Protein & Cell

REVIEW

Minor fibrillar collagens, variable regions alternative splicing, intrinsic disorder, and tyrosine sulfation

Ming Fang[1,3], Reed Jacob[2,3], Owen McDougal[2,3], Julia Thom Oxford[1,3] ✉

[1] Department of Biological Sciences, Boise State University, Boise, ID 83725, USA
[2] Department of Chemistry and Biochemistry, Boise State University, Boise, ID 83725, USA
[3] Biomolecular Research Center, Boise State University, Boise, ID 83725, USA
✉ Correspondence: joxford@boisestate.edu

ABSTRACT

Minor fibrillar collagen types V and XI, are those less abundant than the fibrillar collagen types I, II and III. The alpha chains share a high degree of similarity with respect to protein sequence in all domains except the variable region. Genomic variation and, in some cases, extensive alternative splicing contribute to the unique sequence characteristics of the variable region. While unique expression patterns in tissues exist, the functions and biological relevance of the variable regions have not been elucidated. In this review, we summarize the existing knowledge about expression patterns and biological functions of the collagen types V and XI alpha chains. Analysis of biochemical similarities among the peptides encoded by each exon of the variable region suggests the potential for a shared function. The alternative splicing, conservation of biochemical characteristics in light of low sequence conservation, and evidence for intrinsic disorder, suggest modulation of binding events between the surface of collagen fibrils and surrounding extracellular molecules as a shared function.

KEYWORDS   minor fibrillar collagens, variable regions, alternative splicing, fibrillogenesis, heparan sulfate binding sites, intrinsic disorder, tyrosine sulfation

INTRODUCTION

Collagens are the predominant proteins of the extracellular matrix in animals. They play diverse roles in cell behavior, cell adhesion, differentiation, cell migration, and in tissue repair. To date, there are 29 different collagens (Kadler et al., 2007; Söderhäll et al., 2007), each of which is composed of three polypeptide alpha chains forming either heterotrimers or homotrimers. Depending on their functions and domain homology, collagens can be grouped into fibril-forming, fibril associated collagens with interrupted triple helices (FACITs), network forming, beaded-filament-forming, anchoring, or transmembrane collagens (Kadler et al., 2007). Collagen types I, II, III, V and XI as well as the recently discovered XXIV and XXVII, are fibril-forming collagens. Historically, the fibrillar collagens are further divided into major (I, II and III) and minor (V and XI) fibrillar collagens based on their abundance in the collagen fibrils.

The importance of minor fibrillar collagens has been shown by their profound influence in connective tissue formation and function. For example, mutations in collagen type V genes can cause Ehlers-Danlos syndrome with joint hypermobility (Michalickova et al., 1998; Richards et al., 1998; Schwarze et al., 2000; Wenstrup et al., 2000), whereas mutations in collagen type XI genes can cause Marshall and Stickler syndromes, each resulting in altered facial appearance, eye abnormalities, joint alterations and hearing loss (Sirko- Osadsa et al., 1998; Pihlajamaa et al., 1998; Annunen et al., 1999; McGuirt et al., 1999; Melkoniemi et al., 2000). Mutations in Col11a1 have recently been shown to give rise to fibrochondrogenesis, a severe autosomal-recessive, short-limbed skeletal dysplasia (Tompson et al., 2010). Three different collagen type V alpha chains have been described in humans: Col5a1, Col5a2 and Col5a3. Similarly, in collagen type XI, there are three alpha chains, Col11a1, Col11a2 and

Col11a3. However, Col11a3 is actually Col2a1 (alpha1 chain, type II), one of the major fibril-forming collagens. Minor fibril-forming collagens are composed of several domains, including the C-propeptide (NC1), a major helix (COL1) and minor helix (COL2) separated by a short sequence (NC2), and the N-terminal domain (NC3). The NC1 domain is well conserved among fibril-forming collagens, whereas the NC3 domain varies in size, primary structure and subdomain compositions (Ricard-Blum and Ruggiero, 2005). In the minor fibrillar collagens, NC3 contains either a cysteine-rich repeat domain (CRR)/von Willebrand factor type C domain, characterized by the presence of 10 cysteines (Col5a2 and Col11a3), or a thrombospondin N-terminal-like (TSPN) domain adjacent to a variable region (VR) (Col5a1, Col5a3, Col11a1 and Col11a2). Despite the differences in NC3 domains, Col5a2 shares similar expression patterns with other minor fibrillar collagen alpha chains with which it associates. At early development stages, minor fibrillar collagens are expressed in a variety of different tissues; whereas at later development stages, more restricted skeleton related expression patterns are observed (Andrikopoulos et al., 1992; Oxford et al., 1995; Tsumaki and Kimura, 1995; Yoshioka et al., 1995b; Lui et al., 1996; Sugimoto et al., 1998; Imamura et al., 2000; Roulet et al., 2007). Both Col24a1 and Col27a1, exhibit TSPN and VR domains common to collagen types V and XI; however, they differ from the other minor fibrillar collagen alpha chains with respect to interruptions within the major helical region (Ricard-Blum and Ruggiero, 2005).

Type V and XI collagens form heterotypic collagen fibrils by cooperating with type I and II collagens, respectively (Ricard-Blum and Ruggiero, 2005, Fernandes et al., 2007). Additional complexity arises due to the fact that the alpha chains of collagen types V and XI can also form heterotrimers with each other in a tissue dependent manner (Fichard et al., 1995). Therefore, it has been proposed that type V and XI collagens are not actually distinct collagens, but rather belong to the same family, collagen type V/XI (Fichard et al., 1995). In agreement with this and the structural differences among fibrillar collagens, the alpha chains of fibrillar collagens are classified into three distinct subclasses, in which Col5a1, Col5a3, Col11a1 and Col11a2 belong to the same clade (Zhang et al., 2007). In this review, we focus on the minor fibril-forming collagens with specific emphasis on the variable regions of Col5a1, Col5a3, Col11a1 and Col11a2.

Alternative splicing in the VR

Alternative splicing (AS) is a commonly utilized pathway for higher organisms to enrich the functional proteome without increasing the number of genes (Stamm et al., 2005; Romero et al., 2006; Ben-Dov et al., 2008). It is a process mediated by the spliceosome composed of ribonucleoproteins (RNPs) (Wahl et al., 2009). Recent studies indicate that genes undergoing AS often function in cell differentiation and devel-

opment. The outcomes of AS are either the addition/deletion of a new protein domain or the introduction of a stop codon. Consequently, this would lead to a potential change in the function of the protein (Stamm et al., 2005). Historically, the VR of the collagen molecule is defined as a region of low homology among different collagen alpha chains. Collagen type XI alpha chain genes exhibit extensive AS in the VR in human, rat (Oxford et al., 1995; Tsumaki and Kimura, 1995; Lui et al., 1996) and other characterized species, indicating a possible role in cell differentiation and development. Although no evidence exists for AS in the genes encoding collagen type V alpha chains in any species reported, preliminary data from our laboratory indicates that AS does take place within the paralogs of zebrafish type V alpha chains. The alternatively spliced VRs are located from exon 6 to exon 8 for both *Col11a1* and *Col11a2* (Oxford et al., 1995; Tsumaki and Kimura, 1995; Lui et al., 1996). In Col11a1, there are two different forms of exon 6, referred to as exon 6A and exon 6B, which encode amino acids with distinct negative and positive charges under physiological conditions, respectively (Oxford et al., 1995). Although different combinations of exons in the variable regions of *Col11a1* and *Col11a2* could theoretically result in many splicing-forms, the most abundant detected so far in Col11a1 are the combinations exon 6A-7-8, exon 6B-7, and exon 7 alone (Oxford et al., 1995; Morris et al., 2000); in the case of Col11a2, the combinations exon 6, exon 7 and exon 8 can be either included or excluded (Tsumaki and Kimura, 1995; Lui et al., 1996). The most abundantly detected form of Col11a2 results from the exclusion of exons 6, 7, and 8 at all times other than early development. Sequence comparison between human, chimp, mouse, rat, chicken and zebrafish shows some conserved sequence for all of the exons encoding the VR in Col5a1, Col5a3, Col11a1 and Col11a2 (Fig. 1). However, the degree of conservation is relatively low in comparison to other regions of the same proteins. Supplemental Fig. 1 (Fig. S1) shows sequence alignment; and accession numbers are indicated in supplemental Table 1 (Table S1).

Isoform-dependent expression patterns

Much attention has been focused on the temporal and spatial expression patterns for minor fibrillar collagens (Table 1). In early developmental stages of the mouse embryo, *Col5a1* expression is observed in the aorta, heart, branchial arches, developing mesenchyme and neuroepithelium; late in development, its expression is limited to bones, vertebral column, cornea, tendons and ligaments (Imamura et al., 2000; Roulet et al., 2007). For *Col5a2*, early expression is found in the peritoneal membranes, intestinal and craniofacial mesenchyme. At later stages, the expression is restricted to some ossification centers, joints, tendons, the atrioventricular valve of the heart, and head structures including mandibular arch, optic nerve region, otic capsule and calvarial mesenchyme
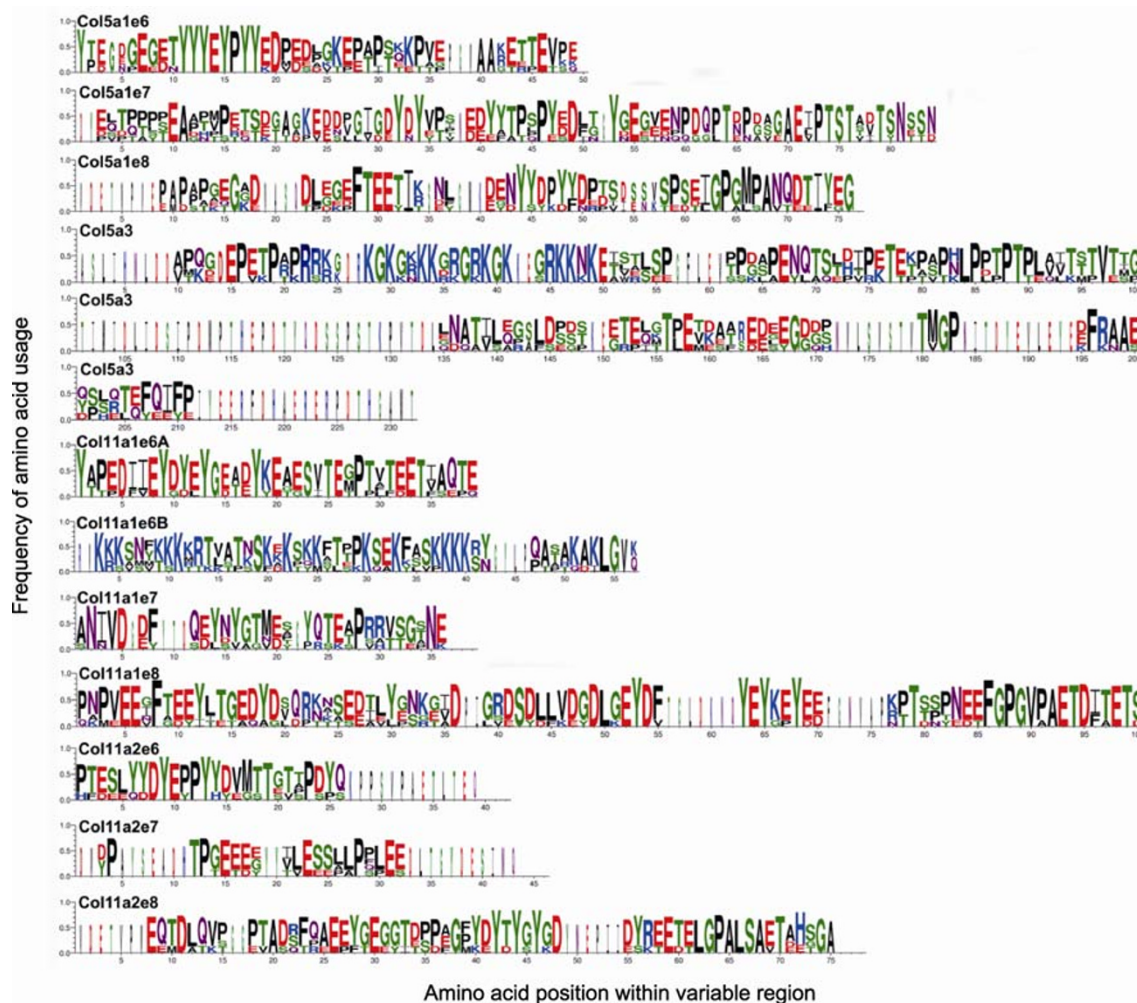
Figure 1. Conservation in the variable regions of Col5a1, Col5a3, Col11a1 and Col11a2 among human, chimpanzee, mouse, rat, chicken and zebrafish. Alignments of peptides encoded by exons 6, 7 and 8 of *Col5a1*, *Col5a3*, *Col11a1* and *Col11a2*, were carried out and are displayed as LOGOS to highlight the conserved sequences. Negatively charged glutamic acid and aspartic acid (D and E) are shown in red, glycine, tyrosine, serine, threonine (G, Y, S and T) are shown in green, proline, alanine, leucine, methionine, valine, phenylalanine, isoleucine (P, A, L, M, V, F, and I) are shown in black, glutamine and asparagines (Q and N) are shown in purple, and lysine and arginines (K and R) are shown in blue. Relative height of single letter abbreviation in each sequence position indicates the relative frequency of amino acids used among the six species compared. Exon-intron boundaries in zebrafish sequence are not well defined based on current annotation of *Col5a3* and therefore, sequence encoded by *Col5a3* VR was compared in total. Exon assignments were made according to the human sequences, and sequences from other species are in agreement with the human in most cases. However, for *Col5a1*, exons 4, 5 and 6 were used from chicken for the alignment of Col5a1e6, Col5a1e7 and Col5a1e8, respectively; whereas for Chimp, exons 7, 8 and 9 were used in the same alignment. Alignment is shown in supplemental Fig. 1. Accession numbers are given in the supplemental Table 1.

(Andrikopoulos et al., 1995). Additionally, the expression of *Col5a2* has also been reported in the bovine vitreous (Mayne et al., 1993). The expression of *Col5a3* has been shown as a subset of that of *Col5a1* (Imamura et al., 2000). The expression of *Col11a1* tends to be broader than that of the collagen type V genes discussed above. *Col11a1* has been detected in limbs, vertebrae, mandibular bones, otic vesicle, the atrioventricular valve of the heart, neuroepithelium of the brain, liver, kidney, lung, muscle and intestine (Lui et al., 1995; Oxford et al., 1995;

Yoshioka et al., 1995b). Additional expression sites for *Col11a1* include the mesothelial layer of the cornea, and perioptic mesenchyme within the sclera. Compared to the expression of *Col11a1*, *Col11a2* has a more restricted, cartilage-related expression pattern in developing limbs and the axial skeleton (Sugimoto et al., 1998). Common localization patterns of the alpha chains of collagen types V and XI in mesenchymal tissues at the early development stages indicate the possibility of the formation of hybrid V/XI triple helical molecules.

Table 1 Summary of the spatiotemporal expression patterns of the minor fibrillar collagens

Gene	7.5	8.5–9.5	10.5	11.5	12.5–13.5	14.5–15.5	16.5–18.5
Col5a1	Not detectable	Dorsal aorta; heart	Nerves; blood vessels; neuroepithelium; trabeculated wall of the common ventricular chamber of the heart; dorsal aorta wall; mesonephrotic tubules; branchial arches; mesenchyme	Developing mesenchyme (cephalic mesenchyme and the trunk region around somites)	Vertebral column; craniofacial tissues	Perichondrium; corneal stroma; tendon	Developing skin; odontoldasts; vertebrae; Vertebrae; mandibular arch
Col5a2					Peritoneal membranes; intestinal mesenchyme craniofacial mesenchyme	Intestinal mesenchyme; Primary ossified regions; perichondrium; joints; tendon; diaphragm; peritoneal membrane; atrioventricular valve; mandibular arch; tongue; optic nerve region; otic capsule; calvarial mesenchyme	
Col5a3	Mammary glands; placenta; uterus; heart; lung; brain			Superficial fascia; epimysia; lining on the developing muscles; epineureum of some nerves			
Col11a1			Notochord; sclerotome of the somites	Sclerotome-derived condensations; notochord Nasal; mesenchymal condensation; limbs	All cartilaginous tissues	Mesothelial layer of the cornea; condensation of perioptic mesenchyme within sclera; calvarium; limbs; tendons	
Col11a2			Notochord sclerotome of the somites	Sclerotome-derived condensations; notochord Nasal; mesenchymal condensation; limbs (isoforms w/o exon 6–8)	All cartilaginous tissues (isoforms w/o exon 6–8)	calvarium; limbs (isoforms w/exon 6–8)	

As a result of AS, *Col11a1* and *Col11a2* have several different isoforms that demonstrate differential expression patterns within the expression sites of total *Col11a1* and *Col11a2*. For *Col11a1*, the predominantly expressed isoforms in non-cartilaginous tissues including brain, muscle, calvaria, heart, kidney, liver and lung, are exon 6A–7–8 and exon 7–8. In contrast, in cartilaginous tissues, the expression patterns are more complicated (Oxford et al., 1995). The localization of Col11a1 isoforms in cartilage-derived tissues was investigated in the rat humerus, where exon 6B–7 was found to be restricted to the periphery of the cartilage covering the perichondrium of the diaphysis, and the expression of exon 7–8 was prevalent in prechondrogenic mesenchyme and immature cartilage (Morris et al., 2000). Moreover, the expression of exon 7–8 diminished as chondrocytes underwent hypertrophy, indicating a potential association with chondrocyte maturation. In the case of *Col11a2*, the expression of the isoform containing exons 6–8 is detected in non-chondrogenic tissues, including muscle, brain, heart, eye, periosteum and calvarial mesenchymal layer, but not in cartilaginous tissues (Lui et al., 1996; Sugimoto et al., 1998). Although *Col11a1* and *Col11a2* have different expression patterns, some cartilaginous tissues are enriched with VR domains lacking amino acids bearing a negative charge at physiological pH. In the case of the cartilage that is immediately adjacent to the perichondrium at locations of mineralization, the variable region is comprised of amino acids bearing a positive charge at physiological pH within the NC3 domain, as seen in the amino acid sequence encoded by exon 6B of Col11a1. Exons 6–7–8 of Col11a2 encode acidic sequences which are omitted from Col11a2 in cartilage.

Functions of the minor fibrillar collagens

Although the function of the minor fibrillar collagens are not fully understood, targeted mutations provide some insight. The first targeted minor fibrillar collagen was mouse *Col5a2*. The heterozygous mutant (exon 6 deletion) has an apparently normal phenotype, whereas the homozygous mutants feature spinal lordosis and kyphosis with a marginal survival rate (Andrikopoulos et al., 1995). The primary cause of death in the homozygous mutants is thought to be the effects of respiratory problems caused by spinal abnormalities, which result from the loss of tensile strength in the ligaments. Addi-

tionally, mutants also exhibit skin fragility and collapse of the corneal stroma caused by abnormal assembly of type I collagen fibrils (Andrikopoulos et al., 1995).

A truncated form of Col5a1 was used to demonstrate that alteration in Col5a1 leads to modification of the fibril morphology synthesized by chicken corneal fibroblasts (Marchant et al., 1996). Deletion of exons 3 and 4 of *Col5a1* in mice is associated with embryonic lethality and a lack of mesenchymal collagen fibrils (Wenstrup et al., 2004). Together, these studies further support the theory that type V collagen chains are critical in the regulation of fibrillogenesis (Birk, 2001). However, *in vitro* studies showed that a heterotrimer of collagen V is able to regulate the diameter of collagen I fibrils but not the heterotrimer, indicating that the regulation of the fibril diameter by collagen V occurs in a stoichiometric-dependent manner (Chanut-Delalande et al., 2001). Also, collagen V can form heterotrimers with collagen XI depending on the tissue content (Fichard et al., 1995). Therefore, the role of minor fibrillar collagens in fibril assembly may rely on the specific collagen alpha chains expressed in a given tissue.

The *cho/cho* mouse is a homozygous mutant of *Col11a1* (Li et al., 1995). These mice display autosomal recessive chondrodysplasia, and die at birth with a wide range of cartilage abnormalities. The phenotype is caused by a reading frame shift at +570 nucleotides resulting in a premature stop codon downstream from the deletion (Li et al., 1995). Although the phenotype of this mutant has been well documented (Li et al., 1995), the mechanism is not yet clear. Similar to type V collagen, one of the roles for Col11a1 is that it may be able to regulate the diameter of collagen fibrils. This is supported by evidence that thickened, shorter collagen fibrils have been observed in the *cho/cho* mouse (Li et al., 1995; Fernandes et al., 2007). Additionally, *in vitro* studies show that collagen XI is crucial for normal collagen type II fibril assembly (Blaschke et al., 2000). A model has been proposed in which the NC3 domain of Col11a1 acts as the regulator for the uniformity of thin collagen fibrils in developing cartilage (Fernandes et al., 2007; Gregory et al., 2000). A "10+4" model has been described for thin collagen fibrils containing collagen types II, IX and XI where collagen type XI is found within the core of four microfibrils surrounded by collagen type II. Collagen type IX is found associated with the surface of the heterotypic fibril (Holmes and Kadler, 2006). The model shown in Fig. 2 combines known details of the structure of the minor fibrillar collagen type XI within a heterotypic collagen fibril showing the NC3 domains on the surface of the fibril.

The VRs of minor fibrillar collagen alpha chains are predicted to be at the surface of the collagen fibril, thus contributing to the biochemical nature of the fibrillar surface. In this location, the NC3 domain can contribute to regulation of fibrillar growth by a mechanism that relies upon steric hindrance and/or molecular interactions with other constituents of the extracellular matrix. Predicted surface properties and potential interactions are enabled by the minor fibrillar collagen constituents at the surface of a collagen fibril (Zhu et al., 1999; Xu et al., 2005; Leitinger and Hohenester, 2007; Kadler et al., 2008). Worthy of mention, the expression of Col11a1 has also been detected in certain type of cancers (Fischer et al., 2001; Halsted et al., 2008), indicating a possible role for Col11a1 in carcinogenesis and metastasis.

A similar approach has been used to study the function of Col11a2, producing a mutant (deletion of exons 27-28; premature stop codon in the insertion) resulting in a milder phenotype compared to the c*ho/cho* mouse. However, the mutant does exhibit some phenotypic traits of nonocular Stickler syndrome and otospondylomegaepiphyseal dysplasia (OSMED) in humans, including deafness and a receding snout in the mouse, with disorganized chondrocytes in the growth plates of long bones also observed (Li et al., 2001).

Although different mutants produce diverse phenotypes, homozygous mutants of both *Col5a1* and *Col11a1* show more severe abnormalities leading to neonatal or embryonic death compared to those of *Col5a2* and *Col11a2*, suggesting the crucial roles of these two genes in early development. Dose-dependency is indicated by heterozygous mutants, shedding some light on the mechanism. Dermal fibroblasts from patients with Ehlers-Danlos syndrome with haploinsufficiency of *Col5a1* exhibit decreased Col1a1 fibrillar assembly in the extracellular matrix (Wenstrup et al., 2004). In mutants of *Col5a2*, the collagen composition of the cornea shows larger fibrillar diameters compared to wild type, exhibiting a dose-dependent relationship with respect to the genetic composition of *Col5a2* (Andrikopoulos et al., 1995).

The relationship of the *cho/+* heterozygous mutant to the progression of osteoarthritis has also been explored and found to be caused by an increased interaction between Col2a1 and discoidin domain receptor 2 (DDR2) (Xu et al., 2005). DDR2 belongs to receptor tyrosine kinases (RTKs), functioning as a collagen receptor independent of β1 integrins (Leitinger and Hohenester, 2007). It is thought that low levels of Col11a1 cause an altered cartilage collagen fibril structure leading to over-exposure of Col2a1 to DDR2, which, in turn, up-regulates the expression of matrix metalloproteinase 13 (MMP-13) and hence, degeneration of cartilage that is associated with disease (Xu et al., 2005). This indicates that the functions of collagens may rely on the other molecules that interact with them, within the extracellular matrix and on the cell surface. A recent review summarized the collagen receptors, including integrin, DDR1 and DDR2, and glycoprotein VI among others (Leitinger and Hohenester, 2007).

To date, most observations of the effect of mutations have been made either at the time of death or after birth using mouse models. Limited information is currently available to address the potential roles that minor fibrillar collagens play during early embryonic development. Recently, transient expression of Col5a1 has been detected in mouse epidermal cells during early development (Roulet et al., 2007). Thus,
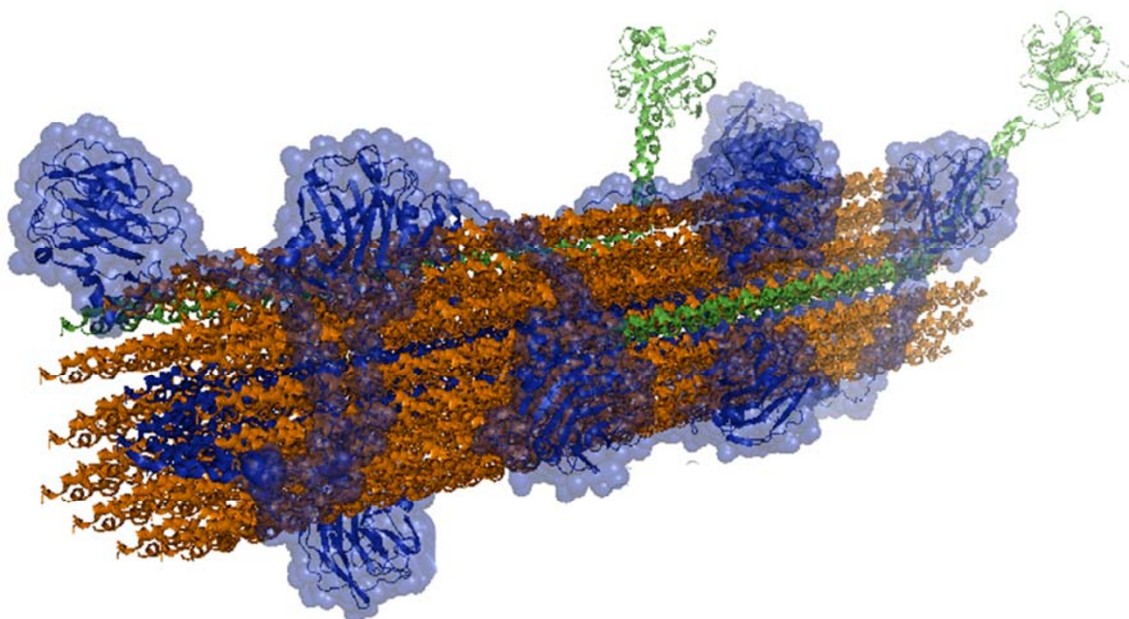
Figure 2.   Model of heterotypic thin collagen fibril. The minor fibrillar collagen type XI (blue) is shown in the core of the fibril with the major fibrillar collagen type II (orange) making up the bulk of the fibril. The fibril associated collagen with interrupted triple helix (FACIT) collagen type IX (green) is shown associated with the surface of the collagen fibril. The structure of collagen type II was created using the PyMOL Molecular Graphics System, Version 1.3 (Schrödinger, LLC). The structure of ColXI was created as a homology model built from the ColIX (2UUR) structure template (McDougal et al., 2011).

careful re-examination of the minor fibrillar collagens is needed to understand the spatiotemporal expression patterns during early development, particularly in the established mutants.

The alternatively spliced isoforms in collagen type XI genes have not been studied *in vivo* or in transgenic mutant studies. Varied, overlapping or perhaps opposing roles are suggested by studies in which isoforms generated by AS have been analyzed. For example, *FGF8* (fibroblast growth factor 8) produces two evolutionarily conserved isoforms by AS: FGF8a and FGF8b. In Xenopus, FGF8b is a mesoderm inducer, whereas FGF8a is involved in posterior neural development (Fletcher et al., 2006). Another study in a mouse model system demonstrated a partial compensatory effect from FGF8a and FGF8b deficient embryos (Guo and Li, 2007), indicating overlapping functions of these two molecules.

Based on *in vitro* studies, different splice-forms of *Col11a1* have also been shown to influence the processing of Col11a1 itself. Acceleration of the processing of pro-α1(XI) by BMP-1 has been observed by the presence of either p6B (encoded by exon 6B) or p8, compared to p7 and p6A (Medeck et al., 2003). Either retention of or release of N-terminal propeptides could affect the gradients of morphogens that bind to the NC3 domain. Because the NC3 domain of Col11a1 is slowly processed, the expression of exon 6B and exon 8 may be

related to a higher availability of certain molecules that bind to p6B or p8 in tissues where they are expressed. Studies of collagen type V show an influence on cell behavior. The N-terminal domain of Col5a4 has been found to modulate Schwann cell migration by utilizing syndecan-3 (Erdman et al., 2002), which is a transmembrane heparan sulfate proteoglycan. Although not specified as to which alpha chain or which domain is involved, Luparello and colleagues demonstrated that type V collagen can induce apoptosis in a cell line-dependent manner (Pucci-Minafra et al., 2000; Luparello and Sirchia, 2005). Heparan sulfate (HS) binding sites have also been identified in the NC3 domain of Col5a3 exhibiting specific adherence to osteosarcoma cells (Yamaguchi et al., 2005). Interestingly, an HS binding site has been described in the spliced region of *Col11a1*, encoded by exon 6B, which is expressed in chondrocytes (Warner et al., 2006). Although exclusion of exon 6B would naturally abolish this binding site as demonstrated in the isoform containing exon 6A–7–8, its function is not clear. Collagen type XIV does not belong to the minor fibrillar collagens; however, its alternatively spliced F3 domain behaves similarly to exon 6B in Col11a1. It has been shown that the F3 domain of Col14a1 has both heparin binding and cell binding properties (Imhof and Trueb, 2001). Thus, it is possible that the inclusion of exon 6B in *Col11a1* may have a similar function in cell surface association.

© Higher Education Press and Springer-Verlag Berlin Heidelberg 2012

Sequence similarity in the VRs

Characterized with low homology among different alpha chains, the VRs share patterns of amino acid residues in polypeptide chains among human, chimpanzee, mouse, rat, chicken and zebrafish. For Col5a1, a "YYYEYPYYED" is highly conserved among all the species compared. The peptide encoded by exon 8 (p8) in Col5a1 has conserved "FTEETI" and "IGPGMPANQDT" fragments among different species. As for Col5a3, the exon-intron boundaries of the zebrafish sequence are not well defined based on the current available annotations. Therefore, the alignment for this molecule was done without separating the exons. With some exceptions for the zebrafish sequences, the "KGKGKKRGRKGKGRKKNK" and "TMGP" as well as "FRAAE" and "FQIFP" are conserved among species. Similarly, there are several different conserved segments among species in Col11a2: "DYEPPYY" is conserved in p6; while for p8, "EQTDLQ" and "YDYTYGYGD" as well as "DYREETELGPALSAET" are conserved among species. In the case of Col11a1, alignment and comparison of different species indicate that tyrosines (Y) and lysines (K) are highly conserved in the peptides encoded by exon 6A and exon 6B, respectively. Higher similarity has been observed for p8 in Col11a1, specifically for the sequence signatures "DGDLGEYDF," "YEYKEYE" and "FGPGVPAETD" which align well among these species. The conserved sequences corresponding to p6 in Col5a1, p6 and p8 in Col11a2 as well as p6A and p8 in Col11a1 are tyrosine-rich; while p6 in Col5a3 and p6B in Col11a1 are lysine-rich (Fig. 1) demonstrating further similarities among the collagen V/XI alpha chains.

Biochemical properties

As demonstrated by sequence alignment, peptides in different minor fibrillar collagens have similar biochemical composition. Although high overall similarity is not exhibited, they do possess similar biochemical properties and amino acid composition. To demonstrate this, theoretical isoelectric points were calculated. As summarized in Table 2, all the conserved fragments from VRs are acidic and negatively charged under physiological conditions, except those encoded by exon 6 in Col5a3 and by exon 6B in Col11a1. Although there is no direct evidence describing the role of the variable regions, the overall biochemical properties described in part by the isoelectric point may contribute to the function. In the case of Col11a2, this acidic segment is excluded from cartilaginous tissues (Sugimoto et al., 1998), indicating potentially strong biological relevance for the regulated inclusion and exclusion of this domain. Similarly for Col11a1, exon 6A is favored in various tissues, whereas exon 6B is restricted to the cartilage underlying the perichondrium of the diaphysis (Morris et al., 2000). In summary, we predict a functional role

dictated by the acidic and basic isoforms of the minor fibrillar collagens.

## Isoform-dependent post-translation modification

Sulfation is one of the most prevalent modifications of extracellular molecules, playing a role in protein-protein interactions, and enhancing the affinity of the binding partners (Bowman and Bertozzi, 1999). Receptor-ligand interactions for a number of signaling molecules and their receptors are favored through tyrosine sulfation (Huang et al., 1989; Leyte et al., 1991; Seibert et al., 2002). Extracellular matrix molecules that undergo tyrosine sulfation include fibromodulin, osteoadherin, lumican, Ucma (Onnerfjord et al., 2004; Surmann-Schmitt et al., 2008) and fibronectin (Paul and Hynes, 1984; Liu and Lipmann, 1985). As shown by alignment, Col5a1, Col11a1 and Col11a2 are rich in tyrosines within the variable regions. Using SulfoSite (Chang et al., 2009) to predict the possible sulfation modifications, most of the tyrosine residues are predicted to undergo sulfation, although the probability varies from residue to residue (Fig. 3). Experimental evidence exists for tyrosine sulfation in the case of the NC3 domain of Col5a1, by Fessler and colleagues (Fessler et al., 1986). In Col11a1, it was demonstrated that the splice variant containing exon 6A–7–8 is more widely expressed in tissues including non-cartilaginous ones (Chen et al., 2001). Tyrosine sulfation may therefore play an important role in the function of Col11a1 in non-cartilaginous tissues.

Intrinsic disorder

Although cumulative evidence supports the idea that the different splice-forms have different functions in a time- or tissue-dependent manner (Stamm et al., 2005; Ben-Dov et al., 2008), the connection between AS and the functional changes that result remain to be addressed. A recent study attempted to connect the AS with intrinsic disorder (Romero et al., 2006), which may be important in modulating the conformation of polypeptide segments in proteins. To examine whether the minor fibrillar collagens have intrinsic disorder, particularly in the VRs, we analyzed the human Col11a1, Col11a2, Col5a1 and Col5a3 by using VLXT PONDR (Fig. 3) (Romero et al., 2001). Not surprisingly, the major triple helices of all of the collagens analyzed were predicted to possess intrinsic disorder, since these regions are rich in G-X-Y repeats and match the criteria for intrinsic disorder. Interestingly, the VRs were also found to have intrinsic disorder, indicating the potential for a functional role of this region in minor fibrillar collagens. Moreover, as discussed in previous sections, the VR projects away from the triple helix in collagen fibrils. Together, it is possible that extracellular molecules interact with the VR thus inducing a conformational change, regulating the diameter of the collagen fibrils, or alternatively, providing a mechanism for molecular interactions. Whether

Table 2 Prediction of isoelectric point (pI) for each peptide in the variable regions of minor fibrillar collagens

Fragments		Sequence	Average pI
Col5a1	p6	YTEGDGEGETYYYEYPYYEDPEDLGKEPTPSKKPVEAAKETTEVPE	4.08
	p7	ELTPTPTEAAPMPETSEGAGKEEDVGIGDYDYVPSEDYYTPSPYDDLTYGEGEENPDQPTDPG AGAEIPTSTADTSNSSN	3.20
	p8	PAPPPGEGADDLEGEFTEETIRNLDENYYDPYYDPTSSPSEIGPGMPANQDTIYEG	3.36
Col5a3	p6	APQGEPETPRPRRKGKGKGRKKGRGRKGKGRKKNKEIWTSSPPPDSAENQ	11.39
	p7	TSTDIPKTETPAPNLPPTPTPLVVTSTVTTGLNATILE	4.43
	p8	RSLDPDSGTELGTLETKAAREDEEGDDSTMGPDFRAAEYPSRTQFQIFP	4.13
Col11a1	p6A	YAPEDIIEYDYEYGEAEYKEAESVTEGPTVTEETIAQTE	3.61
	p6B	KKKSNFKKKMRTVATKSKEKSKKFTPPKSEKFSSKKKKSYQASAKAKLGVK	10.77
	p7	ANIVDDFQEYNYGTMESYQTEAPRHVSGTNE	4.24
	p8	PNPVEEIFTEEYLTGEDYDSQRKNSEDTLYENKEIDGRDSDLLVDGDLGEYDFYEYKEYEDKPT SPPNEEFGPGVPAETDITETS	3.75
Col11a2	p6	PTESLYYDYEPPYYDVMTTGTTPDYQ	3.33
	p7	PTPGEEEEILESSLLPPLEE	3.63
	p8	EQTDLQVPPTADRFQAEEYGEGGTDPPEGPYDYTYGYGDDYREETELGPALSAETAHSGA	3.73
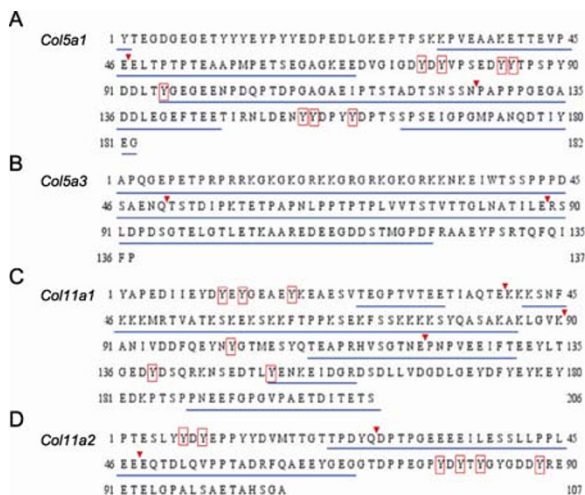
Figure 3. Prediction of sulfation modification and intrinsic disorder within the variable region of human minor fibrillar collagens. Prediction of sulfation modification was carried out using SulfoSite and intrinsic disorder of the variable region was identified by VLXT for the human minor fibrillar collagens. (A) Col5a1; (B) Col5a3; (C) Col11a1; and (D) Col11a2. Red squared tyrosines (Y) are those predicted to be sulfation modification sites. Underlined letters represent predicted intrinsic disorder; red arrows point to the exon boundaries in the variable region for each collagen.

the interaction of HS with this region functions in this way or not has yet to be established.

Regulation of transcription and AS

While the spatiotemporal expressions of different isoforms of type XI collagens, especially *Col11a1*, have been docu-

mented in different organisms, the mechanism determining specific expression patterns is still elusive. Studies of promoters of these collagen genes reveal that the common feature is a lack of a TATA box and an enrichment in GC content, except for *Col5a2* (Table 3) (Fichard et al., 1995; Lee and Greenspan, 1995; Yoshioka et al., 1995a; Tsumaki et al., 1996; Kinoshita et al., 1997; Tsumaki et al., 1998; Penkov et al., 2000; Tanaka et al., 2002; Matsui et al., 2003; Matsuo et al., 2003; Nagato et al., 2004; Sakata-Takatani et al., 2004). The transcription factors identified for collagen types V and XI are quite similar, such as SP1 and CBF/NF (Table 3). Although CBF/NF has been recognized as a ubiquitous nuclear factor (Li et al., 1992), it may be involved in modulating the expression of Col1a2 in mesenchymal-origin tissues (Tanaka et al., 2004), indicating the complexity of the regulation of collagen genes. While co-expressed with collagen type XI chains, collagen type II is regulated by Sox9, which is a high mobility group (HMG) domain transcription factor (Bi et al., 1999). Similar to the Col2a1 chondrocyte-specific enhancer, two HMG protein binding consensus sites have been identified in Col11a2 (Bridgewater et al., 1998). Another study proved that these consensus sites of Col11a2 are indeed recognized by Sox9 (Bi et al., 1999). Currently, there is no direct evidence that Sox9 regulates the expression of *Col11a1*. However, seven conserved *Sox* pair sequences in the promoter of *Col11a1* from both human and mouse have been identified by computational analysis (Davies et al., 2007). Additionally, Lincoln and colleagues have shown that Sox9 as well as Col5a1 and Col11a1 is required for heart valve development (Lincoln et al., 2006; 2007). Therefore, it is possible that Sox9 regulates the expression of *Col11a1* and *Col5a1*. Most recently, the transactivation of *Col11a1* by Lef1 was reported (Kahler et al., 2008); however the mechanistic details of regulation still remain to be investigated. Further, the responsible *trans*-acting factors that modulate

Table 3 Summary of the transcriptional factors for each minor fibrillar collagen gene

Gene	Proximal promoter	Transcription factor(s) identified
Col11a1	TATA box lack, GC rich	AP2, SP1, FP9C, CBF/NF-Y
Col11a2	TATA box lack, GC rich	SP1, EWS/ERG, NT2
Col5a1	TATA box lack, CCAAT box lack, GC rich	SP1, CBF
Col5a2	Similar to *Col11a1*, CCAAT box lack, TATA-like element	FP-A, FP-B, PBX1/2, PREP1, HOXB1
Col5a3	TATA box lack, GC rich	CBF/NF-Y

AS in *Col11a1* and *Col11a2* are not currently known. Therefore, it is highly likely that other *cis*-acting elements and nuclear factors that specifically regulate the finely controlled expression of minor fibrillar collagen molecules remain to be discovered.

AS involves the recognition of the basic splicing elements, including 5′ donor site, 3′ acceptor site and the auxiliary *cis*-elements: exonic/intronic splicing enhancer (ESE/ISE) as well as silencer (ESS/ISS) by the spliceosome (Ladd and Cooper, 2002; Matlin et al., 2005). To analyze the splice exon-intron junction sequences in the VR of minor fibrillar collagens, 20 nucleotides (nts) from both 5′ donor site and 3′ acceptor site were aligned among 4 different species (Supplementary Fig. 2 and supplementary Table 2). As is true for many alternatively spliced exons genome-wide, exons within the variable region of *Col5a1*, *Col5a3*, *Col11a1* and *Col11a2* have exon-intron sequences that are not consistent with the classic 5′ donor site (AGguragu) and 3′ acceptor site (yyyyyyynagG). The constitutively expressed exons display non-consensus junction sites in *Col5a1* and *Col5a3*. Interestingly, the acceptor sites of intron 6A–6B and intron 6–7 in *Col11a1* and *Col11a2*, respectively, exhibit a potential hairpin structure (Supplementary Fig. 2). Because of the inconsistency of the donor and acceptor sites in the VR of minor fibrillar collagens, the other splicing elements, such as ESE/ISS as well as ISE/ISS, may play roles in regulating the alternatively spliced *Col11a1* and *Col11a2*. Recently, the importance of auxiliary *cis*-elements has been recognized in the determination of exon splicing patterns (Matlin et al., 2005). In *Col2a1*, exon 2 is alternatively spliced; the inclusion or exclusion of exon 2 is regulated by the combination of a 5′ weak donor site and a stem-loop structure in the intron 2 (McAlinden et al., 2005). Another study has shown that TIA-1 (T-cell-restricted intracellular antigen-1) can interact with the conserved stem-loop structure to regulate the AS (McAlinden et al., 2007). Exon7 of *Col11a1* possesses a *cis*-acting function in the modulation of the AS (Chen et al., 2001). Tsumaki and colleagues found that both the deletion of an upstream 290 bp in the proximal promoter and the intron 1 sequence of *Col11a2* could affect tissue expression pattern (Tsumaki et al., 1996). Additional studies suggested that there are at least three different *cis*-acting elements involved in the modulation of the tissue-specific expression patterns of *Col11a2* (Tsumaki et al., 1998). Similar to collagen type XI, fibronectin is also found in cell type- and developmental stage-specific

splice-forms derived by AS (Pankov and Yamada, 2002). Recent studies show that integrin-mediated signaling can modulate the pre-mRNA splicing of fibronectin through a non-stress c-Jun N-terminal kinase (JNK) signaling pathway (Pelisch et al., 2005). Interestingly, Col5a1 and Col3a1 have been recognized as regulators of integrin organization (Zoppi et al., 2004), indicating a possible signaling pathway within the extracellular matrix. Additionally, Srebrow and colleagues showed that collagen type IV and laminin can inhibit extra domain I inclusion within fibronectin (Srebrow et al., 2002). BMP-2 and TGF-beta 1 are able to bind to Col2a1 (Zhu et al., 1999), and these two molecules, also display different effects in the promotion of the expression of collagen type II isoforms; BMP-2 promotes the expression of Col2a1b, while TGF-beta 1 favors the expression of Col2a1a (Valcourt et al., 2003).

Possible role of microRNAs

The transcription level regulation of minor fibrillar collagen genes is not thoroughly understood, and it is possible that new nuclear factors will be discovered in the future. However, there may be another mechanism regulating the AS of type XI collagen genes. Although current understanding of the role of microRNA (miRNA) in regulating AS is limited, evidence supports the regulation of components of the spliceosome (Makeyev and Maniatis, 2008). It is possible that miRNAs regulate the expression of *Col11a1*, given the fact that miRNAs have a profound effect on vertebrate development post-transcriptionally (Makeyev and Maniatis, 2008). To predict possible miRNAs involved in the regulation of the expression of minor fibrillar collagens, we used Pictar (Krek et al., 2005) and TargetScan (Grimson et al., 2007) to analyze the 3′ UTR of all human collagen type V and XI genes. Although the collagen genes analyzed were predicted to be the targets of many different miRNAs, all are potential targets of the miR-29a/b/c family except *Col11a2*, suggesting a possible co-regulation mechanism. MiR-29s have been shown to participate in many different pathways, including methyl-transferase 3A/3B in cancer (Fabbri et al., 2007), BACE/beta-secretase in Alzheimer's disease, p85alpha and CDC42 in promoting apoptosis (Park et al., 2009). Noteworthy, previous studies by others have shown that miR-29 family members regulate the expression of a broad range of collagen genes (Sengupta et al., 2008). In addition to the minor fibrillar collagen genes, *Col1a1*, *Col2a1* and *Col3a1* also
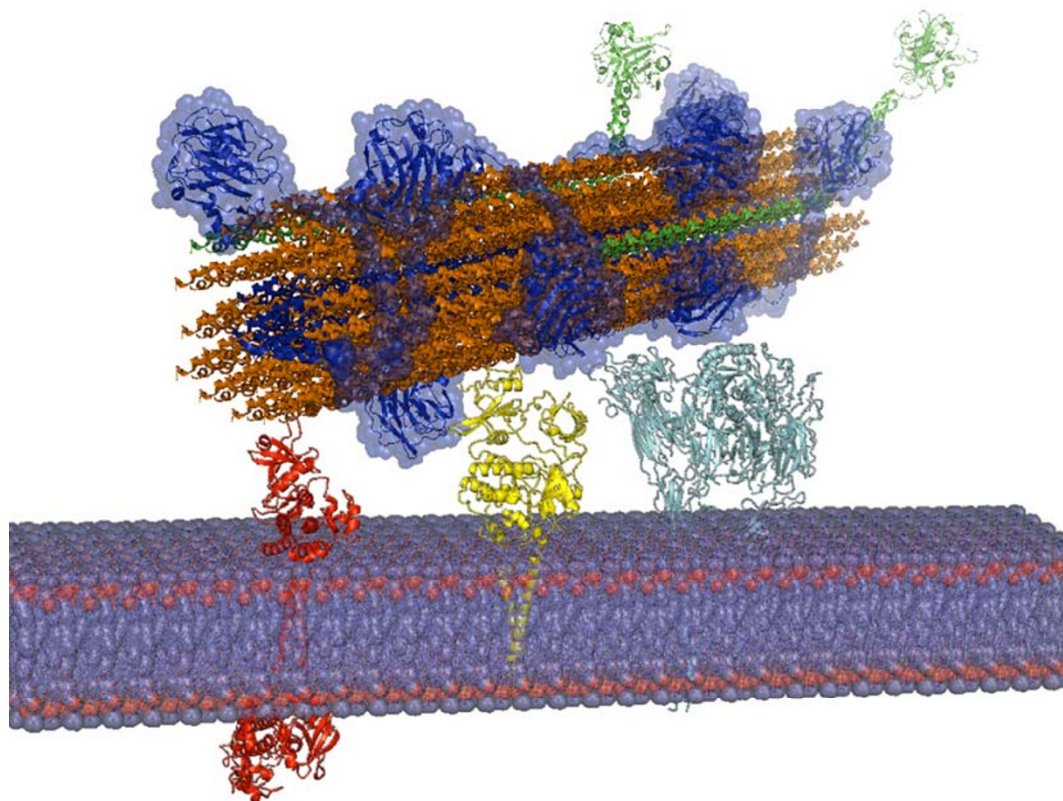
Figure 4.   Potential interactions between collagen fibril surface domains and cell surface receptors. Examples of cell surface receptors that interact with extracellular matrix collagen fibrils include DDR2 (red) and integrins (yellow). TGFβR (cyan) is shown as an example of growth factor receptors that may be influenced by interaction between growth factors and the extracellular matrix molecules, resulting in either the inhibition of cell signaling events via sequestration, or alternatively, the facilitated delivery of the signaling molecule to its destination on the cell surface receptor. TGFβR, transforming growth factor beta receptor; DDR2, discoidin domain receptor 2. Structures with accession numbers are as follows: ColIX (2UUR), TGFβR (1B6C, 2L5S), DDR2 (2WUH, 2SRC), integrin (1JV2, 2K9J), and the lipid bilayer was obtained from http://www.umass.edu/microbio/rasmol/bilayers.htm. (Heller et al., 1993). The structure of ColII was created using The PyMOL Molecular Graphics System, Version 1.3, Schrödinger, LLC. The structure of ColXI was created as a homology model built from the ColIX (2UUR) structure template (McDougal et al., 2011).

possess the consensus sites for the miR-29a/b/c family. The targets of miRNAs can alter the expression of downstream genes themselves if they are upstream signaling molecules, transcription factors (Makeyev and Maniatis, 2008), or *trans*-acting elements. Kato and colleagues found that miR-192 represses the expression of the E-box repressor, resulting in an enhanced expression level of *Col1a2* in kidney cells (Kato et al., 2007). This phenomenon potentially increases the complexity with which miRNAs may exert their effect on the target genes. It is also possible that miRNAs regulate the enzymes that degrade collagens in the extracellular matrix. Therefore, direct prediction through 3′ UTR of a given gene will not tell the entire story.

CONCLUSION

The spatiotemporal expression patterns of minor fibrillar col-

lagens are not well documented in early embryonic developmental stages. Demonstrated by the homozygous mutants of *Col11a1* and *Col5a1*, embryonic lethality indicates that minor fibrillar collagens are crucial during development. Although their roles in fibrillogenesis have been emphasized, other potential functional aspects are not as well understood, and require a concentrated research effort. Noncollagenous domains such as the VRs may also contribute to the function of minor fibrillar collagens. The fibril surface location of the specialized noncollagenous domains places them ideally for interaction with cell surface receptors in addition to other matrix molecules, as illustrated in Fig. 4. AS determination of extra heparan sulfate binding sites within the VR of Col11a1 provides evidence for their specialization among the splice-forms with respect to unique functions. Detailed expressional and more extensive functional studies are needed. Although the alpha chains of collagen type V and XI do not

show high sequence identity within the variable region, the distribution of charged side chains, such as tyrosine and lysine, is conserved. This indicates that Col5a3 may harbor similar binding sites for heparan sulfate as demonstrated for Col11a1. The prediction of sulfation modification to tyrosines also suggests that Col11a1 and Col5a1 may undergo similar post-transcriptional modification. Different isoforms resulting from AS are differentially expressed during development for Col11a1, Col11a2 and Col1a3 (Col2a1). Although this is a strong indication that specific isoforms play novel functional roles, direct evidence is still lacking. The transcriptional regulation is similar for all minor fibrillar collagen genes; however, the AS tends to be more diverse. Although the regulation of Col2a1 has been described extensively, less information is available for Col11a1 and Col11a2. It is tempting to speculate that other factors contribute to the expressional regulation of minor fibrillar collagens, such as the mir-29 family, which has been suggested as a master regulator for collagen genes. This suggestion is supported by the findings that the mir-29 microRNA family has diverse roles in regulating cell survival related genes, in lung cancer, breast cancer and Alzheimer's disease. Col11a1 has also been related to certain cancer types. Recently a relation between AS and intrinsic disorder has been described (Romero et al., 2006). The increased flexibility provided by intrinsic disorder may mechanistically connect the functions of different splicing-forms to their function. However, whether the predicted intrinsic disorder in the VRs of minor fibrillar collagens does contribute to their functions is not clear at present, and must await further study.

ACKNOWLEDGEMENTS

ABBREVIATIONS

AS, alternative splicing; BMP-1, bone morphogenetic protein 1; DDR2, discoidin domain receptor 2; FACITs, fibril associated collagen with interrupted helices; FGF8, fibroblast growth factor 8; HMG, high mobility group; HS, heparan sulfate; JNK, Janus kinase; MMP-13, matrix metalloproteinase 13; TSPN, thrombospondin amino terminal domain; RTKs, receptor tyrosine kinases; TIA-1, T-cell-restricted intracellular antigen-1; TGFβ1, transforming growth factor beta 1; TGFβR, transforming growth factor beta receptor

REFERENCES

Andrikopoulos, K., Liu, X., Keene, D.R., Jaenisch, R., and Ramirez, F. (1995). Targeted mutation in the col5a2 gene reveals a regulatory role for type V collagen during matrix assembly. Nat Genet 9, 31–36.

Andrikopoulos, K., Suzuki, H.R., Solursh, M., and Ramirez, F. (1992). Localization of pro-alpha 2(V) collagen transcripts in the tissues of the developing mouse embryo. Dev Dyn 195, 113–120.

Annunen, S., Körkkö, J., Czarny, M., Warman, M.L., Brunner, H.G., Kääriäinen, H., Mulliken, J.B., Tranebjaerg, L., Brooks, D.G., Cox, G.F., et al. (1999). Splicing mutations of 54-bp exons in the COL11A1 gene cause Marshall syndrome, but other mutations cause overlapping Marshall/Stickler phenotypes. Am J Hum Genet 65, 974–983.

Ben-Dov, C., Hartmann, B., Lundgren, J., and Valcárcel, J. (2008). Genome-wide analysis of alternative pre-mRNA splicing. J Biol Chem 283, 1229–1233.

Bi, W., Deng, J.M., Zhang, Z., Behringer, R.R., and de Crombrugghe, B. (1999). Sox9 is required for cartilage formation. Nat Genet 22, 85–89.

Birk, D.E. (2001). Type V collagen: heterotypic type I/V collagen interactions in the regulation of fibril assembly. Micron 32, 223–237.

Blaschke, U.K., Eikenberry, E.F., Hulmes, D.J.S., Galla, H.-J., and Bruckner, P. (2000). Collagen XI nucleates self-assembly and limits lateral growth of cartilage fibrils. J Biol Chem 275, 10370–10378.

Bowman, K.G., and Bertozzi, C.R. (1999). Carbohydrate sulfotransferases: mediators of extracellular communication. Chem Biol 6, R9–R22.

Bridgewater, L.C., Lefebvre, V., and de Crombrugghe, B. (1998). Chondrocyte-specific enhancer elements in the Col11a2 gene resemble the Col2a1 tissue-specific enhancer. J Biol Chem 273, 14998–15006.

Chang, W.C., Lee, T.Y., Shien, D.M., Hsu, J.B., Horng, J.T., Hsu, P.C., Wang, T.Y., Huang, H.D., and Pan, R.L. (2009). Incorporating support vector machine for identifying protein tyrosine sulfation sites. J Comput Chem 30, 2526–2537.

Chanut-Delalande, H., Fichard, A., Bernocco, S., Garrone, R., Hulmes, D.J.S., and Ruggiero, F. (2001). Control of heterotypic fibril formation by collagen V is determined by chain stoichiometry. J Biol Chem 276, 24352–24359.

Chen, Y., Sumiyoshi, H., Oxford, J.T., Yoshioka, H., Ramirez, F., and Morris, N.P. (2001). Cis-acting elements regulate alternative splicing of exons 6A, 6B and 8 of the α1(XI) collagen gene and contribute to the regional diversification of collagen XI matrices. Matrix Biol 20, 589–599.

Davies, S.R., Chang, L.-W., Patra, D., Xing, X., Posey, K., Hecht, J., Stormo, G.D., and Sandell, L.J. (2007). Computational identification and functional validation of regulatory motifs in cartilage-expressed genes. Genome Res 17, 1438–1447.

Erdman, R., Stahl, R.C., Rothblum, K., Chernousov, M.A., and Carey, D.J. (2002). Schwann cell adhesion to a novel heparan sulfate binding site in the N-terminal domain of alpha 4 type V collagen is mediated by syndecan-3. J Biol Chem 277, 7619–7625.

Fabbri, M., Garzon, R., Cimmino, A., Liu, Z., Zanesi, N., Callegari, E.,

Liu, S., Alder, H., Costinean, S., Fernandez-Cymering, C., et al. (2007). MicroRNA-29 family reverts aberrant methylation in lung cancer by targeting DNA methyltransferases 3A and 3B. Proc Natl Acad Sci U S A 104, 15805–15810.

Fernandes, R.J., Weis, M., Scott, M.A., Seegmiller, R.E., and Eyre, D.R. (2007). Collagen XI chain misassembly in cartilage of the chondrodysplasia (cho) mouse. Matrix Biol 26, 597–603.

Fessler, L.I., Brosh, S., Chapin, S., and Fessler, J.H. (1986). Tyrosine sulfation in precursors of collagen V. J Biol Chem 261, 5034–5040.

Fichard, A., Kleman, J.P., and Ruggiero, F. (1995). Another look at collagen V and XI molecules. Matrix Biol 14, 515–531.

Fischer, H., Stenling, R., Rubio, C., and Lindblom, A. (2001). Colorectal carcinogenesis is associated with stromal expression of COL11A1 and COL5A2. Carcinogenesis 22, 875–878.

Fletcher, R.B., Baker, J.C., and Harland, R.M. (2006). FGF8 splice-forms mediate early mesoderm and posterior neural tissue formation in Xenopus. Development 133, 1703–1714.

Gregory, K.E., Oxford, J.T., Chen, Y., Gambee, J.E., Gygi, S.P., Aebersold, R., Neame, P.J., Mechling, D.E., Bächinger, H.P., and Morris, N.P. (2000). Structural organization of distinct domains within the non-collagenous N-terminal region of collagen type XI. J Biol Chem 275, 11498–11506.

Grimson, A., Farh, K.K.-H., Johnston, W.K., Garrett-Engele, P., Lim, L.P., and Bartel, D.P. (2007). MicroRNA targeting specificity in mammals: determinants beyond seed pairing. Mol Cell 27, 91–105.

Guo, Q., and Li, J.Y.H. (2007). Distinct functions of the major Fgf8 spliceform, Fgf8b, before and during mouse gastrulation. Development 134, 2251–2260.

Halsted, K.C., Bowen, K.B., Bond, L., Luman, S.E., Jorcyk, C.L., Fyffe, W.E., Kronz, J.D., and Oxford, J.T. (2008). Collagen α1(XI) in normal and malignant breast tissue. Mod Pathol 21, 1246–1254.

Heller, H., Schaefer, M., and Schulten, K. (1993). Molecular dynamics simulation of a bilayer of 200 lipids in the gel and in the liquid-crystal phases. J Phys Chem 97, 8343–8360.

Holmes, D.F., and Kadler, K.E. (2006). The 10+4 microfibril structure of thin cartilage fibrils. Proc Natl Acad Sci U S A 103, 17249–17254.

Huang, S.C., Yu, D.H., Wank, S.A., Mantey, S., Gardner, J.D., and Jensen, R.T. (1989). Importance of sulfation of gastrin or cholecystokinin (CCK) on affinity for gastrin and CCK receptors. Peptides 10, 785–789.

Imamura, Y., Scott, I.C., and Greenspan, D.S. (2000). The pro-alpha3(V) collagen chain. Complete primary structure, expression domains in adult and developing tissues, and comparison to the structures and expression domains of the other types V and XI procollagen chains. J Biol Chem 275, 8749–8759.

Imhof, M., and Trueb, B. (2001). Alternative splicing of the first F3 domain from chicken collagen XIV affects cell adhesion and heparin binding. J Biol Chem 276, 9141–9148.

Kadler, K.E., Baldock, C., Bella, J., and Boot-Handford, R.P. (2007). Collagens at a glance. J Cell Sci 120, 1955–1958.

Kadler, K.E., Hill, A., and Canty-Laird, E.G. (2008). Collagen fibrillogenesis: fibronectin, integrins, and minor collagens as organizers and nucleators. Curr Opin Cell Biol 20, 495–501.

Kahler, R.A., Yingst, S.M.C., Hoeppner, L.H., Jensen, E.D., Krawczak, D., Oxford, J.T., and Westendorf, J.J. (2008). Collagen 11a1 is indirectly activated by lymphocyte enhancer-binding factor 1 (Lef1) and negatively regulates osteoblast maturation. Matrix Biol 27, 330–338.

Kato, M., Zhang, J., Wang, M., Lanting, L., Yuan, H., Rossi, J.J., and Natarajan, R. (2007). MicroRNA-192 in diabetic kidney glomeruli and its function in TGF-β-induced collagen expression via inhibition of E-box repressors. Proc Natl Acad Sci U S A 104, 3432–3437.

Kinoshita, A., Greenwel, P., Tanaka, S., Di Liberto, M., Yoshioka, H., and Ramirez, F. (1997). A transcription activator with restricted tissue distribution regulates cell-specific expression of alpha1(XI) collagen. J Biol Chem 272, 31777–31784.

Krek, A., Grün, D., Poy, M.N., Wolf, R., Rosenberg, L., Epstein, E.J., MacMenamin, P., da Piedade, I., Gunsalus, K.C., Stoffel, M., et al. (2005). Combinatorial microRNA target predictions. Nat Genet 37, 495–500.

Ladd, A., and Cooper, T. (2002). Finding signals that regulate alternative splicing in the post-genomic era. Genome Biol 3, reviews0008.0001 reviews0008.0016.

Lee, S., and Greenspan, D.S. (1995). Transcriptional promoter of the human alpha 1(V) collagen gene (COL5A1). Biochem J 310, 15–22.

Leitinger, B., and Hohenester, E. (2007). Mammalian collagen receptors. Matrix Biol 26, 146–155.

Leyte, A., van Schijndel, H.B., Niehrs, C., Huttner, W.B., Verbeet, M.P., Mertens, K., and van Mourik, J.A. (1991). Sulfation of Tyr1680 of human blood coagulation factor VIII is essential for the interaction of factor VIII with von Willebrand factor. J Biol Chem 266, 740–746.

Li, S.-W., Takanosu, M., Arita, M., Bao, Y., Ren, Z.-X., Maier, A., Prockop, D.J., and Mayne, R. (2001). Targeted disruption of Col11a2 produces a mild cartilage phenotype in transgenic mice: comparison with the human disorder otospondylomegaepiphyseal dysplasia (OSMED). Dev Dyn 222, 141–152.

Li, X.-Y., Mantovani, R., Hooft van Huijsduijnen, R., Andre, I., Benoist, C., and Mathis, D. (1992). Evolutionary variation of the CCAAT-binding transcription factor NF-Y. Nucleic Acids Res 20, 1087–1091.

Li, Y., Lacerda, D.A., Warman, M.L., Beier, D.R., Yoshioka, H., Ninomiya, Y., Oxford, J.T., Morris, N.P., Andrikopoulos, K., Ramirez, F., et al. (1995). A fibrillar collagen gene, Col11a1, is essential for skeletal morphogenesis. Cell 80, 423–430.

Lincoln, J., Florer, J.B., Deutsch, G.H., Wenstrup, R.J., and Yutzey, K.E. (2006a). ColVa1 and ColXIa1 are required for myocardial morphogenesis and heart valve development. Dev Dyn 235, 3295–3305.

Lincoln, J., Florer, J.B., Deutsch, G.H., Wenstrup, R.J., and Yutzey, K.E. (2006b). ColVa1 and ColXIa1 are required for myocardial morphogenesis and heart valve development. Dev Dyn 235, 3295–3305.

Lincoln, J., Kist, R., Scherer, G., and Yutzey, K.E. (2007). Sox9 is required for precursor cell expansion and extracellular matrix organization during mouse heart valve development. Dev Biol 305, 120–132.

Liu, M.C., and Lipmann, F. (1985). Isolation of tyrosine-O-sulfate by

Pronase hydrolysis from fibronectin secreted by Fujinami sarcoma virus-infected rat fibroblasts. Proc Natl Acad Sci U S A 82, 34–37.

Lui, V.C., Kong, R.Y., Nicholls, J., Cheung, A.N., and Cheah, K.S. (1995). The mRNAs for the three chains of human collagen type XI are widely distributed but not necessarily co-expressed: implications for homotrimeric, heterotrimeric and heterotypic collagen molecules. Biochem J 311, 511–516.

Lui, V.C., Ng, L.J., Sat, E.W., Nicholls, J., and Cheah, K.S. (1996). Extensive alternative splicing within the amino-propeptide coding domain of alpha2(XI) procollagen mRNAs. Expression of transcripts encoding truncated pro-alpha chains. J Biol Chem 271, 16945–16951.

Luparello, C., and Sirchia, R. (2005). Type V collagen regulates the expression of apoptotic and stress response genes by breast cancer cells. J Cell Physiol 202, 411–421.

Makeyev, E.V., and Maniatis, T. (2008). Multilevel regulation of gene expression by microRNAs. Science 319, 1789–1790.

Marchant, J.K., Hahn, R.A., Linsenmayer, T.F., and Birk, D.E. (1996). Reduction of type V collagen using a dominant-negative strategy alters the regulation of fibrillogenesis and results in the loss of corneal-specific fibril morphology. J Cell Biol 135, 1415–1426.

Matlin, A.J., Clark, F., and Smith, C.W.J. (2005). Understanding alternative splicing: towards a cellular code. Nat Rev Mol Cell Biol 6, 386–398.

Matsui, Y., Chansky, H.A., Barahmand-Pour, F., Zielinska-Kwiatkowska, A., Tsumaki, N., Myoui, A., Yoshikawa, H., Yang, L., and Eyre, D.R. (2003). COL11A2 collagen gene transcription is differentially regulated by EWS/ERG sarcoma fusion protein and wild-type ERG. J Biol Chem 278, 11369–11375.

Matsuo, N., Yu-Hua, W., Sumiyoshi, H., Sakata-Takatani, K., Nagato, H., Sakai, K., Sakurai, M., and Yoshioka, H. (2003). The transcription factor CCAAT-binding factor CBF/NF-Y regulates the proximal promoter activity in the human α1(XI) collagen gene (COL11A1). J Biol Chem 278, 32763–32770.

Mayne, R., Brewton, R.G., Mayne, P.M., and Baker, J.R. (1993). Isolation and characterization of the chains of type V/type XI collagen present in bovine vitreous. J Biol Chem 268, 9381–9386.

McAlinden, A., Havlioglu, N., Liang, L., Davies, S.R., and Sandell, L.J. (2005). 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 280, 32700–32711.

McAlinden, A., Liang, L., Mukudai, Y., Imamura, T., and Sandell, L.J. (2007). Nuclear protein TIA-1 regulates COL2A1 alternative splicing and interacts with precursor mRNA and genomic DNA. J Biol Chem 282, 24444–24454.

McDougal, O.M., Mallory, C., Warner, L.R., and Oxford, J.T. (2011). Predicted structure and binding motifs of collagen α1(XI). J BioInformatics BioTech (In press).

McGuirt, W.T., Prasad, S.D., Griffith, A.J., Kunst, H.P.M., Green, G.E., Shpargel, K.B., Runge, C., Huybrechts, C., Mueller, R.F., Lynch, E., et al. (1999). Mutations in COL11A2 cause non-syndromic hearing loss (DFNA13). Nat Genet 23, 413–419.

Medeck, R.J., Sosa, S., Morris, N., and Oxford, J.T. (2003). BMP-1-mediated proteolytic processing of alternatively spliced isoforms of collagen type XI. Biochem J 376, 361–368.

Melkoniemi, M., Brunner, H.G., Manouvrier, S., Hennekam, R., Superti-Furga, A., Kääriäinen, H., Pauli, R.M., van Essen, T., War-

man, M.L., Bonaventure, J., et al. (2000). Autosomal recessive disorder otospondylomegaepiphyseal dysplasia is associated with loss-of-function mutations in the COL11A2 gene. Am J Hum Genet 66, 368–377.

Michalickova, K., Susic, M., Willing, M.C., Wenstrup, R.J., and Cole, W.G. (1998). Mutations of the alpha2(V) chain of type V collagen impair matrix assembly and produce ehlers-danlos syndrome type I. Hum Mol Genet 7, 249–255.

Morris, N.P., Oxford, J.T., Davies, G.B.M., Smoody, B.F., and Keene, D.R. (2000). Developmentally regulated alternative splicing of the α1(XI) collagen chain: spatial and temporal segregation of isoforms in the cartilage of fetal rat long bones. J Histochem Cytochem 48, 725–741.

Nagato, H., Matsuo, N., Sumiyoshi, H., Sakata-Takatani, K., Nasu, M., and Yoshioka, H. (2004). The transcription factor CCAAT-binding factor CBF/NF-Y and two repressors regulate the core promoter of the human pro-α3(V) collagen gene (COL5A3). J Biol Chem 279, 46373–46383.

Onnerfjord, P., Heathfield, T.F., and Heinegård, D. (2004). Identification of tyrosine sulfation in extracellular leucine-rich repeat proteins using mass spectrometry. J Biol Chem 279, 26–33.

Oxford, J.T., Doege, K.J., and Morris, N.P. (1995). Alternative exon splicing within the amino-terminal nontriple-helical domain of the rat pro-α 1(XI) collagen chain generates multiple forms of the mRNA transcript which exhibit tissue-dependent variation. J Biol Chem 270, 9478–9485.

Pankov, R., and Yamada, K.M. (2002). Fibronectin at a glance. J Cell Sci 115, 3861–3863.

Park, S.-Y., Lee, J.H., Ha, M., Nam, J.-W., and Kim, V.N. (2009). miR-29 miRNAs activate p53 by targeting p85α and CDC42. Nat Struct Mol Biol 16, 23–29.

Paul, J.I., and Hynes, R.O. (1984). Multiple fibronectin subunits and their post-translational modifications. J Biol Chem 259, 13477–13487.

Pelisch, F., Blaustein, M., Kornblihtt, A.R., and Srebrow, A. (2005). Cross-talk between signaling pathways regulates alternative splicing: a novel role for JNK. J Biol Chem 280, 25461–25469.

Penkov, D., Tanaka, S., Di Rocco, G., Berthelsen, J., Blasi, F., and Ramirez, F. (2000). Cooperative interactions between PBX, PREP, and HOX proteins modulate the activity of the alpha 2(V) collagen (COL5A2) promoter. J Biol Chem 275, 16681–16689.

Pihlajamaa, T., Prockop, D.J., Faber, J., Winterpacht, A., Zabel, B., Giedion, A., Wiesbauer, P., Spranger, J., and Ala-Kokko, L. (1998). Heterozygous glycine substitution in the COL11A2 gene in the original patient with the Weissenbacher-Zweymüller syndrome demonstrates its identity with heterozygous OSMED (nonocular Stickler syndrome). Am J Med Genet 80, 115–120.

Pucci-Minafra, I., Carella, C., Cirincione, R., Chimenti, S., Minafra, S., and Luparello, C. (2000). Type V collagen induces apoptosis of 8701-BC breast cancer cells and enhances m-calpain expression. Breast Cancer Res 2, E008.

Ricard-Blum, S., and Ruggiero, F. (2005). The collagen superfamily: from the extracellular matrix to the cell membrane. Pathol Biol (Paris) 53, 430–442.

Richards, A.J., Martin, S., Nicholls, A.C., Harrison, J.B., Pope, F.M., and Burrows, N.P. (1998). A single base mutation in COL5A2 causes Ehlers-Danlos syndrome type II. J Med Genet 35,

846–848.

Romero, P., Obradovic, Z., Li, X., Garner, E.C., Brown, C.J., and Dunker, A.K. (2001). Sequence complexity of disordered protein. Proteins 42, 38–48.

Romero, P.R., Zaidi, S., Fang, Y.Y., Uversky, V.N., Radivojac, P., Oldfield, C.J., Cortese, M.S., Sickmeier, M., LeGall, T., Obradovic, Z., et al. (2006). Alternative splicing in concert with protein intrinsic disorder enables increased functional diversity in multicellular organisms. Proc Natl Acad Sci U S A 103, 8390–8395.

Roulet, M., Ruggiero, F., Karsenty, G., and LeGuellec, D. (2007). A comprehensive study of the spatial and temporal expression of the col5a1 gene in mouse embryos: a clue for understanding collagen V function in developing connective tissues. Cell Tissue Res 327, 323–332.

Sakata-Takatani, K., Matsuo, N., Sumiyoshi, H., Tsuda, T., and Yoshioka, H. (2004). Identification of a functional CBF-binding CCAAT-like motif in the core promoter of the mouse pro-α 1(V) collagen gene (Col5a1). Matrix Biol 23, 87–99.

Schwarze, U., Atkinson, M., Hoffman, G.G., Greenspan, D.S., and Byers, P.H. (2000). Null alleles of the COL5A1 gene of type V collagen are a cause of the classical forms of Ehlers-Danlos syndrome (types I and II). Am J Hum Genet 66, 1757–1765.

Seibert, C., Cadene, M., Sanfiz, A., Chait, B. T., and Sakmar, T.P. (2002). Tyrosine sulfation of CCR5 N-terminal peptide by tyrosyl-protein sulfotransferases 1 and 2 follows a discrete pattern and temporal sequence. Proc Natl Acad Sci USA 99, 11031–11036.

Sengupta, S., den Boon, J.A., Chen, I.H., Newton, M.A., Stanhope, S.A., Cheng, Y.-J., Chen, C.-J., Hildesheim, A., Sugden, B., and Ahlquist, P. (2008). MicroRNA 29c is down-regulated in nasopharyngeal carcinomas, up-regulating mRNAs encoding extracellular matrix proteins. Proc Natl Acad Sci USA 105, 5874–2000, 5878.

Sirko-Osadsa, D.A., Murray, M.A., Scott, J.A., Lavery, M.A., Warman, M.L., and Robin, N.H. (1998). Stickler syndrome without eye involvement is caused by mutations in COL11A2, the gene encoding the α2(XI) chain of type XI collagen. J Pediatr 132, 368–371.

Söderhäll, C., Marenholz, I., Kerscher, T., Rüschendorf, F., Esparza-Gordillo, J., Worm, M., Gruber, C., Mayr, G., Albrecht, M., Rohde, K., et al. (2007). Variants in a novel epidermal collagen gene (COL29A1) are associated with atopic dermatitis. PLoS Biol 5, e242.

Srebrow, A., Blaustein, M., and Kornblihtt, A.R. (2002). Regulation of fibronectin alternative splicing by a basement membrane-like extracellular matrix. FEBS Lett 514, 285–289.

Stamm, S., Ben-Ari, S., Rafalska, I., Tang, Y., Zhang, Z., Toiber, D., Thanaraj, T.A., and Soreq, H. (2005). Function of alternative splicing. Gene 344, 1–20.

Sugimoto, M., Kimura, T., Tsumaki, N., Matsui, Y., Nakata, K., Kawahata, H., Yasui, N., Kitamura, Y., Nomura, S., and Ochi, T. (1998). Differential in situ expression of alpha2(XI) collagen mRNA isoforms in the developing mouse. Cell Tissue Res 292, 325–332.

Surmann-Schmitt, C., Dietz, U., Kireva, T., Adam, N., Park, J., Tagariello, A., Önnerfjord, P., Heinegård, D., Schlötzer-Schrehardt, U., Deutzmann, R., et al. (2008). Ucma, a novel secreted cartilage-specific protein with implications in osteogenesis. J Biol

Chem 283, 7082–7093.

Tanaka, K., Tsumaki, N., Kozak, C.A., Matsumoto, Y., Nakatani, F., Iwamoto, Y., and Yamada, Y. (2002). A Krüppel-associated box-zinc finger protein, NT2, represses cell-type-specific promoter activity of the α 2(XI) collagen gene. Mol Cell Biol 22, 4256–4267.

Tanaka, S., Antoniv, T.T., Liu, K., Wang, L., Wells, D.J., Ramirez, F., and Bou-Gharios, G. (2004). Cooperativity between far upstream enhancer and proximal promoter elements of the human α2(I) collagen (COL1A2) gene instructs tissue specificity in transgenic mice. J Biol Chem 279, 56024–56031.

Tompson, S.W., Bacino, C.A., Safina, N.P., Bober, M.B., Proud, V.K., Funari, T., Wangler, M.F., Nevarez, L., Ala-Kokko, L., Wilcox, W.R., et al. (2010). Fibrochondrogenesis results from mutations in the COL11A1 type XI collagen gene. Am J Hum Genet 87, 708–712.

Tsumaki, N., and Kimura, T. (1995). Differential expression of an acidic domain in the amino-terminal propeptide of mouse pro-alpha 2(XI) collagen by complex alternative splicing. J Biol Chem 270, 2372–2378.

Tsumaki, N., Kimura, T., Matsui, Y., Nakata, K., and Ochi, T. (1996). Separable cis-regulatory elements that contribute to tissue- and site-specific alpha 2(XI) collagen gene expression in the embryonic mouse cartilage. J Cell Biol 134, 1573–1582.

Tsumaki, N., Kimura, T., Tanaka, K., Kimura, J.H., Ochi, T., and Yamada, Y. (1998). Modular arrangement of cartilage- and neural tissue-specific cis-elements in the mouse alpha2(XI) collagen promoter. J Biol Chem 273, 22861–22864.

Valcourt, U., Gouttenoire, J., Aubert-Foucher, E., Herbage, D., and Mallein-Gerin, F. (2003). Alternative splicing of type II procollagen pre-mRNA in chondrocytes is oppositely regulated by BMP-2 and TGF-beta1. FEBS Lett 545, 115–119.

Wahl, M.C., Will, C.L., and Lührmann, R. (2009). The spliceosome: design principles of a dynamic RNP machine. Cell 136, 701–718.

Warner, L.R., Brown, R.J., Yingst, S.M.C., and Oxford, J.T. (2006). Isoform-specific heparan sulfate binding within the amino-terminal noncollagenous domain of collagen alpha1(XI). J Biol Chem 281, 39507–39516.

Wenstrup, R.J., Florer, J.B., Cole, W.G., Willing, M.C., and Birk, D.E. (2004). Reduced type I collagen utilization: a pathogenic mechanism in COL5A1 haplo-insufficient Ehlers-Danlos syndrome. J Cell Biochem 92, 113–124.

Wenstrup, R.J., Florer, J.B., Willing, M.C., Giunta, C., Steinmann, B., Young, F., Susic, M. and Cole, W.G. (2000). COL5A1 haploinsufficiency is a common molecular mechanism underlying the classical form of EDS. Am J Hum Genet 66, 1766–1776.

Xu, L., Peng, H., Wu, D., Hu, K., Goldring, M.B., Olsen, B.R., and Li, Y. (2005). Activation of the discoidin domain receptor 2 induces expression of matrix metalloproteinase 13 associated with osteoarthritis in mice. J Biol Chem 280, 548–555.

Yamaguchi, K., Matsuo, N., Sumiyoshi, H., Fujimoto, N., Iyama, K.-I., Yanagisawa, S., and Yoshioka, H. (2005). Pro-α3(V) collagen chain is expressed in bone and its basic N-terminal peptide adheres to osteosarcoma cells. Matrix Biol 24, 283–294.

Yoshioka, H., Greenwel, P., Inoguchi, K., Truter, S., Inagaki, Y., Ninomiya, Y., and Ramirez, F. (1995a). Structural and functional analysis of the promoter of the human alpha 1(XI) collagen gene.

J Biol Chem 270, 418–424.

Yoshioka, H., Iyama, K.-I., Inoguchi, K., Khaleduzzaman, M., Ninomiya, Y., and Ramirez, F. (1995b). Developmental pattern of expression of the mouse alpha 1 (XI) collagen gene (Col11a1). Dev Dyn 204, 41–47.

Zhang, X., Boot-Handford, R.P., Huxley-Jones, J., Forse, L.N., Mould, A.P., Robertson, D.L., Li, L., Athiyal, M., and Sarras, M.P. Jr. (2007). The Collagens of Hydra Provide Insight into the Evolution of Metazoan Extracellular Matrices. J Biol Chem 282, 6792–6802.

Zhu, Y., Oganesian, A., Keene, D.R., and Sandell, L.J. (1999). 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 144, 1069–1080.

Zoppi, N., Gardella, R., De Paepe, A., Barlati, S., and Colombi, M. (2004). Human fibroblasts with mutations in COL5A1 and COL3A1 genes do not organize collagens and fibronectin in the extracellular matrix, down-regulate α2β1 integrin, and recruit alphavbeta3 Instead of α5β1 integrin. J Biol Chem 279, 18157–18168.