Identical G+1 mutations in three different introns of the gene for type III procollagen (COL3A1) that cause aberrant splicing of RNA were found in three probands with life-threatening variants of Ehlers-Danlos syndrome. Because the three mutations were in a gene with multiple and homologous exons, they provided an interesting test for factors that influence aberrant splicing. The G+1 to A mutation in intron 16 caused extensive exon skipping, the G+1 to A mutation in intron 20 caused both use of a cryptic splice site and retention of all the intron sequences, and the G+1 to A mutation in intron 42 caused efficient use of a single cryptic splice site. The different patterns of RNA splicing were not explained by evaluation of potential cryptic splice sites in the introns by either their homology with 5'-splice sites from other genes or by their AG/G values for binding to U1 RNA. Instead, the results suggested that the patterns of aberrant RNA splicing were primarily determined by the relative rates at which adjacent introns were normally spliced.

A series of mutations that alter RNA splicing have been shown to cause a number of genetic diseases (4-23). The genes for the major fibrillar collagens present an unusual target for RNA splicing mutations, since the large triple-helical domains of the proteins are encoded by 42 exons that are similar both in size and nucleotide sequence (24-26). The typical exon is 54 bp in length and the remaining exons of 45, 99, 108, and 162 bp appear to be related in size to the typical 54-bp exon either as multiples of 54 bp or by loss of 9 bp from 54- or 108-bp exons. Also, all of the 42 exons coding for the repeating -Gly-X-Y- sequence of the α chains of the protein. Four RNA splicing mutations have recently been reported in the gene for type I procollagen in probands with osteogenesis imperfecta (20, 21) or EDS-VII (22, 23). Three of these mutations were single base mutations (20, 22, 23) and the fourth was a 19-base pair deletion (21). All four mutations produced efficient exon skipping so that the coding sequences of an exon were eliminated from mRNA transcripts and shorter pro-α chains were synthesized.

Here we report three single base mutations in the gene for type III procollagen that cause aberrant splicing of RNA. All three mutations were found in probands with the type IV variant of EDS (27-29), the most serious form of the disease which often has life-threatening consequences such as rupture of arteries or hollow organs. Each of the mutations was a single base substitution that converted the first G of an intervening sequence to an A. The three mutations, however, had very different effects in terms of the efficiency with which they induced exon skipping or the use of cryptic splice sites.

**Materials and Methods**

**The Probands**—Proband I (+1 IVS16) was a 36-year-old pregnant female who had thin skin with abnormally prominent superficial blood vessels. She had a minimal degree of joint hypermobility and a history of two surgical procedures for correction of patellar dislocations. At approximately 33 weeks of pregnancy, she was found to have hypertension, pre-eclampsia, and gestational diabetes. She had a premature rupture of membranes and underwent cesarean section. The infant immediately developed severe bleeding and expired 4 h later. At surgery, the proband's tissues appeared to be unusually friable. Her brother died at the age of 20 of a ruptured cervical artery sustained during karate practice. There was no other family history of EDS or related genetic disease of connective tissue.

Proband II (+1 IVS20) was a male who died of a ruptured aneurysm of the thoracic aorta at the age of 34. He was admitted to an emergency ward because of acute onset of back and abdominal pain which radiated to the groin. He was in shock and underwent emergency surgery. He had massive left retroperitoneal hemorrhage without any obvious bleeding point. He continued to bleed postoperatively and expired within 24 h. At autopsy he had a dissecting aneurysm of the proximal descending thoracic aorta with massive mediastinal and left pleural hemorrhage. The patient had a previous history of easy bruising, nose bleeds, and bleeding from gums after brushing of his teeth. His father died at the age of 43 and his brother at the age of 35 of ruptured aneurysms of the thoracic aorta. The proband did not have thin skin that made venous patterns prominent, and he did not have the ecchymoses and scarring over bony prominences that are usually considered as criteria for the diagnosis of EDS IV (27, 28). Therefore, he probably constituted an example of phenotypic overlap between EDS type IV and familial aneurysms (47).

Proband III (+1 IVS42) was a female who died suddenly from a ruptured dissecting aortic aneurysm at the age of 22. She had thin
and transparent skin with abnormally prominent blood vessels. She had mild hypermobility of the joints, congenital dislocation of the hips, and a torn knee ligament. Both the hip dislocations and the torn knee ligament required surgical repair. She had a history of bouts of abdominal pain and urinary tract infections as well as childhood pyloric stenosis. She had an 11-month-old daughter who did not have any evidence of EDS or related heritable disorder of connective tissue. Also, the family history was negative for a heritable disorder of connective tissue.

Initial Screening with S1 Nucleotide Protection Experiments— Cultures of skin fibroblasts were prepared (30) from the three probands and DNA was extracted with guanidium isothiocyanate and purified by extraction through phenol/chloroform (31). RNA was used for S1 nucleotide protection experiments using five different single-stranded antisense DNA probes coding for sequences of type III procollagen. The uniformly labeled DNA probes were prepared by cloning appropriate fragments (0.7–1.5 kilobases) of the cDNAs for type III procollagen (32) into the filamentous bacteriophage M13. The inserts in M13 were used to synthesize single-stranded antisense DNA labeled with [α-32P]dCTP as described previously (20, 21). The labeled probes were electrophoretically separated from template strands in a 6% polyacrylamide DNA sequencing gel and were electroeluted (31).

To form RNA-DNA hybrids, about 50,000 cpm (about 0.5 fmol) of the single-stranded DNA probe were hybridized with 10–50 μg of total RNA in 80% formamide, 0.25 mM EDTA, and 300 mM NaCl in 25 mM HEPES buffer (pH 7.0) at 56 °C for 3 h. The sample was digested at 40 °C for 30 min with 10–100 units of S1 nuclease (United States Biochemical Corp.) in 500 μl of 50 mM NaCl and 1 mM ZnSO4 in 30 mM NaOAc buffer (pH 5.3). The products were precipitated with ethanol and separated by electrophoresis on a 5% polyacrylamide gel. The sizes of the fragments were estimated by comparison with a DNA sequencing ladder and standard DNA fragments (Bethesda Research Laboratories; Boehringer Mannheim).

PCR with Templates of cDNA and Genomic DNA—To prepare a cDNA template for PCR, total RNA was isolated from cultured skin fibroblasts as described above (30, 31). The RNA was used to synthesize single-stranded cDNA as described by Gubler and Hoffman (33) with a commercial kit (Bethesda Research Laboratories). For the PCR (34), four sets of oligonucleotides based on the cDNA sequence (32) were used as primers so as to generate overlapping products spanning all the exons of the α1(III) chain and, therefore, made it possible to detect any mutations in the sequences (Fig. 1). The primers were 30-mer oligonucleotides which were complementary to the cDNA and 10 nucleotides contained sequences of restriction sites convenient for cloning.

To isolate genomic DNA templates for the PCR, DNA was extracted from 175-cm2 flasks of cultured skin fibroblasts with sodium dodecyl sulfate and protease K (31). Three sets of oligonucleotide primers were used for the PCR. For proband I (+1 IV616), the 5′ primer was identical to coding sequences in exon 15 and the 3′ primer was complementary to coding sequences in exon 17. For proband II (+1 IV520), the 5′ primer was identical to coding sequences in exon 19 and the 3′ primer was complementary to coding sequences in exon 21. For proband III (+1 IV542), the 5′ primer was identical to sequences in intron 41 and the 3′ primer was complementary to sequences in intron 43 (35).

For DNA sequencing, the PCR products were cloned into M13mp18 or M13mp19, and the clones were sequenced with the dideoxynucleotide method (36). Some of the clones were sequenced with standard procedures using radiolabeled nucleotides. Other sequences were determined using fluorescently tagged M13 primers and the fluorescent dye terminator kit (37). The products were separated by denaturing polyacrylamide gel electrophoresis and analyzed on an automated DNA sequencer (model 370A, Applied Biosystems).

RESULTS

Analysis for Potential 5′-Splice Sites—To score normal and potential 5′-splice sites with the method of Shapiro and Senapathy (38), the sequences were searched for -GT- dinucleotides and a score was computed for each -GT- dinucleotide in the context of the two nucleotides preceding and the four nucleotides following it. The scoring method was modified slightly (38) in that the -GT- dinucleotide used to identify potential 5′ sites did not contribute to the score. The frequency (in percent) with which a particular nucleotide occurred at given positions was used. The DNA was used as a database for the EcoRI digestion. The appropriate single-stranded DNA probes were used for S1 nucleotide protection experiments under the conditions described above except that the DNA was end-labeled with [α-32P]dCTP and the DNA was electroeluted (39).

Analysis of Potential 5′-Splice Sites—To determine the RNA splicing patterns, a series of single-stranded DNA sequences were cloned into M13 and sequenced. Nucleotide sequencing of the clone demonstrated that the insert contained the expected coding sequences and the minimum and maximum possible totals for the six positions was then used to give a score according to the formula: score = 100 (t – tmin)/(tmax – tmin), where tmin and tmax are the minimum and maximum possible totals for the six positions.

To calculate ΔG° for binding of U1 RNA to normal and potential 5′-splice sites, the free energy parameters of Freier et al. (39) were used.

CONCLUSIONS

Initial Screening of RNA Using S1 Nucleotide Protection Experiments—To search for mutations in mRNAs from the three probands, five single-stranded cDNA probes spanning all the coding sequences of the human α1(III) chain (32) were prepared and used for S1 nucleotide protection experiments. With RNA from proband I (+1 IV616), one probe was only partially protected (Fig. 2). The probe was an Xhol/Sall fragment of 708 nt (nt 975–1682) in the cDNA extending from exons 15–25; Ref. 32) that was uniformly labeled. About half of the probe was fully protected by the RNA from the proband’s fibroblasts, and about half was cleaved to two fragments (Fig. 2). One of the cleaved fragments was 620 nt. The other was apparently too small to detect on the polyacrylamide gel. For proband II (+1 IV520), the cDNA from the proband’s mRNA was used for the PCR using a 5′ primer identical to codons in exon 15 and a 3′ primer complementary to codons in exon 17 (35, 37). The PCR products were separated on Nonade agarose (FMC BioProducts). The largest fragment corresponded to RNA with unspliced sequences from intron 20. The PCR products were used in the M13 cloning and sequenced. Nucleotide sequencing of the clone demonstrated that the insert contained the expected coding sequences and the minimum and maximum possible totals for the six positions was then used to give a score according to the formula: score = 100 (t – tmin)/(tmax – tmin), where tmin and tmax are the minimum and maximum possible totals for the six positions.
was a uniformly labeled and single-stranded cDNA complementary cDNA probe and RNA from proband I (±l IVS16). The probe control fibroblasts; digestion with 30 units of S1 nuclease. Lane 2, 10 µg of total RNA from proband’s fibroblasts; digestion with 30 units of S1 nuclease. Lane 3, 10 µg of total RNA from proband’s fibroblasts; digestion with 60 units of S1 nuclease. RNA from the proband’s fibroblasts gave two fragments: a fully protected fragment of 708 nt and a cleaved fragment of about 620 nt. The expected smaller fragment of about 34 nt was not detected.

Fig. 2. S1 nuclease probe protection experiment with a cDNA probe and RNA from proband I (±l IVS16). The probe was a uniformly labeled and single-stranded cDNA complementary to nt 975–1682 in the cDNA (32). Lane 1, 10 µg of total RNA from control fibroblasts; digestion with 30 units of S1 nuclease. Lane 2, 10 µg of total RNA from proband’s fibroblasts; digestion with 30 units of S1 nuclease. Lane 3, 10 µg of total RNA from proband’s fibroblasts; digestion with 60 units of S1 nuclease. RNA from the proband’s fibroblasts gave two fragments: a fully protected fragment of 708 nt and a cleaved fragment of about 620 nt. The expected smaller fragment of about 34 nt was not detected.

amid gels. The results, therefore, suggested that the RNA had an aberrant structure about 88 nt from either the 5’- or 3’-end of the probe, i.e. at about nt 1063 (13 nt downstream of the junction between exons 16’ and 17’) or nt 1594 (15 nt upstream of the junction between exons 23 and 24’) in the cDNA (32).

RNAs from probands II and III fully protected the five probes (not shown).

Sequencing of Coding Sequences of the cDNAs—A series of overlapping PCR products were prepared using mRNA-derived cDNA from fibroblasts from the three probands (Fig. 1). The PCR products were subcloned into M13 and sequenced. To further characterize the mutation in proband I (Fig. 1), PCR products 1 and 2 were sequenced. In order to find mutations in probands II and III, multiple clones of all PCR products were sequenced.

With proband I (+1 IVS16), six M13 clones of PCR product 1 showed a deletion of all 18 codons of exon 16. Since exon 16 contains nt 997–1050 of the cDNA (32), the results were consistent with the S1 nuclease experiments if allowance is made for the difficulty of accurately estimating the size of the protected fragments (Fig. 2). Four M13 clones had a normal sequence in the same region.

With proband II (+1 IVS20), alterations in nucleotide sequences were also seen in M13 clones of PCR product 1 (Fig. 1). Of 156 M13 clones that were partially sequenced, 60 had an insertion of the first 24 nt from intron 20, three had an insertion of all 132 nt from intron 20, two lacked all the codons of exon 20, and 91 had a normal sequence. The 24- and 132-nt insertions all had a single-base mutation that converted the first G of the intron 20 to an A and thereby altered the consensus sequence of -GT- that has been found in most of the introns of eukaryotic genes reported to date (38, 40). The 24- and 132-nt insertions did not alter the reading frame, but the 132-nt insertion contained a single-base mutation that altered the consensus sequence of -GT- that has been found in most of the introns of eukaryotic genes reported to date (38, 40). The 24- and 132-nt insertions did not alter the reading frame, but the 132-nt insertion contained a stop codon at position +34 to +36 of intron 20. The results were consistent with the negative results of the S1 nuclease experiments if it is assumed that the presence of intron sequences in the transcript produced looping out of the RNA but not of the labeled DNA probe used to protect the RNA.

With proband III (+1 IVS42), 45 M13 clones from PCR product 4 had an insertion of the first 30 nt from intron 42, and three clones had a normal sequence in the region. The 30-nt insertion had a single-base mutation that converted the first G of intron 42 to an A and thereby altered the consensus sequence of -GT-. The 30-nt insertion did not change the reading frame. Again, the results were consistent with the negative results of the S1 nuclease experiments if it is assumed that the presence of intron sequences produced looping out of the RNA but not of the labeled cDNA probe.

The data from the experiments were used to design primers and to carry out PCRs using genomic DNA as a template. The products of the PCRs were again cloned into M13 for nucleotide sequencing. The results confirmed the G+-1 to A mutation in intron 20 in proband II and the G+-1 to A mutation in intron 42 in proband III. They also demonstrated that proband I had an identical G+-1 to A mutation in the first nucleotide of intron 16.

Pattern of RNA Splicing Induced by the Three Mutations—To define the pattern of RNA splicing produced by the three single-base mutations in the three probands, a further series of S1 nuclease protection experiments were carried out. Because only about half of the genomic structure of the human type III procollagen gene is known (see Refs. 32, 35, and 37), several different strategies were used to prepare single-stranded DNA probes containing the appropriate coding and intervening sequences (see “Materials and Methods”).

To examine RNA splicing in proband I (+1 IVS16), an end-labeled hybrid DNA probe of 865 nt was prepared. The probe contained part of exon 10, all of exons 11–16, intron 16, and the first 50 nt of exon 17 (Fig. 3). RNA from the proband’s fibroblasts protected fragments of 820, 373, 349, and 295 nt. The fragment of 295 nt indicated exon skipping in RNA splicing so that the 18 codons of exon 16 were eliminated from the mRNA. The fragment of 349 nt reflected normal splicing of the RNA. The fragment of 373 nt indicated use of a cryptic 5’-splice site at position +25 of intron 16. The sequence of the cryptic 5’-splice site was -GGTATAA-. The fragment of 820 nt reflected a failure to splice out intron 16 since it contains all 421 nt of intron 16. (Cleavage of the 865-nt probe to a fragment of 820 nt was more easily seen in other gels (not shown).) An additional band of about 312 nt varied widely in intensity from one experiment to another. The fragment of 312 nt indicated a deletion of 37 nt of exon 16 but no potential cryptic splice site was found in exon 16 that could account for the band. Therefore, the origin of this fragment was not apparent. Densitometry of the gels demonstrated that about 71% of the mRNA from the mutated allele (half of the total mRNA) was spliced by exon skipping, about 21% was spliced by insertion of 24 nt from intron 16, about 1% was processed without any splicing of 421 nt of intron 16, and about 8% was accounted for by the undefined deletion (Table I).

In parallel experiments with proband II (+1 IVS20), the probe was 382 nt long (Fig. 4). The probe contained part of exon 19, all of exon 20, all of intron 20, and part of exon 21. RNA from the proband’s fibroblasts generated single-stranded DNA fragments of 112, 136, and 340 nt. The fragment of 112 nt indicated normal RNA splicing. The fragment of 136 nt indicated that the cryptic 5’-splice site at position +25 of the intron was used in splicing of some RNA. The sequence of the cryptic splice site was -TGTTTATT-. The fragment of 340 nt indicated that some of the RNA was spliced so that all 132 nt of intron 20 were retained. The fragment of 58 nt that might reflect exon skipping of exon 20 was not detected. As indicated above, evidence of exon skipping was obtained by sequencing of cDNA clones generated from PCR products in that two cDNA clones of 156 nt had...
RNA Splicing Mutations in Type III Procollagen Gene

FIG. 3. S1 nuclease probe protection experiments with total RNA from proband I (+1 VIS16). A, radiograph of electrophoretic gel. P, probe without S1 nuclease digestion; P', S1 nuclease digestion of probe without RNA. Lanes 1 and 2, control RNA; digestion with 50 and 400 units of S1 nuclease. Lanes 3–6, RNA from the proband’s fibroblasts; digestion with 50, 100, 200, and 400 units of S1 nuclease, respectively. B, schematic representation of probe used for the experiments and the sizes of the protected fragments expected with control RNA and with RNA from the proband.

TABLE I

| Proband | Mutation | RNA splicing from mutated allele | Pattern  | % total
|---------|----------|---------------------------------|----------|--------
| I       | GIVS16   | Exon skipping                   | 71       |
|         |          | Insertion of 24 nt in-frame     | 21       |
|         |          | Insertion of whole intron (421 nt) | 1      |
|         |          | Undefined deletion of 37 nt     | 8        |
| II      | GIVS20   | Exon skipping                   | 1        |
|         |          | Insertion of 24 nt in-frame     | 53       |
|         |          | Insertion of whole intron (132 nt) | 34     |
|         |          | Undefined insertion of 113 nt   | 12       |
| III     | GIVS42   | Insertion of 30 nt in-frame     | 100      |

* Values for probands I and II are means from densitometry of two probe-protection experiments in which 2–4 lanes of each experiment were scanned. The standard deviation of the values was ±6%.

FIG. 4. S1 nuclease probe protection experiments with RNA from proband II (+1 IVS20). A, radiograph of electrophoretic gel. P and P', as in Fig. 3. Lane 1, RNA from proband’s fibroblasts; digestion with 150 units of S1 nuclease. Lane 2, RNA from control fibroblasts; digestion with 150 units of S1 nuclease. S, sequencing ladder of DNA used as size markers (G reaction shown). B, schematic representation of probe and size of protected fragments expected with RNA from control and the proband’s fibroblasts.

from intron 20, about 34% by insertion of all the 132 nt from the intron, and about 12% by an undefined insertion of 113 nt (Table I).

In parallel experiments with proband III (+1 IVS42), the probe was about 1259 nt long (Fig. 5). It contained part of the coding sequences of exon 40, all the coding sequences of exon 41 and exon 42, all the coding sequences of exon 43, and part of the sequences of intron 43. The RNA from the proband’s fibroblasts generated two fragments of 373 and 403 nt. The fragment of 373 nt indicated normal splicing of the RNA. The fragment of 403 nt indicated aberrant RNA splicing so that the first 30 nt of intron 42 was retained in the RNA. The sequence of the cryptic splice site in intron 42 was -AGGTAGAA-. Densitometry of the gel indicated that essentially all of the RNA from the mutated allele was spliced by using the cryptic splice site in intron 42 (Table I).

an in-frame deletion of all the codons from exon 20. An additional band seen in the S1 nuclease experiment corresponded to an insertion of about 113 nt from intron 20, but no potential cryptic splice site was found in intron 20 that could account for the band. Densitometry of the gels indicated that less than 1% of the RNA from the mutant allele was processed by exon skipping, about 53% by insertion of 24 nt
RNA Splicing Mutations in Type III Procollagen Gene

I. Splicing Patterns

A. S 1 nuclease protection experiments with RNA from fibroblasts from proband III (+1 IVS42). A, radiograph of electrophoretic gel from S 1 nuclease protection experiments. P and P', as in Fig. 3. Lanes 1 and 2, RNA from control fibroblasts; digestion with 200 and 400 units of S 1 nuclease. Lanes 3 and 4, RNA from proband's fibroblasts; digestion with 200 and 400 units of S 1 nuclease.

B. A schematic representation of probe and fragments obtained from control RNA and proband's RNA. The 3′ end of the probe contained 34 nt of intron 43 as well as 45 nt of M13 sequences.

Analysis of Potential 5′-Splice Sites—To explore why the three mutations produced different RNA splicing patterns (Table I and Fig. 6, I), an analysis of the potential 5′-splice sites was carried out with the statistical procedure of Shapiro and Senapathy (38) in which potential 5′-splice sites are scored relative to the frequency with which the same bases are found in 542 normal 5′-splice sites of primate genes. In the analysis, the sequence complementary to the 5′ end of U1 RNA scores 100. The normal and the potential 5′-splice sites were also evaluated with the free energy parameters of Freier et al. (39).

Comparison of the statistical scores for the normal 5′-splice sites (Fig. 7 and Table II) indicated that the normal sites in the three introns had higher values than any potential sites in the same introns after sites within 70 nt of the next exon were excluded (41). Comparison of the ΔG°C values did not, however, present as consistent a picture. For introns 16 and 20, the normal 5′ sites had more favorable ΔG°C values than any potential sites in the two introns. This was not true for intron 42 in which four potential sites had more favorable ΔG°C values (−5.8 to −7.1 kcal/mol) than the normal 5′-site (−3.5 kcal/mol).

The data from the three mutations were examined to determine whether they explained why specific cryptic sites were used and why exon skipping occurred in proband I (+1 IVS16) and not in the other two probands. The results indicated that when a cryptic site was used, there was a preference for the most 5′ site with a high statistical score or with a favorable ΔG°C value. In proband I (+1 IVS16), the +9 site in intron 16 was not used, apparently because it had unfavorable values by both parameters (34% and −0.6 kcal/mol). The +25 site was used, whereas the +153 site was not, apparently because the +25 site was more 5′. The same situation appeared to hold for proband II and proband III in that there was preference for cryptic sites near the beginning of the intron.

The data did not, however, provide a simple explanation as to why exon skipping occurred in proband I (+1 IVS16) but not in the other two probands. In particular, intron 16 had two potential 5′-sites with ΔG°C values more favorable (−5.2 and −5.4 kcal/mol) than the two best potential sites in intron...
RNA Splicing Mutations in Type III Procollagen Gene

**FIG. 7.** Scoring of potential 5' sites by the method of Shapiro and Senapathy (38). Solid bars indicate potential sites in exons, and open bars indicate potential sites in introns. For simplicity, sites with scores of less than 20% were eliminated. In the case of intron 41, about 1 kilobase of the sequence are not known and, therefore, are not scored here. Open arrows, normal sites used for splicing of the introns at the 5' end. Solid arrows, major cryptic sites used as the result of the mutations described here. Vertical lines on the x axis indicate spacings of 200 nt.

20 (~2.1 and ~2.7 kcal/mol). Also, one of the sites at position +153 in intron 16 had a statistical score (47%) similar to the statistical score (49%) of the best site at position +25 in intron 20. Therefore, the values predicted that the extent of exon skipping and use of cryptic splice sites should be comparable in proband I (+1 IVS16) and proband II (+1 IVS20). Instead, 70% of the RNA from the mutated allele in proband I was spliced by exon skipping and no more than about 1% was processed by exon skipping in proband II (Table I).

**DISCUSSION**

The results here describe the first RNA splicing mutations in the gene for type III procollagen. All three mutations were found in probands with EDS IV. All the mutations consisted of single base changes that converted the first G in an intervening sequence to an A, and thereby destroyed the sequence of -GT- found as the first two nucleotides of most intervening sequences (38, 40). Two of the mutations occurred adjacent to homologous 54-bp exons coding for the repetitive -Gly-X-Y- amino acid sequences characteristic of collagens. The third mutation was adjacent to a similar 108-bp exon. The consequences of the three single base mutations, however, were very different in terms of the patterns of RNA splicing they generated (Fig. 6,1, and Table I). The mutation in proband I (+1 IVS16) primarily caused exon skipping. The identical mutation in intron 20 in proband II (+1 IVS20) produced little exon skipping, but about 53% of the RNA was spliced through a cryptic site in intron 20 and about 34% was processed without any splicing of the sequences from intron 20. An identical G+1 mutation in proband III (+1 IVS42) produced highly efficient use of a single cryptic site in intron 42.

The observation that all three G+1 to A mutations altered RNA splicing is consistent with a large number of previous observations. Analyses of over 3000 introns indicated that 99.4% (40) or more (38) contained the dinucleotide sequence of -GT- at the 5' end of the intron. Most of the deviations from the -GT- rule were a change of the T to a C (13, 40, 42), and only three introns, all in immunoglobins, lacked a G in
RNA Splicing Mutations in Type III Procollagen Gene

The extensive exon skipping seen with the mutation in proband I (+1 IVS16) suggests the sequence of events illustrated in Fig. 6, IIA. The normal splicing of an intron involves binding of U1 snRNP complex to the 5' end of the intron and binding of a U2 and U5 snRNP to the 3' end (44, 45). The 5' end is then cleaved, a lariat loop is formed by the ligation of the 5' end to an A at the U2 snRNP binding site, and the process is completed by cleavage of the 3' end and joining of the two exons. The exon skipping observed in proband I (+1 IVS16) can be explained by a U2 and U5 snRNP complex binding to the 3' end of intron 10 before a similar complex binds to the 3' end of intron 15. The absence of significant exon skipping in probands II and III suggests that a U2 and U5 snRNP complex binds to the 3' end of the adjacent upstream intron before a similar complex binds to the 3' end of the intron containing the mutation (Fig. 6, IIIB). In proband II (+1 IVS20), the observation that 34% of the RNA was processed without any splicing of intron 20 (Table I), is consistent with the sequence shown in Fig. 6, IIIB. In proband III (+1 IVS42), the sequence of events may differ slightly in that the cryptic site at position +51 may be efficiently cleaved before intron 41 is spliced. The two different sequences of events (Fig. 6, II) depend on the assumption that the introns adjacent to the mutations are normally spliced at different rates relative to the intron containing the mutation. Different rates of splicing of introns from multi-exon genes was demonstrated in vitro for the murine interleukin-3 gene that contains four introns (46). The reasons for a preferential ordering of splicing within the multi-exon transcripts from multi-exon genes is not apparent. However, examination of further mutations in collagen genes or in comparable multi-exon gene constructs may define the factors that are critical.

A technical point of interest was that probe protection experiments with a labeled cDNA probe and S1 nuclease did not detect the abnormal transcripts produced by two of the mutations. The results demonstrated, therefore, that under the conditions employed here S1 nuclease did not cleave a labeled DNA probe in RNA-DNA hybrids in which there was looping out of RNA but not the DNA. At the same time, it was apparent that probe protection experiments with the appropriate genomic probe were necessary to establish the relative abundance of aberrantly spliced RNAs, since data obtained from sequencing 10 or more randomly selected M13 clones of PCR products were frequently misleading.

ACKNOWLEDGMENTS—We gratefully acknowledge the expert technical assistance of Gi-Chung Chen and Karen Engle in tissue culture. We are also grateful to Dr. Leena Ala-Kokko for the normal type III procollagen cDNA clone. We thank Drs. Steven A. Klein, Roger L. Ladda, and Boris G. Kousseff for the clinical data on the probands.

REFERENCES

1. Kuivaniemi, H., Tromp, G., Sabol, C., Ala-Kokko, L., Kontusaari, S., Klein, S., and Prockop, D. J. (1988) J. Cell Biol. 107, 805 (abstr.).
2. Zhao, M., Kontusaari, S., Kuivaniemi, H., Tromp, G., Sabol, C., Ala-Kokko, L., Klein, S. A., Ladda, R. B., Kousseff, B. G., and Prockop, D. J. (1990) Annu. Rev. Genet. 22, 131-181.
3. Marvit, J., DiLella, A. G., Brayton, K., Ledley, F. D., Robson, K. J. H., and Woo, S. L. C. (1987) Nucleic Acids Res. 15, 5613-5617.
4. Antonarakis, S. E., Orkin, S. H., Cheng, T.-C., Scott, A. F., Sexton, J., Trosku, S. P., Charache, S., and Kazazian, H. H., Jr. (1985) Proc. Natl. Acad. Sci. U. S. A. 82, 2689-2693.
5. Treisman, R., Proudfoot, N. J., Shander, M., and Maniatis, T. (1982) Cell 29, 903-911.
6. Marvit, J., DiLella, A. G., Brayton, K., Ledley, F. D., Robson, K. J. H., and Woo, S. L. C. (1987) Nucleic Acids Res. 15, 5013-5028.
7. Hidaka, Y., Kondo, T., Tada, T., and O'catalogue, T. E., Marvit, J., DiLella, A. G., Bravton, K., Ledlev, F. D., Robson, K. J. H., and Woo, S. L. C. (1987) Nucleic Acids Res. 15, 5613-5617.
8. Treisman, R., Proudfoot, N. J., Shander, M., and Maniatis, T. (1982) Cell 29, 903-911.
9. Marvit, J., DiLella, A. G., Brayton, K., Ledley, F. D., Robson, K. J. H., and Woo, S. L. C. (1987) Nucleic Acids Res. 15, 5013-5028.
10. Kuivaniemi, H., Tromp, G., Sabol, C., Ala-Kokko, L., Kontusaari, S., Klein, S., and Prockop, D. J. (1988) J. Cell Biol. 107, 805 (abstr.).
11. Zhao, M., Kontusaari, S., Kuivaniemi, H., Tromp, G., Sabol, C., Ala-Kokko, L., Klein, S. A., Ladda, R. B., Kousseff, B. G., and Prockop, D. J. (1990) Annu. Rev. Genet. 22, 131-181.
12. Marvit, J., DiLella, A. G., Brayton, K., Ledley, F. D., Robson, K. J. H., and Woo, S. L. C. (1987) Nucleic Acids Res. 15, 5613-5617.
13. Hidaka, Y., Kondo, T., Tada, T., and O'Thek, T. E., Marvit, J., DiLella, A. G., Bravton, K., Ledlev, F. D., Robson, K. J. H., and Woo, S. L. C. (1987) Nucleic Acids Res. 15, 5013-5028.
14. Treisman, R., Proudfoot, N. J., Shander, M., and Maniatis, T. (1982) Cell 29, 903-911.
15. Marvit, J., DiLella, A. G., Brayton, K., Ledley, F. D., Robson, K. J. H., and Woo, S. L. C. (1987) Nucleic Acids Res. 15, 5013-5028.
RNA Splicing Mutations in Type III Procollagen Gene

16. Ohno, K., and Suzuki, K. (1988) J. Biol. Chem. 263, 18563-18567
17. Grandchamp, B., Picat, C., Mignotte, V., Wilson, J. H. P., Te Velde, K., Sandkuy, L., Romeo, P. H., Goosens, M., and Nordmann, Y. (1988) Proc. Natl. Acad. Sci. U. S. A. 86, 661-664
18. Grandchamp, B., Picat, C., de Rooij, F., Beaumont, C., Wilson, P., Deybach, J. C., and Nordmann, Y. (1989) Nucleic Acids Res. 17, 6637-6649
19. Vidaud, M., Gattoni, R., Stevenin, J., Vidaud, D., Amselem, S., Chibani, J., Rosa, J., and Goosens, M. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 1041-1045
20. Tromp, G., and Prockop, D. J. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 5154-5158
21. Kuvanieni, H., Sabol, C., Tromp, G., Sippola-Thiehe, M., and Prockop, D. J. (1988) J. Biol. Chem. 263, 11407-11413
22. Weil, D., Bernard, M., Combates, N., Wirtz, M. K., Hollister, D. W., Steinmann, B., and Ramirez, F. (1988) J. Biol. Chem. 263, 8561-8564
23. Weil, D., D'Alessio, M., Ramirez, F., de Wet, W., Cole, W. G., Chan, D., and Bateman, J. F. (1989) EMBO J. 8, 1705-1710
24. Boedtker, H., Finer, M., and Ahd, S. (1985) Ann. N. Y. Acad. Sci. 460, 85-116
25. Ohkubo, H., Vogeli, G., Mudryj, M., Avvedimento, V. E., Sullivan, M., Pastan, I., and de Crombrugge, B. (1986) Proc. Natl. Acad. Sci. U. S. A. 77, 7059-7063
26. Ramirez, F., Bernard, M., Chu, M.-L., Dickson, L., Sangiorgi, F., Weil, D., de Wet, W., Junien, C., and Sobel, M. (1986) Ann. N. Y. Acad. Sci. 460, 117-125
27. Beighton, P. (1970) The Ehlers-Danlos Syndrome. William Heinemann, London
28. McKusick, V. A. (1972) Heritable Disorders of Connective Tissue, 4th Ed., pp. 792-791, D. V. Mosby Company, St. Louis, MO
29. Byers, P. H. (1983) in Clinical Medicine (Spittell, J. A., Jr, ed) pp. 1-41, Harper & Row, Philadelphia
30. Stolle, C. A., Pyeritz, R. E., Myers, J. C., and Prockop, D. J. (1985) J. Biol. Chem. 260, 1937-1944
31. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
32. Ala-Kokko, L., Kontusaari, S., Baldwin, C., Kuivaniemi, H., and Prockop, D. J. (1989) Biochem. J. 260, 509-516
33. Guibler, U., and Hoffman, R. J. (1983) Gene (Amst.) 25, 262-269
34. Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A., and Arnheim, N. (1985) Science 230, 1350-1354
35. Tromp, G., Kuivaniemi, H., Eishiaki, H., and Prockop, D. J. (1989) J. Biol. Chem. 264, 1349-1352
36. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 5463-5467
37. Benson-Chanda, V., Su, M.-W., Weil, D., Chu, M.-L., and Ramirez, F. (1989) Gene (Amst.) 78, 255-265
38. Shapiro, M. B., and Senapathy, P. (1987) Nucleic Acids Res. 15, 7155-7174
39. Freier, S., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T., and Turner, D. H. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 9373-9377
40. Jacob, M., and Gallinaro, H. (1988) Nucleic Acids Res. 17, 2159-2180
41. Wieringe, B., Hofer, E., and Weissman, C. (1984) Cell 37, 915-926
42. Soominen, R., Huotari, M., Sangiorgi, F., Prockop, D. J., and Tryggvason, K. (1989) J. Biol. Chem. 264, 13565-13571
43. Mitchell, J. P., Uralow, G., and Chasin, J. L. (1986) Mol. Cell. Biol. 6, 1926-1935
44. Ruskin, B., and Green, M. R. (1985) Cell 43, 131-142
45. Chabot, B., Black, D. L., LeMaster, D. M., and Steitz, J. A. (1985) Science 230, 1344-1349
46. Lang, K. M., and Spritz, R. A. (1987) Mol. Cell. Biol. 7, 3428-3437
47. Kontusaari, S., Tromp, G., Kuivaniemi, H., Ladda, R. L., and Prockop, D. J. (1990) Am. J. Hum. Genet., in press
Identical G+1 to A mutations in three different introns of the type III procollagen gene (COL3A1) produce different patterns of RNA splicing in three variants of Ehlers-Danlos syndrome. IV. An explanation for exon skipping some mutations and not others.

H Kuivaniemi, S Kontusaari, G Tromp, M J Zhao, C Sabol and D J Prockop

*J. Biol. Chem.* 1990, 265:12067-12074.

Access the most updated version of this article at [http://www.jbc.org/content/265/20/12067](http://www.jbc.org/content/265/20/12067)

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at [http://www.jbc.org/content/265/20/12067.full.html#ref-list-1](http://www.jbc.org/content/265/20/12067.full.html#ref-list-1)