Published in final edited form as:

Connect Tissue Res. 2014 June ; 55(3): 165–176. doi:10.3109/03008207.2014.908860.

Alternative Splicing of Type II Procollagen: IIB or not IIB?

Audrey McAlinden, Ph.D.

Department of Orthopaedic Surgery, Musculoskeletal Research Center, Washington University
School of Medicine, St Louis, Missouri, USA

Abstract

Over two decades ago, two isoforms of the type II procollagen gene (*COL2A1*) were discovered. These isoforms, named IIA and IIB, are generated in a developmentally-regulated manner by alternative splicing of exon 2. Chondroprogenitor cells synthesize predominantly IIA isoforms (containing exon 2) while differentiated chondrocytes produce mainly IIB transcripts (devoid of exon 2). Importantly, this IIA-to-IIB alternative splicing switch occurs only during chondrogenesis. More recently, two other isoforms have been reported (IIC and IID) that also involve splicing of exon 2; these findings highlight the complexities involving regulation of *COL2A1* expression. The biological significance of why different isoforms of *COL2A1* exist within the context of skeletal development and maintenance is still not completely understood. This review will provide current knowledge on *COL2A1* isoform expression during chondrocyte differentiation and what is known about some of the mechanisms that control exon 2 alternative splicing. Utilization of mouse models to address the biological significance of *Col2a1* alternative splicing *in vivo* will also be discussed. From the knowledge acquired to date, some new questions and concepts are now being proposed on the importance of *Col2a1* alternative splicing in regulating extracellular matrix assembly and how this may subsequently affect cartilage and endochondral bone quality and function.

Keywords

type II procollagen; precursor mRNA; alternative splicing; chondrocyte differentiation; cartilage; type XI collagen; extracellular matrix

Introduction

The removal of introns (i.e. splicing) from precursor mRNAs (pre-mRNAs) is a precise, regulated process that is necessary for the expression of most human genes. Pre-mRNA splicing requires a complex array of *cis* elements (the splicing code) embedded in pre-mRNAs and *trans*-acting factors that bind to these elements (1). The basic core of the splicing code for the majority of introns comprises short sequences at or near the exon-intron junction (i.e. 5' and 3' splice sites and a branch point sequence). Essentially, the splicing process is performed by the spliceosome, a ribonucleoprotein megaparticle that assembles around the splice sites (2). In general, strong splice sites that conform well to the

Declaration of interest

The author reports no conflicts of interest. The author alone is responsible for the content and writing of the paper.

known consensus sequence (3) leads to constitutive splicing and inclusion of the exon in the mature mRNA.

However, **alternative splicing** of pre-mRNA can also occur which is the major cellular mechanism for generating protein diversity (4, 5). In most cases, alternatively-spliced exons have what are termed “weak” 3′ or 5′ splice sites meaning that their sequence does not conform well to that of known consensus sequences (3). Alternative splicing permits the expression of multiple mRNA isoforms from an individual gene which often encodes protein isoforms of different functionality (6). Various modes of alternative splicing can occur including: 1) inclusion/exclusion of an entire exon; 2) utilization of alternative 5′ or 3′ splice sites resulting in lengthening or truncation of an exon; 3) different transcription start sites (i.e. alternative promoters); 4) alternative polyadenylation sites; or 5) removal or retention of an intron (4, 5, 7). The degree to which weak sites are used is regulated by both non splice site *cis*-regulatory sequences and *trans*-acting factors. *Cis*-regulatory sequences include exonic splicing enhancers (ESEs), exonic splicing silencers (ESSs), intronic splicing enhancers (ISEs) and intronic splicing silencers (ISSs), depending on their locations and how they affect usage of the splice site (5). *Trans*-acting factors function through binding to *cis* splicing enhancers or silencers and include members of the well characterized Ser/Arg-rich and heterogeneous nuclear ribonucleoprotein (hnRNP) protein families (8, 9). A growing number of studies demonstrate that alternative splicing is utilized to adjust gene output in accordance with physiological requirements. Given that more than 90% of human intron-containing genes undergo alternative splicing, many aspects of biology are affected as a result (5, 6).

During the process of cartilage tissue development, a number of genes expressed by differentiating chondrocytes are subject to alternative splicing as has been previously reviewed (7). For the purpose of this review, specific focus will be placed on alternative splicing of the type II procollagen gene (*COL2A1* in human; *Col2a1* in mouse) (10–12). *COL2A1* encodes a homotrimeric protein of three identical $\alpha 1(\text{II})$ chains that form a triple-helical domain flanked by a 5′ amino (NH_2) propeptide and a 3′ carboxy (COOH) propeptide (Figure 1). Processing of these procollagens to remove the NH_2 and COOH propeptides must occur to permit formation of mature type II collagen molecules. These processed triple-helical collagens then form a covalently cross-linked fibrillar network in the extracellular matrix (ECM) that provides tensile strength to cartilage tissue (13–16). In addition to $\alpha 1(\text{II})$ chains, *COL2A1* also encodes $\alpha 3(\text{XI})$ protein chains that trimerize with $\alpha 1(\text{XI})$ and $\alpha 2(\text{XI})$ chains to form heterotrimeric type XI procollagen (17–19) (Figure 1). This fibrillar collagen is a minor component of the cartilage ECM but plays an important role in controlling type II collagen fibrillogenesis and fibril diameter (20). Therefore, within the context of *COL2A1/COL11A3* alternative splicing, one must consider that alternative splicing will generate different isoforms of type II procollagen as well as type XI procollagen. This will be elaborated on further when discussing mouse models to study alternative splicing *in vivo*.

Alternatively-spliced isoforms of *COL2A1*: IIA and IIB

Over two decades ago, two isoforms of type II procollagen were identified as a result of pre-mRNA alternative splicing of exon 2 (21). Type IIA isoforms contain exon 2 while type IIB isoforms are devoid of exon 2 (Figure 1). Interestingly, alternative splicing of exon 2 was found to be a developmentally-regulated event during chondrogenesis whereby chondroprogenitor cells generate predominantly IIA isoforms while differentiated chondrocytes produce mainly IIB isoforms (22–26). It should be noted that, in these studies, the differential expression patterns of IIA and IIB mRNA isoforms found by *in situ* hybridization cannot distinguish between $\alpha 1(\text{II})$ or $\alpha 3(\text{XI})$ mRNA. Importantly, this developmentally-regulated IIA-to-IIB alternative splicing switch is a unique event restricted to differentiating chondrocytes. Early transient expression of type II procollagen has been found in various non-chondrogenic mouse and human embryonic tissues, but only the IIA isoform is expressed at these sites (23, 25). Expression of IIA mRNA in non-chondrogenic regions is largely conserved between mouse and human and includes epithelial structures that participate in epithelial-mesenchymal interactions (e.g. apical ectodermal ridge of the limb bud, pharyngeal endoderm and bronchial epithelium); these expression patterns suggest a potential role in regulating morphogenesis. As will be discussed later, a morphogenetic role for IIA procollagen has been demonstrated in developing *Xenopus* embryos (27) but not convincingly so in mouse models that have been engineered to either lack exon 2 or persistently express exon 2-containing *Col2a1* isoforms throughout development (28, 29). With respect to post-natal cells/tissue, both IIA and IIB isoforms have been identified in the vitreous of the bovine eye (30, 31) as well as in myelinating and non-myelinating Schwann cells of rat peripheral nerve (32). So, while alternative splicing of exon 2 occurs in these non-chondrocyte cells, there is no apparent developmentally-regulated IIA-to-IIB switch as occurs during chondrogenesis. This suggests that the IIA procollagen isoform is required for post-natal vitreous/nerve cell development or maintenance, but is not necessary or desirable in post-natal cartilage tissue.

In the first series of publications reporting the differential expression patterns of IIA and IIB mRNA isoforms in developing cartilaginous tissues, theories were proposed in an attempt to understand why exon 2-containing isoforms are preferentially expressed in progenitor chondrocytes but not by differentiated chondrocytes. It was suggested that the exon 2-encoded protein domain may be important for i) interactions with other matrix or cell surface components; ii) regulation of morphogenetic processes; iii) controlling type II collagen biosynthesis; or iv) regulating fibrillogenesis (23, 25, 33). As will be discussed, research on exon 2 knock-in transgenic mice has shed light on some of these hypotheses.

Other *COL2A1* isoforms: IIC and IID

More recently, two other *COL2A1* spliced isoforms were identified in rabbit mesenchymal stem/stromal cells (MSCs) during chondrogenic differentiation in aggregate culture (34). Sequencing of cloned cDNA products following amplification of exons 1-7 (encoding most of the *COL2A1* NH₂-propeptide) revealed cDNAs corresponding to IIA and IIB isoforms as expected. However, a cDNA containing only the first 34 nucleotides of exon 2 (named IIC) and another cDNA with the same sequence as IIA, but with an additional three nucleotides

at the 3' end of exon 2 (named IID) were also identified (Figure 1). These isoforms were subsequently also found to be expressed in human differentiating MSCs in this study (34).

Generation of the IIC isoform occurs through utilization of a 5' alternative splice site within exon 2, resulting in a truncated transcript containing the first 34 nucleotides of exon 2 and premature termination codons (PTCs) in exon 6 (Figure 1). Further analysis of IIC mRNA by Southern blotting showed expression only at very early time points during differentiation of rabbit or human MSCs toward the chondrocyte lineage (34). This suggests that IIC mRNA may only be detectable at a time when the abundance of the other *COL2A1* isoforms in chondroprogenitor cells is low. The presence of PTCs renders IIC mRNA a substrate for nonsense-mediated decay (NMD) (35). The fact that no protein was detected by Western blotting using antiserum generated against a putative IIC peptide epitope (34) is further evidence to support NMD. Regardless of the likelihood that no protein product is generated following IIC splicing, this event appears to control the expression levels of IIA/IID and IIB procollagen as will be discussed further in this review.

Generation of the IID isoform occurs via another 5' splice site located just three nucleotides downstream from the IIA splice site in intron 2 (Figures 1 and 2). Utilization of this splice site does not alter the protein reading frame of the mRNA transcript. If IID mRNA is indeed translated, then a procollagen will be generated that is identical to IIA procollagen except for the addition of a tryptophan (in human) or arginine (in mouse or rabbit) at the end of the exon 2-encoded protein domain. This amino acid addition may generate a conformational change and potentially alter the function of the exon 2-encoded domain within the NH₂-propeptide. However, utilizing an alternative transcript quantitative PCR (AT-qPCR) technique that can accurately distinguish between IIA, IIB, IIC and IID isoforms (36), IID isoforms were found to be expressed at very low levels in primary chondrocytes from embryonic cartilage (E12.5, E16.5); IIA mRNAs were by far the predominant exon 2-containing isoforms generated by these cells. Also, during chondrogenic differentiation of human primary bone marrow-derived MSCs, IID mRNA was found to be approximately one third as abundant as IIA using an earlier version of the AT-qPCR assay (34). These findings may be somewhat surprising given that the theoretical "strength score" of the IID 5' splice site is higher than the IIA 5' splice site (Figure 2). It may be that the necessary *trans*-acting splicing factor proteins required to strengthen the IID 5' splice site are not expressed in sufficient levels. Alternatively, the IID 5' splice site may be less accessible than the IIA site due to the presence of RNA secondary structure in the form of a stem-loop in intron 2 (37) (Figure 3). Specifically, six nucleotides of the IID 5' splice site are located in the double stranded intronic stem region whereas only three nucleotides of the IIA 5' splice site are embedded within this site. This stem-loop *cis* element will be elaborated on further in this review.

The exon 2-encoded domain of *COL2A1*

COL2A1 exon 2 encodes a 69 amino acid von Willebrand factor type C (VWC) domain due to the presence of a consensus sequence of 10 cysteines (33, 38). This globular cysteine-rich (CR) domain is present in the NH₂-propeptide of type IIA/D procollagen but is absent from type IIB procollagen (Figure 1). Similar CR domains are also found in the NH₂-propeptide

of fibrillar procollagens types I and III, but there is higher conservation of the type II procollagen CR domain between species when compared to these other fibrillar collagens (38). Homologous CR domains are also present in other extracellular proteins including chordin, connective tissue growth factor and thrombospondins, all of which are all known to interact with TGF- β superfamily growth factors (39–42). In attempts to decipher the function of the *COL2A1* CR domain, a recombinant exon 2-encoded GST fusion protein was shown to interact with BMP-2 and TGF- β *in vitro* whereas a recombinant fusion protein encoding exons 3-8 of the *COL2A1* NH₂-propeptide did not (43). Subsequent studies in *Xenopus* embryos showed that over-expression of full length IIA procollagen mRNA, but not IIB procollagen mRNA or mRNA encoding the IIA CR domain alone, resulted in secondary axis induction (27). This dorsalization phenotype was similar to that induced by injection of chordin mRNA, thereby suggesting that full-length IIA procollagen may function to bind and inactivate BMP-2/4 *in vivo* (27). In addition, findings obtained from analysis of the solution structure of the *COL2A1* IIA domain *in vitro* showed similarities to the fibronectin type I domain, further suggesting the possibility for ligand binding function (44).

Generation of a polyclonal antibody that specifically recognizes the exon 2-encoded CR protein domain of type II procollagen (26) has been a useful reagent to show its expression and localization patterns in embryonic mouse and human cartilage as well as other tissues (31, 32, 43, 45). This antibody revealed the presence of IIA NH₂-propeptides on the surface of type II collagen fibrils extracted from human embryonic cartilage (43). It was concluded from this study that unprocessed type IIA procollagen retaining the NH₂-propeptide (pN-IIA-procollagen) was present in embryonic cartilage. This suggests that perhaps chondroprogenitor cells lack the necessary N-proteinase enzymes required to cleave the IIA NH₂-propeptide and that unprocessed pN-IIA-procollagen may be required early during chondrogenesis to interact with and regulate BMP/TGF- β growth factors. However, there is a possibility that the IIA NH₂-propeptides localized on fetal collagen fibrils may be part of the α 3 chain of type XI collagen heterotrimers since these are also encoded by *COL2A1* (17–19) (Figure 1). It is known that heterotypic fibrils exist in cartilage ECM consisting of type II, type IX and type XI collagens (16) and that the NH₂-propeptide of α 3(XI) chains are normally retained (46, 47). Also, recent work from our laboratory has shown that IIA isoforms of type XI collagen do exist and that the anti-IIA antibody used in the study by Sandell and colleagues (43) can recognize type XI(IIA) collagen isoforms (48). Therefore, if the CR-containing NH₂-propeptide plays a role in growth factor regulation during early embryogenesis, it is not clear whether this function may reside in the CR domain of pN-IIA-procollagen or the CR domain of the IIA α 3(XI) collagen chain. However, as will be discussed further in this review, the lack of any obvious skeletal phenotype in transgenic mice engineered to either lack exon 2 or continuously retain exon 2 throughout development does not support a major role for the IIA CR domain in regulating growth factor activity during chondrogenesis.

Molecular mechanisms regulating *COL2A1* alternative splicing: intronic *cis* elements

To attempt to understand the cellular mechanisms that control exon 2 splicing, mini-gene model systems have been utilized. In insulin-induced ATDC-5 cells, alternative splicing of a murine *Col2a1* mini-gene containing exons 1-4 and full-length intervening introns (49) was found to be similar to that of a truncated *Col2a1* mini-gene with large deletions in introns 1 and 2 (50). It was therefore concluded that important *cis* elements required to regulate exon 2 alternative splicing, at least in differentiating ATDC-5 cells, are present near the 3' and 5' splice site regions of exon 2. Subsequent studies by McAlinden *et al* using a human *COL2A1* mini-gene containing exons 1-3 and full-length intervening introns 1 and 2, showed that exon 2 is alternatively-spliced due to the presence of a weak 5' splice site (37) (Figure 2). Since most alternatively-spliced exons contain weak splice site sequences that have diverged from the known consensus sequence, other signals are normally required to ensure correct assembly of the spliceosome and efficient splicing to occur. As mentioned earlier in this review, these auxiliary *cis*-regulatory elements may be present in an intron or within a regulated exon and function as splicing enhancers or silencers (5). They may be short, single stranded elements or generate RNA secondary structure that may affect the accessibility of a regulated exon (5, 7, 51, 52). Equally important are the accessory *trans*-acting splicing factor proteins that bind to these *cis* elements (5).

With respect to *COL2A1* alternative splicing, an intronic auxiliary *cis* sequence containing an AU-rich element (ARE) located directly adjacent to the weak 5' splice site was found to be important in regulating exon 2 splicing (37). The potential for this AU-rich intronic sequence to form a stem-loop secondary structure was first hypothesized by Ala-Kokko *et al* (11) and later confirmed experimentally by McAlinden and colleagues (37) (Figure 3). Splicing of a *COL2A1* mini-gene devoid of this stem-loop sequence, but still containing the weak 5' splice site, significantly decreased formation of IIA exon 2-containing isoforms (37). This highlights the necessity of the AU-rich stem-loop *cis* element in permitting splicing and inclusion of exon 2. Subsequent studies identified that the *trans*-acting nuclear protein TIA-1 could bind to part of the stem-loop containing the ARE sequence AUUUUAUUUAUUU and function as a splicing enhancer by promoting exon 2 inclusion (53) (Figure 3A). Similar ARE sequences are present in 3' untranslated regions (UTRs) of genes and are targeted by proteins (including TIA-1) to affect stability, decay rate or translation initiation of the transcript. It is interesting that this stem-loop AU-rich *cis* element is found in an intron adjacent to an alternatively spliced exon. Also, we have shown that the sequences of the weak 5' splice site of exon 2 and the adjacent AU-rich element are highly conserved from humans to teleosts (37). Interestingly, in teleosts (e.g. zebrafish), two *col2a1* orthologues exist (*col2a1a* and *col2a1b*) (54, 55) whereas terrestrial vertebrates contain only one gene encoding *Col2a1* (56). Personal communication with Dr. Rodney Dale (Department of Biology, Loyola University, Chicago) has revealed the existence of alternatively-spliced mRNA isoforms of zebrafish *col2a1a* containing or devoid of exon 2 (unpublished observations). These findings raise important questions as to why this alternative splicing event has been conserved throughout evolution. It will be interesting, as more sequence information becomes available, to determine if conservation of the exon 2

splice site sequence and the adjacent ARE sequence exists in primitive jawless vertebrates such as lampreys or hagfishes that are known to possess non-ossified cartilaginous skeletal elements (57, 58). Determining whether or not exon 2 alternative splicing occurs in these species would provide further insight into understanding the biological significance of this event.

Regulatory *cis* elements within *COL2A1* exon 2

In attempts to discover other splicing regulatory *cis* elements that may exist within exon 2 itself, a web based resource was used by our group to identify potential exonic splicing enhancer binding sites (ESEfinder; <http://rulai.cshl.edu/tools/ESE/>) (59, 60). A high-scoring *cis* element was identified (CAGAGGT) as a potential binding site for Ser/Arg-rich splicing factor 1 (SRSF1; also known as ASF/SF2). This heptameric element was subsequently deleted from a *COL2A1* mini-gene construct and the result was enhanced production of exon 2-containing transcripts (34). Closer analysis revealed that the region we had deleted overlapped with an area corresponding to another potential alternative 5' splice site sequence. As introduced previously in this review, this region corresponds to the IIC 5' splice site, utilization of which results in production of the IIC truncated mRNA transcript that was identified as being expressed during rabbit and human MSC chondrogenic differentiation (Figure 1) (34). Whether or not SRSF1 splicing factor protein interacts with its predicted binding site to enhance/suppress spliceosome assembly at the IIC splice site is not yet known. Subsequent studies were carried out by our group to introduce three point mutations into the *COL2A1* mini-gene that would disrupt the IIC splice site without affecting the protein reading frame. Splicing of this mini-gene by various cell lines also resulted in an increase in the ratio of IIA/D : IIB transcripts, similar to the mini-gene deletion mutation (unpublished observations). We have also generated transgenic mice containing the same three silent point mutations; findings from characterization of these mice will be published soon in *Matrix Biology*. While no overt skeletal phenotype was observed in heterozygous or homozygous mutant IIC knock-in mice (*Col2a1-mIIC*), we did detect an increase in the ratio of IIA/D : IIB mRNA transcripts generated by chondrocytes *in vivo*.

We have proposed that the IIC splice site may function to enhance production of IIB procollagen isoforms in differentiated chondrocytes and prevent spliceosome interaction with the IIA/D 5' splice sites. In differentiated chondrocytes, the spliceosome should theoretically favor the IIC splice site over the IIA/D 5' splice sites since the former conforms better to the 5' splice site consensus sequence than the latter two (Figure 2). Therefore, by mutating the IIC splice site *in vivo*, spliceosome complexes are available to bind to the IIA/D 5' splice sites, resulting in the increased levels of exon 2-containing isoforms observed in the *Col2a1-mIIC* knock-in mice (Figure 3B). It is interesting that this IIC splicing event results in a truncated mRNA transcript containing three premature termination codons (PTCs) in exon 6. We hypothesize that IIC mRNA is a substrate for nonsense-mediated decay (NMD). This would explain why it is often difficult to detect this transcript by AT-qPCR (36) and why no protein product was detected using an antibody raised against the putative IIC peptide epitope (34). Therefore, IIC splicing may be an example of "Regulated Unproductive Splicing and Translation" (RUST), a reported cellular mechanism

linking alternative splicing and NMD to regulate the abundance of mRNA transcripts (61). From our studies on the *Col2a1-mIIC* mice, it is apparent that the function of the IIC splicing event is to control levels of IIA and IIB procollagen isoforms that we know are translated to protein.

Potential association of **COL2A1** transcription rate and alternatively-spliced isoform production

While it has been shown that the stem-loop AU-rich element in intron 2 is critical for inclusion of exon 2 in chondroprogenitor cells (McAlinden, 2005), the mechanism by which differentiating chondrocytes excludes exon 2 is not completely known. It may be that maturing chondrocytes do not express appropriate levels of the of the necessary *trans*-acting splicing accessory enhancer proteins (e.g. TIA-1 and others yet to be identified) required to “strengthen” interactions of the spliceosome with the weak 5′ splice site of exon 2 and the AU-rich stem loop. It will be interesting to investigate by microarray or RNA-Seq analysis whether or not chondroprogenitor cells generate a different subset of splicing factor accessory proteins compared to differentiated chondrocytes.

Another theory to explain why differentiating chondrocytes exclude exon 2 relates to the rate of *COL2A1* transcription. It is well known that transcription and splicing processes are co-regulated in the cell (62). Studies have also demonstrated that rates of RNA polymerase II transcription elongation can influence the splicing of alternative exons. Specifically, approaches to modulate RNA pol II transcription rates in cells has shown that a slow-acting RNA pol II induces inclusion of the fibronectin EDI alternatively-spliced exon whereas increased elongation stimulates skipping of the EDI exon (63–66). Therefore, in the case of *COL2A1* exon 2 splicing, lower rates of transcription in progenitor cells may provide sufficient time for splicing of the relatively long intron 1 (~4.1kb) and assembly of the spliceosome complex and other accessory splicing factor proteins at the weak 5′ splice site and AU-rich stem loop site. However, in differentiating chondrocytes when the rate of *COL2A1* transcription is dramatically increased to support rapid synthesis of the cartilage ECM, exon 2 skipping occurs since spliceosome/accessory factor complexes may not be able to assemble in enough time at the weak 5′ splice site of exon 2, but will instead favor the stronger constitutive 5′ splice site of exon 3. Studies on growth factor treatment of de-differentiated murine chondrocytes showed that TGF-β1 enhanced *Col2a1* exon 2 inclusion (IIA isoform production) whereas BMP-2 induced exon 2 skipping (IIB isoform production) (67). This study also reported that overall levels of type II procollagen expression were decreased by TGF-β1 but increased with BMP-2 in this culture system. These findings add support to the hypothesis that low *Col2a1* transcription rates favor exon 2 inclusion while high rates may favor exon 2 skipping. It will be interesting to alter RNA pol II elongation in chondrocytes to test this hypothesis in more detail.

COL2A1 isoforms as reliable markers of chondrocyte differentiation

Type IIA/D or IIB procollagen transcripts are reliable cellular markers to identify chondroprogenitor/de-differentiated chondrocytes or differentiated chondrocytes, respectively. These markers should always be analyzed during *in vitro* chondrogenesis

assays to determine the efficiency of chondrocyte differentiation. For example, a recent study by our group revealed that insulin-induced chondrogenic differentiation of the murine ATDC-5 cell line resulted in higher levels of IID exon 2-containing isoforms than IIB isoforms following 28 days in culture (36). This surprisingly demonstrated that the majority of cells in these cultures were less differentiated than expected regardless of the fact that the cultures stained positive with alcian blue, indicating proteoglycan synthesis. Even more surprising was the observation that IID isoforms were more abundant than IIA isoforms in these ATDC-5 cultures; a finding that would not have been recognized without development of the AT-qPCR assay that can quantitatively distinguish between these two almost identical isoforms (36). As already described in this review, IIA mRNA is the predominant isoform generated by early embryonic murine chondrocytes *in vivo* or by chondroprogenitor cells derived from primary MSCs *in vitro* (34, 36). This raises some interesting questions related to potential differences in splicing mechanisms regulating exon 2 inclusion between primary cells and transformed cell lines.

In the cartilage tissue engineering field, a wide range of biomaterials are being utilized either as a scaffold system or as growth factor delivery vehicles to induce chondrocyte differentiation for cartilage repair or regeneration (68–71). Commonly, type II procollagen (the processed triple helical domain) and aggrecan are analyzed in these tissue engineering approaches to confirm chondrocyte differentiation and cartilage formation. Expression of specific isoforms of type II procollagen is rarely determined in these systems. Our group recently investigated gene and protein expression levels of *COL2A1* isoforms between TGF- β -induced pellet cultures of human MSCs in the presence or absence of poly(ethylene glycol) (PEG) microspheres (72). Although IIB mRNA isoforms were predominant in both culture systems, the ratio of IIA/D : IIB mRNA isoforms was higher in MSC pellets containing PEG microspheres. Also, more widespread localization of the IIA exon 2-encoded protein domain was identified in the engineered cartilage tissue containing the microspheres. Whether or not delayed or abnormal differentiation had occurred in these cultures containing biomaterials is not yet known. Regardless, it is important that the biological consequences of biomaterial interactions with MSCs or chondrocytes are considered and that analysis of appropriate cellular markers to properly define differentiation status should always be carried out.

Mouse models to study *Col2a1* alternative splicing

Studies on *Col2a1* null mice have demonstrated the absolute requirement for a type II collagen matrix in the proper formation of articular cartilage, epiphyseal growth plate cartilage, endochondral bone and intervertebral discs (73, 74). In attempts to address the significance of the developmentally-regulated *Col2a1* alternative splicing switch *in vivo*, two different transgenic mouse models have been generated, one devoid of exon 2 and the other engineered to exclusively express exon 2 (28, 29). Cheah and colleagues generated *Col2a1* exon 2 null mice and reported a partially penetrant phenotype in homozygous knock-out embryos involving abnormal development of head tissues, holoprosencephaly and loss of mid facial structures (28). This phenotype was associated with reduced sonic hedgehog (*Shh*) expression in prechordal mesoderm. Genetic interaction studies revealed that potential IIA function in forebrain and facial development did not involve BMPRI1A or

NODAL-mediated signaling activity. Therefore, it does not appear that the IIA CR domain plays a role in regulating BMP activity as was postulated previously (27, 43). It is also somewhat surprising that no skeletal or cartilage-specific phenotypes have been reported in these mice (Dr. Kathryn Cheah, personal communication). This would suggest that the IIA exon 2-encoded CR domain does not play a significant role in regulating growth factor-driven mechanisms during early stages of cartilage tissue development.

Transgenic mice exclusively expressing the embryonic IIA procollagen isoform of *Col2a1*

Recently, our group utilized a splice site targeting approach to generate transgenic mice containing a knock-in mutation that converted the weak 5' splice site of *Col2a1* exon 2 to a strong, consensus splice site sequence. This mutation permitted efficient splicing of exon 2 and expression of only IIA procollagen isoforms regardless of the status of chondrocyte differentiation (29). In other words, the developmentally-regulated alternative splicing of exon 2 and subsequent production of type IIB procollagen was inhibited in these mice. Heterozygous knock-in (ki/+) and homozygous knock-in (ki/ki) transgenic mice (*Col2a1^{+ex2}*) were viable and surprisingly displayed no obvious skeletal phenotype. Articular cartilage, growth plate cartilage and endochondral bone appeared to develop normally. We had originally hypothesized that persistent mis-expression of embryonic IIA procollagen isoforms would result in some form of skeletal abnormality if indeed the IIA CR domain played an important role in regulating the activity of TGF- β superfamily growth factors. For example, if the CR domain could inhibit BMP-2 activity during chondrogenesis, then a chondrodysplasia phenotype would be expected, similar to that observed in chondrocyte specific *Bmp-2* knock-out mice (75).

There have been interesting reports that a recombinant IIB NH₂-propeptide protein (encoded by human exons 3-8) has anti-angiogenic activity (76) and can also induce cell death of tumor cells and osteoclasts via an integrin-mediated mechanism (77, 78). Normal, healthy articular cartilage is avascular and resistant to tumor invasion. Collectively, these results suggest that IIB NH₂-propeptide may function to maintain the avascular nature of articular cartilage and protect it from tumor cell invasion. However, we did not observe abnormal vascular invasion during pre- and post-natal development of articular cartilage in *Col2a1^{+ex2}* mice devoid of IIB procollagen. One explanation for this may be that the IIA NH₂-propeptide (encoded by exons 2-8) could also function to protect articular cartilage from vessel invasion. Although preliminary studies by Sandell and colleagues showed that a IIA recombinant protein encoded only by exon 2 displayed no anti-angiogenic function, a recombinant IIA protein encoding the entire NH₂ propeptide was not analyzed in this study (Dr. Linda Sandell, personal communication).

Another explanation for the lack of abnormal vascular phenotype in articular cartilage of *Col2a1^{+ex2}* mice is that the IIA NH₂-propeptide may be processed further by matrix metalloproteinases (MMPs) to generate smaller fragments resembling IIB NH₂-propeptides. In fact, Fukui *et al* showed that a recombinant trimeric IIA NH₂-propeptide (encoded by exon 2-8) could be digested *in vitro* by a range of different MMPs (79). Specifically, it was found that cleavage of the recombinant IIA NH₂-propeptide by MMP-7 could potentially

generate IIB-like NH₂-propeptides. Alternatively it may be that, *in vivo*, the IIB NH₂-propeptide is rapidly processed and removed from the ECM and is therefore not present at sufficient levels to function as an anti-vascular agent *in vivo*. This theory is supported from the recent report characterizing a new antibody that specifically recognizes the murine IIB procollagen isoform, but not IIA procollagen (80). While this antibody was able to detect unprocessed type IIB procollagen by Western blotting, no positive signals were detected in the ECM of cartilage tissue sections by immunolocalization suggesting the processed IIB NH₂-propeptide is rapidly removed from the procollagen molecule and cleared from the ECM. With respect to the anti-tumorigenic effects of IIB NH₂-propeptide (78), it was confirmed in these studies that a recombinant IIA NH₂-propeptide protein encoded by exons 2-8 could not induce tumor cell death (Dr. Linda Sandell, personal communication). Therefore, it will be interesting to challenge our *Col2a1^{+ex2}* mice to determine if they could be more susceptible to chondro/osteosarcomas.

***Col2a1* splicing and type XI collagen**

We showed that IIA procollagen derived from post-natal *Col2a1^{+ex2}* mice could be processed similar to IIB procollagen (29). Subsequently, “free” IIA NH₂-propeptides would be removed from the ECM and explains the decrease in IIA CR protein expression patterns (green fluorescent staining) in the territorial and inter-territorial matrix of post-natal articular cartilage tissue in these mice (29) (Figure 4). However, in 10wk old ki/+ and ki/ki *Col2a1^{+ex2}* mice, persistent expression of IIA protein was still found in the superficial zone and in the pericellular matrix surrounding deep zone chondrocytes in articular cartilage as well as in the pericellular matrix of growth plate chondrocytes (Figure 4) (29). Such pericellular localization is reminiscent of type XI collagen staining patterns in epiphyseal cartilage (81, 82). It is known that the NH₂-propeptide encoded by the α3(XI) chain is retained in the type XI collagen heterotrimer (46, 47). Also, we have recently confirmed the presence of IIA isoforms of type XI collagen containing the exon 2-encoded CR domain in the α3(XI) chain in rib cartilage from *Col2a1^{+ex2}* mice (48). Therefore, it is possible that the pericellular expression of IIA protein that persists in mature cartilage tissue in *Col2a1^{+ex2}* mice is due to the presence of IIA isoforms of type XI collagen.

Type XI collagen, through retention of its NH₂-propeptide (pN-procollagen), is an important regulator of collagen fibrillogenesis and fibril diameter (20). Therefore, it is tempting to speculate that the presence of the globular IIA CR domain in the α3(XI) chain will alter the structure of the NH₂-propeptide of type XI collagen thereby affecting its role in controlling fibrillogenesis. Some evidence to support this hypothesis comes from electron microscopy analysis of cartilage ECM in *Col2a1^{+ex2}* mice. We found differences in the ultrastructure of the collagen fibrillar matrix in ki/+ and ki/ki articular cartilage when compared to WT tissue (McAlinden Laboratory, unpublished observations). Specifically, thicker collagen fibrils were found in the pericellular domain of ki/+ and ki/ki tissue in addition to electron dense areas suggesting fibril fusion or aggregation. Similar abnormalities were noted in the territorial and inter-territorial zones as well. Recently, we also reported altered collagen fibril structure in ki/+ and ki/ki rib cartilage matrix in addition to lower levels of hydroxyproline cross-links when compared to WT tissue (48). Taken together, it may

be that alternative splicing of *Col2a1* is necessary to control type XI collagen $\alpha 3$ chain isoform expression to permit proper matrix assembly.

Mutations within alternatively spliced *COL2A1* exon 2

Human mutations in *COL2A1* exon 2 have been identified and shown to result in a form of Stickler syndrome affecting the eye only (83). Typically, the ocular defects include vitreous degeneration, retinal breaks and detachments. However, Stickler syndrome can also result from mutations in other exons of *COL2A1* and, in these cases, patients suffer not only from eye problems but also from range of other abnormalities including joint hypermobility and pain, hearing loss, spinal defects, facial abnormalities and premature osteoarthritis (83–85).

The reason why exon 2 mutations result primarily in ocular disease is because cells of the eye express both IIA and IIB isoforms pre- and post-natally (86). In addition, the adult bovine vitreous has been shown to contain approximately equal levels of IIA and IIB procollagen protein (30). These expression patterns in the eye are different to the developmentally-regulated loss of IIA expression that is known to occur as cartilage tissue matures. Therefore, no other systemic defects are noted in these patients since differentiating and mature cartilaginous tissues primarily express type IIB procollagen, and are therefore not affected by mutations in exon 2.

The majority of *COL2A1* exon 2 mutations reported result in premature stop codons that are expected to cause type IIA procollagen haploinsufficiency (83, 87). Therefore, the eye defects may occur simply due to an overall deficiency in type IIA procollagen levels. Persistence of IIA procollagen throughout post-natal development of the vitreous may be important to regulate production of thin collagen fibrils that are required for proper health of this tissue (31). Thin collagen fibrils are also found in embryonic cartilage extracellular matrices at a time when *COL2A1* IIA isoforms are still expressed (43). Since *COL2A1* also encodes the $\alpha 3$ chain of type XI collagen, it is also possible that the exon 2-encoded IIA domain detected on collagen fibrils in the vitreous (31) is actually part of type XI heterotrimers; it has been shown in other studies that type XI collagen is restricted to thin collagen fibrils (82). Also, we know that type XI collagen retains its NH₂-propeptide (pN-procollagen) (46, 47) and previous work has suggested that NH₂-propeptides of fibrillar collagens may function to regulate fibril diameter and prevent fibril fusion (88). Whether the IIA exon 2-encoded domain is present in type II or type XI pN-procollagen, it is likely that its persistent expression in the eye may be required to maintain structural integrity of the collagen ECM. It will be interesting to examine the eyes of the *Col2a1*^{+ex2} mice that express only IIA isoforms of *Col2a1*. It may be that a balance between IIA and IIB procollagen isoforms is necessary for vitreous health and that sole expression of IIA procollagen will have detrimental effects.

Perspectives and some remaining questions

IIB or not IIB: is IIB procollagen necessary for development of healthy cartilage?

Studies on the *Col2a1*^{+ex2} mice have provided some answers to this question. Inhibition of *Col2a1* alternative splicing is apparently well-tolerated in these mice during cartilage

development and IIA procollagen appears to compensate quite well for the absence of IIB procollagen. Therefore, regulating the expression of type II procollagen isoforms during chondrogenesis via alternative splicing may not be as important as once thought given that type II procollagen synthesis and processing are apparently normal when the IIA procollagen isoform is expressed by differentiated chondrocytes. Also, evidence does not suggest a role for the IIA isoform in controlling growth factor activity. However, as already discussed, it may be that *Col2a1* alternative splicing is necessary to control isoform expression and, subsequently, the regulatory function of type XI collagen. Whether or not the changes in collagen ultrastructure detected in articular cartilage of *ki/+* or *ki/ki Col2a1^{+ex2}* mice translates to a biomechanically-inferior tissue remains to be determined. Toward this end, challenging the cartilage through destabilization of the knee joint (via surgical-induced or non-invasive joint loading approaches) will be informative. Increases in trabecular bone mass have also been found in *Col2a1^{+ex2}* mice (McAlinden laboratory, unpublished observations). This suggests that a proper cartilage template is necessary for endochondral bone development. These findings raise some interesting questions on how an abnormal type II collagen fibrillar matrix may directly affect osteoblast and/or osteoclast activity, for example.

Why is COL2A1 exon 2 programmed to undergo alternative splicing while homologous exon 2 domains in COL1A1 or COL3A1 are not?

Exon 2-encoded cysteine-rich domains are present in the $\alpha 1$ chains of both type I procollagen and type III procollagen and share homology with that found in $\alpha 1/\alpha 3$ chains of type II/XI procollagens (38). Interestingly, *COL1A1* or *COL3A1* exon 2 is not alternatively-spliced and is always included in the final processed mRNA. This is because the 5' splice site sequence of exon 2 in these other fibrillar collagen genes conforms better to the known consensus sequence (3) and is therefore constitutively spliced. The 5' splice site sequence of human *COL1A1* exon 2 is CAGgtgcgg (NCBI Ref Seq: NG_007400.1) and that of *COL3A1* exon 2 is GCTgtgagt (NCBI Ref Seq: NG_007404.1). Uppercase letters represent nucleotides from exon 2 and lowercase letters are from intron 2. Calculation of the 5' splice site strength of exon 2 by MaxEntScan (89) revealed a score of 9.2 for *COL1A1* and 7.4 for *COL3A1*. Compare this with the extremely low splice site strength score (−10.4) of the 5' splice site of *COL2A1* exon 2 (Figure 2).

Given that exon 2 is constitutively spliced and always included in fully processed mRNA encoding *COL1A1* and *COL3A1* suggests that it is required throughout pre- and post-natal development of various connective tissues where these collagens are expressed. However, Bornstein *et al* generated a transgenic mouse lacking the exon 2 domain of *COL1A1* and discovered no overt phenotype (Bornstein, 2002). Closer examination of the dermis by electron microscopy revealed a slight increase in fibril diameter and some irregularities in fibril contours (90), suggesting that the *COL1A1* exon 2 domain plays a role in regulating fibrillogenesis and may could explain why it is a constitutively spliced exon. Why then is exon 2 programmed to be removed from *COL2A1* pre-mRNA either in the eye or in a developmentally-regulated manner in cartilage? One could hypothesize that *COL2A1* exon 2 encodes a redundant protein due to the lack an obvious phenotype in the *Col2a1^{+ex2}* mice that were engineered to persistently express exon 2 (29). Therefore, could removal of exon 2

(which encodes approximately 20% of the entire type II procollagen protein) have arisen by chance and subsequently proven to be advantageous to the cell during periods when vast amounts of type II procollagen secretion is required for skeletal growth? The argument against this comes from our observations of altered collagen ultrastructure in the ribs of *Col2a1^{+ex2}* mice (48) as well as in the articular cartilage (unpublished observations). We therefore propose that *COL2A1* exon 2 removal does happen for a reason which is to regulate specific aspects of type II/XI procollagen fibrillogenesis and subsequent matrix assembly.

Why is the IIA isoform of Col2a1 re-expressed in normal murine articular cartilage?

From our studies on the *Col2a1^{+ex2}* mice, we have also noticed distinct re-expression of IIA procollagen in normal, mature wild type articular cartilage tissue (Figure 4A, D). Chondrocytes in human and mouse osteoarthritic cartilage have been found to re-express type IIA procollagen and it was suggested that the cells switch to generate embryonic exon 2-containing isoforms in an attempt to repair the damaged cartilage tissue (91, 92). However, our findings of IIA procollagen expression by differentiated chondrocytes in normal mature murine articular cartilage suggest that IIA synthesis may not be associated with disease. Instead, it is possible that during normal matrix turnover in post-natal cartilage, rates of type II procollagen synthesis are low enough to permit inclusion of exon 2 and synthesis of type IIA procollagen. It is also tempting to speculate that expression of IIA procollagen in non-diseased, mature articular cartilage may be a mechano-regulated response. A number of alternatively-spliced genes have in fact been shown to be regulated by mechanical stress, although the mechanisms remain unclear (93).

Acknowledgments

Thank you to Dr. Thomas Hering (T.H.) for assistance in calculating splice site strength scores (Figure 2). Thanks also to Soumya Ravindran and Louisa Wirthlin for generating the immunofluorescence data shown in Figure 4. Research involving the *Col2a1^{+ex2}* mice received funding from an NIH R21 grant (AR053513; to A.M.) and a Pilot & Feasibility grant provided from an NIH Musculoskeletal P30 Core Grant (P30 AR057235; to A.M.). Research on *Col2a1-mIIc* mice was funded by an NIH R21 grant (RR025397; to A.M and T.H).

References

1. De Conti L, Baralle M, Buratti E. Exon and intron definition in pre-mRNA splicing. Wiley Interdiscip Rev RNA. 2013; 4(1):49–60. [PubMed: 23044818]
2. Jurica MS, Moore MJ. Pre-mRNA splicing: awash in a sea of proteins. Mol Cell. 2003; 12(1):5–14. [PubMed: 12887888]
3. Burge, CB.; Tuschl, T.; Sharp, PA. Splicing of Precursors to mRNAs by the Spliceosomes. In: Gesteland, R.; Cech, T.; Atkins, J., editors. The RNA World. 2. Cold Spring Harbor: Cold Spring Harbor Press; 1999. p. 525-60.
4. Black DL. Mechanisms of alternative pre-messenger RNA splicing. Annu Rev Biochem. 2003; 72:291–336. [PubMed: 12626338]
5. Kornblihtt AR, Schor IE, Allo M, Dujardin G, Petrillo E, Munoz MJ. Alternative splicing: a pivotal step between eukaryotic transcription and translation. Nat Rev Mol Cell Biol. 2013; 14(3):153–65. [PubMed: 23385723]
6. Kalsotra A, Cooper TA. Functional consequences of developmentally regulated alternative splicing. Nat Rev Genet. 2011; 12(10):715–29. [PubMed: 21921927]

7. McAlinden A, Havlioglu N, Sandell LJ. Regulation of protein diversity by alternative pre-mRNA splicing with specific focus on chondrogenesis. *Birth Defects Res C Embryo Today*. 2004; 72(1): 51–68. [PubMed: 15054904]
8. Lin S, Fu XD. SR proteins and related factors in alternative splicing. *Adv Exp Med Biol*. 2007; 623:107–22. [PubMed: 18380343]
9. Martinez-Contreras R, Cloutier P, Shkreta L, Fisette JF, Revil T, Chabot B. hnRNP proteins and splicing control. *Adv Exp Med Biol*. 2007; 623:123–47. [PubMed: 18380344]
10. Cheah KS, Stoker NG, Griffin JR, Grosveld FG, Solomon E. Identification and characterization of the human type II collagen gene (COL2A1). *Proc Natl Acad Sci USA*. 1985; 82(9):2555–9. [PubMed: 3857598]
11. Ala-Kokko L, Kvist AP, Metsaranta M, Kivirikko KI, de Crombrughe B, Prockop DJ, et al. Conservation of the sizes of 53 introns and over 100 intronic sequences for the binding of common transcription factors in the human and mouse genes for type II procollagen (COL2A1). *Biochem J*. 1995; 308(Pt 3):923–9. [PubMed: 8948452]
12. Metsaranta M, Toman D, de Crombrughe B, Vuorio E. Mouse type II collagen gene. Complete nucleotide sequence, exon structure, and alternative splicing. *J Biol Chem*. 1991; 266(25):16862–9. [PubMed: 1885613]
13. Prockop DJ, Kivirikko KI. Collagens: molecular biology, diseases, and potentials for therapy. *Annu Rev Biochem*. 1995; 64:403–34. [PubMed: 7574488]
14. Eyre D. Collagen of articular cartilage. *Arthritis Res*. 2002; 4(1):30–5. [PubMed: 11879535]
15. Canty EG, Kadler KE. Procollagen trafficking, processing and fibrillogenesis. *J Cell Sci*. 2005; 118(Pt 7):1341–53. [PubMed: 15788652]
16. Eyre DR, Weis MA, Wu JJ. Articular cartilage collagen: an irreplaceable framework? *Eur Cells Mat*. 2006; 12:57–63.
17. Burgeson RE, Hollister DW. Collagen heterogeneity in human cartilage: identification of several new collagen chains. *Biochem Biophys Res Commun*. 1979; 87(4):1124–31. [PubMed: 465027]
18. Furuto DK, Miller EJ. Different levels of glycosylation contribute to the heterogeneity of alpha 1(II) collagen chains derived from a transplantable rat chondrosarcoma. *Arch Biochem Biophys*. 1983; 226(2):604–11. [PubMed: 6639071]
19. Eyre, DR. Type XI or 1a2a3a collagen. In: Mayne, R.; Burgeson, RE., editors. *Structure and function of collagen types*. New York: Academic Press; 1987. p. 261–81.
20. Blaschke UK, Eikenberry EF, Hulmes DJ, Galla HJ, Bruckner P. Collagen XI nucleates self-assembly and limits lateral growth of cartilage fibrils. *J Biol Chem*. 2000; 275(14):10370–8. [PubMed: 10744725]
21. Ryan MC, Sandell LJ. Differential expression of a cysteine-rich domain in the amino-terminal propeptide of type II (cartilage) procollagen by alternative splicing of mRNA. *J Biol Chem*. 1990; 265(18):10334–9. [PubMed: 2355003]
22. Sandell LJ, Morris N, Robbins JR, Goldring MB. Alternatively spliced type II procollagen mRNAs define distinct populations of cells during vertebral development: differential expression of the amino-propeptide. *J Cell Biol*. 1991; 114(6):1307–19. [PubMed: 1894696]
23. Ng LJ, Tam PP, Cheah KS. Preferential expression of alternatively spliced mRNAs encoding type II procollagen with a cysteine-rich amino-propeptide in differentiating cartilage and nonchondrogenic tissues during early mouse development. *Dev Biol*. 1993; 159(2):403–17. [PubMed: 8405667]
24. Sandell LJ, Nalin AM, Reife RA. Alternative splice form of type II procollagen mRNA (IIA) is predominant in skeletal precursors and non-cartilaginous tissues during early mouse development. *Dev Dyn*. 1994; 199(2):129–40. [PubMed: 8204907]
25. Lui VC, Ng LJ, Nicholls J, Tam PP, Cheah KS. Tissue-specific and differential expression of alternatively spliced alpha 1(II) collagen mRNAs in early human embryos. *Dev Dyn*. 1995; 203(2):198–211. [PubMed: 7655082]
26. Ogenesian A, Zhu Y, Sandell LJ. Type IIA procollagen amino propeptide is localized in human embryonic tissues. *J Histochem Cytochem*. 1997; 45(11):1469–80. [PubMed: 9358849]

27. Larrain J, Bachiller D, Lu B, Agius E, Piccolo S, De Robertis EM. BMP-binding modules in chordin: a model for signalling regulation in the extracellular space. *Development*. 2000; 127(4): 821–30. [PubMed: 10648240]
28. Leung AW, Wong SY, Chan D, Tam PP, Cheah KS. Loss of procollagen IIA from the anterior mesendoderm disrupts the development of mouse embryonic forebrain. *Dev Dyn*. 2010; 239(9): 2319–29. [PubMed: 20730911]
29. Lewis R, Ravindran S, Wirthlin L, Traeger G, Fernandes RJ, McAlinden A. Disruption of the developmentally-regulated Col2a1 pre-mRNA alternative splicing switch in a transgenic knock-in mouse model. *Matrix Biol*. 2012; 31(3):214–26. [PubMed: 22248926]
30. Bishop PN, Reardon AJ, McLeod D, Ayad S. Identification of alternatively spliced variants of type II procollagen in vitreous. *Biochem Biophys Res Commun*. 1994; 203(1):289–95. [PubMed: 8074668]
31. Reardon A, Sandell L, Jones CJ, McLeod D, Bishop PN. Localization of pN-type IIA procollagen on adult bovine vitreous collagen fibrils. *Matrix Biol*. 2000; 19(2):169–73. [PubMed: 10842100]
32. D'Antonio M, Michalovich D, Paterson M, Droggiti A, Woodhoo A, Mirsky R, et al. Gene profiling and bioinformatic analysis of Schwann cell embryonic development and myelination. *Glia*. 2006; 53(5):501–15. [PubMed: 16369933]
33. Ryan MC, Sieraski M, Sandell LJ. The human type II procollagen gene: identification of an additional protein-coding domain and location of potential regulatory sequences in the promoter and first intron. *Genomics*. 1990; 8(1):41–8. [PubMed: 2081599]
34. McAlinden A, Johnstone B, Kollar J, Kazmi N, Hering TM. Expression of two novel alternatively spliced COL2A1 isoforms during chondrocyte differentiation. *Matrix Biol*. 2008; 27(3):254–66. [PubMed: 18023161]
35. Brogna S, Wen J. Nonsense-mediated mRNA decay (NMD) mechanisms. *Nat Struct Mol Biol*. 2009; 16(2):107–13. [PubMed: 19190664]
36. McAlinden A, Shim KH, Wirthlin L, Ravindran S, Hering TM. Quantification of type II procollagen splice forms using alternative transcript-qPCR (AT-qPCR). *Matrix Biol*. 2012
37. McAlinden A, Havlioglu N, Liang L, Davies SR, Sandell LJ. Alternative splicing of type II procollagen exon 2 is regulated by the combination of a weak 5' splice site and an adjacent intronic stem-loop cis element. *J Biol Chem*. 2005; 280(38):32700–11. [PubMed: 16076844]
38. Bornstein P. The NH(2)-terminal propeptides of fibrillar collagens: highly conserved domains with poorly understood functions. *Matrix Biol*. 2002; 21(3):217–26. [PubMed: 12009328]
39. Murphy-Ullrich JE, Poczatek M. Activation of latent TGF-beta by thrombospondin-1: mechanisms and physiology. *Cytokine Growth Factor Rev*. 2000; 11(1–2):59–69. [PubMed: 10708953]
40. Zhang JL, Huang Y, Qiu LY, Nickel J, Sebald W. von Willebrand factor type C domain-containing proteins regulate bone morphogenetic protein signaling through different recognition mechanisms. *J Biol Chem*. 2007; 282(27):20002–14. [PubMed: 17483092]
41. Fujisawa T, Huang Y, Sebald W, Zhang JL. The binding of von Willebrand factor type C domains of Chordin family proteins to BMP-2 and Tsg is mediated by their SD1 subdomain. *Biochem Biophys Res Commun*. 2009; 385(2):215–9. [PubMed: 19450552]
42. Abreu JG, Ketpura NI, Reversade B, De Robertis EM. Connective-tissue growth factor (CTGF) modulates cell signalling by BMP and TGF-beta. *Nat Cell Biol*. 2002; 4(8):599–604. [PubMed: 12134160]
43. Zhu Y, Oganessian A, Keene DR, Sandell LJ. Type IIA procollagen containing the cysteine-rich amino propeptide is deposited in the extracellular matrix of prechondrogenic tissue and binds to TGF-beta1 and BMP-2. *J Cell Biol*. 1999; 144(5):1069–80. [PubMed: 10085302]
44. O'Leary JM, Hamilton JM, Deane CM, Valeyev NV, Sandell LJ, Downing AK. Solution structure and dynamics of a prototypical chordin-like cysteine-rich repeat (von Willebrand Factor type C module) from collagen IIA. *J Biol Chem*. 2004; 279(51):53857–66. [PubMed: 15466413]
45. Zhu Y, McAlinden A, Sandell LJ. Type IIA procollagen in development of the human intervertebral disc: regulated expression of the NH(2)-propeptide by enzymic processing reveals a unique developmental pathway. *Dev Dyn*. 2001; 220(4):350–62. [PubMed: 11307168]
46. Thom JR, Morris NP. Biosynthesis and proteolytic processing of type XI collagen in embryonic chick sterna. *J Biol Chem*. 1991; 266(11):7262–9. [PubMed: 2016327]

47. Wu JJ, Eyre DR. Structural analysis of cross-linking domains in cartilage type XI collagen. Insights on polymeric assembly. *J Biol Chem.* 1995; 270(32):18865–70. [PubMed: 7642541]
48. McAlinden A, Traeger G, Hansen U, Weis MA, Ravindran S, Wirthlin L, et al. Molecular properties and fibril ultrastructure of types II and XI collagens in cartilage of mice expressing exclusively the alpha1(IIA) collagen isoform. *Matrix Biol.* 2013 [Epub ahead of print].
49. Hatano H, Sarkar G, Bolander ME. Development of a cellular model to study alternative splicing of type II collagen gene. *J Orthop Res.* 2002; 20(3):516–9. [PubMed: 12038625]
50. Nishiyama T, Hatano H, Kurosaka M, Bolander M, Sarkar G. Cis-Acting Intronic Elements that Regulate Cartilage-Specific Alternative splicing of the Type II Collagen (Col2) Pre-mRNA Lie at or Near Splice Site Junction Sequences Flanking Exon 2 of the Gene. *J Bone Miner Res.* 2003; 18(9):1716–22. [PubMed: 12968682]
51. Libri D, Piseri A, Fiszman MY. Tissue-specific splicing in vivo of the beta-tropomyosin gene: dependence on an RNA secondary structure. *Science.* 1991; 252(5014):1842–5. [PubMed: 2063196]
52. Blanchette M, Chabot B. A highly stable duplex structure sequesters the 5' splice site region of hnRNP A1 alternative exon 7B. *RNA.* 1997; 3(4):405–19. [PubMed: 9085847]
53. McAlinden A, Liang L, Mukudai Y, Imamura T, Sandell LJ. Nuclear protein TIA-1 regulates COL2A1 alternative splicing and interacts with precursor mRNA and genomic DNA. *J Biol Chem.* 2007; 282(33):24444–54. [PubMed: 17580305]
54. Woods IG, Wilson C, Friedlander B, Chang P, Reyes DK, Nix R, et al. The zebrafish gene map defines ancestral vertebrate chromosomes. *Genome Res.* 2005; 15(9):1307–14. [PubMed: 16109975]
55. Duran I, Mari-Beffa M, Santamaria JA, Becerra J, Santos-Ruiz L. Actinotrichia collagens and their role in fin formation. *Dev Biol.* 2011; 354(1):160–72. [PubMed: 21420398]
56. Dale RM, Topczewski J. Identification of an evolutionarily conserved regulatory element of the zebrafish col2a1a gene. *Dev Biol.* 2011; 357(2):518–31. [PubMed: 21723274]
57. Ota KG, Kuratani S. Expression pattern of two collagen type 2 alpha1 genes in the Japanese inshore hagfish (*Eptatretus burgeri*) with special reference to the evolution of cartilaginous tissue. *J Exp Zoolology Part B, Molecular and Developmental Evolution.* 2010; 314(2):157–65.
58. Zhang G, Miyamoto MM, Cohn MJ. Lamprey type II collagen and Sox9 reveal an ancient origin of the vertebrate collagenous skeleton. *Proc Natl Acad Sci USA.* 2006; 103(9):3180–5. [PubMed: 16492784]
59. Smith PJ, Zhang C, Wang J, Chew SL, Zhang MQ, Krainer AR. An increased specificity score matrix for the prediction of SF2/ASF-specific exonic splicing enhancers. *Hum Mol Genet.* 2006; 15(16):2490–508. [PubMed: 16825284]
60. Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR. ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Res.* 2003; 31(13):3568–71. [PubMed: 12824367]
61. Lareau LF, Brooks AN, Soergel DA, Meng Q, Brenner SE. The coupling of alternative splicing and nonsense-mediated mRNA decay. *Adv Exp Med Biol.* 2007; 623:190–211. [PubMed: 18380348]
62. Kornblihtt AR, de la Mata M, Fededa JP, Munoz MJ, Nogues G. Multiple links between transcription and splicing. *RNA.* 2004; 10(10):1489–98. [PubMed: 15383674]
63. Kadener S, Cramer P, Nogues G, Cazalla D, de la Mata M, Fededa JP, et al. Antagonistic effects of T-Ag and VP16 reveal a role for RNA pol II elongation on alternative splicing. *EMBO J.* 2001; 20(20):5759–68. [PubMed: 11598018]
64. Kadener S, Fededa JP, Rosbash M, Kornblihtt AR. Regulation of alternative splicing by a transcriptional enhancer through RNA pol II elongation. *Proc Natl Acad Sci USA.* 2002; 99(12):8185–90. [PubMed: 12060763]
65. de la Mata M, Alonso CR, Kadener S, Fededa JP, Blaustein M, Pelisch F, et al. A slow RNA polymerase II affects alternative splicing in vivo. *Mol Cell.* 2003; 12(2):525–32. [PubMed: 14536091]
66. Howe KJ, Kane CM, Ares M Jr. Perturbation of transcription elongation influences the fidelity of internal exon inclusion in *Saccharomyces cerevisiae*. *RNA.* 2003; 9(8):993–1006. [PubMed: 12869710]

67. Valcourt U, Gouttenoire J, Aubert-Foucher E, Herbage D, Mallein-Gerin F. Alternative splicing of type II procollagen pre-mRNA in chondrocytes is oppositely regulated by BMP-2 and TGF-beta1. *FEBS Lett.* 2003; 545(2-3):115-9. [PubMed: 12804760]
68. Vinatier C, Mrugala D, Jorgensen C, Guicheux J, Noel D. Cartilage engineering: a crucial combination of cells, biomaterials and biofactors. *Trends Biotechnol.* 2009; 27(5):307-14. [PubMed: 19329205]
69. Spiller KL, Maher SA, Lowman AM. Hydrogels for the Repair of Articular Cartilage Defects. *Tissue Eng Part B Rev.* 2011; 5(8):636-47.
70. Johnstone B, Alini M, Cucchiari M, Dodge GR, Eglin D, Guilak F, et al. Tissue engineering for articular cartilage repair--the state of the art. *Eur Cell Mater.* 2013; 25:248-67. [PubMed: 23636950]
71. Tuan RS, Chen AF, Klatt BA. Cartilage regeneration. *J Am Acad Orthop Surg.* 2013; 21(5):303-11. [PubMed: 23637149]
72. Ravindran S, Roam JL, Nguyen PK, Hering TM, Elbert DL, McAlinden A. Changes of chondrocyte expression profiles in human MSC aggregates in the presence of PEG microspheres and TGF-beta3. *Biomaterials.* 2011; 32(33):8436-45. [PubMed: 21820171]
73. Li SW, Prockop DJ, Helminen H, Fassler R, Lapvetelainen T, Kiraly K, et al. Transgenic mice with targeted inactivation of the Col2 alpha 1 gene for collagen II develop a skeleton with membranous and periosteal bone but no endochondral bone. *Genes Dev.* 1995; 9(22):2821-30. [PubMed: 7590256]
74. Aszodi A, Chan D, Hunziker E, Bateman JF, Fassler R. Collagen II is essential for the removal of the notochord and the formation of intervertebral discs. *J Cell Biol.* 1998; 143(5):1399-412. [PubMed: 9832566]
75. Shu B, Zhang M, Xie R, Wang M, Jin H, Hou W, et al. BMP2, but not BMP4, is crucial for chondrocyte proliferation and maturation during endochondral bone development. *J Cell Sci.* 2011; 124(Pt 20):3428-40. [PubMed: 21984813]
76. Sandell LJ, Wang Z, Franz C, Bryan J, Siegel A, Mecham R, et al. Live-cell Imaging of Endothelial Cell Tube Formation: Inhibition by Chondrostatin. *FASEB J.* 2008;101.4.
77. Hayashi S, Wang Z, Bryan J, Kobayashi C, Faccio R, Sandell LJ. The type II collagen N-propeptide, PIIBNP, inhibits cell survival and bone resorption of osteoclasts via integrin-mediated signaling. *Bone.* 2011; 49(4):644-52. [PubMed: 21708300]
78. Wang Z, Bryan J, Franz C, Havlioglu N, Sandell LJ. Type IIB procollagen NH(2)-propeptide induces death of tumor cells via interaction with integrins alpha(V)beta(3) and alpha(V)beta(5). *The J Biol Chem.* 2010; 285(27):20806-17.
79. Fukui N, McAlinden A, Zhu Y, Crouch E, Broekelmann TJ, Mecham RP, et al. Processing of type II procollagen amino propeptide by matrix metalloproteinases. *J Biol Chem.* 2002; 277(3):2193-201. [PubMed: 11705992]
80. Patra D, Delassus E, McAlinden A, Sandell LJ. Characterization of a murine type IIB procollagen-specific antibody. *Matrix Biol.* 2013 Nov 7. [Epub ahead of print].
81. Petit B, Ronziere MC, Hartmann DJ, Herbage D. Ultrastructural organization of type XI collagen in fetal bovine epiphyseal cartilage. *Histochemistry.* 1993; 100(3):231-9. [PubMed: 8244774]
82. Keene DR, Oxford JT, Morris NP. Ultrastructural localization of collagen types II, IX, and XI in the growth plate of human rib and fetal bovine epiphyseal cartilage: type XI collagen is restricted to thin fibrils. *J Histochem Cytochem.* 1995; 43(10):967-79. [PubMed: 7560887]
83. Donoso LA, Edwards AO, Frost AT, Ritter R 3rd, Ahmad N, Vrabec T, et al. Clinical variability of Stickler syndrome: role of exon 2 of the collagen COL2A1 gene. *Surv Ophthalmol.* 2003; 48(2):191-203. [PubMed: 12686304]
84. Snead MP, McNinch AM, Poulson AV, Bearcroft P, Silverman B, Gomersall P, et al. Stickler syndrome, ocular-only variants and a key diagnostic role for the ophthalmologist. *Eye (Lond).* 2011; 25(11):1389-400. [PubMed: 21921955]
85. Richards AJ, Snead MP. The influence of pre-mRNA splicing on phenotypic modification in Stickler's syndrome and other type II collagenopathies. *Eye (Lond).* 2008; 22(10):1243-50. [PubMed: 18309338]

86. Du F, Acland GM, Ray J. Differential splicing of type II procollagen mRNA in canine retina. *Anim Biotechnol.* 1998; 9(1):15–20. [PubMed: 9676231]
87. McAlinden A, Majava M, Bishop PN, Perveen R, Black GC, Pierpont ME, et al. Missense and nonsense mutations in the alternatively-spliced exon 2 of COL2A1 cause the ocular variant of Stickler syndrome. *Human Mutat.* 2008; 29(1):83–90.
88. Fleischmajer R, Olsen BR, Timpl R, Perlish JS, Lovelace O. Collagen fibril formation during embryogenesis. *Proc Natl Acad Sci USA.* 1983; 80(11):3354–8. [PubMed: 6574488]
89. Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J Comp Biol.* 2004; 11(2–3):377–94.
90. Bornstein P, Walsh V, Tullis J, Stainbrook E, Bateman JF, Hormuzdi SG. The globular domain of the proalpha 1(I) N-propeptide is not required for secretion, processing by procollagen N-proteinase, or fibrillogenesis of type I collagen in mice. *J Biol Chem.* 2002; 277(4):2605–13. [PubMed: 11705995]
91. Salminen H, Vuorio E, Saamanen AM. Expression of Sox9 and type IIA procollagen during attempted repair of articular cartilage damage in a transgenic mouse model of osteoarthritis. *Arthritis Rheum.* 2001; 44(4):947–55. [PubMed: 11315934]
92. Aigner T, Zhu Y, Chansky HH, Matsen FA 3rd, Maloney WJ, Sandell LJ. Reexpression of type IIA procollagen by adult articular chondrocytes in osteoarthritic cartilage. *Arthritis Rheum.* 1999; 42(7):1443–50. [PubMed: 10403272]
93. Liu H, Tang L. Mechano-regulation of alternative splicing. *Curr Genomics.* 2013; 14(1):49–55. [PubMed: 23997650]

Figure 1. Alternatively-spliced isoforms of *COL2A1*

Exons 1, 2, 3 and intervening introns 1 and 2 of *COL2A1* precursor mRNAs (pre-mRNAs) are shown. Splicing of pre-mRNA results in production of IIA, IID, IIB and IIC mature mRNA transcripts as shown. Type II and type XI procollagen products containing unprocessed amino (NH₂) and carboxy (COOH) propeptides are also shown. The exon 2-encoded (IIA/IID) CR domain of type II/XI procollagen is denoted by the pink line in the NH₂ propeptide. The blue and green lines in the NH₂ propeptide of type XI procollagen represent the $\alpha 1$ (XI) and $\alpha 2$ (XI) chains, respectively. It is predicted that the truncated IIC mRNA containing premature termination codons (PTCs) in exon 6 will not be translated. Chondroprogenitor cells are known to generate predominantly exon 2-containing *COL2A1* transcripts (IIA/IID). An alternative splicing switch occurs as chondrocyte differentiation takes place. Subsequently, differentiated chondrocytes synthesize predominantly IIB transcripts, devoid of exon 2. The IIC splicing event occurs in progenitor and differentiated chondrocytes. However, IIC mRNA levels as measured by AT-qPCR (36) are low or undetectable which may be due to transcript degradation via nonsense-mediated decay mechanisms.

Figure 2. Sequence and strength scores of IIA, IID and IIC 5' splice sites

Top panel shows the splicing events leading to production of IIA, IID or IIC transcripts and the positions of 5' and 3' splice sites. Middle panel shows the splice site strength scores of the 5' and 3' splice site sequences using MaxEntScan (89) (http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html). The strength of the 3' splice sites were scored using sequences at positions [-20 to +3] with the AG consensus at [-2, -1]. The 5' splice site strength was scored using sequences at [-3 to +6] with the GT consensus at positions [+1, +2]. Bottom panel shows the actual sequence of the IIA, IID and IIC human 5' splice sites [-3 to +6]. A typical consensus (Con) 5' splice site sequence for human U2 introns (3) is also included. Splicing occurs between the uppercase and lowercase nucleotides; only those nucleotides in uppercase are included in the final mature mRNA.

Figure 3. Molecular mechanisms regulating *COL2A1* exon 2 splicing

Panel A shows the presence of the regulatory stem-loop *cis* element in intron 2 of *COL2A1* pre-mRNA (not drawn to scale). This 38 nucleotide stem-loop contains an AU-rich repeat element (ARE) (37). TIA-1 was found to bind to this ARE and enhance inclusion of exon 2 in a *COL2A1* mini-gene system (53). It is predicted that TIA-1 and other enhancer splicing factor proteins are required to strengthen the interaction of the spliceosome with the weak 5' splice sites of exon 2. Panel B shows the consequence of either deleting the alternative IIC 5' splice site *in vitro* (34) or mutating this splice site *in vivo* (McAlinden Laboratory, unpublished observations). As a result of IIC splicing inhibition, levels of IIB mRNA (devoid of exon 2) decrease while levels of IIA/D transcripts (containing exon 2) increase; IIA being the predominant isoform generated. We hypothesize that the IIC splice site functions to control levels of IIA/D and IIB in differentiated chondrocytes. Therefore, by abolishing the IIC splice site, available spliceosome complexes and required accessory factor (AF) splicing proteins may be re-routed to form at the IIA/D weak 5' splice sites.

Figure 4. Immunolocalization of type II collagen triple-helical domain (red) and the exon 2-encoded CR domain (green) in the proximal tibiae of P70 wild type and *Col2a1*^{+ex2} knock-in mice

Dual staining patterns are shown in wild type (+/+) (A, D), heterozygous knock-in (ki/+) (B, E) and homozygous knock-in (ki/ki) (C, F) articular and growth plate cartilage tissue of *Col2a1*^{+ex2} mice. Panels D, E and F are higher magnification images of the white boxed regions of articular cartilage shown in panels A, B and C, respectively. Note the persistent staining of the exon 2-encoded CR domain in the articular cartilage and growth plate cartilage of ki/+ and ki/ki 10 week old mice (B, C, E, F). This staining pattern may be due to the unprocessed IIA NH₂ propeptide of type II/XI procollagen and/or processed IIA NH₂ propeptide from type II procollagen that has not yet been removed from the ECM. Note the positive staining of the IIA CR domain in the superficial zone of wild type articular cartilage as well as in the pericellular domain of deeper zone chondrocytes (A, D). Scale bars = 100 μm (A, B, C) or 50 μm (D, E, F). Image taken from a recent Matrix Biology publication by McAlinden and colleagues (29).

Table 1**Abbreviations**

ARE	adenylate-uridylate-rich element/AU-rich element
AT-qPCR	alternative transcript quantitative real time PCR
<i>Col2a1/COL2A1</i>	type II procollagen gene (mouse/human)
<i>Col2a1</i> ^{+ex2}	transgenic knock-in mouse exclusively expressing exon 2-containing isoforms of <i>Col2a1</i>
<i>Col2a1-mIIC</i>	transgenic knock-in mouse containing an inhibiting mutation in the IIC 5' splice site in <i>Col2a1</i> exon 2
CR	cysteine-rich
ECM	extracellular matrix
ki/ki	homozygous knock-in
ki/+	heterozygous knock-in
MSC	mesenchymal stem/stromal cells
NMD	nonsense-mediated decay
pN-IIA-procollagen	type II procollagen containing the NH2 propeptide
PTC	premature termination codon
RUST	regulated unproductive splicing and translation
VWC	von Willebrand factor type C