Available online http://arthritis-research.com/content/3/6/342

Review

Articular cartilage and changes in arthritis
Noncollagenous proteins and proteoglycans in the extracellular matrix of cartilage
Peter J Roughley

Genetics Unit, Shriners Hospital for Children, Montreal, Quebec, Canada

Correspondence: Peter J Roughley, Genetics Unit, Shriners Hospital for Children, 1529 Cedar Avenue, Montreal, Quebec, H3G 1A6, Canada. Tel: +1 514 282 7156; fax: +1 514 842 5581; e-mail: proughley@shriners.mcgill.ca

Received: 27 March 2001
Revisions requested: 23 May 2001
Revisions received: 21 June 2001
Accepted: 14 August 2001
Published: 13 September 2001

Abstract
Cartilage contains numerous noncollagenous proteins in its extracellular matrix, including proteoglycans. At least 40 such molecules have been identified, differing greatly in structure, distribution, and function. Some are present in only selected cartilages or cartilage zones, some vary in their presence with a person’s development and age, and others are more universal in their expression. Some may not even be made by the chondrocytes, but may arise by absorption from the synovial fluid. In many cases, the molecules’ function is unclear, but the importance of others is illustrated by their involvement in genetic disorders. This review provides a selective survey of these molecules and discusses their structure, function, and involvement in inherited and arthritic disorders.

Keywords: cartilage, extracellular matrix, protein, proteoglycan

Introduction
The extracellular matrix of articular cartilage contains a large variety of noncollagenous proteins. Many of these are listed in Table 1, and while this list is by no means exhaustive, it does include those that have been studied in the most detail. It is impossible to give any common feature that unites this group of molecules, as they vary greatly in structure and function, and in some cases it is not clear that they are even made by the chondrocytes. Many of the molecules are proteoglycans, bearing glycosaminoglycan chains, whereas others are glycoproteins or even nonglycosylated proteins. Some of the molecules represent degradation products of larger precursors that accumulate because of their interaction with other matrix components. Many of the molecules play a structural role, whereas others may be involved in regulating cell function. In addition, many of the molecules vary in their abundance and structure with anatomical site or the person’s age, and many are not unique to cartilage. The importance of many of the molecules to cartilage function is illustrated in Table 2, which shows their association with pathology when they are produced in a mutant form.

Proteoglycans of the cartilage extracellular matrix

Aggregating proteoglycans
Among the noncollagenous proteins of cartilage, aggrecan has undoubtedly received the greatest attention, because of its high abundance in cartilage, its close association with the ability of the tissue to resist compression, and its modification in many cartilage disorders. Aggrecan belongs to the family of aggregating proteoglycans that form large, multimolecular complexes with hyaluronan [1]. The family also includes versican, neurocan, and brevican,

CILP = cartilage intermediate-layer protein; CMP = cartilage matrix protein; COMP = cartilage oligomeric matrix protein; CS = chondroitin sulfate; CS1/CS2 = chondroitin-sulfate-attachment regions of aggrecan; G1/G2/G3 = globular regions of aggrecan; IL-1 = interleukin-1; PRELP = proline- and arginine-rich end leucine-rich repeat protein; SLRP = small leucine-rich repeat proteoglycan.
though of these only versican has been shown to be expressed in cartilage, and at much lower levels than aggre-
can. All the members of the family have an amino-terminal globular domain, which is responsible for interaction with
hyaluronan, and a carboxy-terminal globular domain, which
has lectin-like homology. These features have resulted in the
family being termed hyalectans or lecticans.

Aggrecan has an additional globular domain (G2) that is
separated from the amino-terminal globular domain (G1)
by a short, interglobular domain [2]. The G2 domain is
separated from the carboxy-terminal globular domain (G3)
by a keratan sulfate attachment domain and two chon-
droitin sulfate (CS) attachment domains (CS1 and CS2).
Over 100 CS and keratan sulfate chains may be present
in the three glycosamino-glycan attachment domains,
though it is not clear at present whether all potential
attachment sites are always occupied or whether variation
may occur among individuals. The high CS and keratan
sulfate content of aggrecan and its ability to interact with
hyaluronan are essential features for normal articular
cartilage function, as they provide the rheological prop-
erties necessary for resisting compression. The function of
the G3 domain of aggrecan is unclear. Its lectin-like prop-
erties suggest the possibility of interaction with other com-
ponents of the extracellular matrix [3], though it has also
been suggested that it is involved in intracellular trafficking
during aggrecan synthesis. Mutations in the aggrecan
gene that prevent core protein synthesis form the basis of
chondrodysplasias in mice (cartilage matrix deficiency)
and chicks (nanomelia) [3]. In addition, impaired glyco-
saminoglycan sulfation on aggrecan causes the chondro-
dysplastic phenotypes associated with the brachymorphic
mouse and diastrophic dysplasia in humans.

An interesting feature of the human aggrecan gene is the
existence of polymorphism in the region encoding the
CS1 domain. This region is composed of repeat
sequences, which may range in number from 13 to 33 [4].
Individuals with the shortest alleles will have the lowest
proportion of CS on their aggrecan molecules, and may
be at risk for cartilage degeneration due to impaired
aggrecan function. Irrespective of such polymorphism, the
glycosaminoglycan composition of aggrecan varies con-
siderably during juvenile development, as both the size
and sulfation pattern of the CS and keratan sulfate
change, though the functional consequence of this
change is unclear. In addition, size heterogeneity is gener-
atized in the aggrecan core protein by the action of pro-
teinases, with those fragments bearing a G1 domain being
selectively retained in the tissue matrix. Proteolysis ulti-
mately results in the accumulation of free G1 domains that
have a long half-life in the tissue [5]. Many proteinases are
able to degrade aggrecan if they gain access to the carti-
lage matrix, but most physiological and pathological
degradation of articular cartilage is associated with the

| Table 1 |
| Proteoglycans (PGs) and proteins of the cartilage extracellular matrix |

<table>
<thead>
<tr>
<th>Proteoglycans</th>
<th>Proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggregating</td>
<td>Structural</td>
</tr>
<tr>
<td>Aggrecan</td>
<td>COMP (Thrombospondin-5)</td>
</tr>
<tr>
<td>Versican</td>
<td>Thrombospondin-1 and -3</td>
</tr>
<tr>
<td>Link protein</td>
<td>COMP (Matrilin-1)</td>
</tr>
<tr>
<td>Leucine-rich repeat</td>
<td>Matrilin-3</td>
</tr>
<tr>
<td>Biglycan (DS-PGII)</td>
<td>C-type lectin</td>
</tr>
<tr>
<td>Decorin (DS-PGIII)</td>
<td>Fibronectin</td>
</tr>
<tr>
<td>Epiphican (DS-PGIII)</td>
<td>PRELP</td>
</tr>
<tr>
<td>Fibromodulin</td>
<td>Chondroadherin</td>
</tr>
<tr>
<td>Lumican</td>
<td>Tenascin-C</td>
</tr>
<tr>
<td>Other</td>
<td>Fibrillin</td>
</tr>
<tr>
<td>Perlecan</td>
<td>Elastin</td>
</tr>
<tr>
<td>Szp/Lubricin</td>
<td>Regulatory</td>
</tr>
<tr>
<td>gp-39/YKL-40</td>
<td></td>
</tr>
<tr>
<td>Matrix gln protein</td>
<td></td>
</tr>
<tr>
<td>Pleiotrophin</td>
<td></td>
</tr>
<tr>
<td>Chondromodulin-1/SCGP</td>
<td></td>
</tr>
<tr>
<td>Chondromodulin-II</td>
<td></td>
</tr>
<tr>
<td>CD-RAP</td>
<td></td>
</tr>
<tr>
<td>Growth factors</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>Chondrocalcin</td>
</tr>
<tr>
<td>PArP</td>
<td></td>
</tr>
<tr>
<td>Lysozyme</td>
<td></td>
</tr>
<tr>
<td>Phospholipase A2</td>
<td></td>
</tr>
<tr>
<td>Proteinases and inhibitors</td>
<td></td>
</tr>
</tbody>
</table>

CD-RAP = cartilage-derived retinoic acid responsive protein; CILP = cartilage intermediate layer protein; CMP = cartilage matrix protein; COMP = cartilage oligomeric matrix protein; DS-PG(I, II, III) = dermatan sulfate proteoglycan (I, II, III); gln = gamma-carboxyglutamic acid; gpl = glycoprotein; PArP = proline- and arginine-rich protein; PRELP = proline- and arginine-rich end leucine-rich repeat protein; Szp = superficial zone protein.

| Table 2 |
| Genetic disorders and the mutant cartilage matrix proteoglycans and proteins with which they are associated |

<table>
<thead>
<tr>
<th>Mutant matrix molecule</th>
<th>Associated disorder</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggrecan (impaired core protein synthesis)</td>
<td>Cartilage matrix deficiency (in mice)</td>
</tr>
<tr>
<td></td>
<td>Nanomelia (in chickens)</td>
</tr>
<tr>
<td></td>
<td>Brachymorphism (in mice)</td>
</tr>
<tr>
<td>Aggrecan (impaired glycosaminoglycan sulfation)</td>
<td>Diastrophic dysplasia</td>
</tr>
<tr>
<td></td>
<td>Achondrogenesis 1B</td>
</tr>
<tr>
<td></td>
<td>Atelosteogenesis type II</td>
</tr>
<tr>
<td>Perlecan</td>
<td>Chondrodystrophic myotonia</td>
</tr>
<tr>
<td></td>
<td>Dyssegmental dysplasia</td>
</tr>
<tr>
<td>Szp</td>
<td>Camptodactyly–arthropathy– coxa vara–pericarditis syndrome</td>
</tr>
<tr>
<td>COMP</td>
<td>Pseudoachondroplasia</td>
</tr>
<tr>
<td></td>
<td>Multiple epiphyseal dysplasia</td>
</tr>
</tbody>
</table>

1 Human disorder unless indicated otherwise. COMP = cartilage oligomeric protein; Szp = superficial zone protein.
propeptide from type II collagen. Propeptide removal is
have short, amino-terminal propeptides that are removed
functional consequence [14]. Decorin and biglycan also
matrix with age, but it is not clear whether this is of any
nonglycanated biglycan accumulates in the cartilage
terminal region bearing the dermatan sulfate chains. Such
proteolytic processing that results in removal of the amino-
in its intact form at all ages, biglycan exhibits age-related
they are present throughout life. Whereas decorin remains
and biglycan have been found in articular cartilage, and
( DS-PGI), decorin ( DS-PGII), and epiphycan ( DS-PGIII) –
sulfate proteoglycans ( also called DS-PGs) – biglycan
cartilage has been shown to contain three dermatan
sulfate chains or keratan sulfate chains. Human
disulfide-bonded domains [1]. The family may be divided
adjacent leucine-rich repeats bordered at each end by
depolymerization of hyaluronan and involves concomitant
importance of link protein in proteoglycan aggregate
function is demonstrated by the impaired cartilage develop-
ment observed in the link-protein-null mouse [ 13 ].

The interaction of aggrecan with hyaluronan is stabilized
by the presence of link proteins. As with aggrecan, these
proteins undergo proteolytic modification throughout life
and can be used as an indicator of proteinase action. They
provide evidence of the action of matrix metalloproteinase
throughout juvenile development, and the participation of
additional agents in the adult [ 10,11 ]. The link proteins are
not susceptible to cleavage by the aggrecanase produced
under cytokine stimulation of cartilage [12], and there is
no evidence that any of the proteolytically modified link
proteins have impaired function. Link protein can be lost
from the cartilage matrix during periods of tissue degen-
eration, but such loss is most likely due to depolymerization
of hyaluronan and involves concomitant loss of aggrecan.
The importance of link protein in proteoglycan aggregate
function is demonstrated by the impaired cartilage develop-
ment observed in the link-protein-null mouse [13].

Small leucine-rich repeat proteoglycans
The small leucine-rich repeat proteoglycans ( SLRPs) are
characterized by a central domain composed of a series of
adjacent leucine-rich repeats bordered at each end by
disulfide-bonded domains [1]. The family may be divided
into two subfamilies, depending on the presence of
dermatan sulfate chains or keratan sulfate chains. Human
cartilage has been shown to contain three dermatan
sulfate proteoglycans ( also called DS-PGs) – biglycan
( DS-PGI), decorin ( DS-PGII), and epiphycan ( DS-PGIII) –
and in all of these, the dermatan sulfate chains are in the
amino-terminal region of the core proteins. Only decorin
and biglycan have been found in articular cartilage, and
they are present throughout life. Whereas decorin remains
in its intact form at all ages, biglycan exhibits age-related
proteolytic processing that results in removal of the amino-
terminal region bearing the dermatan sulfate chains. Such
nonglycanated biglycan accumulates in the cartilage
matrix with age, but it is not clear whether this is of any
functional consequence [14]. Decorin and biglycan also
have short, amino-terminal propeptides that are removed
in the extracellular matrix by procollagen-C proteinase, the
same enzyme responsible for removing the carboxy
propeptide from type II collagen. Propeptide removal is
incomplete in adult cartilage [15], but again, the functional
consequence, if any, is unclear.

Human articular cartilage contains two potential keratan
sulfate proteoglycans, fibromodulin and lumican. Like
decorin and biglycan, fibromodulin is present in articular
cartilage throughout life, though it contains keratan sulfate
chains only in the fetus and juvenile [16]. In the adult, it
exists as a glycoprotein devoid of keratan sulfate. In con-
trast, lumican is not present in articular cartilage of the
fetus or young juvenile [17]; in the adult, it is present in
predominantly a glycoprotein form. It is unclear whether
the presence or absence of keratan sulfate influences the
function of these proteoglycans in cartilage. All SLRPs
have all been shown to interact with the fibrillar collagens
of the extracellular matrix, though their site and strength of
interaction may vary. The importance of these molecules in
matrix organization is illustrated by the abnormalities asso-
ciated with SLRP-null mice [18–21], though these abnor-
malities are perhaps less severe than might have been
expected and it is possible that there is a functional redu-
dancy between some family members. Unlike aggrecan,
the SLRPs of the cartilage matrix appear relatively resis-
tant to extensive proteolytic modification and do not show
a ready sensitivity towards cytokine-induced damage [12].
Fragments have, however, been observed in the matrix of
arthritic cartilage.

Other proteoglycans
The cartilage matrix also contains the proteoglycan per-
lecanc. This is somewhat surprising, because perlecanc is
commonly thought of as a basement membrane proteo-
glycan [1], yet articular cartilage is devoid of basement
membranes. Basement membrane perlecanc is character-
ized by the presence of heparan sulfate chains in its
amino-terminal region, though it has been reported that
cartilage perlecanc may exist in a nonglycanated form [22].
The perlecanc core protein is extremely large and might be
expected to be a good candidate for proteolytic process-
ning, but at present there is no information available on
structural changes with either age or arthritis. The impor-
tance of perlecanc to cartilage function is demonstrated by
the perlecanc-null mouse [23], in which severe chondro-
dysplasia is a major part of the phenotype in addition to
basement membrane defects affecting heart and brain
development. In the human, mutations in the perlecanc
gene have been associated with Schwartz–Jampel
syndrome (chondrodystrophic myotonia) [24], and have
recently been reported in dysegmental dysplasia. At
present, the function of perlecanc in cartilage, and in partic-
ular in the growth plates, is unknown.

A final proteoglycan associated with cartilage has been
termed superficial zone protein [25]. It is synthesized by
the superficial chondrocytes of articular cartilage and by
synoviocytes, and has an attachment site for a CS chain. It
is identical to the precursor protein of megakaryocyte-stimulating factor, and probably is the same as a protein originally described as lubricin, which is responsible for the lubrication and frictionless motion of the cartilage surface. While some superficial zone protein may be retained in the extracellular matrix, most is destined for secretion into the synovial cavity. The synthesis of this protein is impaired in the arthritic joint, where alternative splicing has been reported, and production is downregulated by the presence of inflammatory cytokines such as IL-1. Gene defects in this protein have been associated with camptodactyly–arthropathy–coxa vara–pericarditis syndrome [26]. In addition to its role as a lubricant, the protein may play a role in regulating synovial cell proliferation, as this syndrome and various forms of arthritis are associated with synovial hyperplasia. In the case of camptodactyly–arthropathy–coxa vara–pericarditis syndrome, hyperplasia occurs in the absence of inflammation.

### Proteins of the cartilage extracellular matrix

#### Structural proteins

The extracellular matrix of cartilage contains numerous proteins that are neither collagens nor proteoglycans [27], and several of these are thought to play a structural role in the matrix. Cartilage oligomeric matrix protein (COMP) is perhaps the best studied of these proteins. It belongs to the thrombospondin family and has been termed thrombospondin-5, and is structurally more closely related to thrombospondins 3 and 4 than to thrombospondins 1 and 2 [28]. Other members of the thrombospondin family have been detected in cartilage, though not at the same level or widespread distribution as COMP. This protein is present in all cartilages, being most abundant in growth plate during development, but also in mature articular cartilage. It exists as a disulfide-bonded pentamer linked near its amino-terminal region, and the projecting carboxy-terminal regions are suggested to interact with collagen. The need for COMP in cartilage is best illustrated by the presence of pseudoachondroplasia or multiple epiphyseal dysplasia in individuals bearing a mutation in the COMP gene [29].

A phenotype of multiple epiphyseal dysplasia can also arise by mutations in a type IX collagen gene, and this may indicate an association between COMP and type IX collagen. During cartilage turnover, COMP undergoes degradation, and fragments are released into the synovial fluid. An increase in such fragments has been observed in the synovial fluid of patients suffering from joint trauma and those in the early stages of primary osteoarthritis [30], and it has been suggested that elevated levels of COMP in synovial fluid may serve as a marker of such disorders.

Cartilage matrix protein (CMP) is also thought to serve a structural role in the extracellular matrix [31]. It belongs to the matrilin family and has also been termed matrilin-1. Matrilin-3 has also been detected in some cartilages. CMP exists in the cartilage matrix as a disulfide-bonded trimer, joined near the carboxy terminus of its subunits. While CMP is present in skeletal cartilages during development, it is most abundant in extraskeletal cartilages in the adult and is deficient in articular cartilage. This protein is known to interact with both type II collagen and aggrecan, though its precise function remains unclear. Indeed, CMP-null mice do not exhibit any obvious skeletal phenotype and appear to develop normally [32], which may imply a functional redundancy between CMP and matrilin-3. Although CMP is not detected in normal articular cartilage, it is produced by the chondrocytes of arthritic cartilage [33].

Articular cartilages have a matrix protein that is most abundant in the mid-zone of the tissue but deficient in the deepest and superficial zones [34]. On the basis of this localization, the protein has been termed cartilage intermediate-layer protein (CILP). CILP is more abundant in adult than in juvenile articular cartilage, but the relevance of the site- and age-related distribution to function is unknown. CILP production has also been reported to be increased in osteoarthritic cartilage. Interestingly, the transcript from the CILP gene encodes two proteins. The amino-terminal portion of the message encodes CILP, while the carboxy-terminal portion encodes nucleotide pyrophosphohydrolase (NTPPHase) [35]. The initial translation product contains both proteins, which are separated by proteolytic cleavage within the chondrocytes. The relevance of this phenomenon and the function of CILP are at present unknown, and CILP does not appear to have a close structural relationship to any other protein yet described.

Other structural proteins are thought to be involved in cell–matrix interactions rather than matrix–matrix interactions. Among these, fibronectin deserves particular mention. Fibronectin is present in many tissues and exists as a disulfide-bonded dimer joined at the carboxy terminus of its subunits [36]. Fibronectin can exist in multiple isoforms, because of alternative splicing of its gene, and chondrocytes appear to produce a characteristic splice variant [37]. The abundance of fibronectin increases about 10-fold in osteoarthritic cartilage [38], though the functional significance of this is unclear. However, it is interesting that fibronectin fragments, resulting from proteolytic degradation, are able to propagate degradation of aggrecan at the same sites as expected for the action of aggrecanase [39]. It has been suggested that the fibronectin fragments that may accumulate in the arthritic joint may stimulate the local production of inflammatory cytokines, such as IL-1, that upregulate aggrecanase expression.

Another molecule of interest is proline- and arginine-rich end leucine-rich repeat protein (PRELP), which is closely related in protein structure and gene organization to fibromodulin and lumican but is devoid of keratan sulfate chains. PRELP shows selective distribution among cartilagenous tissues and is not present in fetal and young
juvenile human cartilage [40]. The unique amino-terminal region of PRELP may facilitate interaction with heparan sulfate proteoglycans on cell membranes [41]. A final protein worthy of note is chondroadherin, which bears neither glycosaminoglycan chains nor N-linked oligosaccharides and, in common with elastin, may be devoid of carbohydrate. Chondroadherin also belongs to the family of leucine-rich repeat proteins [42] and, in common with PRELP, is thought to play a role in mediating cell–matrix interactions.

Regulatory proteins
Several proteins in the extracellular matrix are thought to influence cell proliferation or metabolism rather than play a structural role in the matrix (see Table 1), but a discussion of their properties is beyond the scope of this review. However, one of these proteins, termed gp-39, deserves special recognition. It is related to the chitinase family but does not have enzymic activity. It is not detected in normal articular cartilage, but is produced by chondrocytes in culture and is present in arthritic cartilage [43]. As such, it may reflect situations in which rapid tissue remodelling is occurring and may be indicative of the capacity of chondrocytes to recognize an abnormal environment and initiate a repair response.

Other proteins
This category includes proteinases and their inhibitors, degradation products of collagen, and basic proteins that associate with the extracellular matrix. Two products of collagen degradation have been reported to accumulate in cartilage [27]. One is chondrocalcin, which represents the carboxy-propeptide of type II collagen, and the second is proline–arginine-rich protein, which represents the amino-propeptide domain of the α2(III) chain of type XI collagen. It is possible that these molecules are not merely innocent bystanders but are involved in feedback regulation of collagen synthesis. The abundance of chondrocalcin in cartilage is often used as an indication of new collagen synthesis. Finally, lysozyme [44] and phospholipase A2 [45] are worthy of mention. Both are cationic proteins that may owe their presence in the cartilage matrix to the high content of anionic aggrecan. In the case of lysozyme, it is likely that much of it is not produced by the chondrocytes but rather is absorbed from the synovial fluid.

Conclusion
It is apparent from this brief review that the extracellular matrix of cartilage contains many noncollagenous proteins and proteoglycans whose precise functions are only just beginning to be understood. These molecules may serve a structural or regulatory role, and in some cases may do both, as degradation products of some of the structural molecules are known to influence the chondrocyte. The recognition of genetic disorders in which synthesis of the matrix molecules is perturbed has aided greatly in our understanding of their functional role, but the reason for many site- and age-related restrictions in expression remains unclear. The role of many of the molecules in the arthritic joint is also unclear, as in many cases they may be pawns of the disease, undergoing destruction, yet in others they may be actively involved in propagating destruction or initiating repair. This is an area where there is still a wealth of information to be mined.

Acknowledgement
The author is funded by the Shriners of North America.

References


