Published in final edited form as:

Matrix Biol. 2008 April ; 27(3): 254–266.

Expression of Two Novel Alternatively Spliced COL2A1 Isoforms During Chondrocyte Differentiation

Audrey McAlinden^{b,1,2}, Brian Johnstone^{a,2}, John Kollar^a, Najam Kazmi^a, and Thomas M. Hering^{a,*}

^aDepartment of Orthopaedics, Case Western Reserve University, 10900 Euclid Ave., Cleveland, OH 44106

^bDepartment of Orthopaedic Surgery, Washington University School of Medicine 4940 Parkview Place, St Louis, MO, 63110

Abstract

Alternative splicing of the type II procollagen gene (COL2A1) is developmentally-regulated during chondrogenesis. Type IIA procollagen (+ exon 2) is synthesized by chondroprogenitor cells while type IIB procollagen (- exon 2) is synthesized by differentiated chondrocytes. Here, we report expression of two additional alternatively spliced COL2A1 isoforms during chondrocyte differentiation of bone marrow derived mesenchymal stem cells (MSCs). One isoform, named IIC, contains only the first 34 nucleotides of exon 2 by use of an alternative 5' splice site, resulting in a premature termination codon and possible nonsense-mediated decay of IIC mRNA. Low levels of the IIC isoform were detected by RT-PCR and Southern analysis of COL2A1 cDNA amplified from differentiating rabbit and human MSCs. A second novel transcript, named IID, arises by use of another 5' alternative splice site in intron 2. The IID isoform contains exon 2 plus 3 nucleotides, resulting in the insertion of an additional amino acid. The IID isoform was co-expressed with the IIA isoform during chondrogenesis, and was approximately one-third as abundant. Deletion of the IIC alternative 5' splice site from a COL2A1 mini-gene construct resulted in a significant increase in the IIA:IIB ratio. A mutant mini-gene that inhibited production of the IID isoform, however, had differential effects on the production of the IIA and IIB isoforms: this was apparently related to the differentiation status of the cell type used. These data suggest that COL2A1 mRNA abundance and other aspects of chondrocyte differentiation may be regulated by the use of these previously undetermined alternative splice sites.

Keywords

procollagen type II; COL2A1; alternative splicing; chondrogenesis; mesenchymal stem cell; chondrocyte; cartilage

* To whom correspondence should be addressed: Department of Orthopaedics, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106, Tel: 216-368-1375, Fax: 216-368-1332, E-mail: thomas.hering@case.

^{1,2}Contributed equally

¹Current address: Department of Orthopaedics & Rehabilitation, Oregon Health and Science University, 3181 S. W. Sam Jackson Park Road, Portland, OR 97239

The nucleotide sequence for *O. cuniculus* COL2A1 mRNA (IID isoform) has been submitted to the GenBank database with accession number AF027122.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

Sequencing of the human genome and subsequent genome-wide bioinformatic analyses has revealed that more than 50% of human genes are alternatively spliced to produce different mRNAs derived from the same gene (Modrek et al. 2002). These mRNAs may encode proteins with distinct biological properties or they may contain altered reading frames that render them non-functional and subject to degradation by the intracellular surveillance mechanism called nonsense-mediated decay (NMD1) (Lewis et al. 2003; Lareau et al. 2004).

One of the most important alternative splicing events identified in connective tissue biology is the developmentally-regulated processing of the type II procollagen gene (COL2A1). Type II collagen is the major collagen component of cartilage extracellular matrix and is first synthesized as a procollagen molecule of three identical $\alpha 1(\text{II})$ chains, containing an amino and carboxyl propeptide (Ala-Kokko et al. 1995). The amino and carboxyl propeptides are subject to cleavage prior to the formation of type II collagen fibrils. During COL2A1 alternative splicing, exon 2 is either included in or excluded from mature mRNA producing either type IIA procollagen (+ exon 2) in chondroprogenitor cells or type IIB procollagen (-exon 2) in differentiated chondrocytes. Transcription of the type IIA embryonic isoform occurs during development of other tissues (Ng et al. 1993; Sandell et al. 1994; Lui et al. 1995; Oganessian et al. 1997), but the developmental switch to the type IIB isoform occurs almost exclusively during chondrogenesis. Although it is a well-described event, the importance of this COL2A1 splicing switch in promoting correct development of cartilage tissue is not completely understood.

COL2A1 alternatively-spliced exon 2 encodes a 69 amino acid, cysteine-rich (CR), von Willebrand factor C-like domain in the amino propeptide of COL2A1. Homologues of this CR domain are present in other fibrillar collagen amino propeptides and extracellular matrix proteins including thrombospondins, connective tissue growth factor, and chordin (Larrain et al. 2000; Garcia Abreu et al. 2002). Previous studies have suggested that the CR exon 2-encoded domain may interact with and regulate specific growth factors belonging to the transforming growth factor (TGF)- β superfamily (Zhu et al. 1999; Larrain et al. 2000). Thus, it is possible that the type IIA procollagen isoform may provide the required biological function to ensure proper development of tissues including cartilage. The splicing switch to type IIB procollagen during chondrocyte differentiation suggests that the function of exon 2 is no longer required in mature cartilage. Therefore, the primary function of this type IIB protein isoform may be related to maintaining structural integrity of the extracellular matrix.

To date, there is very little information on mechanisms that control COL2A1 alternative splicing and subsequent production of the type IIA and IIB protein isoforms. However, one recently published study identified an essential intronic regulatory motif that is required for correct splicing of COL2A1 pre-mRNA (McAlinden et al. 2005). The present study has identified two novel alternatively-spliced COL2A1 mRNA transcripts that are expressed during chondrocyte differentiation. One transcript, named IIC, is formed by use of a 5' alternative splice site in exon 2 resulting in an mRNA containing only the first 34 nucleotides of exon 2. This splicing event results in introduction of a premature stop codon rendering it a potential target for mRNA degradation by NMD. A second novel transcript, named IID, was also identified. This isoform is produced by use of another 5' alternative splice site located in intron 2 directly adjacent to the 5' splice site that produces the type IIA procollagen isoform. This alternative splicing event results in the insertion of an additional codon without altering the reading frame. Splicing of mutant mini-genes lacking either the IIC or IID alternative 5' splice site resulted in altered levels of IIA and IIB mRNA compared to levels produced by splicing of the wild type mini-gene. Interestingly, the differentiation status of the cells had a profound effect on IIA and IIB production in the context of the IID splice site mutation. Here,

opposing effects were noted when the mutant IID mini-gene was spliced by mature chondrocytes (RCS cells) compared to the other less differentiated chondrocytes used in the present study. Identification of these novel spliced mRNA transcripts during chondrogenesis provides new, intriguing insights into the potential mechanisms governing COL2A1 gene expression at the post-transcriptional level.

2. Results

2.1 Identification of COL2A1 mRNA IIC and IID isoforms

RNA was prepared from rabbit bone marrow-derived MSC day 7 aggregate cultures, and was used for RT-PCR using primers spanning exons 1-7 (primers 1 and 2 in Fig. 1), with the expectation that the IIA and IIB isoforms could be identified in differentiating cells. Following PCR, collagen IIA and IIB amplification products were observed on ethidium bromide-stained gels. These DNA fragments were ligated into a TA cloning vector (pCRII, Invitrogen) and transformed into *E. coli*. Plasmids representing different cloned PCR products were purified and cDNA inserts were sequenced. As expected the majority of cloned cDNAs corresponded to the collagen IIB splice variant (exon 1 spliced to exon 3). One cloned cDNA, however, was found to contain only the first 34 nucleotides of exon 2 (hence designated exon 2C) between exons 1 and 3 (Fig. 1). Other cloned inserts resembled the published human IIA sequence, with the insertion of exon 2, but contained an additional three nucleotides (GTC in the rabbit COL2A1 gene) before the exon 3 junction (hence designated exon 2D) (Fig. 1).

2.2 Temporal differences in expression of COL2A1 alternative isoforms during chondrogenesis in vitro

Rabbit and human MSC aggregate cultures were prepared to determine the temporal expression of alternatively spliced COL2A1 mRNAs during chondrogenesis in vitro. PCR primers spanning the alternatively spliced exons were used to amplify cDNA from successive stages of chondrogenesis, and products were identified by Southern hybridization to exon-exon spanning oligonucleotides. This analysis reveals the presence of alternatively spliced mRNAs, and is non-quantitative. As with any endpoint analysis of PCR samples, low and high abundance cDNA may yield the same amount of reaction product. The preferred amplification of a particular cDNA isoform is a measure of the “uniqueness” of that isoform in the sample. For example, the IIC and IID isoforms are easily detected at day 0 and 1 of rabbit MSC differentiation, since there is little competition from the IIB isoform, which predominates at later time points.

Northern blot analysis was performed to verify the differentiation of rabbit MSCs (Fig. 2A), and increasingly abundant COL2A1 mRNA was observed from days 6-14 in the rabbit MSC cultures. Southern blot analysis was carried out to identify bands corresponding to IIA, IIB, IIC, and IID mRNA isoforms (Fig. 2B) (Fig. 2C-F). Four different sized cDNA bands were identified in Southern blots with the ABCD(R) probe (Fig. 2C). These bands were further identified using probes D, B, and C (Fig. 2D-F). The IID form was seen from day 1-6 (Fig. 2D), the B isoform from day 3-14 (Fig. 2E) and the C isoform band was detected at day 0, and was less apparent at days 1-6 (Fig. 2F), likely due to competitive amplification of the B isoform at later times.

Northern blot analysis detected increasingly abundant COL2A1 expression from day 7-21 in human MSC aggregate cultures (Fig. 2G). Initiation of chondrogenic differentiation in the human MSC cultures appeared to be delayed relative to the more rapidly differentiating rabbit MSC cultures. Human MSCs, however, proceeded through a very similar progression of alternatively spliced forms, showing predominantly IIC initially, when IIB was less abundant, followed by IIA/D and IIB. RNA harvested from each pellet culture time point was used to

amplify cDNA using primers 1 and 2 (Table 1 and Fig. 1) spanning exons 1-7. PCR products (Fig. 2H) were analyzed by Southern blotting (Fig. 2I-L). Southern blotting with the ABCD (H) probe revealed the abundant IIA and IIB isoforms (Fig. 2I). Oligonucleotides designed to detect isoforms A, B, and C individually were also used. Hybridization to detect the A isoform individually revealed a doublet (Fig. 2J). The B isoform was confirmed as an abundant species from day 7-21 (Fig 2K). Hybridization with the C probe showed a band consistent in size with the C isoform at day 7 (Fig 2L). This band was not detectable at later time points, likely due to consumption of primers and dNTPs by amplification of the A and B isoforms. There was crossreactivity between the C oligonucleotide and the more abundant IIA and IIB PCR products, which could not be completely eliminated following hybridization and washing.

2.3 Cloning and sequencing the human IIC isoform

We confirmed the IIC splicing event in another experiment by generating PCR products that excluded the more abundant IIB isoform, which also enabled cloning and sequencing of the human IIC product. The samples used for this analysis were from a different human MSC differentiation experiment that examined chondrogenesis from day 0 to 14, and was previously analyzed for total COL2A1 expression (Hering et al. 2004). In this experiment, COL2A1 expression over time was similar to the results shown in Fig. 2, but covered a shorter time course. Primers 3 and 4 were designed to amplify between exon 2 and exon 6 (Table 1 and Fig. 1). The most abundant product over this time course was a singlet or doublet band near the expected size of the IIA (335 bp) or IID (338 bp) product (Fig. 3A). The IIC isoform was identified by Southern blot analysis using probes ABCD and C (Fig. 3B and 3D). The human IIC product was evident at day 14. The IIC PCR product from this reaction was subsequently enriched by preparative agarose gel electrophoresis and cloned into the pCRII vector (Invitrogen). The cloned fragment was sequenced and confirmed to be the human ortholog of the previously isolated rabbit IIC isoform.

2.4 Characterization of Human IIA and IID PCR products

Two bands were typically observed at the position where IIA cDNA was expected to migrate using primers that amplified a region between exons 1 and 7 (Fig. 2) or between exons 2 and 6 (Fig. 3). To clarify what this doublet band represents, different amounts of IIA and IID cDNA templates were amplified by PCR using the primers spanning exons 2-6. The PCR products of IIA and IID template mixtures was also found to migrate as a doublet on an agarose gel (Fig. 4). Amplification of only the IIA or IID template, however, produced a single, faster migrating band. This result suggests that the upper band of the doublet may represent a conformationally distinct heteroduplex PCR product containing mismatched IIA and IID cDNA strands. Alternatively, cDNA containing a hairpin fold or other secondary structure may have been produced. Of interest to our analysis, this doublet apparently indicates coexpression of the IIA and IID isoforms.

2.5 Identification of conserved IIC and IID splice sites in multiple species alignment of exon 2

A multi-species comparison of COL2A1 sequences between exon 2 and the end of exon 3 is presented in Fig. 5. These include sequences obtained in this study (*H. sapiens* (1) and *O. cuniculus*), as well as sequences from other species obtained from the Genbank database. It is apparent that there is significant conservation of the GT residues representing the alternative splice site utilized to produce the IIC alternative isoform, which is present in all species for which sequence is available except *G. gallus*. It became apparent that when this comparison was performed, there was variability in the length of exon 2 (e.g. IID contains three nucleotides more than the IIA isoform, e.g. GTT in *H. sapiens*; GTC in *O. cuniculus* and *M. musculus*, GTG in *X. laevis*) between sequences compared from different species. This additional triplet

results in the insertion of a Trp residue between Gly-98 and Gln-99 in the human pro- α 1(II) chain sequence (for human IIA sequence see: Entrez Protein database Acc. No. P02458; for conceptual translation of human IID sequence see EAW57968) or an Arg residue in the corresponding position in the rabbit sequence (for rabbit IIA sequence see: Acc. # D83229; rabbit IID sequence (this work) see: Acc. No. AF027122). Further analysis revealed that the length of exon 2 is also variable within species. Human and mouse isoforms of exon 2 with or without the additional three bases were identified. Due apparently to its relatively low abundance in tissues expressing COL2A1, the IIC isoform was not found in the database for any species examined. The human IID sequence obtained in this study appears to be the first experimental verification of the previously derived conceptual IID isoform sequence.

2.6 Quantitation of COL2A1 isoform expression during chondrogenesis by real-time PCR

The Southern blot analyses shown in Figs. 2 and 3 were not quantitative, but are indicative of the predominant spliced product in each sample. In this assay, the total amount PCR product produced after multiple PCR cycles can not be directly correlated with mRNA abundance. Quantitative real time PCR analyses were used to determine the percentage of the total COL2A1 mRNA contributed by each isoform at each timepoint during pellet culture chondrogenesis. Relative COL2A1 mRNA values were determined by the $\Delta\Delta C_T$ method. In addition, absolute values (copy number) were calculated relative to dilutions of cloned cDNA standards. Relative and absolute values for total COL2A1 mRNA in each sample were first determined with an exon 5-6 probe-primer set (IIABCD assay), which was used to determine a correction factor for RNA measurement or cDNA synthesis. Absolute determinations for each isoform were corrected using this factor, summed and plotted as a stacked bar graph (Fig. 6).

Data shown are from one of three pellet culture differentiation series showing similar profiles. The absolute determination for each isoform was converted to a percentage of relative COL2A1 mRNA levels (all isoforms). In this assay the IIC isoform was below the level of detection at all time points. The IIA, IID, and IIB isoforms, however, were each within the range of the assay standard curves constructed using cloned inserts for each alternatively spliced isoform. Interestingly, levels of IID isoform expression were found to be approximately one-third of the IIA mRNA expression levels at all time points. As anticipated, the ratio of IIB to (IIA + IID) increased progressively with time in pellet culture.

2.7 Deletion of the IIC alternative 5'splice site in exon 2 increases type IIA mRNA isoform production

We have previously used a wild-type human COL2A1 mini-gene to study splicing regulation of exon 2 (McAlinden et al. 2005). We used this model system (Fig. 7) to investigate if the newly-identified alternative splice sites are functional in regulating expression of the IIA and IIB splice forms. A deletion mutant mini-gene ($\Delta E2$ -c5'ss) was generated to investigate the effect of deleting the 10 nucleotide sequence GGGCAGAGGT from exon 2 (Fig. 8) upon COL2A1 alternative splicing. Interestingly, this sequence was originally targeted because it was predicted to be a potential binding site for a ubiquitously expressed exon splicing enhancer (ESE) protein ASF/SF2. This was determined by entering the human exon 2 sequence into the web based ESEfinder program (<http://exon.cshl.org/ESE>) (Cartegni et al. 2003). The present study identified a novel, minor splice product, termed IIC, containing the first 34 nucleotides of exon 2, which is formed by use of an alternative 5' splice site with the sequence AG ↓ gtaaa. Interestingly, this sequence conforms better to the 5' splice site consensus sequence than the type IIA 5' splice site sequence (TG ↓ gttgt; human sequence) (Fig. 5A). The last four nucleotides of the region deleted to create the mutant mini-gene $\Delta E2$ -c5'ss (AGGT) corresponds to part of the IIC alternative 5' splice site.

The ratio of IIA and IIB spliced isoforms resulting from splicing the wild type mini-gene is somewhat different for each cell type, as was previously established (McAlinden et al. 2005). In the T/C28 chondrocytes and RCS cells, there are higher ratios of the IIB/IIA isoforms, presumably reflective of chondrocyte differentiation. In this study, the RCS cells represent the most mature chondrocyte phenotype compared to the other chondrocyte cell lines used; this was also proven in a previous study (McAlinden et al. 2005). Splicing of the $\Delta E2$ -c5'ss mini-gene results in a significant change in the COL2A1 splicing pattern regardless of the cell type in which the mini-gene construct is expressed (Fig. 8). Here the type IIA isoform was the predominant transcript produced and there was a significant decrease in the level of the IIB mRNA transcript produced. In light of our discovery of the novel IIC splice form, however, we conclude that this cis element may actually be an alternative 5' splice site, which may serve to regulate synthesis of correct levels of IIA/D and IIB spliced isoforms. In addition, we analyzed splicing of another mutant mini-gene in which the 5' splice site of exon 2 was completely disrupted such that IIA or IID transcripts would never be synthesized. We hypothesized that disruption of the IIA/D splice sites may potentially result in increased use of the alternative 5' splice site and hence production of the IIC isoform. In all cell types tested, there was a complete inhibition of IIA/D production as expected. However, we did not detect synthesis of the IIC isoform when the IIA/D splice sites were removed (results not shown). This suggests that the IIC splicing event is not in direct competition with the IIA/D splice site, and may be independently regulated.

2.8 Mutation of the IID 5' splice site alters the ratio of IIA:IIB spliced isoforms

We produced a mutant mini-gene that disrupted the IID 5' splice site (mIID; Fig. 9). We interpret the disappearance of the minor upper band (also see Fig. 4) after splicing of the mIID mini-gene to indicate that we had inhibited synthesis of IID mRNA, but retained production of the IIA isoform (Fig. 9). Densitometric quantitation of the bands corresponding to the IIA and IIB isoforms showed that there was a decrease in the ratio of IIA:IIB mRNA transcripts derived from splicing of the mIID mini-gene in HEK-293, MC615 and TC28 cells. This change in ratio was mainly due to an increase in the levels of IIB isoform production. In the RCS cells, however, the opposite effect was seen where there was a significant increase in the ratio of IIA:IIB spliced isoforms as a result of splicing the mIID mini-gene. Taken together, it is clear the production of the IID isoform occurs in all cell types tested, albeit at different levels. This result suggests that the IID splicing event may play a role in regulating the switch from collagen IIA to IIB.

3. Discussion

It has been well documented that during development of cartilage, chondroprogenitor cells express the embryonic type IIA procollagen isoform containing exon 2, which encodes a cysteine-rich globular domain in the amino propeptide of the procollagen molecule. With further chondrocytic differentiation there is an increase in the production of the IIB isoform, and a decrease in the IIA isoform (Ryan et al. 1990). The present study provides new information on this developmentally-regulated COL2A1 alternative splicing event. Specifically, we have identified expression of two novel alternatively-spliced mRNA transcripts of type II procollagen during chondrocyte differentiation of bone marrow-derived mesenchymal stem cells. These two mRNA transcripts, that we have named IIC and IID, are formed by the use of two different 5' alternative splice sites. We suspect that the IIC isoform has previously remained undetected due to its low steady-state levels in tissues expressing COL2A1. The IID isoform, many examples of which have already been placed in DNA databases, may have been assumed to be a species-specific difference or a polymorphism. Further study of these transcripts, recognized now to be alternatively spliced mRNAs, will

likely provide new insights into the regulatory mechanisms governing COL2A1 gene expression post-transcriptional level.

The IIC transcript was detected in both rabbit and human MSC aggregate cultures using primers that were designed to amplify both the type IIA (+ exon 2) and type IIB (- exon 2) isoforms. In both rabbit and human MSCs, the IIC variant was apparent at early times in culture. We observed an apparent delay in the differentiation of human relative to the rabbit MSC cultures, which may be age (Zheng et al. 2007) and/or species specific. Further analysis confirmed this transcript to be a third alternatively-spliced isoform consisting of the first 34 nucleotides of exon 2. This truncated mRNA, named IIC, is formed by use of an alternative 5' splice site in exon 2, which conforms better to a 5' splice site consensus sequence (Burge et al. 1999) than that used to create the IIA spliced isoform.

Steady-state levels of IIC mRNA during chondrogenic differentiation of MSCs are very low. This may not necessarily indicate, however, that the IIC transcript is produced in low abundance by the splicing apparatus. Interestingly, the splicing event that produces the type IIC mRNA transcript results in the introduction of a downstream premature termination codon (PTC). Consequently, the PTC-containing mRNA is a likely target for degradation by nonsense-mediated decay (NMD). NMD is an mRNA surveillance mechanism that selectively degrades nonsense mRNAs (Frishmeyer et al. 1999) and has been shown to occur in the nucleus (Buhler et al. 2002) as well as the cytoplasm (Atkin et al. 1995). We were not able to detect a IIC protein product in MSC cultures, using an antibody raised against a putative IIC peptide epitope (result not shown). Although IIC mRNA may be degraded by the NMD mechanism, the splicing event itself may influence the ratio of the other alternatively spliced transcripts.

One common outcome of alternative splicing is down-regulation of gene expression by synthesis of non-functional mRNA isoforms of the active gene product. This can be accomplished by various mechanisms including production of alternatively-spliced mRNA transcripts containing a PTC to induce its degradation by NMD (Morrison et al. 1997; Lareau et al. 2004). Recently, large scale computational analyses of alternative transcripts from expressed sequence tags (ESTs) (Lewis et al. 2003; Baek et al. 2005) and quantitative microarray profiling (Pan et al. 2006) has demonstrated that as many as 35% of mammalian genes produce at least one transcript that contains a PTC, thus rendering it a target for NMD. Consequently, it was proposed that gene expression can be regulated at the level of post-transcription by the coupling of alternative splicing and NMD processes through synthesis of these non-translatable spliced isoforms; this process has been termed "regulated unproductive splicing and translation" (RUST) (Lewis et al. 2003). Several recent studies have added to a growing list of experimentally characterized cases that validate the proposed broad role of RUST gene regulation in addition to a general mechanistic link between NMD and pre-mRNA splicing in mammalian cells (Lejeune et al. 2005). For example, polypyrimidine tract binding protein (PTB), an important regulator of mRNA processing, was shown to autoregulate its protein level through RUST (Wollerton et al. 2004). Similarly, other proteins involved in regulating pre-mRNA alternative splicing (Tra2- β , TIAR and SC35) have also shown to be regulated by expression of PTC-containing transcripts (Sureau et al. 2001; Stoilov et al. 2002; Le Guiner et al. 2003). This mode of regulation, however, is not restricted to genes involved in mRNA processing since other genes encoding FGFR-2, AUF1, ABCC4 and SSAT, for example, have also been shown to be regulated by producing alternatively-spliced transcripts containing a PTC (Wilson et al. 1999; Jones et al. 2001; Lamba et al. 2003; Hyvonen et al. 2006).

With respect to the present study, the production of low levels of the COL2A1 type IIC alternative transcript during chondrocyte differentiation may represent a RUST-related mechanism necessary to control production of correct levels of type IIA and type IIB

transcripts. This is supported by our functional analyses using a COL2A1 mini-gene devoid of the 5' alternative splice site in exon 2 required to produce the IIC isoform. Splicing of this mutant mini-gene ($\Delta E2$ -c5'ss) resulted in a significant change in the ratio of IIA:IIB mRNA transcripts. This was mainly due to a decrease in the level of IIB mRNA synthesized (Fig. 8). However, when all cell types in the present study were transfected with a mutant mini-gene construct in which the regular 5' splice site of exon 2 was completely abrogated, there was no apparent increase in the IIC spliced isoform. This suggests that there is a low, constant rate of splicing activity at the IIC alternative splice site, which is not affected by the inhibition of IIA/IID mRNA production. Indeed, if the IIC isoform is functional only at the mRNA level and thereafter degraded by NMD, then it would not be particularly beneficial to the cell to over-utilize the IIC alternative splice site. Alternatively, it may be that the IIC spliced isoform is so rapidly degraded by NMD, that it is never detected in the in vitro model system used in our experiments. A recent study from one of our laboratories has identified an essential cis element in intron 2 of COL2A1 pre-mRNA that is necessary for correct splicing of IIA and IIB transcripts (McAlinden et al. 2005). This cis element, as well as other as yet unidentified cis elements within or spanning the regulated exon 2 of COL2A1, will likely serve as binding sites for specific trans-acting splicing factor proteins, some of which may be expressed in a developmentally-regulated manner. In the present study, we have identified an additional mechanism for correct production of IIA and IIB isoforms and one that also provides evidence for a link between alternative splicing and NMD during chondrocyte differentiation. In addition, the 5' alternative splice site in exon 2 that is used to produce the PTC-containing IIC isoform is highly conserved between species (Fig. 5) which may relate to the functional importance of this mechanism during chondrogenesis.

DNA sequencing of cDNA clones derived from human and rabbit MSCs during chondrogenic differentiation revealed the presence of exon 2-containing transcripts with an additional 3 nucleotides (GTT in the human sequence; GTC in the rabbit sequence) (Fig. 1) at the end of exon 2 that does not disrupt the protein reading frame. We named this transcript IID and showed by quantitative real time PCR that it is co-expressed with the IIA isoform during in vitro differentiation of human MSCs and is approximately one-third as abundant as the IIA isoform (Fig. 6).

Although there is only a 3 nucleotide difference between the IIA and IID isoforms, cDNA samples containing mixtures of IIA and IID templates (Figs. 2, 3, and 9) showed distinct doublet bands. PCR amplification using a mixture of templates with and without an insertion has been previously shown to result in production of heteroduplex PCR products showing anomalous electrophoretic migration (Separovic et al. 1994). We observed that, as expected, PCR products spanning exons 2-6 generated from purified IIA or IID templates migrated similarly on 2% agarose gels (Fig. 4). Amplification from mixed IIA and IID templates, however, resulted in the appearance of a second, slower-migrating cDNA product (Fig. 4). It is possible that this trailing band represents a heteroduplex PCR product resulting from the annealing of IIA and IID single strands during PCR. Consistent with this finding, we observed that an oligonucleotide probe specific for the IID form was reactive predominantly with the upper band (Fig. 2D) and a probe for the IIA form was reactive with both bands of the doublet (Fig. 2J). Alternatively, this difference in migration might also be explained by formation of hairpin secondary structure in the human IID spliced mature mRNA due to existence of an axis of symmetry around the G residue shown in bold font (5'-TGG**TTG** G CAACCA-3'). The underlined T nucleotide marks the end of exon 2 in the IID isoform and the underlined G residue marks the beginning of exon 3. The extra GTT nucleotides found in the IID isoform is shown in italics. This structure may also be present in cDNA which may explain the differential migration pattern seen after PCR and electrophoresis of the reverse-transcribed mRNA derived from the COL2A1 mini-gene. If our interpretation of this finding is correct, then the appearance of a doublet following PCR, when analyzing a cDNA sample such as those described here,

may be an indication of the co-expression of both the IIA and IID isoforms. Our data (Fig. 2) suggests that IID may predominate early in differentiation. The possible switch from expression of the IID to IIA form warrants further investigation.

We further confirmed expression of IID mRNA by mutating a region of the 5' splice site sequence that would be utilized to produce this novel isoform. Consequently, the slower-migrating band which appears to represent a hybrid of IIA and IID cDNAs (Fig. 4) disappeared after splicing of the mutant mini-gene (mIID) (Fig. 9). Interestingly, data obtained from splicing of the mIID mini-gene showed that inhibition of IID mRNA formation affected the production of the IIA and IIB mRNA isoforms. The cell types studied that normally favor production of the IIA/IID isoforms (HEK-293, MC615, T/C28), produced more of the IIB isoform when the IID splice site was disrupted. This suggests that both the IIA and IID splice sites are required to promote efficient spliceosome formation at the 3' end of exon 2, which may be particularly relevant since these IIA/IID 5' splice sites do not conform to the general consensus sequence and are referred to as "weak" splice sites. Therefore, it is not altogether surprising that exon skipping, hence production of the IIB isoform, would be upregulated in the absence of the IID splicing event.

The opposite effect, however, was found in RCS cells where splicing of the mIID COL2A1 mini-gene, devoid of the IID splice site, resulted in an increase in IIA mRNA synthesis and a decrease in the level of IIB mRNA production. RCS cells were used in this study as a source of differentiated chondrocytes. We confirmed this, as well as the de-differentiated nature of the MC615 and TC28 chondrocytes, in a previous study by measuring endogenous levels of aggrecan and type I collagen mRNA (McAlinden et al. 2005). This dramatic change in splicing pattern in RCS chondrocytes is likely due to differences in splicing factor protein profiles. It would appear likely that there would be a distinct subset of trans-acting splicing factors expressed in the nucleus of mature chondrocytes that enable them to favor the exon skipping event to produce predominantly IIB mRNA. It is possible, therefore, that part of the IID splice site sequence functions as an inhibitory cis element that serves as a binding site for a splicing factor repressor protein (expressed only in mature chondrocytes), which functions to inhibit exon 2 splicing (IIA production) and thereby promote exon skipping (IIB production). Therefore, mutation of this site would reverse this splicing repression, as was shown in Fig. 9. This interesting possibility is currently being examined in our laboratories.

These studies have demonstrated not only the presence of novel IIC and IID mRNA isoforms of COL2A1 but also that the alternative splice sites used to produce these novel isoforms may function to regulate the production of the well-described IIA and IIB COL2A1 isoforms that we know are translated to proteins. In this respect, it is possible that the overall chondrocyte phenotype may be affected if the mini-gene modifications described here were stably incorporated into the endogenous COL2A1 gene to either disrupt the IIC or IID splicing events. The seminal studies described in this manuscript have provided a solid foundation to justify doing knock-in mutational analyses. We are currently designing strategies to produce these transgenic mice.

Unlike the type IIC mRNA, we predict that the IID mRNA transcript could be translated to produce a protein containing an extra tryptophan (in human) or arginine (rabbit and mouse) amino acid at the end of the globular cysteine-rich (CR) domain encoded by exon 2, and may result in a significant functional change in the protein. Tryptophan residues are highly hydrophobic, containing an aromatic side chain, while arginine residues are amphipathic, with part of the side chain nearest to the backbone being hydrophobic and the side chain end being positively charged. The presence of either amino acid within a protein interaction domain could potentially alter ligand binding affinities. For example, the von Willebrand Factor type A domain of integrin $\alpha 2\beta 1$ has been shown to bind with greater affinity to collagens, laminin and

E-cadherin due to a glutamic acid to tryptophan gain-of-function mutation (Aquilina et al. 2002). A proline to arginine mutation in FGFR1 and FGFR3 results in an enhancement of FGF binding affinity (Ibrahimi et al. 2004). Therefore, if type IID procollagen protein is synthesized, it will likely contain a globular CR domain within the amino propeptide having conformational properties different from the type IIA exon 2-encoded domain. Previous studies have suggested that the type IIA CR domain can bind to and potentially regulate TGF- β superfamily growth factors during development (Zhu et al. 1999; Larrain et al. 2000). It is therefore possible that the type II procollagen CR domain may have a different binding affinity for growth factors or other matrix molecules when the IID isoform is expressed. Studies are currently underway to determine whether IID mRNA can indeed be translated to protein or whether the splicing event is an RNA-based regulatory mechanism that functions to keep splicing and translation of the other IIA and IIB COL2A1 transcripts in check.

4. Materials and methods

4.1 In vitro chondrogenesis of rabbit and human bone marrow derived MSCs

Human and rabbit bone marrow-derived mesenchymal stem cells (MSCs) were placed in pellet cultures to promote chondrogenic differentiation as we have previously described (Johnstone et al. 1998; Yoo et al. 1998). Briefly, human bone-marrow derived MSCs were plated at a density of 10×10^6 nucleated cells/100 mm dish after fractionation on a Percoll gradient. Rabbit bone marrow-derived MSCs were plated directly without fractionation. Adherent cells formed colonies that covered approximately 80-90% of the culture dish by 14-20 days in monolayer culture. The cells were then trypsinized, counted and aliquots containing 2×10^5 cells were centrifuged at $500 \times g$ for 5 minutes in 15 ml polypropylene tubes containing 0.5 ml of a defined medium consisting of DMEM-high glucose with ITS+ Premix: insulin (6.25 $\mu\text{g/ml}$), transferrin (6.25 $\mu\text{g/ml}$), selenous acid (6.25 $\mu\text{g/ml}$) and linoleic acid (5.35 $\mu\text{g/ml}$) in addition to bovine serum albumin (1.25 $\mu\text{g/ml}$), pyruvate (1 mM), ascorbate 2-phosphate (37.5 $\mu\text{g/ml}$), dexamethasone (10^{-7} M) and TGF β -1 (10 ng/ml, recombinant human). Resulting cell pellets were incubated at 37° C, in a 5% CO₂ atmosphere. Within 24 hours of incubation, the cells formed non-adherent aggregates. Medium was replaced at 2-3 day intervals and cell aggregates were harvested at various time points up to 21 days.

4.2 RNA isolation

RNA was isolated from pooled rabbit and human MSC pellets using the TRIZOL reagent (Life Technologies Inc.) or the RNeasy kit (Qiagen). RNA samples were assayed by A₂₆₀ and A₂₈₀ measurements to determine purity, concentration and yield.

4.3 Northern blot analysis

RNA samples (15 μg) were electrophoresed through 1% agarose gels containing formaldehyde, transferred to nitrocellulose membranes and were hybridized with a ³²P-labeled human COL2A1 probe (a 3.8 kb EcoRI fragment containing exons 45 to 54), washed and exposed to X-ray film as previously described (Hering et al. 1994). Films were scanned and densitometry was performed using NIH image software.

4.4 RT-PCR amplification and cloning of COL2A1 mRNA isoforms

RNA isolated from either rabbit and human pre-aggregate MSCs, MSC pellets was used as a template for reverse transcription using oligo d(T)-primers. Resulting rabbit or human cDNA was amplified by PCR using a human-specific primer pair (Primers 1 and 2 in Fig. 1 and Table 1) to amplify a region (exon 1-exon 7) of COL2A1 containing all previously known splice variants within this interval (Su et al. 1989). The expected sizes of resulting PCR products were: 432 bp (IIA form) and 225 bp (IIB form). These primers also amplified a 259 bp product

identified as the IIC form. To prevent competition from the more abundant IIB isoform, and to expedite the detection and isolation of human IIC and IID clones, additional primers (primers 3 and 4 in Fig. 1 and Table 1) were designed to amplify a region from the first 34 nucleotides of exon 2 to exon 6. The anticipated sizes of PCR products were 168 bp (IIC form), 335 bp (IIA form) and 338 bp (IID form). The rabbit IID sequence from exon 1-7 has been deposited to the GenBank data base under the accession number AF027122. Total PCR reaction products amplified from human or rabbit cDNAs were ligated into a TA cloning vector (pCRII, Invitrogen) and transformed into *E. coli*. Plasmids representing different cloned PCR products (designated as clones 97-3/4-3 (IIA plasmid), 97-1/2-32 (IIB plasmid), 222-4 (IIC plasmid), and 223-6 (IID plasmid) were purified, assayed for DNA concentration by A_{280} , and cDNA inserts were sequenced. Sequence data was analyzed and alignments were performed using MacVector software (Accelrys, Inc., San Diego, CA). Plasmid 97-3/4-3 (IIA) and 223-6 (IID) were used in template mixtures for PCR amplification to characterize the products of exon 2-6 PCR reactions. Template mixtures containing different ratios of the IIA and IID plasmids were amplified using primers 3 and 4 (Table 1 and Fig. 1). Reaction conditions were: 94°C/10 min/1 cycle; 94°C/1 min, 59°C/1 min, 72°C/1 min $\times$ 30 cycles; 72°C/10 min/1 cycle in a 20 μ l reaction volume. All four plasmids were used as standards for absolute determination of mRNA copy number using quantitative real time PCR.

4.5 Southern blot analysis

Oligonucleotide probes were radiolabeled at the 5'-termini by T4 polynucleotide kinase using a standard protocol (Ausubel et al. 1991). Oligonucleotides were complementary to the COL2A1 sequence internal to the PCR amplifying primers (Fig. 1 and Table 1), and were used sequentially to detect cDNA representing all isoforms, or specifically the IIA, IIB, IIC and IID isoforms of COL2A1 by applying either SSC or TMAC-based hybridization protocols (Ausubel et al. 1991). The C(R) probe was inadvertently synthesized with a single bp deletion (asterisk in Table 1) which had no detectable effect upon Southern hybridization to the rabbit IIC PCR product. Following agarose electrophoresis, membrane transfer, hybridization and washing, membranes were exposed to Kodak BioMax film for varying periods of time at -70°C using a radiation intensifying screen.

4.6 Real-time Quantitative PCR analysis

Custom designed Taqman assays specific for each human collagen II isoform (IIA, IIB, IIC and IID) were purchased from Applied Biosystems Inc. (Table 2). For each assay, a forward and reverse primer specified the amplicons, the lengths of which were 68 bp (IIA), 65 bp (IIB), 65 bp (IIC) and 71 bp (IID). MGB (Minor Groove Binding) probe oligonucleotides complementary to either the sense (IIA) or antisense (IIB, IIC and IID) strand were labeled with the FAMTM fluorescent dye. The probe oligos spanned exons 2A-3 (IIA), exons 1-3 (IIB), exons 2C-3 (IIC), and exon 2D-3 (IID). An assay (Assay ID: Hs01060346_g1) to detect all COL2A1 mRNA isoforms (termed IIABCD) at the common exon 5-6 splice junction, and an assay for human 18S rRNA (Assay ID: Hs99999901_s1) were also purchased from ABI. Assays were performed using an ABI 7500 Real-Time PCR instrument. Standard dilutions of each of the cloned IIA, IIB, IIC and IID PCR products described above were produced, ranging from 3.0E+3 to 3.0E+8 copies per aliquot assayed, with copy number calculated relative to each plasmid molecular mass and molar concentration determined from the A_{280} value.

We initially performed a relative assay using the $\Delta\Delta C_T$ method with the IIABCD Taqman assay. 18S rRNA was used as an endogenous control for all samples, and all relative quantities were determined relative to the day 0 value (defined as 1.0). Following the relative quantification for total collagen II mRNA, absolute quantification was performed for total collagen II mRNA using the same IIABCD primer-probe as for the relative quantification, and the IIB plasmid standard curve.

The IIA, IIB, IIC and IID plasmid standards were used for absolute quantification of each individual isoform in the MSC pellet culture cDNA samples. We confirmed that 3.0E+6 copies of each standard had the same C_t using an assay (IIABCD) that would react equally well with all four isoforms to confirm the accuracy of our DNA quantification and dilutions.

To test the specificity of each Taqman assay, each plasmid standard was assayed with the IIA, IIB, IIC and IID probe/primer sets. We determined that the IIA and IIC assays required higher annealing temperatures (66°C and 62°C, respectively) than the manufacturer-recommended 60°C to achieve 98% or greater specificity for the “target” isoform. The IIB and IID assays showed essentially no cross-reactivity with the other two COL2A1 isoforms.

Absolute quantification provided copy number per aliquot, but did not account for inaccuracies in measuring RNA concentration or differences in efficiency of cDNA synthesis. Therefore, a factor $F = \text{ABCD}(\text{rel}) / \text{ABCD}(\text{abs})$ was calculated for each sample assayed. This factor is a number which, if multiplied by the absolute value would equal the relative value. The A, B, C and D isoforms were each assayed separately versus their own standard curves to derive absolute copy numbers of each isoform per sample. Each of the IIA, IIB, IIC, IID, and IIABCD absolute values were multiplied by the factor calculated above, wherein $A(F) + B(F) + C(F) + D(F) = \text{ABCD}(F)$. Each corrected value was then plotted as a stacked bar graph, showing the relative amounts of each isoform at each time point.

4.7 Synthesis of mutant COL2A1 mini-genes

A mutant mini-gene construct was synthesized by deleting a 10 nucleotide sequence (GGGCA-GAGGT) in exon 2 that contained the first 4 nucleotides (underlined) of the 5' alternative splice site (AGGTATAA) responsible for the IIC spliced isoform detected in the present study. This mini-gene was named $\Delta E2\text{-c5'ss}$ (Fig. 9). Another mutant mini-gene, mIID, was constructed to disrupt the IID 5' splice site sequence while retaining the 5' splice site sequence that produces the IIA isoform (Fig. 10). In each case, nucleotide deletions or substitutions were generated using the QuikChange™ site-directed mutagenesis kit (Stratagene). Briefly, a complementary primer pair (purified by SDS-PAGE; Invitrogen) devoid of the nucleotide sequence of interest or containing the desired nucleotide substitution was added to approximately 20ng of substrate DNA. Substrate DNA was prepared by sub-cloning a wild-type 2.6kb fragment of the COL2A1 mini-gene into pSP73 vector (Promega) using BamHI and EcoRV restriction sites. This 2.6kb fragment contains the specific region of exon 2 and intron 2 that was to be mutated. By doing this, the construct size was reduced from ~11.3kb (size of COL2A1 mini-gene in pcDNA3) to ~5kb (the total size of pSP73 containing the 2.6kb COL2A1 mini-gene fragment) in order to increase the efficiency of the *in vitro* mutagenesis procedure. PCR mutagenesis was carried out over 18 cycles [95°C, 30 s; 55°C, 1 min; 68°C, 5.30 min] and resulting PCR products were digested with DpnI (1μl) for 1 h at 37°C to digest parental, methylated DNA. An aliquot (1μl) 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 2.6kb fragment was re-ligated back into the COL2A1-pcDNA3 construct to create the mutant COL2A1 mini-gene. Transfection of these mutant mini-genes and detection of spliced mRNA isoforms was done as described previously (McAlinden et al. 2005).

4.8 Detection of spliced mRNA isoforms derived from the COL2A mini-gene

Total RNA was harvested from each cell line 48 h after transfection using the QIAGEN RNeasy kit. Approximately 1μg of RNA was reverse transcribed using random primers in a final volume of 20μl and resulting cDNA was diluted to 80μl with sterile water. 10μl of cDNA was used for semi-quantitative PCR in the presence of [$\gamma\text{-}^{32}\text{P}$]dCTP (10mCi/ml, 3000Ci/mmol; Amersham Biosciences). The primers, pcDNA3-COL2A1-Exon1 (5'-CAAGCTTACATGATTC-GC-3')

and sp6 (termed P1 and P2 in Fig. 8) were used to amplify cDNA derived only from the exogenously-transfected COL2A1 mini-gene. The linear range for amplification using these primers was established, and PCR was carried out for 20 cycles: 95°C, 30s; 55°C, 30s; 72°C, 30s. To each reaction was added 10 µl of 6x loading dye [30% glycerol, 0.025% (w/v) bromophenol blue, 0.025% (w/v) cyanol blue] and 7µl was electrophoresed at 200V through 6% polyacrylamide gels. Gels were dried and exposed to PhosphorImaging screens for 1 h and then scanned on the STORM™ 840 PhosphorImager (Amersham Biosciences). Bands corresponding to the type IIA (386bp) and IIB (179bp) mRNA isoforms were viewed and quantified using ImageQuant™ software whenever necessary.

Acknowledgements

We thank Diane Kocka, Bharath Surapaneni, and Tru Huynh for their outstanding technical contributions. This work was supported by National Institutes of Health grants AR 46196 (TMH), HD049834 (TMH), AR44390 (BJ), and AR036994 (LJS/AM). Dr. Audrey McAlinden was the recipient of an Arthritis Investigator Award from the Arthritis Foundation.

References

- Ala-Kokko L, Kvist AP, Metsaranta M, Kivirikko KI, de Crombrughe B, Prockop DJ, Vuorio E. 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):923–929. [PubMed: 8948452]
- Aquilina A, Korda M, Bergelson JM, Humphries MJ, Farndale RW, Tuckwell D. A novel gain-of-function mutation of the integrin alpha2 VWFA domain. *Eur J Biochem* 2002;269:1136–1144. [PubMed: 11856343]
- Atkin AL, Altamura N, Leeds P, Culbertson MR. The majority of yeast UPF1 co-localizes with polyribosomes in the cytoplasm. *Mol Biol Cell* 1995;6:611–625. [PubMed: 7545033]
- Ausubel, FM.; Brent, R.; Kingston, RE.; Moore, DD.; Seidman, JG.; Smith, JA.; Struhl, D., editors. *Current Protocols in Molecular Biology*. John Wiley and Sons, Inc; New York: 1991.
- Baek D, Green P. Sequence conservation, relative isoform frequencies, and nonsense-mediated decay in evolutionarily conserved alternative splicing. *Proc Natl Acad Sci U S A* 2005;102:12813–12818. [PubMed: 16123126]
- Buhler M, Wilkinson MF, Muhlemann O. Intranuclear degradation of nonsense codon-containing mRNA. *EMBO Rep* 2002;3:646–651. [PubMed: 12101097]
- Burge, CB.; Tuschl, T.; Sharp, PA. Splicing of Precursors to mRNAs by the Spliceosomes. Cech, T., editor. *The RNA World*, Cold Spring Harbor Press; Cold Spring Harbor: 1999. p. 525–560.
- Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR. ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Res* 2003;31:3568–3571. [PubMed: 12824367]
- Frischmeyer PA, Dietz HC. Nonsense-mediated mRNA decay in health and disease. *Hum Mol Genet* 1999;8:1893–1900. [PubMed: 10469842]
- Garcia Abreu J, Coffinier C, Larrain J, Oelgeschlager M, De Robertis EM. Chordin-like CR domains and the regulation of evolutionarily conserved extracellular signaling systems. *Gene* 2002;287:39–47. [PubMed: 11992721]
- Hering TM, Kazmi NH, Huynh TD, Kollar J, Xu L, Hunyady AB, Johnstone B. Characterization and chondrocyte differentiation stage-specific expression of KRAB zinc-finger protein gene ZNF470. *Exp Cell Res* 2004;299:137–147. [PubMed: 15302581]
- Hering TM, Kollar J, Huynh TD, Varelas JB, Sandell LJ. Modulation of extracellular matrix gene expression in bovine high density chondrocyte cultures by ascorbic acid and enzymatic resuspension. *Arch Biochem Biophys* 1994;314:90–98. [PubMed: 7944410]
- Hyvonen MT, Uimari A, Keinänen TA, Heikkinen S, Pellinen R, Wahlfors T, Korhonen A, Narvanen A, Wahlfors J, Alhonen L, Janne J. Polyamine-regulated unproductive splicing and translation of spermidine/spermine N1-acetyltransferase. *RNA* 2006;12:1569–1582. [PubMed: 16809818]

- Ibrahimi OA, Zhang F, Eliseenkova AV, Linhardt RJ, Mohammadi M. Proline to arginine mutations in FGF receptors 1 and 3 result in Pfeiffer and Muenke craniosynostosis syndromes through enhancement of FGF binding affinity. *Hum Mol Genet* 2004;13:69–78. [PubMed: 14613973]
- Johnstone B, Hering TM, Caplan AI, Goldberg VM, Yoo JU. In vitro chondrogenesis of bone marrow-derived mesenchymal progenitor cells. *Exp Cell Res* 1998;238:265–272. [PubMed: 9457080]
- Jones RB, Wang F, Luo Y, Yu C, Jin C, Suzuki T, Kan M, McKeehan WL. The nonsense-mediated decay pathway and mutually exclusive expression of alternatively spliced FGFR2IIIb and -IIIc mRNAs. *J Biol Chem* 2001;276:4158–4167. [PubMed: 11042206]
- Lamba JK, Adachi M, Sun D, Tammur J, Schuetz EG, Allikmets R, Schuetz JD. Nonsense mediated decay downregulates conserved alternatively spliced ABCC4 transcripts bearing nonsense codons. *Hum Mol Genet* 2003;12:99–109. [PubMed: 12499391]
- Lareau LF, Green RE, Bhatnagar RS, Brenner SE. The evolving roles of alternative splicing. *Curr Opin Struct Biol* 2004;14:273–282. [PubMed: 15193306]
- 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;127:821–830. [PubMed: 10648240]
- Le Guiner C, Gesnel MC, Breathnach R. TIA-1 or TIAR is required for DT40 cell viability. *J Biol Chem* 2003;278:10465–10476. [PubMed: 12533540]
- Lejeune F, Maquat LE. Mechanistic links between nonsense-mediated mRNA decay and pre-mRNA splicing in mammalian cells. *Curr Opin Cell Biol* 2005;17:309–315. [PubMed: 15901502]
- Lewis BP, Green RE, Brenner SE. Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proc Natl Acad Sci U S A* 2003;100:189–192. [PubMed: 12502788]
- 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;203:198–211. [PubMed: 7655082]
- 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;280:32700–32711. [PubMed: 16076844]
- Modrek B, Lee C. A genomic view of alternative splicing. *Nat Genet* 2002;30:13–19. [PubMed: 11753382]
- Morrison M, Harris KS, Roth MB. smg mutants affect the expression of alternatively spliced SR protein mRNAs in *Caenorhabditis elegans*. *Proc Natl Acad Sci U S A* 1997;94:9782–9785. [PubMed: 9275202]
- Nalin AM, Greenlee TK Jr, Sandell LJ. Collagen gene expression during development of avian synovial joints: transient expression of types II and XI collagen genes in the joint capsule. *Dev Dyn* 1995;203:352–362. [PubMed: 8589432]
- Nemeth GG, Heydemann A, Bolander ME. Isolation and analysis of ribonucleic acids from skeletal tissues. *Anal Biochem* 1989;183:301–304. [PubMed: 2483036]
- 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;159:403–417. [PubMed: 8405667]
- Oganesian A, Zhu Y, Sandell LJ. Type IIA procollagen amino propeptide is localized in human embryonic tissues. *J Histochem Cytochem* 1997;45:1469–1480. [PubMed: 9358849]
- Pan Q, Saltzman AL, Kim YK, Misquitta C, Shai O, Maquat LE, Frey BJ, Blencowe BJ. Quantitative microarray profiling provides evidence against widespread coupling of alternative splicing with nonsense-mediated mRNA decay to control gene expression. *Genes Dev* 2006;20:153–158. [PubMed: 16418482]
- 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;265:10334–10339. [PubMed: 2355003]
- 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;114:1307–1319. [PubMed: 1894696]

- 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;199:129–140. [PubMed: 8204907]
- Separovic E, Nadin-Davis SA. Secondary structure within PCR target sequences may facilitate heteroduplex production. *PCR Meth Appl* 1994;3:248–251.
- Stoilov P, Meshorer E, Gencheva M, Glick D, Soreq H, Stamm S. Defects in pre-mRNA processing as causes of and predisposition to diseases. *DNA Cell Biol* 2002;21:803–818. [PubMed: 12489991]
- Su MW, Lee B, Ramirez F, Machado M, Horton W. Nucleotide sequence of the full length cDNA encoding for human type II procollagen. *Nucleic Acids Res* 1989;17:22–9473.
- Sureau A, Gattoni R, Dooghe Y, Stevenin J, Soret J. SC35 autoregulates its expression by promoting splicing events that destabilize its mRNAs. *EMBO J* 2001;20:1785–1796. [PubMed: 11285241]
- Urabe K, Kim HJ, Sarkar G, Bronk JT, Bolander ME. Determination of the complete cDNA sequence of rat type II collagen and evaluation of distinct expression patterns of types IIA and IIB procollagen mRNAs during fracture repair in rats. *J Orthop Sci* 2003;8:585–590. [PubMed: 12898314]
- Wilson GM, Sun Y, Sellers J, Lu H, Penkar N, Dillard G, Brewer G. Regulation of AUF1 expression via conserved alternatively spliced elements in the 3' untranslated region. *Mol Cell Biol* 1999;19:4056–4064. [PubMed: 10330146]
- Wollerton MC, Gooding C, Wagner EJ, Garcia-Blanco MA, Smith CW. Autoregulation of polypyrimidine tract binding protein by alternative splicing leading to nonsense-mediated decay. *Mol Cell* 2004;13:91–100. [PubMed: 14731397]
- Yoo JU, Barthel TS, Nishimura K, Solchaga L, Caplan AI, Goldberg VM, Johnstone B. The chondrogenic potential of human bone-marrow-derived mesenchymal progenitor cells. *J Bone Joint Surg Am* 1998;80:1745–1757. [PubMed: 9875932]
- Zheng H, Martin JA, Duwayri Y, Falcon G, Buckwalter JA. Impact of aging on rat bone marrow-derived stem cell chondrogenesis. *J Gerontol A Biol Sci Med Sci* 2007;62:136–148. [PubMed: 17339639]
- 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;144:1069–1080. [PubMed: 10085302]
- 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;144:1069–1080. [PubMed: 10085302]

The abbreviations used are

NMD	nonsense mediated decay
MSC	mesenchymal stem cells
COL2A1	collagen $\alpha 1$ (II) gene
RT-PCR	reverse transcriptase-polymerase chain reaction
RUST	regulated unproductive splicing and translation
EST	expressed sequence tag
PTC	premature termination codon
Pre-mRNA	

CR primary mRNA transcript
cysteine-rich domain

Fig. 1.

Alternatively spliced isoforms of the type II procollagen gene (COL2A1). Exon structure of four mRNA COL2A1 spliced isoforms identified in the present study. Two novel isoforms were found, named IIC and IID. The IIC isoform contains a 34 nucleotide-long, truncated fragment of exon 2 (2C) while isoform IID contains an extra 3 nucleotides (GTT in human; GTC in rabbit) at the end of exon 2 (2D). PCR primers used to amplify fragments for Southern analysis and cloning are indicated by arrows. Location of probes to identify specific isoforms by Southern hybridization are indicated by bars.

Fig. 2.

COL2A1 gene expression and Southern blot analysis of COL2A1 exon 1-7 PCR products during in-vitro chondrogenesis of rabbit and human MSCs. Northern blot analyses of rabbit and human COL2A1 expression were performed, and blots were quantified by densitometry of radiographic films (A and G, respectively). The same RNA samples were used for RT-PCR spanning exon 1 to 7, and PCR products were identified by Southern blotting (rabbit, C-F and human, I-L). Samples and 123 bp ladder size markers were separated on 1% agarose gels, and stained with EtBr (gel B and gel H). A blot of rabbit PCR products was hybridized sequentially with oligonucleotide probes specific for all four spliced forms (ABCD(R)), and for splice variant D(R), B(R), and C(R) (in panels C-F). Human PCR products were hybridized with

probes for all four spliced forms (ABCD(H)) and for splice variant A(H), B(H), and C(H) (in panels I-L).

Fig. 3. Southern blot analysis of COL2A1 exon 2-6 PCR products. Human MSC chondrogenesis time course cDNAs were amplified using primers spanning exons 2-6, resulting in PCR products derived only from the IIA, IID and IIC isoforms. Samples and size markers were electrophoresed on a 1% agarose gel which was stained with EtBr (A), and following membrane transfer was hybridized sequentially with human (H)-specific oligonucleotide probes for splice variants ABCD(H), A(H), and C(H) (panels B-D). Arrows indicate COL2A1 IIA and/or IID, and IIC rtPCR products. Band labeled with asterisk in (C) is unidentified. PCR products from the day 14 sample were used to clone and sequence the human IIC isoform.

Fig. 4.

Anomalous migration of IIA and IID PCR products from mixed cDNA templates. PCR amplification was carried out using the indicated mixtures (copies $\times 10^{-9}$) of plasmids containing either IIA or IID insert cDNA (lanes 2-6). Samples were electrophoresed on a 2.0% agarose gel, and DNA was visualized using EtBr. Size (bp) of DNA markers in lane 1 and lane 7 are shown.

Fig. 5.

Alternative COL2A1 5' splice sites. Location and sequence of three 5' splice sites utilized to produce the IIA, IIC or IID mRNA isoform of COL2A1 (A). Uppercase letters represent exon sequences while lowercase letters correspond to intronic sequences. Arrow denotes the site of cleavage by the spliceosome complex. The alternative splice sites within exon 2 are conserved between species (B). Extent of exons 2C, 2A and 2D are indicated by the arrows. Single asterisk indicates 5' terminus of intron resulting from utilization of the IIC splice site. Double asterisk indicates the additional three nucleotides at the 3' end of exon 2D. Unshaded residues are not completely conserved between indicated species. Genbank accession numbers for indicated sequences are as follows: *H. sapiens* (1) with exon 2D, (T. Hering, this work); *H. sapiens* (2) with exon 2A, NM_001844 (also see BU552941); *O. cuniculus* with exon 2D AF027122; *C. familiaris*, AF023169; *R. norvegicus*, L48440 plus additional exon 2 sequence (204 bp) from (Urabe et al. 2003); *M. musculus* with exon 2D, NM_031163 (*M. musculus* 2A sequence is also found in BY732248, BQ568641, W89304); *G. gallus*, M74435; *X. laevis*, BC048221; *D. rerio*, BC 059180.

Fig. 6.

Quantitative real-time PCR analysis of COL2A1 isoform expression during in vitro chondrogenesis of human MSC pellet cultures. Percentage of total COL2A1 mRNA contributed by each of the A, B and D isoforms were determined by absolute quantification at each time point. The total height of each stacked bar is the quantity of total ABCD mRNA relative to the calibrator value defined as 1.0 at day 0. The relative quantity of the IIC isoform was not above background at each time point using cDNA sufficient for analysis of the other isoforms.

Fig. 7.

Alternative splicing of a COL2A1 mini-gene. Mechanism of alternative splicing to remove intron 1 and intron 2 from COL2A1 pre-mRNA to produce the two major isoforms, IIA (+exon 2) or IIB (-exon 2) is shown. Numbers represent the nucleotide size of each exon and intron. This region of human genomic DNA consisting of exons 1-3 and full-length intervening introns 1 and 2 was amplified to synthesize a COL2A1 mini-gene as previously described (McAlinden et al. 2005). PCR using the primer pairs (P1 and P2) specific to amplify the exogenously-transfected mini-gene resulted in cDNA products of either 386 bp (IIA) or 179 bp (IIB). Figs. 9 and 10 shows results of splicing of wild type and specific mutant COL2A1 mini-genes relevant to the present study.

Fig. 8.

Alternative splicing of a COL2A1 mini-gene devoid of the IIC 5' alternative splice site in exon 2 (Δ E2-c5'ss). Wild type or deletion mutant COL2A1 mini-genes were transfected into different cell lines and resulting IIA and IIB mRNA isoforms were detected by radiolabeling RT-PCR products, gel electrophoresis and PhosphorImaging. The boxed region shows the 10 nucleotides that were omitted to create the deletion mutant. The last 4 nucleotides of the deleted region make up part of the IIC alternative 5' splice site. The site of cleavage by the spliceosome complex is shown by the arrow. Nucleotides shown in lowercase represent the remaining region of the IIC 5' splice site that was not deleted. The PhosphorImage shows an increase in IIA and a decrease in the IIB isoform due to deletion of the IIC splice site in all cells tested: HEK-293 (human embryonic kidney cells), MC615 (mouse vertebral chondrocyte cell line), T/C28 (human costal chondrocyte cell line) and RCS (rat chondrosarcoma cell line).

Fig. 9.

Alternative splicing of a mutant COL2A1 mini-gene devoid of the IID 5' splice site. Top panel shows the location and sequence of the wild type (WT) IIA or IID 5' splice site (5'ss) and the arrows show the site of exon-intron cleavage by the spliceosome complex. Sequences in underlined italics refer to the mutations introduced to synthesize the mutant mini-gene, named mIID. Here, the last two of the 6 intronic nucleotides (tt) that make up the IID 5' splice site sequence were mutated. Gel PhosphorImages show the radiolabeled IID, IIA and IIB transcripts derived from splicing of the WT or mIID mini-gene in all cell lines: HEK-293, MC615, T/C28 and RCS.

Table 1

PCR primers and hybridization probes for amplification and identification of alternatively spliced COL2A1cDNA fragments

Oligo	Sequence
1(H)	5'-CTGCTCGTCGCCGCTGTCCTT-3'
2(H)	5'-AAGGGTCCCAGGTTCTCCATC-3'
3(H)	5'-GGCTGGCAGCTGTGTGCAGGA-3'
4(H)	5'-CACCTTTGTCACCACGATCCC-3'
B(R)	5'-GTCCTGGTTTCCGGACATCCTGGC-3'
C(R)	5'-GTCCTGGTTTCC*TCTTCCCCTCCT-3'
D(R)	5'-CTGGTTTCCGACCACTGGCAGTGG-3'
ABCD(R)	5'-CCCAGGGGGTCCTTTAGGTCC-3'
A(H)	5'-GTCCTGGTTGCCCACTGGGAGTGG-3'
B(H)	5'-GTCCTGGTTGCCGGACATCCTGGC-3'
C(H)	5'-GTCCTGGTTGCCCTCTGCCCATCC-3'
ABCD(H)	5'-CCCAGGAGGTCCTTTGGGTCC-3'

Table 2

Oligonucleotides for Real-Time PCR; Taqman primers and probes designed to span splice junctions unique to COL2A1 variants IIA, IIB, IIC, and IID

Assay	Forward Primer	Probe Oligo	Reverse Primer
IIA	5'-CCCCATCTGCCCAACTGA-3'	5'-CCAGTGGGCAACC-3'	5'-GTTCTCCTTTCTGTCCCTTTGGT-3'
IIB	5'-CGCTGTCCTTCGGTGTCA-3'	5'-CTGGTTGCCGGACATC-3'	5'-GTTCTCCTTTCTGTCCCTTTGGT-3'
IIC	5'-GAGGCTGGCAGCTGTGT-3'	5'-TTGCCCTCTGCCCATCC-3'	5'-GTTCTCCTTTCTGTCCCTTTGGT-3'
IID	5'-CCCCATCTGCCCAACTGA-3'	5'-CCTGGTTGCCAACCAC-3'	5'-GTTCTCCTTTCTGTCCCTTTGGT-3'