

THE STRUCTURE AND FUNCTION OF CARTILAGE PROTEOGLYCANS

P.J. Roughley*

Genetics Unit, Shriners Hospital for Children and Department of Surgery, McGill University, Montreal, Quebec, Canada

Abstract

Cartilage contains a variety of proteoglycans that are essential for its normal function. These include aggrecan, decorin, biglycan, fibromodulin and lumican. Each proteoglycan serves several functions that are determined by both its core protein and its glycosaminoglycan chains. This review discusses the structure/function relationships of the cartilage proteoglycans, and the manner in which perturbations in proteoglycan structure or abundance can adversely affect tissue function.

Key Words: Proteoglycan, aggrecan, link protein, hyaluronan, decorin, biglycan, fibromodulin, lumican, cartilage.

Introduction

Cartilagenous tissues are present throughout the body at numerous sites and are classified histologically as being hyaline, elastic or fibrocartilagenous in nature, depending on their molecular composition. Elastic cartilage is associated with the ear and the larynx, whereas fibrocartilage is associated with the menisci of the knee and the intervertebral discs. Hyaline cartilage is the predominant form of cartilage, and is most commonly associated with the skeletal system, where it forms the anlage for many bones in the embryo, the growth plate *via* which many bones increase their size during juvenile development, and the bearing surface upon which bones articulate in movable joints. It is not clear whether articular cartilage can be defined as a distinct tissue in the juvenile, as it forms a continuum with the hyaline cartilage of the central epiphysis, which is destined to be resorbed. Indeed, it is still somewhat of a mystery why articular cartilage survives at all and is not also replaced by bone.

At first sight, articular cartilage seems ill-equipped to deal with the constant mechanical insults to which it is exposed during normal joint function. It has no vascular system to provide nutrition and is only sparsely populated with cells, which are required to maintain the integrity of the vast extracellular matrix in regions that may be very remote. Some of these concerns are not applicable to the growth plate, which is penetrated by blood vessels and which possesses a higher cell density. However, the growth plate does have the additional problem of maintaining a precise columnar organization of its maturing cells to enable linear growth. The nutritional and repair problems of articular cartilage are magnified in the intervertebral discs, which in the lumbar spine are the largest avascular and least cellular tissues in the body. Nature's solution to these problems resides in the unique structure of the extracellular matrix. In articular cartilage and intervertebral disc it provides the ability to resist compressive loading so protecting the cells from injury by mechanical trauma. In the growth plate, it provides the scaffold that maintains the unique cellular organization responsible for long bone elongation.

All hyaline cartilages are characterized by their high content of the proteoglycan aggrecan, which exists in the form of proteoglycan aggregates in association with hyaluronan (HA) and link protein (LP). The proteoglycan aggregates are responsible for the turgid nature of the cartilage and in articular cartilage they provide the osmotic properties needed to resist compressive loads. In common with all other connective tissues, cartilages also contain a variety of small leucine-rich repeat proteoglycans (SLRPs), which on a weight basis are a minor component

*Address for correspondence:

P.J. Roughley
Genetics Unit,
Shriners Hospital for Children,
1529 Cedar Avenue,
Montreal, Quebec,
H3G 1A6, Canada

Telephone Number: 1-514-282-7156

FAX Number: 1-514-842-5581

E-mail: proughley@shriners.mcgill.ca

of the tissue but on a molar basis may rival the abundance of aggrecan. Decorin, biglycan, fibromodulin and lumican are the major SLRPs present in cartilage, where they help maintain the integrity of the tissue and modulate its metabolism.

Abbreviations

CS	chondroitin sulphate
CS1/2	CS-rich domains of aggrecan
DS	dermatan sulphate
EGF	epidermal growth factor
G1-3	globular domains of aggrecan
GAG	glycosaminoglycan
HA	hyaluronan (hyaluronic acid)
HAS	hyaluronan synthase
IGD	interglobular domain
KS	keratan sulphate
LP	link protein
LRR	leucine-rich repeat
MMP	matrix metalloproteinase
SLRP	small leucine-rich repeat proteoglycan
TGF	transforming growth factor
TNF	tumor necrosis factor.

Aggrecan

Aggrecan is a modular proteoglycan with multiple functional domains. Its core protein consists of three globular regions (Doege *et al.*, 1991), termed G1, G2 and G3 (Fig. 1), each containing cysteine residues that participate in disulphide bond formation (Sandy *et al.*, 1990). The G1 and G2 regions are separated by a short interglobular domain (IGD), and the G2 and G3 regions are separated by a long glycosaminoglycan (GAG)-attachment region, which consists of adjacent domains rich in keratan sulphate (KS) and chondroitin sulphate (CS). The G1 region is at the amino terminus of the core protein, and can be further sub-divided into three functional domains, termed A, B1 and B2, with the B-type domains being responsible for the interaction with HA (Watanabe *et al.*, 1997). The G2 region also possesses two B-type domains, but does not appear to interact with HA (Fosang and Hardingham, 1989), and at present its function is unknown. The G3 region resides at the carboxy terminus of the core protein and contains a variety of distinct structural domains. It is essential for normal post-translational processing of the aggrecan core protein and subsequent aggrecan secretion (Zheng *et al.*, 1998).

The human aggrecan gene, which resides at chromosome 15q26 (Korenberg *et al.*, 1993), consists of 19 exons (Vallmu *et al.*, 1995), with each exon encoding a distinct structural domain of the core protein. Exons 3-6 encode the G1 region and exons 8-10 encode the G2 region. The CS-rich domain and much of the KS-rich domain are encoded by the large exon 12. The G3 region is encoded by exons 13-18. It can give rise to different aggrecan transcripts due to alternative splicing (Fülöp *et al.*, 1993), though it is not clear whether this is of any functional consequence.

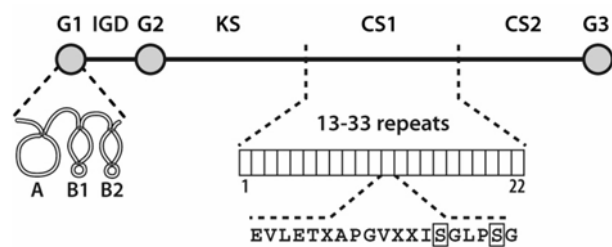


Figure 1. The structural domains of aggrecan. The aggrecan core protein is depicted with three disulphide-bonded globular domains (G1-3), an interglobular domain (IGD), and attachment regions for keratan sulphate (KS) and chondroitin sulphate (CS1 and CS2). The G1 domain is composed of three functional subdomains responsible for the interaction with link protein (A) or hyaluronan (B1 and B2). The CS1 domain exhibits length polymorphism due to a variable number of 19 amino acid tandem repeats. Each repeat contains two serine residues (S) that can act as attachment sites for chondroitin sulphate.

The GAG-attachment region is composed of three domains responsible for the attachment of KS and CS. The KS-attachment domain resides adjacent to the G2 region and is composed largely of repeat motifs whose number varies between species (Barry *et al.*, 1994). The neighbouring CS-attachment domain is divided into two subdomains – the CS1 and CS2 domains. The CS1 domain lies adjacent to the KS-rich domain and is also composed largely of repeat motifs whose number varies between species. In addition, the human CS1 domain exhibits size polymorphism between individuals due to a variable number of 19 amino acid repeats (Fig. 1) (Doege *et al.*, 1997). This results in the aggrecan molecules of different individuals being able to bear different numbers of CS chains. Irrespective of the number of CS chains present, their structure varies throughout life due to changes in length and sulphation pattern (Roughley and White, 1980), though the functional consequence of this change is not clear. The GAG-attachment region also possesses sites for the attachment of O-linked oligosaccharides (Nilsson *et al.*, 1982), which with age may become substituted with KS (Santer *et al.*, 1982). KS may also be present within the G1 region, the IGD and the G2 region, attached to either O-linked or N-linked oligosaccharides (Barry *et al.*, 1995). The KS chains also show age-related changes in structure (Brown *et al.*, 1998).

Aggrecan molecules do not exist in isolation within the extracellular matrix, but as proteoglycan aggregates (Hascall, 1988). Each aggregate is composed of a central filament of HA with up to 100 aggrecan molecules radiating from it, with each interaction able to be stabilized by the presence of a link protein (Morgelin *et al.*, 1988). Proteoglycan aggregate structure is influenced by three parameters – the length of the HA, the proportion of link protein, and the degree of aggrecan processing. Two molecular forms have been described for proteoglycan aggregates extracted directly from cartilage without the use of dissociative agents (Manicourt *et al.*, 1991). These

forms appear to differ in their link protein content, and may have different functional characteristics (Buckwalter *et al.*, 1994). It is the large size of the proteoglycan aggregates and their entrapment by the collagen framework of the tissue that results in aggrecan retention in the extracellular matrix.

Aggrecan molecules rarely exist in an intact form in the proteoglycan aggregates of the cartilage matrix, but instead are subject to extracellular proteolytic processing of their core proteins (Fig. 2). This results in the accumulation of fragments that bear the G1 regions and the loss of those that do not by diffusion from the tissue, ultimately yielding proteoglycan aggregates that are enriched in aggrecan G1 regions rather than more intact molecules. The G1 regions may accumulate in the cartilage matrix for many years (Maroudas *et al.*, 1998). Aggrecanases and matrix metalloproteinases (MMPs) are associated with aggrecan proteolysis (Sztrolovics *et al.*, 1997). The aggrecanases (Abbaszade *et al.*, 1999; Tortorella *et al.*, 1999) are of particular interest because of their selectivity for aggrecan. Five aggrecanase cleavage sites have been described in aggrecan (Tortorella *et al.*, 2000), with one residing in the IGD domain and four in the CS2 domain.

The GAG-attachment region provides the high anionic charge density needed for the unique osmotic properties of aggrecan. Normal cartilage function depends on a high aggrecan content, high GAG substitution and large aggregate size. Loss of cartilage integrity in arthritis is associated with impaired aggrecan function due either to proteolytic cleavage of the aggrecan core protein, which decreases aggrecan charge, or to cleavage of the HA, which decreases aggregate size. It has also been suggested that aggrecan charge and hence function could be affected by size polymorphism within the CS1 domain, as those individuals with the shortest core protein length would possess aggrecan with the lowest CS substitution. Such individuals might be at risk for premature cartilage degeneration. CS2 domain processing by aggrecanases would result in aggrecan fragments enriched in the CS1 domain and therefore enhance any influence that CS1 domain polymorphism may have on aggrecan function. While size polymorphism in the aggrecan CS1 domain has been associated with both articular cartilage and intervertebral disc degeneration (Horton *et al.*, 1998; Kawaguchi *et al.*, 1999), the reason for the linkage is not clear as it is not always the shorter CS1 domains that have been associated with disease. It is possible that the presence of one short aggrecan allele may be of little functional consequence, and only individuals with two short alleles would be at risk. Such individuals represent less than 1% of the population and have not been the focus of any study reported to date.

Mutations in the aggrecan gene leading to chondrodysplasias have been described in the human, mouse and chicken. In the human, a single base pair insertion in exon 12 causes a frameshift and results in a form of spondyloepiphyseal dysplasia (Gleghorn *et al.*, 2005). In the mouse, a 7 bp deletion in exon 5 causes a frameshift and results in a premature termination codon arising in exon 6 (Watanabe *et al.*, 1994). In the chicken, a

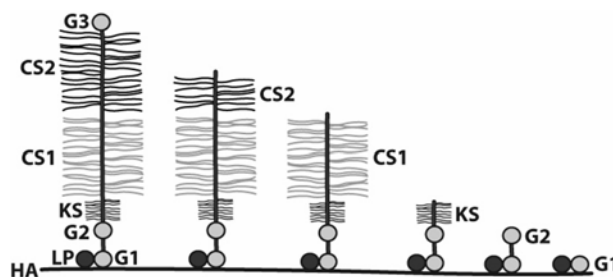


Figure 2. Proteolytic processing of aggrecan. Aggrecan molecules are depicted interacting with hyaluronan (HA) and link protein (LP) to form proteoglycan aggregates. Proteolytic cleavage of the aggrecan core protein generates one fragment which remains bound to HA and LP and is retained in the tissue and a second fragment which is not. The size of the retained fragment varies depending on the site of cleavage, with the smallest fragment representing the free G1 domain.

premature stop codon is present within exon 10 encoding the CS-attachment region (Li *et al.*, 1993), resulting in decreased message accumulation and underproduction of a truncated aggrecan. It is likely that the absence of a G3 region impairs secretion of the mutant aggrecan molecules. In the human, chondrodystrophic phenotypes have also been associated with the undersulphation of aggrecan due to gene defects in a sulphate transporter (Superti-Furga *et al.*, 1996). These disorders illustrate the importance of aggrecan content and charge in embryonic cartilage development and growth.

Link Protein

Link proteins have a structure analogous to that of the G1 region of aggrecan, possessing A, B1 and B2 domains and are thus members of the hyalactan family, the family of proteoglycans to which aggrecan belongs (Neame and Barry, 1993). The A domain has been shown to be responsible for interaction with the G1 region of aggrecan, whereas the B domains interact with HA (Grover and Roughley, 1994). Intact human cartilage link protein can exist in two molecular forms (Fig. 3), termed LP1 and LP2, which differ by the presence of two or one N-linked oligosaccharide chains, respectively, in their N-terminal regions (Roughley *et al.*, 1982). Proteolytic modification of the link proteins can occur in the region between the two oligosaccharide chains to yield a product termed LP3 (Mort *et al.*, 1985). When the link proteins are present in a proteoglycan aggregate this site tends to be the only one that is accessible to most proteinases. Additional proteolysis does take place *in vivo* within the A domain, to produce LP fragments. This type of processing has been mimicked *in vitro* by the action of free radicals. Proteolytic modification of link protein takes place throughout life (Mort *et al.*, 1983), with the modified link proteins accumulating in the proteoglycan aggregates. Matrix metalloproteinases have been implicated in such degradation (Nguyen *et al.*, 1993), but the link proteins are resistant to aggrecanases (Sztrolovics *et al.*, 1999).

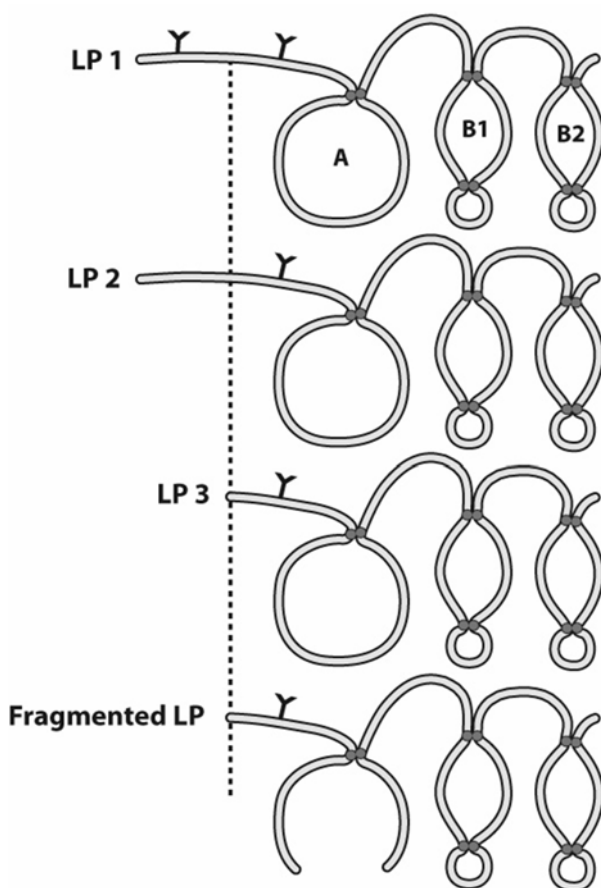


Figure 3. The structural domains of link protein. Link protein is depicted with three disulphide-bonded domains responsible for the interaction with aggrecan (A) or hyaluronan (B1 and B2). The intact link protein may possess two (LP1) or one (LP2) N-linked oligosaccharides in its amino terminal region. Proteolytic cleavage may occur within the amino terminal region to yield a truncated link protein (LP3). Cleavage may also occur within the A domain to yield fragmented LP.

Mammals possess four link protein genes (Spicer *et al.*, 2003), of which one is predominantly expressed in cartilage. Each LP gene is adjacent to a hyalactan gene, though it does not appear that adjacent gene products are necessarily co-expressed, as the cartilage LP gene is not adjacent to that of aggrecan. The cartilage LP serves several functions in the proteoglycan aggregate (Fig. 4). First by virtue of its ability to interact with both HA and the G1 domain of aggrecan it stabilizes the proteoglycan aggregate and prevents its dissociation under physiological conditions. Second, it participates in a phenomenon termed “delayed aggregation”, as newly secreted aggrecan has little ability to interact with HA and requires a LP-mediated conformational change in its G1 domain to promote aggregate formation (Melching and Roughley, 1990). Third, together with the G1 domain of aggrecan, LP forms a protein coat covering the surface of HA. This coat helps protect the HA from undesirable degradation by both hyaluronidases and free radicals.

The human cartilage link protein gene, which resides at chromosome 5q13-14 (Osborne-Lawrence *et al.*, 1990),

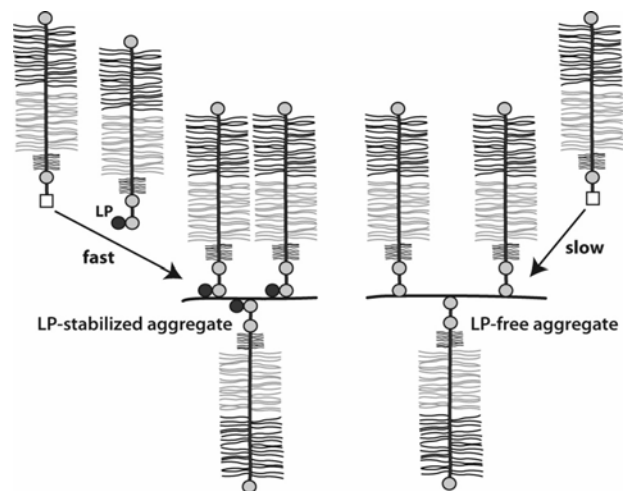


Figure 4. The formation of proteoglycan aggregates. Newly synthesized aggrecan molecules do not interact well with hyaluronan, but require a conformational change within the G1 region to facilitate interaction. The rate of the conformational change is enhanced by the interaction of link protein (LP) with the nascent aggrecan. LP-stabilized aggregates contains densely packed aggrecan molecules which are tightly bound, whereas LP-free aggregates contain more widely spaced aggrecan which is more susceptible to dissociation.

possesses 5 exons, with exon 3 encoding the A domain and exons 4 and 5 encoding the B domains (Dudhia *et al.*, 1994). The importance of link protein in maintaining the structure of the proteoglycan aggregates has been demonstrated in knockout mice, which exhibit cartilage defects resulting in dwarfism and craniofacial abnormalities (Watanabe and Yamada, 1999). To date no human disorder has been directly linked to a mutation in link protein, but its importance in maintaining cartilage function is not in question.

Hyaluronan

HA is a non-sulphated GAG characterized by its large length and unique mode of synthesis (Fraser *et al.*, 1997), which occurs at the plasma membrane of the cell *via* a hyaluronan synthase (HAS) (Weigel *et al.*, 1997). Mammals possess three HAS genes, termed HAS1, HAS2 and HAS3, each residing at a different chromosome location (Spicer *et al.*, 1997). Because of its mode of synthesis, HA is extruded directly into the extracellular space (Fig. 5), and is present as a coat surrounding all chondrocytes. It is likely that proteoglycan aggregate formation initially occurs in this pericellular location. The mechanism whereby the proteoglycan aggregates are released from this environment and move to the more remote parts of the extracellular matrix is not clear. In cartilage, HA size decreases and its abundance increases with age (Holmes *et al.*, 1988). The decrease in size appears to be by degradation rather than synthesis, and could involve the action of hyaluronidases or free radicals.

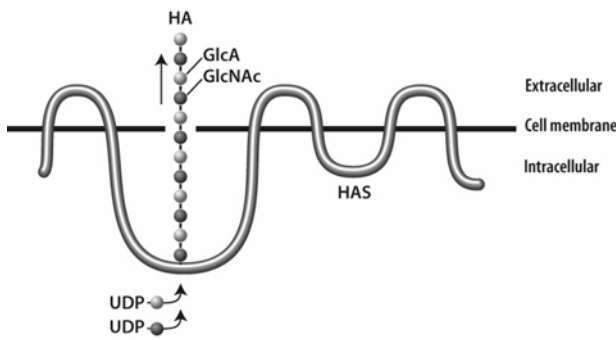


Figure 5. Hyaluronan synthesis. Hyaluronan (HA) is a polysaccharide consisting of repeating disaccharide units of glucuronic acid (GlcA) and N-acetyl glucosamine (GlcNAc). It is synthesized by a hyaluronan synthase (HAS) residing in the plasma membrane of the cell utilizing cytoplasmic UDP-bound monosaccharides. The growing HA chain is extruded directly into the extracellular space *via* a pore in the cell membrane.

While each of the three HA synthases presumably produces HA of an identical composition, there are differences in the chain length of the product and the ease with which it can be released from the cell surface (Itano *et al.*, 1999). Thus, the properties of the extracellular matrix could differ depending on the HAS being utilized. HAS1, HAS2 and HAS3 messages are all expressed in cartilage (Recklies *et al.*, 2001). HAS2 message has been reported to have the highest expression levels and HAS3 message the lowest levels in human chondrocytes, though the different messages exhibit variation in their regulation by growth factors and cytokines. HAS2 appears to be by far the most important synthase during embryonic development, as mice lacking expression of this gene die during mid-gestation (Camenisch *et al.*, 2000). In contrast, those lacking the HAS1 or HAS3 genes are viable and fertile and exhibit no observable consequence on cartilage. However, it is not clear from these knockout experiments whether absence of the HAS2 gene has a direct influence on any cartilage type, as foetal development does not proceed. Recent studies on cartilage-specific HAS2 knockout mice demonstrate that the absence of HAS2 gene expression in cartilage is catastrophic; resulting in a severely abnormal skeleton and perinatal death (Fig. 6). This underlies the essential role of the hyaluronan and proteoglycan aggregates in both normal cartilage function and the chondrogenic differentiation associated with endochondral bone formation. It also illustrates that, at least in cartilage, HAS1 or HAS3 expression cannot compensate for a deficiency in HAS2.

SLRPs

The SLRPs are not structurally related to aggrecan, but belong to the large family of leucine-rich repeat proteins, which are characterized by multiple adjacent domains bearing a common leucine-rich motif (Hocking *et al.*, 1998). They can be divided into several sub-families based

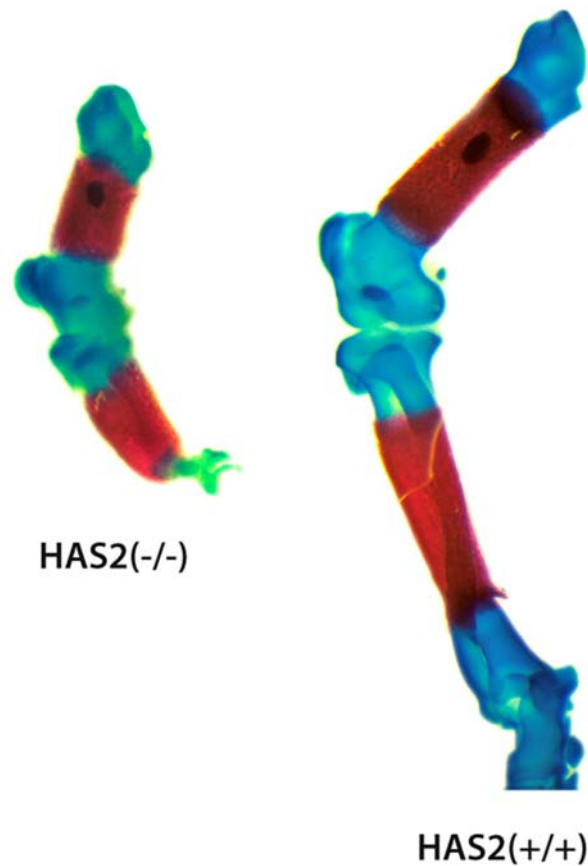


Figure 6. The consequence of HAS2 ablation in cartilage. The figure shows the femur, tibia and knee joint from transgenic mice lacking HAS2 expression in their cartilagenous tissues (HAS2^{-/-}) and wild type mice with normal HAS2 expression (HAS2^{+/+}). The skeletal elements have been stained with alizarin red and alcian blue to identify bone and cartilage, respectively.

on their gene organization, the number of leucine-rich repeats and the type of GAG chain substituent. In the case of decorin, biglycan, fibromodulin and lumican there are ten leucine-rich repeats, which are flanked by disulphide-bonded domains (Fig. 7). Decorin and biglycan are part of one subfamily based on their 8 exon gene structure, whereas fibromodulin and lumican are part of a second subfamily based on their 3 exon gene structure. The genes for decorin and lumican reside close to one another at human chromosome 12q22 (Danielson *et al.*, 1993; Grover *et al.*, 1995), while the gene for fibromodulin is at 1q32 (Sztrolovics *et al.*, 1994) and that for biglycan is at Xq28 (McBride *et al.*, 1990). Decorin and biglycan can also be classified as dermatan sulphate (DS)-proteoglycans, whereas fibromodulin and lumican are keratan sulphate (KS)-proteoglycans. The mature forms of fibromodulin and lumican present in the cartilage matrix correspond to removal of only the signal peptide, whereas for decorin and biglycan an additional amino acid sequence of 14 and 21 amino acids, respectively, is removed (Roughley *et al.*, 1996b; Scott *et al.*, 2000). These additional sequences have been considered as propeptides, though it is not clear whether their removal has any functional consequence.

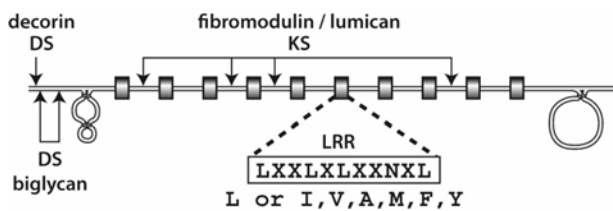


Figure 7. The structural domains of SLRPs. The core protein of decorin, biglycan, fibromodulin or lumican is depicted with two disulphide bonded domains flanking ten leucine-rich repeat (LRR) domains. Each LRR domain is composed of the sequence LXXXLXXNXL, where L is predominantly leucine but may also be I, V, A, M, F or Y, and X may be any amino acid. In the case of decorin or biglycan, one or two chondroitin/dermatan sulphate (DS) chains, respectively, reside in the amino terminal region. In the case of fibromodulin and lumican, one to four keratan sulphate chains (KS) may reside between the LRR domains.

The four SLRPs all possess N-linked oligosaccharide chains within their central leucine-rich repeats (Neame *et al.*, 1989; Plaas *et al.*, 1990). In the case of fibromodulin and lumican, it is these N-linked oligosaccharides that may be modified to KS (Fig. 7). Decorin and biglycan possess attachment sites for CS/DS within the extreme amino terminus of their core proteins. In decorin, there is one such site at amino acid residue 4 of the mature core protein, whereas in biglycan there are two sites at amino acids 5 and 10 in the human (Roughley and White, 1989). In most connective tissues including cartilage the CS is modified to DS. Biglycan may also exist in non-glycanated forms devoid of DS chains, with the abundance of such forms accumulating with age (Roughley *et al.*, 1993). These non-glycanated forms of biglycan appear to be the result of proteolysis within the amino terminal region of the core protein. Non-glycanated forms of fibromodulin and lumican can also accumulate with age (Grover *et al.*, 1995; Roughley *et al.*, 1996a), due to the absence of KS synthesis. *In vitro*, decorin (Imai *et al.*, 1997) and fibromodulin (Heathfield *et al.*, 2004) have been shown to be degraded by matrix metalloproteinases, and it is predicted that such cleavages can occur *in vivo*.

The function of the SLRPs depends on both their core protein and GAG chains. Their core proteins allow the SLRPs to interact with the fibrillar collagen that forms the framework of the tissue (Vogel *et al.*, 1984). In so doing they help regulate fibril diameter during its formation and possibly fibril-fibril interaction in the extracellular matrix. They also appear to limit access of the collagenases to their unique cleavage site on each collagen molecule, and in so doing may help protect the fibrils from proteolytic damage (Fig. 8). While the interaction of decorin, fibromodulin and lumican with collagen fibrils is universally accepted, that of biglycan is not. At least *in vitro*, biglycan interaction depends upon the environmental conditions, and this might contribute to the different tissue locations that have been observed for decorin and biglycan.

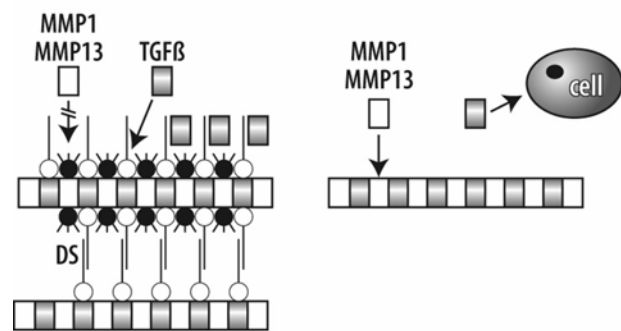


Figure 8. Interaction and function of SLRPs. Decorin (open circles) and fibromodulin or lumican (solid circles) are depicted interacting with different regions of a collagen fibril to form a surface coat. The dermatan sulphate chains (DS) of decorin may self-associate to facilitate fibril-fibril interaction or may interact with TGFβ sequestering the growth factor in the extracellular matrix. SLRP coat formation can impede the access of collagenases (MMP1 and MMP13) to the fibril surface and retard its degradation. In the absence of SLRPs the collagen fibril would be more susceptible to degradation and TGFβ would be freely able to interact with its cell surface receptors to modulate cell metabolism.

Fibromodulin and lumican interact with the same region of the collagen molecule (Svensson *et al.*, 2000), which is distinct from the site at which decorin interacts (Hedbom and Heinegård, 1993). For decorin, interaction involves amino acid sequences present within the leucine-rich repeats (Svensson *et al.*, 1995). Molecular modelling predicts that the SLRPs possess a “horse-shoe” conformation that is able to accommodate a single collagen molecule at the surface of the collagen fibrils within its concave face (Scott, 1996). However, X-ray diffraction analysis of decorin crystals indicate that it exists as a dimer with interlocking concave faces (Scott *et al.*, 2004). It is not clear whether decorin dimers represent the functional form of the molecule in solution and how this impacts their interaction with other molecules.

Interaction of the SLRPs is not confined to the fibrillar collagens. Decorin, biglycan, fibromodulin and lumican have been reported to interact with many other macromolecules, including types VI, XII and XIV collagen, fibronectin and elastin, and growth factors such as EGF, TGFβ and TNFα. At least in the case of decorin, it appears that interaction with fibrinogen occurs in a zinc-dependent manner (Dugan *et al.*, 2003), and it is possible that other family members also act as Zn-binding proteins, and that this influences their conformation and function. The GAG chains have been associated with the interaction with several growth factors, and enable the SLRPs to provide a sink for growth factor accumulation within the extracellular matrix. In this manner the SLRPs can help modulate chondrocyte metabolism by regulating growth factor access to the cells.

The level of SLRP synthesis varies with age, and may be influenced by growth factors. While the precise consequence of the age-related changes in SLRP

abundance is unclear, it is clear that depletion in SLRP production can influence tissue properties. This is best illustrated by the abnormal phenotypes arising in “knockout” mice. Absence of decorin results in lax, fragile skin, in which collagen fibril morphology is irregular with fusion of adjacent fibrils appearing to have occurred (Danielson *et al.*, 1997). Absence of biglycan results in an osteoporosis-like phenotype, with animals having a reduced growth rate and a decreased bone mass (Xu *et al.*, 1998). Absence of lumican produces both skin laxity and corneal opacity, with an increased proportion of abnormally thick collagen fibrils, and delayed corneal epithelial wound healing (Chakravarti *et al.*, 1998). Absence of fibromodulin produces no change in the appearance of the mice, but does result in an abnormal collagen fibril organization in tendons (Svensson *et al.*, 1999). This work clearly shows that collagen fibril architecture is impaired in tissues in which SLRPs are deficient, and that the abnormal phenotype is specific for each proteoglycan. Of these SLRP genes, only that for decorin has currently been linked to a human disorder, with a frameshift mutation being reported in congenital stromal dystrophy of the cornea (Bredrup *et al.*, 2005). It is also evident that the absence of GAG synthesis on the SLRPs can have detrimental consequences, as deficiency in CS/DS substitution of decorin due to mutation in a glycosyl transferase gene has been associated with the progeroid form of Ehlers-Danlos syndrome (Götte and Kresse, 2005).

Conclusion

The cartilage proteoglycans are essential for normal tissue function, and perturbation in their abundance or structure can have dire consequences on cartilage formation in the embryo, growth plate function in the juvenile, and articular cartilage function in the adult.

Acknowledgements

The author wishes to thank the Shriners of North America, the Canadian Institutes of Health Research and the Arthritis Society of Canada for financial support.

References

- Abbaszade I, Liu RQ, Yang F, Rosenfeld SA, Ross OH, Link JR., Ellis DM, Tortorella MD, Pratta MA, Hollis JM, Wynn R, Duke JL, George HJ, Hillman MC, Murphy K, Wiswall BH, Copeland RA, Decicco CP, Bruckner R, Nagase H, Itoh Y, Newton RC, Magolda RL, Trzaskos JM (1999) Cloning and characterization of *ADAMTS11*, an aggrecanase from the ADAMTS family. *J Biol Chem* **274**: 23443-23450.
- Barry FP, Neame PJ, Sasse J, Pearson D (1994) Length variation in the keratan sulfate domain of mammalian aggrecan. *Matrix* **14**: 323-328.
- Barry FP, Rosenberg LC, Gaw JU, Koob TJ, Neame PJ (1995) *N*- and *O*-linked keratan sulfate on the hyaluronan binding region of aggrecan from mature and immature bovine cartilage. *J Biol Chem* **270**: 20516-20524.
- Bredrup C, Knappskog PM, Majewski J, Rodahl E, Boman H (2005) Congenital stromal dystrophy of the cornea caused by a mutation in the decorin gene. *Invest Ophthalmol Vis Sci* **46**: 420-426.
- Brown GM, Huckerby TN, Bayliss MT, Nieduszynski IA (1998) Human aggrecan keratan sulfate undergoes structural changes during adolescent development. *J Biol Chem* **273**: 26408-26414.
- Buckwalter JA, Pita JC, Muller FJ, Nessler J (1994) Structural differences between two populations of articular cartilage proteoglycan aggregates. *J Orthop Res* **12**: 144-148.
- Camenisch TD, Spicer AP, Brehm-Gibson T, Biesterfeldt J, Augustine ML, Calabro A, Kubalak S, Klewer SE, McDonald JA (2000) Disruption of hyaluronan synthase-2 abrogates normal cardiac morphogenesis and hyaluronan-mediated transformation of epithelium to mesenchyme. *J Clin Invest* **106**: 349-360.
- Chakravarti S, Magnuson T, Lass JH, Jepsen KJ, LaMantia C, Carroll H (1998) Lumican regulates collagen fibril assembly: skin fragility and corneal opacity in the absence of lumican. *J Cell Biol* **141**: 1277-1286.
- Danielson KG, Fazio A, Cohen I, Cannizzaro LA, Eichstetter I, Iozzo RV (1993) The human decorin gene: Intron-exon organization, discovery of two alternatively spliced exons in the 5' untranslated region, and mapping of the gene to chromosome 12q23. *Genomics* **15**: 146-160.
- Danielson KG, Baribault H, Holmes DF, Graham H, Kadler KE, Iozzo RV (1997) Targeted disruption of decorin leads to abnormal collagen fibril morphology and skin fragility. *J Cell Biol* **136**: 729-743.
- Doerge KJ, Sasaki M, Kimura T, Yamada Y (1991) Complete coding sequence and deduced primary structure of the human cartilage large aggregating proteoglycan, aggrecan. Human-specific repeats, and additional alternatively spliced forms. *J Biol Chem* **266**: 894-902.
- Doerge KJ, Coulter SN, Meek LM, Maslen K, Wood JG (1997) A human-specific polymorphism in the coding region of the aggrecan gene – Variable number of tandem repeats produce a range of core protein sizes in the general population. *J Biol Chem* **272**: 13974-13979.
- Dudhia J, Bayliss MT, Hardingham TE (1994) Human link protein gene: Structure and transcription pattern in chondrocytes. *Biochem J* **303**: 329-333.
- Dugan TA, Yang VWC, McQuillan DJ, Höök M (2003) Decorin binds fibrinogen in a Zn²⁺-dependent interaction. *J Biol Chem* **278**: 13655-13662.
- Fosang AJ, Hardingham TE (1989) Isolation of the N-terminal globular protein domains from cartilage proteoglycans. Identification of G2 domain and its lack of interaction with hyaluronate and link protein. *Biochem J* **261**: 801-809.
- Fraser JRE, Laurent TC, Laurent UBG (1997) Hyaluronan: Its nature, distribution, functions and turnover. *J Intern Med* **242**: 27-33.
- Fülöp C, Walcz E, Valyon M, Glant TT (1993) Expression of alternatively spliced epidermal growth factor-like domains in aggrecans of different species.

- Evidence for a novel module. *J Biol Chem* **268**: 17377-17383.
- Gleghorn L, Ramesar R, Beighton P, Wallis G (2005) A mutation in the variable repeat region of the aggrecan gene (*AGC1*) causes a form of spondyloepiphyseal dysplasia associated with severe, premature osteoarthritis. *Am J Hum Genet* **77**: 484-490.
- Götte M, Kresse H (2005) Defective glycosaminoglycan substitution of decorin in a patient with progeroid syndrome is a direct consequence of two point mutations in the galactosyltransferase I (b4GalT-7) gene. *Biochem Genet* **43**: 65-77.
- Grover J, Roughley PJ (1994) The expression of functional link protein in a baculovirus system: Analysis of mutants lacking the A, B and B' domains. *Biochem J* **300**: 317-324.
- Grover J, Chen XN, Korenberg JR, Roughley PJ (1995) The human lumican gene – Organization, chromosomal, location, and expression in articular cartilage. *J Biol Chem* **270**: 21942-21949.
- Hascall VC (1988) Proteoglycans: the chondroitin sulfate/keratan sulfate proteoglycan of cartilage. *ISI Atlas Sci Biochem* **1**: 189-198.
- Heathfield TF, Onnerfjord P, Dahlberg L, Heinegård D (2004) Cleavage of fibromodulin in cartilage explants involves removal of the N-terminal tyrosine sulfate-rich region by proteolysis at a site that is sensitive to matrix metalloproteinase-13. *J Biol Chem* **279**: 6286-6295.
- Hedbom E, Heinegård D (1993) Binding of fibromodulin and decorin to separate sites on fibrillar collagens. *J Biol Chem* **268**: 27307-27312.
- Hocking AM, Shinomura T, McQuillan DJ (1998) Leucine-rich repeat glycoproteins of the extracellular matrix. *Matrix Biol* **17**: 1-19.
- Holmes MWA, Bayliss MT, Muir H (1988) Hyaluronic acid in human articular cartilage. Age-related changes in content and size. *Biochem J* **250**: 435-441.
- Horton WE, Lethbridge-Çejku M, Hochberg MC, Balakir R, Precht P, Plato CC, Tobin JD, Meek L, Doege K (1998) An association between an aggrecan polymorphic allele and bilateral hand osteoarthritis in elderly white men: data from the Baltimore Longitudinal Study of Aging (BLSA). *Osteoarthritis Cart* **6**: 245-251.
- Imai K, Hiramatsu A, Fukushima D, Pierschbacher MD, Okada Y (1997) Degradation of decorin by matrix metalloproteinases: Identification of the cleavage sites, kinetic analyses and transforming growth factor- β 1 release. *Biochem J* **322**: 809-814.
- Itano N, Sawai T, Yoshida M, Lenas P, Yamada Y, Imagawa M, Shinomura T, Hamaguchi M, Yoshida Y, Ohnuki Y, Miyauchi S, Spicer AP, McDonald JA, Kimata K (1999) Three isoforms of mammalian hyaluronan synthases have distinct enzymatic properties. *J Biol Chem* **274**: 25085-25092.
- Kawaguchi Y, Osada R, Kanamori M, Ishihara H, Ohmori K, Matsui H, Kimura T (1999) Association between an aggrecan gene polymorphism and lumbar disc degeneration. *Spine* **24**: 2456-2460.
- Korenberg JR, Chen XN, Doege K, Grover J, Roughley PJ (1993) Assignment of the human aggrecan gene (*AGC1*) to 15q26 using fluorescence *in situ* hybridization analysis. *Genomics* **16**: 546-548.
- Li H, Schwartz NB, Vertel BM (1993) cDNA cloning of chick cartilage chondroitin sulfate (aggrecan) core protein and identification of a stop codon in the aggrecan gene associated with the chondrodystrophy, nanomelia. *J Biol Chem* **268**: 23504-23511.
- Manicourt DH, Pita JC, Thonar EJM, Howell DS (1991) Proteoglycans nondissociatively extracted from different zones of canine normal articular cartilage: Variations in the sedimentation profile of aggregates with degree of physiological stress. *Connect Tissue Res* **26**: 231-246.
- Maroudas A, Bayliss MT, Uchitel-Kaushansky N, Schneiderman R, Gilav E (1998) Aggrecan turnover in human articular cartilage: Use of aspartic acid racemization as a marker of molecular age. *Arch Biochem Biophys* **350**: 61-71.
- McBride OW, Fisher LW, Young MF (1990) Localization of PGI (biglycan, BGN) and PGII (decorin, DCN, PG-40) genes on human chromosomes Xq13-qter and 12q, respectively. *Genomics* **6**: 219-225.
- Melching LI, Roughley PJ (1990) Studies on the interaction of newly secreted proteoglycan subunits with hyaluronate in human articular cartilage. *Biochim Biophys Acta* **1035**: 20-28.
- Morgelin M, Paulsson M, Hardingham TE, Heinegård D, Engel J (1988) Cartilage proteoglycans. Assembly with hyaluronate and link protein as studied by electron microscopy. *Biochem J* **253**: 175-185.
- Mort JS, Poole AR, Roughley PJ (1983) Age-related changes in the structure of proteoglycan link proteins present in normal human articular cartilage. *Biochem. J* **214**: 269-272.
- Mort JS, Caterson B, Poole AR, Roughley PJ (1985) The origin of human cartilage proteoglycan link protein heterogeneity and fragmentation during ageing. *Biochem J* **232**, 805-812.
- Neame PJ, Barry FP (1993) The link proteins. *Experientia* **49**: 393-402.
- Neame PJ, Choi HU, Rosenberg LC (1989) The primary structure of the core protein of the small, leucine-rich proteoglycan (PG I) from bovine articular cartilage. *J Biol Chem* **264**: 8653-8661.
- Nguyen Q, Murphy G, Hughes CE, Mort JS, Roughley PJ (1993) Matrix metalloproteinases cleave at two distinct sites on human cartilage link protein. *Biochem J* **295**: 595-598.
- Nilsson B, De Luca S, Lohmander S, Hascall VC (1982) Structures of N-linked and O-linked oligosaccharides on proteoglycan monomer isolated from the Swarm rat chondrosarcoma. *J Biol Chem* **257**: 10920-10927.
- Osborne-Lawrence SL, Sinclair AK, Hicks RC, Lacey SW, Eddy RL, Byers MG, Shows TB, DUBY AD (1990) Complete amino acid sequence of human cartilage link protein (*CRTL1*) deduced from cDNA clones and chromosomal assignment of the gene. *Genomics* **8**: 562-567.

- Plaas AHK, Neame PJ, Nivens CM, Reiss L (1990) Identification of the keratan sulfate attachment sites on bovine fibromodulin. *J Biol Chem* **265**: 20634-20640.
- Recklies AD, White C, Melching L, Roughley PJ (2001) Differential regulation and expression of hyaluronan synthases in human articular chondrocytes, synovial cells and osteosarcoma cells. *Biochem J* **354**: 17-24.
- Roughley PJ, White RJ (1980) Age-related changes in the structure of the proteoglycan subunits from human articular cartilage. *J Biol Chem* **255**: 217-224.
- Roughley PJ, White RJ (1989) Dermatan sulphate proteoglycans of human articular cartilage. The properties of dermatan sulphate proteoglycans I and II. *Biochem J* **262**: 823-827.
- Roughley PJ, Poole AR, Mort JS (1982) The heterogeneity of link proteins isolated from human articular cartilage proteoglycan aggregates. *J Biol Chem* **257**: 11908-11914.
- Roughley PJ, White RJ, Magny MC, Liu J, Pearce RH, Mort JS (1993) Non-proteoglycan forms of biglycan increase with age in human articular cartilage. *Biochem J* **295**: 421-426.
- Roughley PJ, White RJ, Cs-Szabó G, Mort JS (1996a) Changes with age in the structure of fibromodulin in human articular cartilage. *Osteoarthritis Cart* **4**: 153-161.
- Roughley PJ, White RJ, Mort JS (1996b) Presence of pro-forms of decorin and biglycan in human articular cartilage. *Biochem J* **318**: 779-784.
- Sandy JD, Flannery CR, Boynton RE, Neame PJ (1990) Isolation and characterization of disulfide-bonded peptides from the three globular domains of aggregating cartilage proteoglycan. *J Biol Chem* **265**: 21108-21113.
- Santer V, White RJ, Roughley PJ (1982) O-linked oligosaccharides of human articular cartilage proteoglycans. *Biochim Biophys Acta* **716**: 277-282.
- Scott IC, Imamura Y, Pappano WN, Troedel JM, Recklies AD, Roughley PJ, Greenspan DS (2000) Bone morphogenetic protein-1 processes probiglycan. *J Biol Chem* **275**: 30504-30511.
- Scott JE (1996) Proteodermatan and proteokeratan sulfate (decorin, lumican/fibromodulin) proteins are horseshoe shaped. Implications for their interactions with collagen. *Biochemistry* **35**: 8795-8799.
- Scott PG, McEwan PA, Dodd CM, Bergmann EM, Bishop PN, Bella J (2004) Crystal structure of the dimeric protein core of decorin, the archetypal small leucine-rich repeat proteoglycan. *Proc Natl Acad Sci USA* **101**: 15633-15638.
- Spicer AP, Seldin MF, Olsen AS, Brown N, Wells DE, Doggett NA, Itano N, Kimata K, Inazawa J, McDonald JA (1997) Chromosomal localization of the human and mouse hyaluronan synthase genes. *Genomics* **41**: 493-497.
- Spicer AP, Joo A, Bowling RA (2003) A hyaluronan binding link protein gene family whose members are physically linked adjacent to chondroitin sulfate proteoglycan core protein genes – The missing links. *J Biol Chem* **278**: 21083-21091.
- Superti-Furga A, Rossi A, Steinmann B, Gitzelmann R (1996) A chondrodysplasia family produced by mutations in the *diastrophic dysplasia sulfate transporter* gene: Genotype/phenotype correlations. *Am J Med Genet* **63**: 144-147.
- Svensson L, Heinegård D, Oldberg Å (1995) Decorin-binding sites for collagen type I are mainly located in leucine-rich repeats 4-5. *J Biol Chem* **270**: 20712-20716.
- Svensson L, Aszódi A, Reinholt FP, Fässler R, Heinegård D, Oldberg Å (1999) Fibromodulin-null mice have abnormal collagen fibrils, tissue organization, and altered lumican deposition in tendon. *J Biol Chem* **274**: 9636-9647.
- Svensson L, Närlid I, Oldberg Å (2000) Fibromodulin and lumican bind to the same region on collagen type I fibrils. *FEBS Lett* **470**: 178-182.
- Sztrolovics R, Chen XN, Grover J, Roughley PJ, Korenberg JR (1994) Localization of the human fibromodulin gene (FMOD) to chromosome 1q32 and completion of the cDNA sequence. *Genomics* **23**: 715-717.
- Sztrolovics R, Alini M, Roughley PJ, Mort JS (1997) Aggrecan degradation in human intervertebral disc and articular cartilage. *Biochem J* **326**: 235-241.
- Sztrolovics R, White RJ, Poole AR, Mort JS, Roughley PJ (1999) Resistance of small leucine-rich repeat proteoglycans to proteolytic degradation during interleukin-1-stimulated cartilage catabolism. *Biochem J* **339**: 571-577.
- Tortorella MD, Burn TC, Pratta MA, Abbaszade I, Hollis JM, Liu R, Rosenfeld SA, Copeland RA, Decicco CP, Wynn R, Rockwell A, Yang F, Duke JL, Solomon K, George H, Bruckner R, Nagase H, Itoh Y, Ellis DM, Ross H, Wiswall BH, Murphy K, Hillman MC, Hollis GF (1999) Purification and cloning of aggrecanase-1: A member of the ADAMTS family of proteins. *Science* **284**: 1664-1666.
- Tortorella MD, Pratta M, Liu RQ, Austin J, Ross OH, Abbaszade I, Burn T, Arner E (2000) Sites of aggrecan cleavage by recombinant human aggrecanase-1 (ADAMTS-4). *J Biol Chem* **275**: 18566-18573.
- Valhmu WB, Palmer GD, Rivers PA, Ebara S, Cheng JF, Fischer S, Ratcliffe A (1995) Structure of the human aggrecan gene: Exon-intron organization and association with the protein domains. *Biochem J* **309**: 535-542.
- Vogel KG, Paulsson M, Heinegård D (1984) Specific inhibition of type I and type II collagen fibrillogenesis by the small proteoglycan of tendon. *Biochem J* **223**: 587-597.
- Watanabe H, Yamada Y (1999) Mice lacking link protein develop dwarfism and craniofacial abnormalities. *Nature Genet* **21**: 225-229.
- Watanabe H, Kimata K, Line S, Strong D, Gao L, Kozak CA, Yamada Y (1994) Mouse cartilage matrix deficiency (*cmd*) caused by a 7 bp deletion in the aggrecan gene. *Nature Genet* **7**: 154-157.
- Watanabe H, Cheung SC, Itano N, Kimata K, Yamada Y (1997) Identification of hyaluronan-binding domains of aggrecan. *J Biol Chem* **272**: 28057-28065.
- Weigel PH, Hascall VC, Tammi M (1997) Hyaluronan synthases. *J Biol Chem* **272**: 13997-14000.
- Xu T, Bianco P, Fisher LW, Longenecker G, Smith E, Goldstein S, Bonadio J, Boskey A, Heegaard AM, Sommer

B, Satomura K, Dominguez P, Zhao C, Kulkarni AB, Gehron Robey P, Young MF (1998) Targeted disruption of the biglycan gene leads to an osteoporosis-like phenotype in mice. *Nature Genet* **20**: 78-82.

Zheng J, Luo W, Tanzer ML (1998) Aggrecan synthesis and secretion – A paradigm for molecular and cellular coordination of multiglobular protein folding and intracellular trafficking. *J Biol Chem* **273**: 12999-13006.

Discussion with Reviewers

C.W. Archer: You state that changes occur in KS chains that are age-related. What are these changes and what is the significance in relation to tissue functionality?

Author: The chain length of KS on aggrecan increases with age, particularly during the period of juvenile growth. During this time there is also a change in the composition of the KS chains, with galactose sulphation increasing and non-reducing terminal sialylation varying. These changes occur concomitantly with a decrease in CS chain length and it has been suggested that the increase in KS length and sulphation serve to maintain the charge density of the aggrecan molecules needed to provide the osmotic properties essential for aggrecan function. This is particularly true in the adult, where proteolytic processing results in the removal of much of the CS2 domain, but retention of the more proximal KS-rich domain. It has been postulated that these changes in GAG substitution may be a consequence of the avascular nature of mature cartilage which results in a low oxygen tension for chondrocyte metabolism. This may limit the ability of the cells to provide the oxidation needed for glucuronic acid production hence limiting CS synthesis, while not perturbing KS synthesis which then has the ability to compensate.

C.W. Archer: In relation to SLRPs and ageing, are the non-glycanated forms non-functional?

Author: Non-glycanated forms of biglycan, fibromodulin and lumican accumulate with age in the cartilage ECM, and represent the predominant forms of these SLRPs in the adult. In the case of fibromodulin and lumican, the absence of KS is due to its lack of synthesis. This is not an inherent property of the mature chondrocytes but rather a result of their ECM environment, as isolated adult chondrocytes produce only the proteoglycan forms of the SLRPs. The absence of the KS chains does not interfere with the core protein interactions of fibromodulin or

lumican, which interact stably with the collagenous framework of the tissue. It is not clear whether the absence of KS is of any functional consequence in the adult or whether its presence is essential during cartilage growth. In the case of biglycan, the absence of CS/DS is due to proteolytic cleavage within the amino terminal region of the core protein and all newly synthesized biglycan is produced in the proteoglycan form. The absence of the amino terminal region does not affect the interaction of the biglycan core protein with other ECM components as the non-glycanated forms are stably retained in the tissue. It is not clear whether the loss of the CS/DS chains has any functional consequence on the cartilage once the initial ECM interactions have taken place.

V. Duance: Could you give a little more information on the different structural domains of the aggrecan G3 region and speculate on the potential consequences of the alternative splicing?

Author: The G3 region of aggrecan is encoded by 6 exons, the first two of which each encode an EGF-like domain, the central three encode a lectin-like domain, and the final one encodes a CRP-like domain. However, the G3 region does not have a unique structure, as alternative splicing can result in the absence or presence of the EGF-like domains and the CRP-like domain. The major transcript present in human cartilage lacks both EGF-like domains but possesses the CRP-like domain, and the abundance of the different splice variants shows little change with age. The G3 region is essential for the synthesis, glycosylation and secretion of aggrecan, particularly the intracellular trafficking of the aggrecan core protein from the ER to the Golgi and the subsequent synthesis of CS chains. In the absence of a functional G3 region aggrecan synthesis is arrested in the ER and the core protein targeted for degradation. The presence of the lectin-like domain is essential for effective aggrecan secretion, but the contribution of the other G3 domains in this process is unclear. Following secretion, the G3 domain allows the aggrecan molecules to interact with other ECM components, such as tenascins and fibulins, *via* the lectin-like domain, and these interactions do seem to be influenced by the presence or absence of the EGF-like and CRP-like domains. However, while the G3 domain may aid in cartilage ECM assembly, its continued presence on the aggrecan molecule does not appear to be essential in the mature cartilage ECM, as aggrecan molecules extracted from the ECM are commonly devoid of the G3 region due to its proteolytic removal.