Alternative splicing of the type II procollagen gene (COL2A1) is developmentally regulated during chondrogenesis. Chondroprogenitor cells produce the type IIA procollagen isoform by splicing (including) exon 2 during pre-mRNA processing, whereas differentiated chondrocytes synthesize the type IIB procollagen isoform by exon 2 skipping (exclusion). Using a COL2A1 mini-gene and chondrocytes at various stages of differentiation, we identified a non-classical consensus splicing sequence in intron 2 adjacent to the 5′ splice site, which is essential in regulating exon 2 splicing. RNA mapping confirmed this region contains secondary structure in the form of a stem-loop. Mutational analysis identified three cis elements within the conserved double-stranded stem region that are functional only in the context of the natural weak 5′ splice site of exon 2; they are 1) a uridine-rich enhancer element in all cell types tested except differentiated chondrocytes; 2) an adenine-rich silencer element, and 3) an enhancer cis element functional in the context of secondary structure. This is the first report identifying key cis elements in the COL2A1 gene that modulate the cell type-specific alternative splicing switch of exon 2 during cartilage development.

Alternative precursor mRNA (pre-mRNA) processing is an important mechanism to increase proteomic diversity in eukaryotes. Through this process two or more mRNA molecules are generated from a single gene, leading to the synthesis of proteins that differ in structure and/or biological function (1). Numerous reports have shown that some alternative splicing events are cell type-specific or developmentally regulated (2–7). Constitutive removal of non-coding introns from pre-mRNA in the nucleus occurs via a complex set of reactions at exon-intron junctions called splice sites. These splice site sequences are recognized by specific small nuclear ribonucleoproteins and accessory protein factors that make up the spliceosome complex (8). Two bona fide intronic sequences are also required for constitutive splicing to occur in addition to the 5′ and 3′ splice sites; they are the branch point sequence and the polypyrimidine tract sequence, both present upstream of the 3′ splice site (8–10). The information content present in these canonical splicing signals is generally not sufficient to ensure correct assembly of the spliceosome, especially in the case of regulated exons. Therefore, additional signals exist in the form of auxiliary cis elements (11–14), which can be present either within the exon or in the flanking introns. Subsequently, splicing can be affected in a positive or negative manner by trans-acting enhancer or silencer splicing factor proteins that bind to these cis elements (15–20). In addition, other regulatory cis elements exist that are functional in the context of RNA secondary structure conformations (21–23).

Although it has been recently estimated that more than half of all human genes generate more than one mRNA due to alternative splicing, information on the molecular processes governing cell-type or developmentally regulated alternative splicing is limited. In this respect, the process of chondrogenesis is an attractive model to study alternative splicing since a number of important cartilage molecules are spliced during chondrocyte differentiation (24). In particular, the cartilage extracellular matrix proteins type II collagen (25), type XI collagen (26), fibronectin (27) and tenascin C (28) are all alternatively spliced during cartilage development where a specific exon(s) encoding potential binding domains are spliced (included) in mRNAs expressed by chondroprogenitor cells but are skipped (excluded) from mRNAs expressed by differentiated chondrocytes. Of these molecules, type II collagen represents the simplest model and the best described alternative splicing event that occurs during chondrogenesis. Type II collagen is the major collagen component of cartilage extracellular matrix and is synthesized as a procollagen molecule of three identical α chains, α1(II), containing an amino and carboxyl propeptide (29). The amino and carboxyl propeptides are subject to cleavage resulting in mature, homotrimeric collagen fibers that form stable fibrils in the extracellular matrix. Only one of the 54 exons encoding COL2A1 is alternatively spliced, producing two mRNA isoforms, type IIA and type IIB procollagen (25). The type IIA procollagen mRNA isoform contains the regulated, cassette exon (exon 2) and is synthesized by chondroprogenitor cells, whereas type IIB procollagen mRNA, devoid of exon 2, is synthesized by differentiated chondrocytes. Transcription of type IIA procollagen occurs in other cell types during embryonic development (30–33), but the developmentally regulated splicing switch from type IIA to type IIB procollagen apparently only occurs during chondrogenesis. In addition, the phenotype of a differentiated chondrocyte is defined by its expression of the type IIB procollagen isoform. Thus, the COL2A1 alternative splicing event essentially defines the process of chondrocyte differentiation and, as such, is an excellent model to study key mechanisms that control cartilage development.

Exon 2 encodes a cysteine-rich (CR) von Willebrand factor C-like domain within the amino propeptide of type II procollagen. Homologues of this CR domain are present in other fibrillar collagen amino propeptides as well as in extracellular matrix proteins including thrombospondins, connective tissue growth factor, and chordin (34, 35). Pre-
Various studies have shown that the CR exon 2-encoded domain of type II procollagen may have an important biological function during development by binding to growth factors such as bone morphogenetic proteins, similar to the function of chordin (35, 36). The presence of the type IIA procollagen isoform in other non-cartilaginous embryonic tissues such as heart, lung, kidney, and eye (30, 32, 37, 38) also suggests an important function for the CR domain during developmental processes. Furthermore, it has been reported that the immature IIA procollagen isoform is re-expressed during cartilage degradation, as seen in osteoarthritis (39), suggesting an additional function for the exon 2-encoded CR domain during tissue repair.

Two studies have been published on COL2A1 alternative splicing at the pre-mRNA level. One report (40) showed that a murine Col2a1 mini-gene was correctly spliced during insulin-dependent chondrocyte differentiation of murine ATDC-5 cells. The same group subsequently showed that deleting large portions of introns 1 and 2 still resulted in correct splicing of the Col2a1 mini-gene in ATDC-5 cells (41). However, to date, there are no reports of specific cis elements or trans-acting splicing factor proteins that are important in regulating human COL2A1 exon 2 alternative splicing.

Using a human COL2A1 mini-gene as a model system, the present study is the first to identify functional cis elements in intron 2 of COL2A1 pre-mRNA that modulate splicing of exon 2. RNA mapping analysis showed that a non-classical consensus splicing region adjacent to the 5′ splice site of exon 2 contains RNA secondary structure in the form of a stem-loop. This is the first study to experimentally show the existence of a stem-loop directly adjacent to a weak 5′ splice site of an exon that is regulated in a tissue-specific manner during development. We report that the double-stranded stem sequence, which is 100% conserved between species, contains both enhancer and silencer cis elements that are functional in regulating type II procollagen isoform expression during chondrocyte differentiation. From the data reported in the present study, we have proposed a model whereby the secondary structure of the stem-loop functions to mask the weak 5′ splice site. Functionally, it is the interaction of enhancer and/or silencer trans-acting splicing factor proteins with cis elements in the stem-loop region that determines the pattern of exon 2 splicing at a specific phase of cartilage development.

**MATERIALS AND METHODS**

**Construction of a Human COL2A1 Mini-gene**—A human COL2A1 mini-gene was constructed spanning exons 1–3, including full-length intron 1 and intron 2 sequences (Fig. 1). Three separate fragments of the mini-gene were synthesized by PCR from human genomic DNA (Clontech) using the elongase amplification system (Invitrogen). Each fragment was amplified using the primer pairs listed in **TABLE ONE** containing specific restriction enzyme sites for sequential cloning into pcDNA3 vector (Invitrogen). The cloned mini-gene (~5.9 kb) is under transcriptional control of the cytomegalovirus promoter. The DNA construct was sequenced to confirm correct orientation and the absence of mutations.

**Transient Transfections of the COL2A1 Mini-gene**—The following cell lines were transfected with the COL2A1 mini-gene: C3H 10T1/2 murine mesenchymal cells (ATCC), MC615 murine vertebral chondrocytes (a gift from Dr. Frederic Mallein-Gerin, Lyon, France), T/C2812 chondrocytes from human costal cartilage (a gift from Dr. Mary Goldring, Harvard University), RCS (LTC) rat chondrosarcoma cells (42), and HEK-293 human embryonic kidney cells. COL2A1 mini-gene in pcDNA3 vector was transfected into each of the cell lines using FuGENE 6 reagent (Roche Applied Science) following the manufacturer’s protocol. Briefly, 1–3 μg of the mini-gene construct was transfected into each cell line at a ratio of 1:4 (μg/μl) DNA:FuGENE for 5 h in serum-free medium. Serum-containing medium was then added (final concentration, 10% fetal bovine serum), and the cells were cultured for a further 48 h until RNA isolation.

**Analysis of Spliced mRNA Isoforms Derived from the COL2A1 Mini-gene**—Total RNA was harvested from each cell line 48 h after transfection using the Qiagen RNeasy kit. Approximately 1 μg of RNA was reverse-transcribed using random primers in a final volume of 20 μl, and the resulting cDNA was diluted to 80 μl with sterile water. 10 μl of cDNA was used for quantitative PCR in the presence of [α-32P]dCTP (10 mCi/ml, 3000 Ci/mmol; Amersham Biosciences). The primers, pcDNA3-COL2A1-Exon1 (5′-CAAGCCTACATGATCGCCGAC-3′) and 5′ EcoRV-Xhol-3′, were used to amplify cDNA derived only from the exogenously transfected COL2A1 mini-gene (Fig. 1). The linear range for these primer pairs was established, and PCR was carried out for 20 cycles: 95 °C for 30 s; 55 °C for 30 s; 72 °C for 30 s. 10 μl of 6% loading dye (30% glycerol, 0.025% (w/v) bromophenol blue, 0.025% (w/v) cyanol blue) was added to each reaction, and 7 μl was electrophoresed at 200 V through 6% polyacrylamide gels. pBR322 DNA digested with MspI was used as a size marker. Gels were dried and exposed to PhosphorImager screens (Amersham Biosciences) for 1 h and then scanned on the STORM™ 840 PhosphorImager (Amersham Biosciences). Bands corresponding to the type IIA (~390 bp) and IIB (~180 bp) mRNA isoforms were quantified using ImageQuant™ software. From these values, ratios of IIA:IIB mRNA spliced products were calculated for each cell type.

**Detection of Aggrecan and Type I Collagen mRNA**—Primers were designed (TABLE TWO) to amplify aggrecan or type I collagen from total RNA extracted from each of the five cell lines (HEK-293, C3H 10T1/2, MC615, T/C2812, and RCS). RT-PCR was carried out in the linear range as determined for each primer pair. Briefly, 2 μg of RNA was reverse-transcribed using random primers in a total reaction volume of 20 μl. An equal volume of water was added to the RT reaction, and 7 μl was used for PCR in the presence of [α-32P]dCTP (10 mCi/ml, 3000 Ci/mmol; Amersham Biosciences) in a total volume of 50 μl. PCR products (6 μl) were electrophoresed through 6% polyacrylamide gels.

---

**TABLE ONE**

| Primer pairs for amplification of the human COL2A1 mini-gene |   |
|---------------------------------------------------------------|---|
| **Construction of a Human COL2A1 Mini-gene** |   |
| Three sets of forward (F) and reverse (R) primer pairs were used to amplify fragments 1, 2, and 3 of the COL2A1 mini-gene (Fig. 1). Restriction enzyme sites at the 5′ and 3′ ends of each amplified product are shown in bold and also underlined in the primer sequence. Numbers in parentheses denote the nucleotide positions of the region amplified from genomic DNA based on the numbering of the published COL2A1 sequence (accession number L10347). |   |
| **Primer pairs** | **Size** |
| Fragment | Primer pairs | Size (kb) |
| 1. Ex1(1)-In1(2550) | F, CCCGAAGCTTACATGATCGCCGAC | ~2.5 |
| 5′-HindIII-BamHI-3′ | R, GCCGTAACCCGATCCCTAG | ~2.6 |
| 2. In1(2531)-In2(5132) | F, CTAGGGATCCGGTACG |   |
| 5′-BamHI-EcoRV-3′ | R, GATAGGATATCTTGTATGGATCTGGGAGG | ~0.8 |
| 3. In2(5132)-Ex3(5908) | F, AATACAAATATCTTATCTGCCCTGAGAG |   |
| 5′-EcoRV-Xhol-3′ | R, CCCCTCGAGCTTTGTTGCTGTTGCGTCCGAA |   |

---

**Regulation of COL2A1 Exon 2 Alternative Splicing**

---
Gels were dried and exposed to a PhosphorImager screen (Amersham Biosciences) and scanned on a STORM™ 840 PhosphorImager (Amersham Biosciences). Bands corresponding to aggrecan, type I collagen, or β-actin were quantified using ImageQuant™ software. Levels of aggrecan and type I collagen mRNA in each cell type were expressed relative to β-actin.

Conservation Analysis of COL2A1 Genomic Sequence—The May 2004 genomic assembly of the human COL2A1 gene (chr12: 46,679,680, 700–746,778) was accessed on the UCSC Genome Browser (genome.ucsc.edu). The species conservation tracks showing the pairwise alignments were obtained through the conservation link. Twelve pairwise-aligned sequence blocks derived from Blastz alignments (43) were scored by phastCons (44). The resulting annotation alignment was confirmed by DNA sequencing and ligated into the COL2A1 mini-gene construct using specific primers containing ClaI restriction sites. The genes was synthesized devoid of large regions (300–500 bp) of intronic sequence of intron 2 using a commercially available kit (U. S. Biochemical Corp.). Reactions were carried out on the original pcDNA3 construct containing the 265-bp cDNA sequence encoding exon 2 and the first 58 nucleotides of intron 2 using a commercially available kit (U. S. Biochemical Corp.). Reactions were carried out in the presence of [α-32P]dATP (10 mCi/ml, 3000 Ci/mmol; Amersham Biosciences), and samples were electrophoresed in parallel with reverse-transcribed RNA digests.

Synthesis of Mutant COL2A1 Mini-genes—A series of mutant mini-genes was synthesized devoid of large regions (300–500 bp) of intron sequence. PCR was carried out to amplify two separate fragments of the mini-gene using specific primers containing ClaI restriction sites. The resulting fragments were gel-purified and ligated and a third PCR was done to amplify the ligated fragment devoid of the intronic sequence of interest. We named the mutant mini-gene with a 370-nucleotide deletion in intron 2 (adjacent to the 5′ splice site of exon 2) deletion mutant 1 (Del 1). Primers used to synthesize Del 1 were PCR 1 (forward primer +2542 BamHI, 5′-CTAGGGGATCCGGGTACGCCG-3′, and reverse primer, −4406 ClaI, 5′-CATTGATATTTACGAGCTAC-3′), PCR 2 (+4778 ClaI, 5′-CCATCGATATTTACGAGCTAC-3′), and reverse primer, −5140 EcoRV, 5′-GATAGGATATTTACGAGCTAC-3′), and PCR 3 (forward primer +2542 BamHI and reverse primer −5140 EcoRV). The deletion mutant cDNA fragment was confirmed by DNA sequencing and ligated into the COL2A1 mini-gene construct using BamHI and EcoRV restriction sites to replace the wild-type 2.6-kb fragment.

### Table Two

| Primers for amplification of aggrecan and type I collagen mRNA | Primer pairs |
|---------------------------------------------------------------|-------------|
| mRNA                                                          |             |
| Human aggrecan                                                | F, AGTGTCATTTGCTAGCGCCAGCCA |
| Mouse aggrecan                                                | R, GTCGTAATAAGGCGG |
| Rat aggrecan                                                  | F, AGGAATCCCTAGCTGTCGAGAGGAT |
| Human type I collagen                                         | R, ACACCTTGCTCTTGAGTGCTGTTGA |
| Mouse type I collagen                                         | F, ATCGCTGAGTACATCTCACAGA |
| Rat type I collagen                                            | R, CTCAGATGTCCATGAGCTCATTCA |
|                                                                 | F, AACGACCAGGGAGAGG |
|                                                                 | R, CATTGCTCAGGGCCAGTTCCAA |
|                                                                 | F, CAGATGTCCATGCTAGTGA |
|                                                                 | R, ACCCTCTCCACACAGACCT |

Each forward (F) and reverse (R) primer pair was used to amplify aggrecan or type I collagen from human (HEK-293, T/C28i2), mouse (C3H10T1/2, MC615), or rat (RCS) cells by RT-PCR. PCR cycle numbers were in the linear range as determined for each primer pair. cDNA was analyzed by PhosphorImager analysis and expressed relative to β-actin. The sequence of each primer is shown in the 5′-3′ direction.

### Notes

- Units of mRNA/microgram (μg) (Invitrogen), S1 nuclease at 2 or 20 units/μg (Invitrogen), or RNase V1 at 0.002 or 0.02 units/μg (Ambion) was added to 1 μg of radiolabeled RNA fragment in a total volume of 20 μl and digested for 15 min at 30 °C. A control aliquot of RNA without the addition of RNases was processed simultaneously with the digested samples. After digestion, RNA was purified by phenol/chloroform extraction, and 1 pmol was reverse-transcribed for 1 h at 37 °C using Superscript™ II RNase H− reverse transcriptase (Invitrogen) with a sense primer that hybridized to a region in exon 2 (5′-GTGAAGACGTGAAAGACTGCCCTCA-3′) (Fig. 6). Samples were then treated with RNase H (0.5 units) for 20 min at 37 °C. After a final phenol/chloroform extraction, RNA was resuspended in gel loading buffer (95% formamide, 18 mM EDTA, 0.025% SDS, 0.025% (w/v) xylene cyanol, 0.025% (w/v) bromphenol blue), and 3 μl was electrophoresed through 6% urea, polyacrylamide denaturing gel. The gels were dried and exposed to a PhosphorImager screen (Amersham Biosciences) overnight and scanned on a STORM™ 840 PhosphorImager (Amersham Biosciences). To localize the sites of RNA digestion, the dideoxy chain termination reaction was carried out on the original pCDNA3 construct containing the 265-bp cDNA sequence encoding exon 2 and the first 58 nucleotides of intron 2 using a commercially available kit (U. S. Biochemical Corp.). Reactions were carried out in the presence of 0.5 μl [α-32P]dATP (10 mCi/ml, 3000 Ci/mmol; Amersham Biosciences), and samples were electrophoresed in parallel with reverse-transcribed RNA digests.
Other mutant mini-gene constructs were synthesized by either substituting or deleting nucleotides near or within the apparent stem-loop sequence in intron 2. TABLE THREE lists all of the COL2A1 mutant mini-genes analyzed in the present study. Mutations were introduced using the QuikChange™ site-directed mutagenesis kit (Stratagene). Briefly, a complementary primer pair (purified by SDS-PAGE; Invitrogen) containing the desired nucleotide substitution or devoid of the nucleotide sequence of interest was added to ~20 ng of substrate DNA. Substrate DNA was prepared by sub-cloning the wild-type 2.6-kb fragment of the COL2A1 mini-gene into pSPT3 vector (Promega) using BamHI and EcoRV restriction sites. This 2.6-kb fragment contains the specific region of intron 2 that was to be mutated. By doing this, the construct size was reduced from ~11.3 kb (size of COL2A1 mini-gene in pcDNA3) to ~5 kb (the COL2A1 mini-gene 2.6-kb fragment in pSPT3) to increase the efficiency of the in vitro mutagenesis procedure. PCR mutagenesis was carried out over 18 cycles (95 °C for 30 s; 55 °C for 1 min; 68 °C for 5.30 min), and the resulting PCR products were digested with DpnI (1 μl) for 1 h at 37 °C to digest parental, methylated DNA. An aliquot (1 μl) of digested DNA was transformed into XL-1 Blue Supercompetent Cells (Stratagene), and resulting colonies were screened for the presence of the correct mutation. Mutant colonies were selected, and the 2.6-kb fragment was re-ligated back into the COL2A1-pcDNA3 construct to create the mutant COL2A1 mini-gene. Transfections of these mutant mini-genes were done as described previously.

### RESULTS

#### Alternative Splicing of the Human COL2A1 Mini-gene in Different Cell Types

Fig. 1 shows the human COL2A1 mini-gene containing exon 1, the regulated cassette exon (exon 2), exon 3, and full-length intervening intron 1 and intron 2. This genomic DNA fragment (~5.9 kb) was cloned into pcDNA3 vector between T7 and sp6 RNA polymerase transcription initiation sites. Cloning of the mini-gene into pcDNA3 was done using the restriction enzyme sites shown. This human mini-gene contains the necessary bona fide sequences to ensure splicing (i.e. removal of introns 1 and 2) in vivo by any cell type. Transfection of this mini-gene construct and subsequent pre-mRNA splicing by cells used in the present study resulted in production of IIA and/or IIB mRNA isoforms that were distinguished by size difference based on the inclusion (IIA) or exclusion (IIB) of exon 2. RT-PCR using the specific primer pair (P1 and P2 in Fig. 1)-amplified cDNA fragments of ~390 and 180 bp that corresponded to the IIA and IIB mRNA spliced isoforms, respectively.

Five different cell lines were selected to analyze splicing of the COL2A1 mini-gene. Human embryonic kidney (HEK-293) cells were chosen as a source of non-chondrocytes. C3H 10T1/2 cells were included as a source of chondroprogenitors as these cells can be induced to undergo differentiation in the correct culture environment (48, 49). MC615 and T/C2812 cells are transformed chondrocytes isolated from mouse vertebrae and human costal cartilage, respectively; these cells were expected to be in a de-differentiated state in the culture conditions used in the present study. Finally, rat chondroasarcoma (RCS) cells were chosen as a source of differentiated chondrocytes (42). To confirm the differentiation status of these cells, RT-PCR was carried out on RNA isolated from each of the cell lines to analyze the levels of aggrecan and type I collagen mRNA. Aggrecan is a chondrocyte marker, whereas type I collagen is a marker of de-differentiated cells in culture. Fig. 2 shows the levels of aggrecan and type I collagen mRNA relative to β-actin

### Table Three

| Mutant mini-gene | Type and site of mutation |
|------------------|---------------------------|
| Del 1            | Deletion in intron 2 (+11 to +380) |
| +5′/5′           | Nucleotide substitutions, TGTA→AAAGT (+3 to +6) |
| SL-Del-1         | Deletion of stem loop (+7 to +43) |
| +5′/5′/SL-Del-1  | Combination of nucleotide substitution and stem loop deletion |
| SL-Del 2         | Partial deletion of stem loop (+7 to +18) |
| SL-Del 3         | Partial deletion of stem loop (+29 to +41) |
| CCC-1            | Nucleotide substitutions, TTT→CCC (+8 to +10) |
| CCC-2            | Nucleotide substitutions, TTT→CCC (+8 to +10) and TTT→CCC (+12 to +14) |
| CCC-3            | Nucleotide substitutions, TTT→CCC (+8 to +10), TTT→CCC (+12 to +14), and TTT→CCC (+16 to +18) |
| GGG-1            | Nucleotide substitutions, AAA→GGG (+35 to +37) |
| GGG-2            | Nucleotide substitutions, AAA→GGG (+35 to +37) and AAA→GGG (+31 to +33) |
| CCC-17/GGG-1     | Nucleotide substitutions, TTT→CCC (+8 to +10) and AAA→GGG (+35 to +37) |

FIGURE 1. Construction and alternative splicing of the human wild-type COL2A1 mini-gene. The top panel shows the COL2A1 mini-gene containing exons (E) 1–3 with full-length intervening introns (I). Three separate fragments of the mini-gene were amplified from human genomic DNA and ligated into pcDNA3 vector using the restriction enzyme sites shown. The middle panel shows the two alternative pre-mRNA splicing mechanisms to remove either introns 1 and 2 (resulting in IIA mRNA) or introns 1 and 2 and exon 2 (resulting in IIB mRNA). Numbers represent the nucleotide size of each exon and intron in the mini-gene. The bottom panel shows the IIA and IIB mature mRNA transcripts and the approximate nucleotide sizes of each after PCR amplification using the specific primers, P1 and P2. P1 = pcDNA3-COL2A1-Exon1 primer (see “Materials and Methods”), and P2 = SP6 primer.
mRNA for each cell line, and the ratio of aggrecan/type I collagen expression is also shown. As expected, RCS cells expressed the highest ratio of aggrecan/type I collagen, although undifferentiated C3H 10T1/2 and MC615 cells expressed the lowest ratios. The aggrecan/type I collagen ratio value shown for T/C28I2 cells suggests that these cells are in an intermediate stage of differentiation. Therefore, based on this knowledge, we would expect to see different patterns of COL2A1 mini-gene splicing where levels of the type IIA mRNA isoform derived from the mini-gene would exceed those of the type IIB mRNA isoform in chondroprogenitor cells and vice versa in chondrocytes. Fig. 3 shows that the mini-gene spliced products amplified by RT-PCR are consistent with the differentiation status of these cells. The C3H 10T1/2 and MC615 cells contained the highest ratio of IIA/IIB mRNA isoforms; T/C28I2 cells contained a lower IIA/IIB ratio in comparison, whereas the RCS cells spliced the mini-gene to produce more of the IIB mRNA isoform. To confirm the efficacy of using this COL2A1 mini-gene as a model system to study regulation of exon 2 alternative splicing, the ratio of endogenous type I collagen and type IIB collagen isoforms in C3H 10T1/2, MC615, T/C28, and RCS cells was found to be similar to that derived from the COL2A1 mini-gene (results not shown). The non-chondrocyte HEK-293 cells spliced the mini-gene to produce ~2-fold more of the type IIA mRNA isoform than the type IIB mRNA isoform; levels of endogenous type II collagen were undetectable in these cells.

Conservation of COL2A1 Sequence

Fig. 4 shows a sequence alignment comparison of a region of COL2A1 genomic sequence between human, chimp, mouse, rat, dog, zebrafish, and fugu (puffer fish). The alignment shows intronic sequence spanning the 3' and 5' splice sites of exon 2 as well as the first and last 20 nucleotides of exon 2. In all species, the 3' splice site is shown to conform to the classical consensus sequence, (\(5'\)GTGAGGTTGGAGG\(3'\)), where the arrow denotes the exon-intron junction, and highly conserved nucleotides are shown in underlined bold font. Upstream of the 3' splice site is a long polypyrimidine tract sequence which may be an important feature in the regulation of exon 2 splicing. The 5' splice site sequence of the COL2A1 gene (\(5'\)GTGAGGTTGGAGG\(3'\)) does not conform to the classical consensus sequence (\(5'\)GTGAGGTTGGAGG\(3'\)), and is, thus, referred to as a “weak” 5' splice site. This is the first report showing that a weak 5' splice site is present adjacent to the alternatively spliced exon in the type II procollagen gene from a number of different species. It was previously reported that a region directly downstream of the 5' splice site may contain RNA secondary structure in the form of a stem-loop (29). The nucleotides predicted to form the double-stranded RNA of the stem-loop are 100% conserved between all species analyzed (Fig. 4). Therefore, there is a high likelihood that this region of intron 2 contains regulatory cis elements involved in pre-mRNA splicing regulation due to 1) the location, adjacent to an alternatively spliced exon, 2) the conservation between species, and 3) the potential of secondary structure formation.

Altered Splicing of a COL2A1 Mini-gene Devoid of Intron 2 Sequence

To test for the presence of functional intronic splicing cis elements, a series of mutant COL2A1 mini-genes was synthesized devoid of large regions of introns 1 or 2. One of these deletion mutants, named Del 1, showed a marked difference in alternative splicing patterns compared with splicing of the wild-type COL2A1 mini-gene. Fig. 5 shows that the region of intron 2 deleted from the Del 1 mini-gene is a 370-bp fragment from intron 2 nucleotide numbers +11 to +380 directly adjacent to the 5'
splice site of exon 2. In the Del 1 mutant, the intronic 5’ splice site nucleotides were not deleted, so that splicing to either include or exclude exon 2 could still potentially occur. The phosphorimage in Fig. 5 shows that all cell types processed the Del 1 mutant in a similar way, producing only the type IIB isoform. This suggests that the weak 5’ splice site alone is not sufficient to promote exon 2 splicing (inclusion). In addition, the deleted region of intron 2 may also contain essential regulatory elements that modulate the distinct splicing patterns of exon 2 in cells at various stages of chondrocyte differentiation.

Analysis of RNA Secondary Structure in Intron 2 Adjacent to the 5’ Splice Site

Using the Zuker Mfold program to predict secondary structure, it was found that the region in intron 2 directly adjacent to the 5’ splice site of exon 2 (intron 2 nucleotide number 4 to +41) is likely to contain RNA secondary structure in the form of a stem-loop. Fig. 6 shows the location of this stem-loop and the predicted Mfold structure and ΔG (−7.7 kcal). To experimentally determine the presence of an RNA stem-loop structure in intron 2, we performed RNase mapping analysis. RNase T1 digests single-stranded sites, preferentially adjacent to guanine residues, and S1 nuclease digests single-stranded RNA with no particular specificity, whereas RNase V1 digests sites of double-stranded RNA. The

region of the COL2A1 mini-gene that was probed was a 145-nucleotide fragment containing the last 87 nucleotides of exon 2 and the first 58 nucleotides of intron 2. The phosphorimage of the polyacrylamide
Regulation of COL2A1 Exon 2 Alternative Splicing

FIGURE 6. Analysis of RNA secondary structure in human COL2A1 intron 2 adjacent to the 5′ splice site of exon 2. The top panel shows the position of the predicted stem-loop structure in the human COL2A1 pre-mRNA spanning nucleotides 4401–4438 (numbering is based on a published sequence, accession number L10347). P1 indicates the position of the primer used for in vitro transcription to produce a radiolabeled RNA fragment encoding the last 87 nucleotides of exon 2 and the first 58 nucleotides of intron 2. Bottom panel, left, shows the stem-loop structure as predicted by Mfold with a ΔG of −7.7 kcal. Bottom panel, right, is a Phosphorimage scan of a sequencing polyacrylamide gel showing sites of enzymatic cleavage of the in vitro transcribed RNA probe. Bands corresponding to the single-stranded loop region were seen after digestion with RNase T1 or S1 nuclease digestion sites indicated that a single-stranded loop is present. Digestion sites corresponding to the regions of double-stranded RNA were seen in the lanes corresponding to RNase V1 digestions only. This is indicative of the stem region and confirms that RNA stem-loop secondary structure is present in intron 2 of human COL2A1, directly adjacent to the 5′ splice site.

Effect of a Weak 5′ Splice Site on Alternative Splicing of COL2A1 Exon 2

The COL2A1 genomic sequence alignment in Fig. 4 shows conservation of a potentially weak 5′ splice site in all species. To determine the effect of the 5′ splice site sequence on alternative splicing of exon 2, we synthesized a mutant mini-gene (named +5′ SS) with a four nucleotide substitution in intron 2 (TGTAA → TAAAGT) to create a strong splice site that conforms to the classical consensus sequence. The +5′ SS COL2A1 mini-gene was spliced similarly by all cell types to produce only the type IIA mRNA isoform (Fig. 7, second lane of each gel panel). This suggests that the presence of a weak 5′ splice site is important to confer the differential cell type-specific splicing patterns of COL2A1 shown in the present study.

Effect of Intron 2 (stem-loop) Deletions on Alternative Splicing of COL2A1 Exon 2

To determine whether the stem-loop sequence (Fig. 6), located directly adjacent to the weak 5′ splice site, contains regulatory elements that modulate exon 2 splicing, a series of deletion mutant COL2A1 mini-genes were synthesized (TABLE THREE). Stem-loop deletion 1 mutant mini-gene (SL-Del 1) was produced by deleting most of the stem-loop region (intron 2 nucleotides +7 to +43; Fig. 7) except the first three nucleotides (+4 to +6) that are part of the bona fide splice site sequence. Because the double-stranded stem sequence of the stem-loop showed 100% sequence similarity between species, we also produced deletion mutant mini-genes devoid of the uridine-rich region of the stem (SL-Del 2) or the opposite, adenine-rich region of the stem (SL-Del 3). These mutants are shown diagrammatically in the top panel of Fig. 7. In all of these deletion mutant mini-genes, the 5′ splice site intronic sequence (+1 to +6) that binds to U1 snRNA (8) was intact. The PhosphorImager gel pictures in Fig. 7 show that, compared with wild-type mini-gene splicing, deletion of the stem-loop sequence resulted in a marked inhibition of exon 2 splicing (inclusion), favoring type IIB mRNA production (SL-Del 1; third lane of each gel panel). This splicing pattern was displayed by all cells regardless of the cell type or differentiation status, again suggesting that the weak 5′ splice site alone is not sufficient to induce exon 2 splicing and that a cis enhancer element-promoting spliceosome assembly at the 5′ splice site of exon 2 was removed. A mutant mini-gene combining both the strong 5′ splice site sequence together with deletion of most of the stem-loop region (+5′ SS/SL-Del-1; Fig. 7, fourth lane of each gel panel) showed exclusively exon 2 inclusion to produce only the IIA mRNA isoform. This splicing pattern was shown by all cell types and confirms that the strong splice site sequence compensated for the deleted stem-loop region to promote exon 2 splicing. This suggests that 1) the stem-loop sequence does not contain bona fide nucleotides necessary for constitutive splicing (inclusion) of exon 2 and 2) the cell type-specific COL2A1 splicing patterns shown in the present study are dependent on the presence of both a weak 5′ splice site and the adjacent stem-loop sequence.

Removal of the left side of the stem loop containing the uridine stretch sequence (AUUUUUUUUUUU; SL-Del 2), resulted in an apparent decrease in the ratio of IIA:IIB mRNA transcripts compared with splicing of the wild-type mini-gene in all cell types except RCS cells (Fig. 7, fifth lane of each gel panel). This suggests that an enhancer site is located within this uridine stretch that is not functional in RCS cells. However, by deleting the adenine-rich region of the stem (AUAAAUUAAAU; SL-Del 3) the opposite effect was found whereby all cell types, including the differentiated chondrocyte RCS cells, spliced the mutant mini-gene to produce predominantly type IIA mRNA (Fig. 7, sixth lane of each gel panel). This suggests removal of a splicing
silencer element. Analysis of the intronic sequence after deletion of the 10 nucleotides comprising the adenine-rich side of the stem-loop showed that we had not created an alternative, stronger 5′ splice site. This was confirmed by DNA sequencing of the amplified cDNA corresponding to the IIA mRNA isoform derived from SL-Del 3 mutant mini-gene (data not shown).

**Effect of Stem-loop Nucleotide Substitutions on COL2A1 Exon 2 Alternative Splicing**

**Mutation of the Uridine-rich Stretch**—To specifically localize regulatory cis elements within the RNA stem-loop, nucleotide substitutions were introduced on either side of the conserved double-stranded stem. Because of reports of splicing factor proteins that bind to intronic uridine-rich regions in pre-mRNAs to regulate alternative splicing, we first synthesized three mutant mini-genes named CCC-1, CCC-2, and CCC-3 that contained cytosines in place of the first, second, and third triplet set of uridine nucleotides within the double-stranded stem, respectively (TABLE THREE; Fig. 8). Analysis of the spliced mRNA products derived from the CCC-1 mutant mini-gene suggested that this mutation had either a minor effect or no effect in inhibiting exon 2 splicing (i.e. exon 2 skipping) in HEK-293, MC615, T/C2812, and C3H 10T1/2 cells when compared with splicing of the wild-type mini-gene (Fig. 8). However, splicing of CCC-2 and CCC-3 mini-genes by these cell types showed a trend toward type IIB mRNA production, concomitant with an increasing number of uridine substitutions. This result is in agreement with that from splicing of the uridine-stretch deletion mutant (SL-Del2; Fig. 7) in these cell types, confirming the presence of a functional enhancer cis element, particularly within the second and third uridine triplets.

Interestingly, in RCS cells, splicing of the CCC-1 mutant mini-gene produced a significantly higher ratio of IIA:IIB mRNAs compared with splicing of the wild-type mini-gene. This splicing ratio was also higher than that achieved by splicing of the CCC-1 mini-gene in the other four cell types. In addition, the IIA:IIB mRNA splicing ratio decreased by mutating the second (CCC-2) and third (CCC-3) set of uridines. However, levels of IIA:IIB mRNA derived from these mini-genes did not fall below that derived from the wild-type mini-gene, suggesting that the uridine-rich region is not an enhancer site but, rather, a silencer element.
Regulation of COL2A1 Exon 2 Alternative Splicing

in differentiated chondrocytes. The observation that the CCC-3 splicing result was similar to that of the uridine-stretch deletion mutant (SL Del-2; Fig. 7) in RCS cells also supports the latter statement. In the case of CCC-1 mutant mini-gene splicing, it is possible that this mutation (Del-2; Fig. 7) in RCS cells also supports the latter statement. In the case of uridine-stretch deletion mutant (SL Del-2; Fig. 7) in RCS cells also supports the latter statement.

Another mutant mini-gene was also created (CCC-1/GGG-1) to study the effect of secondary structure on exon 2 splicing. Phosphorimage gels show reverse-transcribed IIA and IIB mRNA transcripts from splicing of wild-type (WT) and CCC mutant mini-genes in each cell type used in the present study. Splicing patterns shown from each mutant mini-gene represent one of at least four replicate experiments for each cell line.

Mutation of the Adenine-rich Stretch—To analyze potential regulatory elements on the opposite side of the stem, two mutant mini-genes were made, named GGG-1 and GGG-2, which contained guanines in place of one or both sets of adenine triplet nucleotides, respectively (TABLE THREE; Fig. 9). Splicing of these mini-genes by all cell types resulted in a higher IIA:IIB mRNA ratio by splicing of the CCC-1 mini-gene compared with wild type in HEK-293 and T/C28 cells is in keeping with disruption of a natural enhancer cis element.

FIGURE 8. Effect of alternative splicing of COL2A1 exon 2 by mutations in the uridine-rich site of the stem-loop. The top panel shows the location of cytosine nucleotides in CCC-1, CCC-2, and CCC-3 mutant mini-genes that replaced the first, second, and third set of uridine triplets in the stem-loop, respectively. Phosphorimage gels show reverse-transcribed IIA and IIB mRNA transcripts from splicing of wild-type (WT) and CCC mutant mini-genes in each cell type used in the present study. Splicing patterns shown from each mutant mini-gene represent one of at least four replicate experiments for each cell line.

FIGURE 9. Effect of adenine mutations and secondary structure compensatory mutations on the alternative splicing of COL2A1 exon 2. The top panel shows the location of guanine nucleotides in GGG-1 and GGG-2 mutant mini-genes that altered either one or both adenine triplets in the stem-loop, respectively. Another mutant mini-gene was also created (CCC-1/GGG-1) to study the effect of secondary structure on exon 2 splicing. Phosphorimage gels show reverse-transcribed IIA and IIB mRNA transcripts from splicing of wild-type (WT) and mutant mini-genes in each cell type used in the present study. Splicing patterns shown from each mutant mini-gene represent one of at least four replicate experiments for each cell line.
DISCUSSION

To identify key cis regulatory elements that modulate alternative splicing of exon 2 during chondrogenesis, we used a human COL2A1 mini-gene as a model system (Fig. 1). A recent report described a murine Col2a1 mini-gene that was successfully spliced by ATDC-5 cells during a 21-day chondrocyte differentiation assay system (40). However, it was not specified if the mini-gene was stably transfected to detect spliced products after 3 weeks in culture. Splicing of our human COL2A1 mini-gene was analyzed at one time point in cells at various stages of differentiation; C3H 10T1/2 cells were used as a source of chondroprogenitors, MC615 and T/C2812 cells were a source of de-differentiated chondrocytes, and RCS cells were a source of differentiated chondrocytes. Variations in IIA- and IB-spliced mRNA isoforms derived from the mini-gene correctly reflected the differentiation status of the cells. This indicated that the COL2A1 mini-gene contained the necessary cis elements required for correct splicing of exon 2 within different cellular contexts. The fact that the HEK-293 cells spliced the mini-gene to produce more of the type IIA isoform suggests that similar cis- and/or trans-acting factors are functional in non-chondrocyte cells as they are in chondroprogenitor and de-differentiated chondrocytes.

In the present study we show that splicing of a mutant mini-gene (Del 1; Fig. 5) devoid of 370 bp of intron 2 adjacent to the 5’ splice site resulted in exon 2 exclusion in all cell types. The importance of intron 2 sequences in regulating exon 2 splicing was also reported by Nishiyama et al. (41), who found that processing of a murine Col2a1 deletion mini-gene devoid of ~92% of intron 2 by ATDC-5 cells resulted in IIA and IB mRNA transcripts in addition to an abnormal increase in splicing intermediates. This mini-gene contained nucleotides downstream of the 5’ splice site that were not present in the Del 1 mutant reported here. Therefore, the absence of exon 2 splicing (inclusion) in the Del 1 mutant mini-gene indicated that key, non-consensus splicing enhancer cis elements are present in intron 2 adjacent to the 5’ splice site. A previous report predicted that a stretch of nucleotides in this region may contain secondary structure in the form of a stem-loop (29). By a series of RNase digestions, we showed that this region of intron 2 (nucleotide numbers 4 to +61) does indeed form a stem-loop. Because of location, sequence similarity between species, and secondary structure conformation, we hypothesized that the stem-loop is important in the regulation of exon 2 alternative splicing. Furthermore, nucleotides +4 to +6 of the stem-loop, which are part of the intronic sequence that interacts with U1 snRNA (8), does not conform to the classical 5’ splice site consensus sequence. We identified this weak 5’ splice site as another region of the type II procollagen gene that is conserved between species. Weak 5’ and/or 3’ splice sites are a common feature of many alternatively spliced exons. Similar to COL2A1, reports have been published of other genes containing regulated exons with weak splice sites that are differentially spliced during development, including fibronectin (27), cardiac troponin T (50), myosin phosphatase-targeting subunit-1 (MYPt-1) (51, 52), and protein 4.1R (5). We showed that the presence of a weak 5’ splice site is necessary for the cell type-specific splicing patterns of COL2A1 since conversion to a strong splice site resulted exclusively in exon 2 splicing (type IIA mRNA) regardless of the cell type and state of differentiation.

The functional significance of the stem-loop sequence in regulating exon 2 splicing was demonstrated by constructing a series of mutant mini-genes that contained either deletions or substitutions in this region. Deletion of the entire stem-loop, except nucleotides +4 to +6 that make up the intronic splice site sequence, resulted in a marked inhibition of exon 2 splicing in all cells (SL-Del 1; Fig. 7). Therefore, the weak 5’ splice site alone is not sufficient to yield exon 2 splicing and, thus, requires additional cis elements present in the stem-loop. The stem-loop does not contain constitutive splicing elements since deletion of this region in combination with a strong splice site adjacent to exon 2 resulted exclusively in exon 2 splicing (+5’ss/SL-Del 1; Fig. 7). Thus, the cis regulatory elements in the stem-loop are functional only in the context of a weak 5’ splice site. The importance of non-consensus intronic splicing sequences downstream of the weak 5’ splice site of a regulated exon has been reported in other genes. For example, pyrimidine-rich regions are required for inclusion of the alternatively spliced K-SAM exon in the FGFR-2 gene (53) or exon 6A in the β-tropomyosin gene (54). Similar to COL2A1 exon 2 splicing, these regulatory sequences are not functional when the 5’ splice site sequence is optimized.

Studies of yeast splicing commitment complexes have identified a number of proteins that bind to non-consensus intronic regions downstream of 5’ splice sites; the functional significance of these interactions in stabilizing the U1 small nuclear ribonucleoprotein-pre-mRNA complex to enhance splicing of an exon was also suggested by the authors (55, 56). Therefore, we hypothesized that enhancer cis elements are present in the stem-loop to promote splicing of exon 2 and that these sites would be more functional in cells that naturally express more of the type IIA isoform. Mutations within the conserved double-stranded region of the stem-loop revealed that the uridine (U)-rich site of the stem-loop is a functional enhancer element in all cell types tested except the differentiated chondrocytes (Fig. 7, SL-Del2; Fig. 8). A U-rich enhancer element was identified downstream of the central regulated exon in the MYPT-1 gene which, like COL2A1 exon 2, contains a weak 5’ splice site and is spliced in a developmentally regulated manner (57). Further studies showed that the splicing factor protein TIA-1 interacts with this MYPT-1 U-rich enhancer (58). By binding to U-rich elements, TIA-1 stabilizes the U1 small nuclear ribonucleoprotein complex association at the 5’ splice site, thereby promoting splicing of the regulated exon (18, 59). Furthermore, Shukla and co-workers (58) also reported that decreased in vivo expression levels of TIA-1 was concomitant with the alternative splicing switch (exclusion) of the regulated MYPT-1 exon. Therefore, TIA-1 is a potential candidate protein that may interact with the U-rich region of the stem-loop downstream of COL2A1 exon 2. A sequence corresponding to part of the COL2A1 U-rich element (AUUUAUUU) is also present within a larger splicing enhancer element identified downstream of the MYPT-1 alternative exon (57). This U-rich stretch is not in the context of secondary structure in the MYPT-1 gene, therefore suggesting the possibility that this element can function as a linear sequence in different cell types. Interestingly, this U-rich element apparently contains some silencing activity in the differentiated chondrocytes used in the present study. This points to the likelihood that different trans-acting splicing factors can bind to the same regulatory cis element and that the expression or regulation of specific splicing factor proteins changes during chondrocyte differentiation. This statement is also supported by splicing of the CCC-1 mutant mini-gene (Fig. 8) in differentiated chondrocytes compared with the other cell types. This mutation did not create an improved intron binding site for U6 snRNA, which displaces U1 snRNA during the constitutive splicing process (8). Here, we predict that this mutation created a novel enhancer site that was recognized by a trans-acting factor present or functional in the RCS cells only.

Mutations in the opposite, adenine (A)-rich region of the double-stranded stem suggested disruption of a functional silencer element in all cell types tested (Fig. 7, SL-Del 3; Fig. 9). The presence of both positive and negative splicing cis elements situated in close proximity within either a regulated exon or an intron has been identified in pre-mRNA encoding a number of proteins including fibronectin, MYPT-1, and tau, for example (57, 60–62). Some of these regulatory sites can function as
Regulation of COL2A1 Exon 2 Alternative Splicing

linear, independent elements, whereas others are functional only in the presence of the adjacent, antagonistic cis element (22, 62). We cannot conclude from the present study that the U-rich and A-rich regions present in the stem-loop can function independently of each other in a linear context to regulate exon 2 splicing. For example, deletion or nucleotide substitution of the A-rich element that resulted in increased exon 2 splicing in all cell types may be due to 1) loss of the silencer activity functional in this region, independent of the U-rich site, 2) disruption of stem-loop secondary structure, thereby allowing the U-rich element to function better as a linear enhancer site, or 3) a combination of the above two scenarios.

The importance of stem-loop secondary structure in regulating COL2A1 exon 2 splicing is supported by the results obtained from the mutant mini-gene CCC-1/GGG-1 (Fig. 9). This mini-gene was constructed to restore secondary structure to the CCC-1 mutant mini-gene, thereby creating a more stable stem-loop structure (ΔG = -13.6 kcal; wild-type stem-loop ΔG = -7.7 kcal). An altered splicing pattern derived from the CCC-1/GGG-1 mini-gene was noted in all cell types whereby exon 2 splicing was inhibited. This suggests the involvement of another enhancer protein or protein complex that binds specifically to double-stranded RNA in the stem-loop-containing adjacent U-A residues. Numerous reports have described the effects of RNA secondary structure on the regulation of splicing (63, 64). In general, with respect to alternative exon splicing, the presence of stem-loop structures negatively regulates splicing by masking the splice site sequence, thereby preventing splicingosome formation at the exon-intron junction. For example, secondary structure was shown to sequester the alternative exon 6B in the chicken β-tropomyosin pre-mRNA, resulting in its exclusion from the final mRNA (65–67). A stem-loop situated downstream of the alternative exon in the human growth hormone gene influenced its splicing since mutations that stabilized the stem-loop resulted in use of an alternative splice site (68). To our knowledge, there are only two reports postulating the presence of stem-loop secondary structure overlapping a weak 5′ splice site of an alternatively spliced exon (58, 69). In both cases it was suggested that the hypothesized stem-loop functioned in masking the weak 5′ splice site of the alternative exon. Importantly, the present study is the first to experimentally show the existence of a stem-loop at the weak 5′ splice site of an alternative exon that is regulated in a tissue-specific manner during development. We are in agreement with the generally accepted function of regulatory sites occupied by trans-acting splicing factors during COL2A1 splicing. In addition, the less conserved single-stranded loop region was not determined in the present study. A fine balance of positive and negative-acting proteins likely regulates the COL2A1 splicing pattern, and concentrations of these factors will undoubtedly change in accordance with the differentiation status of the cells during chondrogenesis.

Acknowledgment—We thank Dr Eric Wagner for critical reading of this manuscript.

REFERENCES

1. Black, D. L. (2003) Annu. Rev. Biochem. 72, 291–336
2. Roessler, J. R., Liittschwager, K., and Leff, S. E. (1993) J. Biol. Chem. 268, 8366–8375
3. Cooper, T. A., and Ordahl, C. P. (1983) J. Biol. Chem. 260, 11140–11148
4. Fujita, H., Ohta, M., Kawasaki, T., and Itoh, N. (1991) Biochem. Biophys. Res. Commun. 174, 946–951
5. Hou, V. C., and Conboy, J. G. (2001) Curr. Opin. Hematol. 8, 74–79
6. Gouding, C., Roberts, G. C., Moreau, G., Nadal-Ginard, B., and Smith, C. W. (1994) EMBO J. 13, 3861–3872
7. Gallego, M. E., Sirand-Pugnet, P., Durosay, P., Clouet d’Orval, B., d’Aubenton-Carafa, Y., Brody, E., Expert–Bezancon, A., and Marie, J. (1996) Biochimie (Paris) 78, 457–465
8. Burge, C. B., Tuschl, T., and Sharp, P. A. (1999) in The RNA World (Gesteland, R., Cech, T., and Atkins, J., eds) 2nd Ed., pp. 525–560, Cold Spring Harbor Laboratory
