculoGold transfection kit (Pharmingen). The viral pools were collected, amplified, and plaque-purified.

The deleted α3(IX) cDNA was transfected into Spodoptera frugiperda (Sf9, Invitrogen) insect cells using BaculoGold transfection kit (Pharmingen). The viral pools were collected, amplified, and plaque-purified.

Expression of mutant recombinant type IX collagen was achieved by co-infecting Trichoplusia ni (High Five, Invitrogen) insect cells with the recombinant virus for the α3(IX) chain containing the Gly-Pro-Pro deletion and viruses for the α1(IX) and α2(IX) chains (43) together with a double promoter virus, 4PHab (28) coding for the α and β subunits of human prolyl 4-hydroxylase. For expression of wild-type collagen, a virus for the wild-type α3(IX) chain (43) was substituted for the mutant α3(IX) virus. Culture conditions were as described (43). Culture medium was collected after 72 h of infection, and the recombinant type IX collagen was precipitated with 25% saturation of ammonium sulfate. The precipitate was dissolved overnight at 4 °C in 0.5 M urea, 0.2 M NaCl, 0.05 M Tris buffer, pH 7.4. For pepsin digestion, the samples were adjusted to pH 2, and digestion was performed at room temperature for 4 h. The undigested controls were incubated without pepsin, and denaturation of the samples was performed by heating at 60 °C for 5 min prior to digestion. After pepsin treatment, the pH was adjusted to 7.5.

Expression of mutant recombinant type IX collagen was achieved by co-infecting Trichoplusia ni (High Five, Invitrogen) insect cells with the recombinant virus for the α3(IX) chain containing the Gly-Pro-Pro deletion and viruses for the α1(IX) and α2(IX) chains (43) together with a double promoter virus, 4PHab (28) coding for the α and β subunits of human prolyl 4-hydroxylase. For expression of wild-type collagen, a virus for the wild-type α3(IX) chain (43) was substituted for the mutant α3(IX) virus. Culture conditions were as described (43). Culture medium was collected after 72 h of infection, and the recombinant type IX collagen was precipitated with 25% saturation of ammonium sulfate. The precipitate was dissolved overnight at 4 °C in 0.5 M urea, 0.2 M NaCl, 0.05 M Tris buffer, pH 7.4. For pepsin digestion, the samples were adjusted to pH 2, and digestion was performed at room temperature for 4 h. The undigested controls were incubated without pepsin, and denaturation of the samples was performed by heating at 60 °C for 5 min prior to digestion. After pepsin treatment, the pH was adjusted to 7.5.
**Characterization of Genomic Clones**—Screening of the human genomic phage library yielded two positive clones, gRB2B1 and gRB5B1, that contained the 3’-portion of the COL9A3 gene. gRB2B1 was approximately 12 kb in size and hybridized with both P1061 and p342. gRB5B1 was approximately 18 kb in size and hybridized only with the 3’-probe p342. A third genomic clone, gRB35, that was isolated previously, contained the 5’-most end of the gene (24). However, after DNA sequencing and restriction mapping of the genomic clones, it was found that the clones did not overlap and thus did not cover the entire gene. To obtain clones covering the entire gene, a human P1 library was screened with two PCR primer pairs designed from the cDNA sequence to amplify the 5’-end or the 3’-end of the gene. The screening yielded three positive clones. The P1 clones were analyzed for the presence of the most 5’-end and 3’-end sequences of the corresponding cDNA by sequencing and PCR amplification. All clones were found to contain the entire coding region, and one clone (P1-C93A) was selected for detailed characterization of the gene (Fig. 1). Nucleotide sequencing of the human COL9A3 gene was performed by direct sequencing of the P1 clone or by sequencing of subclones in plasmids. A total of over 26 kb of the nucleotide sequence was determined. The results indicated that the gene is about 23 kb and contains 32 exons (Fig. 1). Also, over 3 kb of 5’-flanking sequences are presented. The sequences extend to the 3’-end of the next gene, which is called 7-60.² Sites for selected restriction enzymes are shown in Fig. 1.

**Exon Organization and Domain Structure**—The genomic organization of the human COL9A3 gene is indicated in Figs. 1 and 2. The genomic structure was compared with a related gene, the human gene for α2(IX) collagen chain (29). Because the domain structures of α2(IX) and α3(IX) collagen chains are almost identical, it was probable that the exon organization of the corresponding genes would also be conserved. As expected, the overall exon organization of the genes showed considerable similarities. There were, however, some unexpected differences in sizes of the exons encoding for the COL3 domain even though the size of the domain is identical in these genes. The COL3 domain is 411 bp or 137 amino acid residues and is encoded by exons 2 through 10 in both genes. Exon 2 in both genes is a junction exon between NC4 and COL3. Exon 2 in the COL9A3 gene codes for one collagen triplet less than exon 2 in the COL9A2 gene, but the COL3 domain is identical because exon 4 in COL9A3 codes for one additional triplet compared with exon 4 in the COL9A2 gene (Fig. 3). In effect, the deletion of 9 bp coding for one -Gly-X-Y- triplet in exon 2 of the COL9A3 gene is compensated by an insertion of 9 bp coding for an additional triplet in exon 4.

² J. Liu, R. G. Brewton, M. Takanosu, B. M. Wood, and R. Mayne, unpublished observations.
of the cloned PCR product identified a 9-bp deletion in the exon, removing a Gly-Pro-Hyp triplet in the 5'-end of the COL1 domain. The rest of the family members were analyzed for the presence of the mutation. As indicated in Fig. 4, two affected members of the family had the deletion. At the same time, two unaffected members had the deletion, and identical twins who were affected did not have the deletion. Hence, the deletion did not co-segregate with the phenotype in the family.

A second family having a different 9-bp deletion in exon 30 that also removed a Gly-Pro-Hyp triplet in the 5'-end of the COL1 domain. The rest of the family members were analyzed for the presence of the mutation. As indicated in Fig. 4, two affected members of the family had the deletion. At the same time, two unaffected members had the deletion, and identical twins who were affected did not have the deletion. Hence, the deletion did not co-segregate with the phenotype in the family.

Expression and Analysis of Recombinant Type IX Collagen Containing the Gly-X-Y Deletion in the α3(IX) Chain—To study whether α3(IX) chains with the deletion will participate in the formation of correctly folded type IX collagen, the Gly-Pro-Pro deleted chain was generated. Specific primers that contained a generation of correctly folded type IX collagen, the Gly-Pro-Pro deleted (Fig. 6). To study the triple helicity of the recombinant type IX collagen, pepsin treatment was performed on native and on denatured recombinant type IX collagen samples followed by SDS-PAGE analysis under nonreducing conditions. Pepsin was found to digest all of the material when the type IX collagen obtained by pepsinization (39). Pepsin resistance of nonmutated and mutated recombinant human type IX collagen. Nonmutated (lanes 2–5) recombinant protein (Wt) and mutated (lanes 6–9) recombinant protein (α3Δ) were prepared as indicated in Fig. 6 and analyzed by 8% SDS-PAGE under nonreducing conditions. Samples in lanes 3, 5, 7, and 9 were subjected to limited pepsinization (P) for 4 h at room temperature, whereas samples in lanes 2, 4, 6, and 8 were incubated for 4 h at room temperature without pepsin. Samples in lanes 4, 5, 8, and 9 were denatured by heating at 60 °C for 5 min before the incubation at room temperature with (lanes 5 and 9) or without (lanes 4 and 8) pepsin. Lane 1, molecular weight marker. HMW, high molecular weight fragments of type IX collagen obtained by pepsinization (39). Pepsin (arrow) indicates the position of pepsin after electrophoresis.

FIG. 5. Nine-bp deletions in exon 30. Partial nucleotide sequence of exon 30 is indicated. Del A and Del B indicate the deletions in the first and the second families, respectively. Del A1 and Del A2 indicate the two possible deleted sequences in the first family.

FIG. 6. SDS-PAGE analysis of recombinant type IX collagen containing the mutated α3(IX) chain (α3Δ). High Five cells in adherent cultures were infected with three recombinant baculoviruses for the α chains of human type IX collagen and a baculovirus 4PHΔβ for the α and β chains of human prolyl 4-hydroxylase. After 72 h of infection, medium was collected, precipitated with ammonium sulfate, and subjected to cation exchange chromatography. After centrifugation, the material was dissolved overnight at 4 °C in 0.5 m urea, 0.2 m NaCl, 0.05 m Tris buffer, pH 7.4, and analyzed by 8% SDS-PAGE under nonreducing conditions (lane 2) or under reducing conditions (lane 3). Lane 1, molecular weight markers.

FIG. 7. Pepsin resistance of nonmutated and mutated recombinant human type IX collagen. Nonmutated (lanes 2–5) recombinant protein (Wt) and mutated (lanes 6–9) recombinant protein (α3Δ) were prepared as indicated in Fig. 6 and analyzed by 8% SDS-PAGE under nonreducing conditions. Samples in lanes 3, 5, 7, and 9 were subjected to limited pepsinization (P) for 4 h at room temperature, whereas samples in lanes 2, 4, 6, and 8 were incubated for 4 h at room temperature without pepsin. Samples in lanes 4, 5, 8, and 9 were denatured by heating at 60 °C for 5 min before the incubation at room temperature with (lanes 5 and 9) or without (lanes 4 and 8) pepsin. Lane 1, molecular weight marker. HMW, high molecular weight fragments of type IX collagen obtained by pepsinization (39). Pepsin (arrow) indicates the position of pepsin after electrophoresis.
The common structure is shown in A; the proposed effect of the Gly-X-Y deletion is shown in B. The deletion in the COL1 domain of the α3(IX) chain shortens the COL1 domain by one triplet and increases the length of the α1(IX) and α2(IX) chains by three amino acid residues in the NC2 domain. The deletion does not interfere with disulfide bond formation between the α1 and α3 chains or the triple helical structure of the COL2 domain.

Discussion

The results demonstrate that the genomic structure of the COL9A3 gene is similar to that of the human COL9A2 gene (29). However, there are two major differences. One is that the large introns of the COL9A3 gene are found primarily near the 5′-end of the gene, whereas the large introns are at the 5′-end of the gene (29). The COL9A3 gene appears to be the exception in this regard because most other genes for collagens have their large introns at the 5′-end of the gene (30–35). A second difference between the two genes is the size distribution among exons coding for the COL3 domain. The size of the COL3 domain is 137 amino acids in both genes. Exon 2 of the COL9A3 gene codes for one less -Gly-X-Y- triplet than exon 2 of the COL9A2 gene, but the number of -Gly-X-Y- triplets needed to form an in-register triple helix remains the same because exon 4 of the COL9A3 gene codes for an additional triplet compared with exon 4 of the COL9A2 gene.

The number of -Gly-X-Y- triplets in the triple helix of collagens is the same to ensure the formation of an in-register triple helix. However, the COL1 domain of the α3(IX) chain is one triplet shorter than the corresponding domain in the α1(IX) and α2(IX) chains. Surprisingly, we observed here an additional 9-bp deletion coding for a triplet of -Gly-X-Y- in an individual with MED, suggesting that the deletion was disease-causing. However, examination of affected and unaffected members of the family indicated that there was no co-inheritance of the deletion with the disease phenotype. Individuals from a second family segregating a different 9-bp deletion within this domain had no skeletal phenotype. Therefore, the deletions must be neutral variants of the gene.

A likely explanation for this finding is that the NC2 domain can compensate for the size difference within COL1 created by the deletion by independently facilitating the correct register of the individual α chains prior to their folding to form the COL2 triple helical domain. Mechanisms leading to the precise alignment of individual α chains have been principally studied in the fibril forming collagens. In these molecules, the C-terminal propeptides of the α chains associate through noncovalent interactions ensuring the correct register of the α chains. The interaction is stabilized by intramolecular disulfide bonds. This is followed by triple helix formation that progresses from the C-terminus to the N terminus of the molecule. It has been shown that the C-terminal propeptides of fibrillar collagens contain all the necessary information for the correct chain selection and association (see Ref. 40). Accordingly, it has been shown recently that synthetic peptides of the three α chains of type IX collagen consisting of the entire NC1 domain and the C-terminal end of the COL1 domain contain all the necessary information for chain selection and assembly (12). The correct assembly of the α chains is critical because triple helix formation progresses in a zipper-like fashion. That a deletion within the COL1 domain of collagen IX does not cause a clinical phenotype suggests that it does not interfere with the normal folding of the remaining domains of this molecule. To study that possibility, recombinant human type IX collagen containing the deletion in α3 chain was expressed in insect cells. Analysis of the protein indicated that the mutated α3 chain participates in the formation of correctly folded heterotrimeric molecules. Thus, the NC2 domain of collagen IX may function to align α chains prior to COL2 triple helix formation similar to the function of the NC1 domain during COL1 formation (Fig. 8). Internal non-triple helical domains within other FACIT collagens may have similar roles.

Similar 9-bp deletions have been reported in direct repeat sequences of the COL1A1 gene in two unrelated probands with lethal osteogenesis imperfecta (41, 42). The deletions led to the loss of one of three consecutive Gly-Ala-Hypro triplets at positions 868–876. The mechanism by which the deletions caused a lethal phenotype is not entirely clear, but the results suggested that the deletions introduced a shift in the phase of the chains in the triple helix, and the shift was propagated from the site of the deletion toward the N terminus of the molecule. Even though the deletions did not abolish the N-proteinase or collagenase cleavage sites, the deletions might interfere with cross-link formation or prevent the lateral association of molecules to form fibrils. A splice mutation in the COL9A2 gene leading to an in-frame deletion of 12 amino acids in the COL3 domain has been reported in a family with MED (16). This mutation shortens the COL3 domain and interferes with the structure of the NC4 domain. Thus, it is unlikely that the 9-bp deletions identified here introduce a shift in a phase of the chains that propagates from the site of the deletion all the way to the N terminus of the molecule. These findings also suggest that the NC2 domain of type IX collagen may compensate for the size difference of the α chains and thus prevent the propagation of the deletion (Fig. 8).

In the clinical context, the fact that similar deletions can be associated with lethal phenotypes in one collagen molecule yet constitute benign variants in another implies the use of caution in overinterpreting the potential consequences of DNA mutations in the absence of complementary biochemical or cell biological studies.

Acknowledgments—We thank Drs. Carla Borrone, Marta Romangono, and Andrea Superti-Furga for sharing clinical and scientific expertise and Aira Harju for expert technical assistance.

References

1. van der Rest, M., and Mayne, R. (1987) in Structure and Function of Collagen Types (Mayne, R., and Burgess, R. E., eds) pp. 195–221, Academic Press, Orlando, Fl.
2. Ninomiya, Y., Castagnola, P., Gerecke, D., Gordon, M. K., Jaenko, O., LuValle, P., McCarthy, M., Muragaki, Y., Nishimura, I., Oh, S., Rosenblum, N., Sato, N., Sugrue, S., Taylor, R., Vassos, G., Yamaguchi, N., and Olsen, B. R. (1990) in Extracellular Matrix Genes (Sandell, L. J., and Boyd, C. D., eds) pp. 79–114, Academic Press, San Diego, CA.
3. Breslow, R. G., and Mayne, R. (1994) in Extracellular Matrix Assembly and Structure (Vurchenko, P. D., Birk, D. E., and Mecham, R. P., eds) pp. 129–170, Academic Press, San Diego, CA.
4. Gordon, M. K., and Olsen, B. R. (1990) Curr. Opin. Cell Biol. 2, 833–838.
5. Eyre, D. R., Apon, S., Wu, J.-J., Ericsson, L. H., and Walsh, K. A. (1987) PERS
Human COL9A3 Gene

22475

Let. 220, 337–341
6. van der Rest, M., and Mayne, R. (1988) J. Biol. Chem. 263, 1615–1618
7. Wu, J.-J., Woods, P. E., and Eyre, D. R. (1992) J. Biol. Chem. 267, 23007–23014
8. Diab, M., Wu, J. J., and Eyre, D. (1996) Biochem. J. 314, 327–332
9. Vasios, G., Nishimura, I., Konami, H., van der Rest, M., Ninomiya, Y., and Olsen, B. R. (1986) J. Biol. Chem. 261, 2324–2329
10. Vaughan, L., Mendler, M., Huber, S., Bruckner, P., Winterhalter, K. H., Irwin, M. H., and Mayne, R. (1988) J. Cell Biol. 106, 391–397
11. Bruckner, P., Vaughan, L., and Winterhalter, K. H. (1985) Proc. Natl. Acad. Sci. U. S. A. 82, 2608–2612
12. Mechling, D. E., Gambee, J. E., Morris, N. P., Sakai, L. Y., Keene, D. R., Mayne, R., and Bachinger, H. P. (1996) J. Biol. Chem. 271, 13781–13785
13. Nakata, K., Ono, K., Miyazaki, J. I., Olsen, B. R., Muragaki, Y., Adachi, E., Yamamura, K., and Kimura, T. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 2870–2874
14. Fassler, R., Schnegelsberg, P. N. J., Dausman, J., Shinya, T., Muragaki, Y., McCarthy, M. T., Olsen, B. R., and Jaenisch, R. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 5070–5074
15. Briggs, M. D., Choi, H.-C., Warman, M. L., Loughlin, J. A., Wordsworth, P., Sykes, B. C., Irven, C. M. M., Smith, M., Wynne-Davies, R., Lipson, M. H., Biesecker, L. G., Garber, A. P., Lachman, R., Olsen, B. R., Rimoin, D. L., and Cohn, D. H. (1994) Am. J. Hum. Genet. 55, 678–684
16. Muragaki, Y., Mariman, E. C. M., van Beersum, S. E. C., Perälä, M., van Mourik, J. B. A., Warman, M. L., Olsen, B. R., and Hamel, B. C. J. (1996) Nat. Genet. 12, 103–105
17. Spranger, J. (1992) Eur. J. Pediatr. 151, 407–415
18. Briggs, M. D., Hoffman, S. S. G., King, L. M., Olsen, A. S., Mohrenweiser, H., Leroy, J. G., Mortier, G. R., Rimoin, D. L., Lachman, R. S., Gaines, E. S., Cockell, J. A., Knowlton, R. G., and Cohn, D. H. (1995) Nat. Genet. 10, 330–336
19. Cohn, D. H., Briggs, M. D., King, L. M., Rimoin, D. L., Wilcox, W. R., Lachman, R. S., and Knowlton, R. G. (1996) Am. N. Y. Acad. Sci. 785, 188–194
20. Balle, R., Briggs, M. D., Cohn, D. H., Knowlton, R. G., Beighton, P. H., and Ramesar, R. S. (1997) Am. J. Med. Genet. 68, 396–400
21. Deere, M., Blanton, S. H., Scott, C. I., Langer, L. O., Pauli, R. M., and Hecht, J. T. (1995) Am. J. Hum. Genet. 56, 698–704
22. Brewton, R. G., Ouspenskaia, M. V., van der Rest, M., and Mayne, R. (1992) Eur. J. Biochem. 205, 443–449
23. Harvel, R., Sharma, Y. D., Aguilera, A., Ueyama, N., Wu, J.-J., Eyre, D. R., Juricic, L., Chandrasekaran, S., Li, M., Nah, H.-D., Upholt, W. B., and Tanzer, M. L. (1992) J. Biol. Chem. 267, 10070–10076
24. Brewton, R. G., Wood, B. M., Ren, Z.-X., Gong, Y., Tiller, G. E., Warman, M. L., Lee, B., Horton, W. A., Olsen, B. R., Baker, J. R., and Mayne, R. (1995) Genomics 30, 329–336
25. Perälä, M., Elima, K., Mertsaranta, M., Rosati, R., de Crombrugghe, B., and Vuorio, E. (1994) J. Biol. Chem. 269, 5064–5071
26. Birnboim, H. C., and Doly, J. (1979) Nucleic Acids Res. 7, 1513–1523
27. Ganguly, A., Rock, M. J., and Prockop, D. J. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 10325–10329
28. Nakelainen, M., Helasakoji, T., Myllyharju, J., Notbohm, H., Pihlajaniemi, T., Dietz, P. P., and Kivirikko, K. I. (1986) Matrix Biol. 6, 329–338
29. Pihlajamaa, T., Vuorio, M. M., Annunen, S., Perälä, M., Prockop, D. J., and Ala-Kokko, L. (1998) Matrix Biol. 17, 237–241
30. Boedeker, H., Finer, M., and Aho, S. (1985) Ann. N. Y. Acad. Sci. 460, 85–116
31. D’Alessio, M., Bernard, M., Pretorius, P., de Wet, W., and Ramirez, P. (1988) Gene 67, 105–115
32. Toman, D., and de Crombrugghe, B. (1995) Gene 147, 161–168
33. Ala-Kokko, L., Krist, A.-P., Mertsaranta, M., Kivirikko, K. I., de Crombrugghe, B., Prockop, D. J., and Vuorio, E. (1995) Biochem. J. 308, 923–929
34. Takahara, K., Hoffman, G. G., and Greenzpan, D. S. (1995) Genomics 29, 588–597
35. Vuorio, M., Pihlajamaa, T., Vandenberg, P., Prockop, D. J., and Ala-Kokko, L. (1995) J. Biol. Chem. 270, 22873–22881
36. Bell, D. M., Leung, K. K. H., Wheatley, S. C., Ng, L. J., Zhou, S., Ling, K. W., Shanm, M. H., Koopman, P., Tam, P. P. L., and Cheah, K. S. E. (1997) Nat. Genet. 16, 174–178
37. Lefebvre, V., Huang, W., Harley, V. R., Goodfellow, P. N., and de Crombrugghe, B. (1997) Mol. Cell. Biol. 17, 2336–2346
38. Krohn, E., Nakata, K., Bernier, S. M., Hatano, O., Miyashita, T., Rhodes, C. S., and Yamada, Y. (1996) J. Biol. Chem. 271, 4296–4303
39. Reese, C. A., and Mayne, R. (1981) Biochemistry 20, 5443–5448
40. Engel, J., and Prockop, D. J. (1991) Annu. Rev. Biophys. Biophys. Chem. 20, 157–152
41. Hawkins, J., Superti-Furga, A., Steinmann, B., and Dalgleish, R. (1991) J. Biol. Chem. 266, 22570–22574
42. Wallis, G. A., Kadler, K. E., Starman, B. J., and Byers, P. H. (1992) J. Biol. Chem. 267, 25529–25534
43. Pihlajamaa, T., Perälä, M., Vuorio, M. M., Nakelainen, M., Bode, M., Schultheiss, T., Vuorio, E., Timpl, R., Engel, J., and Ala-Kokko, L. (1999) J. Biol. Chem. 274, 22464–22468