LUND UNIVERSITY

Functions of small leucine-rich repeat proteoglycans in connective tissues

Kalamajski, Sebastian

2008

Document Version:
Publisher's PDF, also known as Version of record

Link to publication

Citation for published version (APA):
Kalamajski, S. (2008). *Functions of small leucine-rich repeat proteoglycans in connective tissues*. [Doctoral Thesis (compilation)]. Department of Experimental Medical Science, Lund Univeristy.

Total number of authors:
1

Creative Commons License:
CC BY-NC

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Functions of Small Leucine-rich Repeat Proteoglycans in Connective Tissues

Sebastian Kalamajski

Faculty of Medicine

Department of Experimental Medical Science

2008

With the approval from the Faculty of Medicine at Lund University, this doctoral thesis will be defended in public at Segerfalksalen, BMC, Lund, on Tuesday, April 1st 2008 at 9:00

Faculty opponent: Prof. Mats Paulsson, University of Cologne, Germany

© **Sebastian Kalamajski**

Lund University, Faculty of Medicine Doctoral Dissertation Series 2008:30
ISSN 1652-8220
ISBN 978-91-85897-83-4

Printed by Media-Tryck, Lund University, 2008

Enjoy…

LIST OF PAPERS

1. Kalamajski, S., Aspberg, A., and Oldberg, Å. (2007) The decorin sequence SYIRIADTNIT binds collagen type I. *J Biol Chem* **282**, 16062-16067.

2. Kalamajski, S., and Oldberg, Å. (2007) Fibromodulin binds collagen type I via Glu-353 and Lys-355 in leucine-rich repeat 11. *J Biol Chem* **282**, 26740-26745.

3. Kalamajski, S., Oldberg, Å. (2008) Lumican and fibromodulin bind collagen via homologous sequences in LRR 5-7. *Manuscript*.

4. Kalamajski, S., Liu, C., Rubin, K., Oldberg, Å. (2008) Fibromodulin regulates collagen cross-linking in tendon. *Manuscript.*

5. Oldberg, Å., Kalamajski, S., Salnikov, A. V., Stuhr, L., Mörgelin, M., Reed, R. K., Heldin, N. E., and Rubin, K. (2007) Collagen-binding proteoglycan fibromodulin can determine stroma matrix structure and fluid balance in experimental carcinoma. *Proc Natl Acad Sci U S A* **104**, 13966-13971.

6. Nyström, T., Kalamajski, S., Oldberg, Å., Hultgårdh-Nilsson, A. (2008) Fibromodulin deficiency alters injury-induced vascular lesion formation. *Manuscript*.

ABBREVIATIONS

B-CLL – B-type chronic lymphocytic leukemia
BMP – bone morphogenic protein
CNBr – cyanogen bromide
COP – coat protein
CS – chondroitin sulfate
DDR – discoidin domain receptor
DHLNL – dihydroxylysylnorleucine
DS – dermatan sulfate
EGFR – epidermal growth factor receptor
FACIT – fibril-associated collagens with interrupted triple helix
GAG – glycosaminoglycan
HLKNL - hydroxylysylketonorleucine
HLNL – hydroxylysylnorleucine
HP – hydroxylysylpyrridinoline
HS – heparan sulfate
IFN – interferon
IGF – insulin growth factor
IL – interleukin
KS – keratan sulfate
LDL – low density lipoprotein
LH – lysyl hydroxylase
LKNL – lysylketonorleucine
LOX – lysyl oxidase
LP – lysylpyrridinoline
LRR – leucine-rich repeat
LTBP – latent TGF-β binding protein
MAP – mitogen-activated protein
MAPK – mitogen-activated protein kinase
MIP-2 – macrophage inflammatory protein-2
MMP – matrix metalloproteinase
Multiplexin – mutiple triple helix domains with interruptions
P4H – prolyl 4-hydroxylase
PCNA – proliferating cell nuclear antigen
PDGF – platelet-derived growth factor
PDI – protein disulfide isomerase
PEG – polyethyleneglycol
PTB – phosphotyrosine-binding domain
SH2 – Src homology 2
SLRP – small leucine-rich repeat proteoglycan
TGF-β – transforming growth factor-β
TNF-α – tumor necrosis factor-α
TLL – tolloid-like
TLR – toll-like receptor
uPARAP – urokinase-type plasminogen activator receptor associated protein
VSMC – vascular smooth muscle cell
VSVG – vesicular stomatitis virus G-protein

BACKGROUND

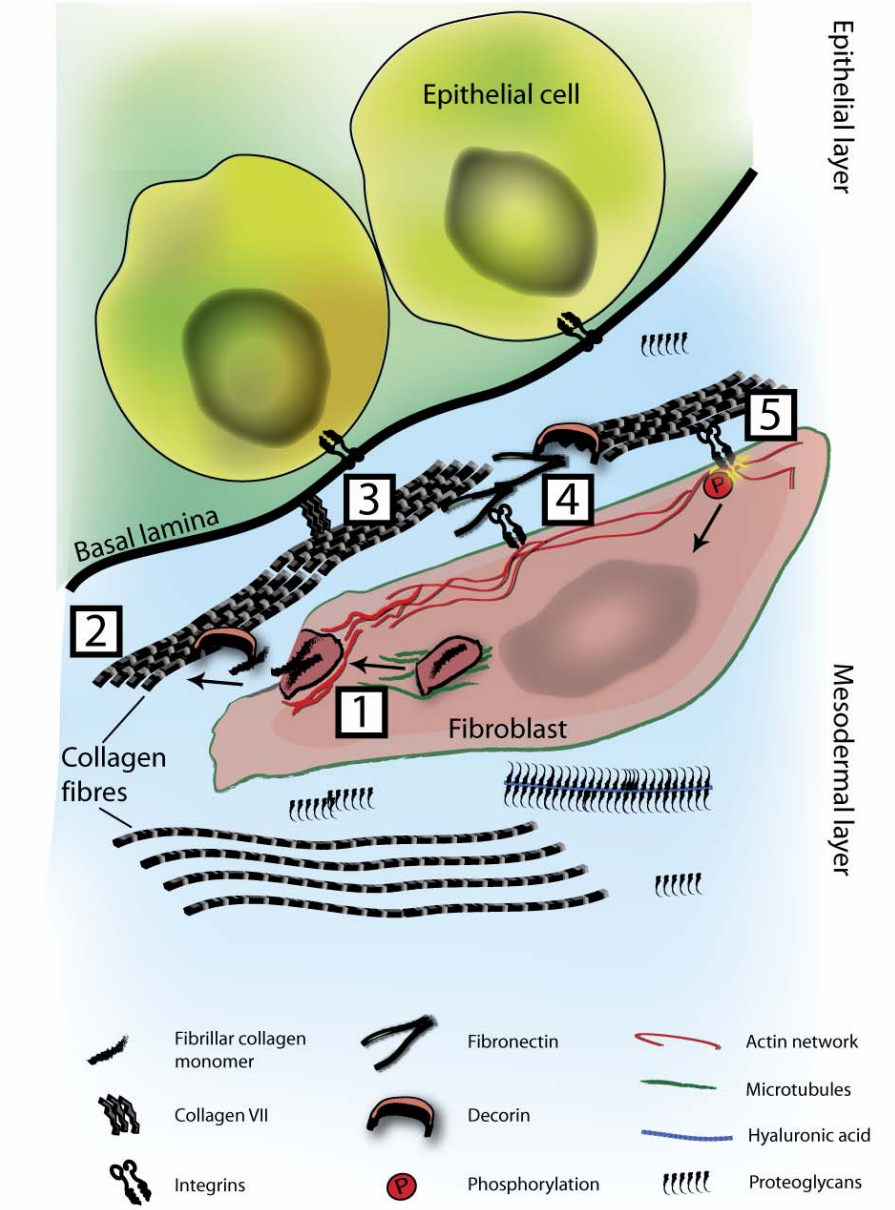
Connective tissues are fundamental for animal development and diversity. Their influence ranges from fine-tuning of cellular proliferation and differentiation, to the more general tasks of protecting from dehydration and infection (skin), or establishing mechanical resistance and strength (bones, cartilage, tendons).

Connective tissue cells produce extracellular matrices adapted to the local functional requirements, be it mineralized hard tissue of a tooth or a soft elastic framework of the skin. The uniqueness of the matrix lies within the relatively low cell density, and development from the mesodermal embryonal layer and the mesenchymal stem cells. Structurally, it is made of fibrous proteins deposited in ground substance – the intercellular viscous compartment filled with hydrophilic carbohydrates and proteoglycans. Another vital part of connective tissues is the fluid that reciprocally transfers nutrition and waste between the cells and the blood. It also transports cell signaling molecules like hormones, cytokines, and paracrine factors, and provides an initial ground for the innate immune response. Collectively, the relative amount and composition of the three matrix components – fibres, ground substance, and tissue fluid – determines the tissue function and pathological response[1].

The structural integrity of extracellular matrix provides mechanical strength and resistance, but can also act as a signaling trigger for cells influenced by mechanical stress. Tendons and bones are able to withstand large tensile or compressive forces due to the high amount of collagen – a major extracellular protein that can assemble into fibrous bundles. Collagen fibres also serve as scaffolding networks in nearly all organs, contributing to their shape and stratified cellular organization. In part, these processes are controlled by factors that are the main subject of this thesis: small leucine-rich repeat proteoglycans (SLRPs). Already in the lower invertebrates, blueprints for tissue rigidity, elasticity, and mechanical resistance preclude the presence of the collagen family, along with proteoglycans and other matrix proteins.

These proteins not only influence the structure, but also the physiological milieu in the tissue. This is indispensable for attachment, survival and migration of cells (cells don't swim aimlessly), as well as differentiation and regulated gene expression. For example, the major matrix proteins (collagens), and proteoglycan-attached carbohydrates, can bind and consequently alter the flow and bioavailablity of growth factors and cytokines[2-5].

In conclusion, the proper function of connective tissues necessitates the presence of collagens and SLRPs, and is linked to their regulated expression and spatial organization. The aim of my investigations was to characterize the interactions between SLRPs and collagen, as well as to analyze the influence of these proteoglycans on connective tissues in cancer and in atherosclerotic models. In the latter case, the primary goal was to examine the influence of fibromodulin on collagen network formation, and how this knowledge could be applied to pre-clinical studies.



Cooperation between cells and matrix. 1/ Collagen is produced by fibroblasts and 2/ secreted and asssembled into fibrils with help of decorin. 3/ Basal lamina can attach to collagen fibres via other collagens. 4/ Integrins bind collagen directly or dependent on fibronectin, where decorin may also be involved. 5/ Cell migration or contraction of collagen can be signalled through collagen-binding integrins.

COLLAGENS

Through the process of evolution, collagen genes have adapted to the new functions required by the fibrous matrix. Representing about 30% of total dry weight of body protein in humans, nearly 30 types of collagens are now described, all distinguished in their structure and varied contributions to the tissue mechanics, architecture, or cell signaling. Basic collagen research is valuable for understanding and treatment of disorders related to defective or metabolically modified collagens, including scurvy, Ehlers-Danlos syndrome, Alport syndrome, osteogenesis imperfecta, osteoarthritis, and osteoporosis. Furthermore, conditions associated with fibrotic response, including cancer and atherosclerosis, may prove suitable for clinical applications of collagen-related research.

The specific function of any connective tissue is largely conferred by its predominant collagen type and the associated proteins. The most abundant protein in the entire human body is collagen type I, belonging to the group of fibril-forming collagens that constitute the bulk of the fibrous matrix. Notably, it is present in bone, tendon, ligament, dentin, skin, and organ capsules. In terms of function, it can provide tensile resistance to the tissues by polymerizing into thick fibers, and can also be calcified – under control of the collagen-producing fibroblasts and the closely related odontoblasts and osteoblasts. This thesis is focused around collagen type I and the proteoglycans (SRLPs) that regulate its fibrillar assembly.

Other fibril-forming collagens include types II, III, V, and XI. Type II collagen differs from type I in being more hydroxylated and glycosylated; it is associated with collagen XI, forms thin fibrils, and is present in cartilage, vitreous body and corneal epithelium, where it is surrounded by a ubiquitous ground substance. The other collagens, type III and V can interact with collagen I, and are distributed in the expansible reticular connective tissues in skin or around smooth muscle cells, or in dense tissues like bone or tendon. Collagen V is more widespread, also found in cornea and skin, and contributes to collagen I fibrillar assembly.

Fibril-associated collagens with interrupted triple helix (FACIT) include type IX that connects the fibrils through binding to the surface of collagen type II in an anti-parallel direction. It also has a cationic flexible N-terminal domain that interacts with other matrix proteins. Other FACIT collagens, type XII and XIV, associate with collagen I, and appear to fulfill a similar purpose – connecting other matrix components and regulate fibril assembly.

Type VI collagen has large terminal globular domains, with extensive post-translational modifications. The chains align into thin fibrils, first into antiparallel dimers, then tetramers, and finally into a network that is formed independently of other collagens. Due to the massive globular domains, the final product resembles a beaded filament.

Type IV and VII collagens are involved in formation of the pericellular basal laminae, the former through interconnection of other lamina components (laminins, entactins), and the latter by anchoring the lamina to the underlying matrix fibers.

Lastly, the closely homologue type VIII and X collagens both form unique hexagonal fibril networks, but are expressed in different tissues – cornea and cartilage, respectively[6, 7].

Collagens, their main tissue distribution, and function

Collagen	Tissue distribution	Structure/Type	Main function
I	Tendon, skin, bone, tooth	Fibrils of varied thickness	Tensile strength
II	Cartilage, intervertebral disk	Thin loose fibrils	Compressive strength
III	Smooth muscle, reticular tissue	Flexible loose fibrils	Expandable matrix
IV	Basal laminae	Network, no visible fibrils	Cell support, filtration
V	Dermis, tendon, cornea, bone	Fibrils	Initial fibril formation
VI	Widespread, except in bone	Beaded thin fibrils	Independent network
VII	Dermis, basal laminae	Fibrils	Anchor for basal laminae
VIII	Endothelium, cornea	Hexagonal network	Stromal network
IX	Cartilage	FACIT	Collagen II function
X	Hypertrophic cartilage	Hexagonal network	Ossification?
XI	Cartilage, intervertebral disk	Fibrils	Collagen II function
XII	Tendon, ligaments	FACIT	Collagen I fibril formation
XIII	Widespread	Transmembrane collagen	Cell adhesion and signaling
XIV	Dermis, tendon, lungs, liver	FACIT	Collagen I fibril formation
XV	Fibroblasts, smooth muscle cells - special basal laminae	Multiplexin, attached chondroitin sulfate	Restin - anti-angiogenic cleavage product
XVI	Fibroblasts, keratinocytes, cartilage	Multiplexin	Fibrillin function in skin Collagen II function
XVII	Dermis-epidermis junction	Transmembrane collagen	Attachment of epidermis to basal lamina
XVIII	Basal laminae	Multiplexin, attached heparan sulfate	Endostatin, anti-angiogenic cleavage product
XIX	In rhabdomyosarcoma	Radial aggregates	
XX	Cornea, skin, tendon, cartilage		Fibrillar function?
XXI	Widespread		Fibril-associated?
XXII	Tissue junctions		Microfibril function?
XXIII	Cancer cells	Transmembrane collagen	
XXIV	Cornea, bone		Collagen I function?
XXV	Neurons	Transmembrane collagen	Found in Alzheimer's amyloid plaques
XXVI	Testis, ovary		
XXVII	Widespread, lots in cartilage		
XXVIII	Skin, calvaria, nerves	Beaded filaments	Basal lamina function?
XXIX	Skin, lung, GI-tract		Epidermal integrity?

Disorders related to collagen mutations

Collagen	Disorder
I	Osteogenesis imperfecta Ehlers-Danlos syndrome
II	Collagenopathy II Chondrodysplasia
III	Ehlers-Danlos syndrome IV Familial aorta aneurysms
IV	Alport syndrome
V	Classical Ehlers-Danlos syndrome
VI	Ulrich myopathy Bethlem myopathy
VII	Epidermolysis bullosa
IX	Collagenopathy IX
XVII	Bullous pemphigoid Epidermolysis bullosa
XXIX	Atopic dermatitis

Collagen synthesis

All collagens have the typical Gly-X-Y repeating amino acid sequence that spans a large part of the protein (here, X is proline, and Y is 4-hydroxyproline). This tandemly repeated motif provides a basis for the interactions that shape the right-handed triple-helical structure, in itself comprised of three left-handed intertwined α-chains. As for collagen type I, it is a heterotrimeric triple helix, made of two α1(I)- and one α2(I)-chain; it is 280 nm long, and flanked by telopeptides and globular propeptides, neither of which resembling the helical Gly-X-Y structure. Thus, procollagen during cellular secretion has five distinct domains: N-propeptide, N-telopeptide, the long triple helix, C-telopeptide, and C-propeptide.

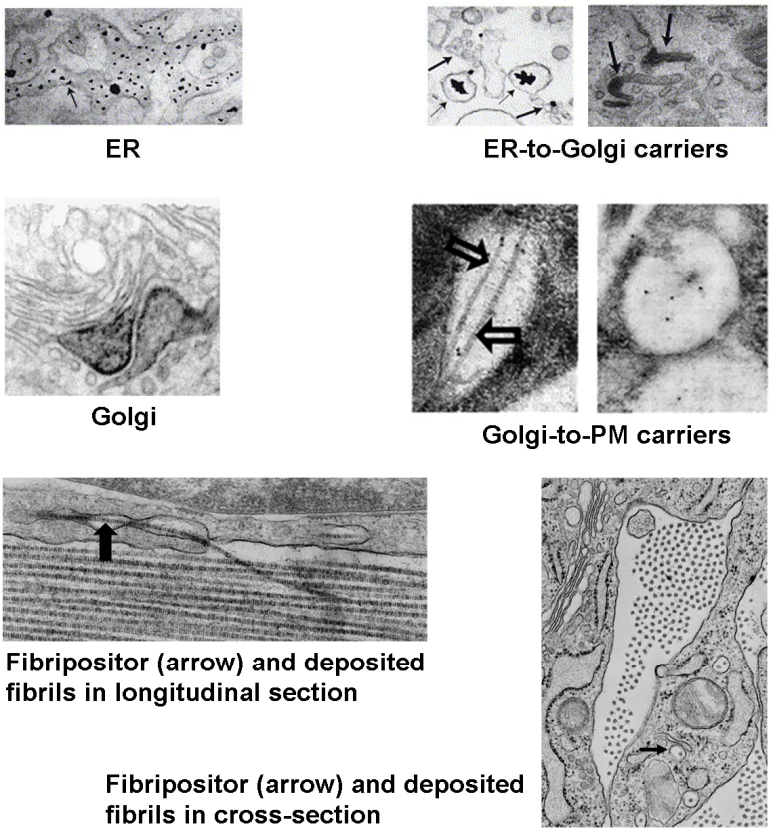
Of functional importance are the short telopeptides – their lysine residues are crucial for the intermolecular cross-linking of collagen. On the other hand, the propeptides protect the protein from intracellular aggregation by rending it water-soluble. When secreted from the cell, they are cleaved off, which triggers the fibril formation. This last function is attributed to the triple helix – the largest and fibrous domain of collagen – the fundamental building block of connective tissues.

The synthesis of collagen follows the common intracellular route, with modifications more or less specific for this protein. The initial transport of procollagen from ER to cis-Golgi is carried out in association with Hsp47 (a collagen-specific chaperone) and cyclophilin B (a protein with peptidyl prolyl cis-trans isomerase activity). During the transport, the association with Hsp47 is lost as Hsp47 is recycled to ER[8, 9]. Considering transport carriers, the classically described COPII- and COPI-coated vesicles are 60-80 nm, too small for incorporation of a 300 nm long procollagen cargo. However, secretion of procollagen is equally efficient with other proteins, and is carried out in larger compartments. Procollagen travels separately from VSVG protein (commonly used for studies of constitutive secretory pathway), and from ts-O45-G and COPI (ER-to-Golgi cargo membrane proteins). Essentially, an early COPII-dependent sorting mechanism concentrates collagen into large carriers that later move along microtubules. These carriers are devoid of ERGIC-53 that normally appears in the classically described vesicles, and COPII association is only apparent in the early formation of the carriers. This complex possibly sorts out procollagen from the unified ER-protruding buds that later separate either into the smaller vesicles or collagen carriers, the former being selectively coated with COPI[10, 11].

In fibroblast Golgi, procollagen is transported by a system of cisternal maturation. Briefly, the Golgi stacks contain large distensions where procollagen is located, without any detached collagen-containing vesicles (COPI-coated). The Golgi cisternae mature as they move from cis-to-trans direction, and procollagen appears denser in the distensions of trans-cisternae – probably concentrated, or assembled into initial polymers. As a whole, the intra-Golgi transport is completed within 10 minutes.

This cisternal maturation model appears to be a default transport pathway for secreted proteins, as seen on live imaging of Golgi in yeast, while the COPI-coated vesicles found in vicinity of the Golgi stack may contain both proteins for anterograde and retrograde transport. However, a complete model of Golgi transport is still elusive[12-15].

Procollagen aggregates can be found in fibroblast cell culture, being about 300 nm x 5-20 nm, and without the typical D-period banding pattern (termed zero-D-array). Reasonably, these collagens are highly concentrated, which may reflect the



Intracellular collagen transport compartments and the deposited fibrils, visualized in electron microscope, using immunogold (the first five pictures) and standard techniques. See the text for references.

status inside the cellular compartment or inside the just described transport carrier. When secreted, and thus diluted, these aggregates may disperse into monomers to allow for a renewed fibril assembly[16]. Similar amorphous aggregates were observed in electron microscopic imaging of cells, being engulfed in 600x100 nm cylindrical vacuoles. These vacuoles are near the cell surface, implying a secretory pathway, although not excluding internalization of collagen[17].

Transport of collagen aggregates from Golgi to plasma membrane (PM) involves Golgi-to-PM carriers – a few μm-long protrusions that are budding off from defined trans-Golgi subdomains as ~1 μm carriers. Within three minutes the budding procedure is completed, and another three minutes lapse between the kinesin-driven transport of the

carrier to the fusion with the plasma membrane. Golgi-to-PM carriers are neither coated with clathrin or the classical Golgi markers (mannosidase II, sialyl transferase). However, procollagen can be cosecreted with VSVG, which implies similar secretion and sorting mechanism as in constitutive secretory pathway; the distinctions are yet unknown[18, 19].

The secreted collagens assemble into fibrils, which then accrue into bigger collagen fibres. Although the assembly is to some degree spontaneous, it is also controlled by fibroblasts. In tendons, these cells deposit collagen in long cytoplasmic recesses, aligned and in contact with a growing or a pre-formed collagen fibril. Fibroblasts can thus form parallel collagen bundles with help of such compartments (termed fibripositors) that are unidirectional, and in which procollagen processing, including removal of propeptides, is initiated[20]. Interestingly, secretion and processing of procollagen appears to be dependent on microtubule transport, while actin filaments regulate the alignment and formation of fibripositors[21]. This entire process may therefore be driven by a preceding cellular sensing of the already arranged collagen network. Notably, in different contexts, fibroblasts arrange the growing fibrils not only in parallel but also in well-defined networks of cornea, or in fused and intertwined matrices during fetal development and scar healing[22, 23].

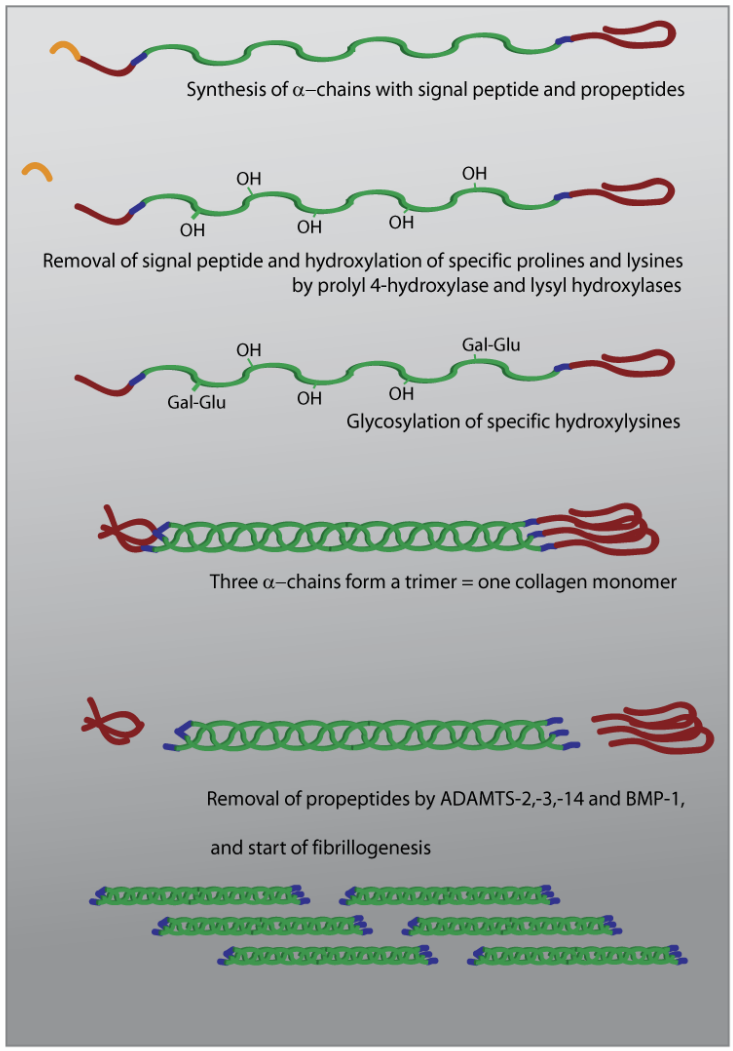
Procollagen processing

Procollagen processing involves enzymatic removal of propeptides from the N- and C-termini, which renders collagen insoluble and amenable for polymerization. This process starts already within the Golgi-to-PM carriers, but continues outside the cell[20]. Removal of propeptides, especially the C-propeptide, triggers the initial assembly of collagen monomers into thin fibril segments. In particular, removal of only the C-propeptide gives about ten-fold thicker fibrils *in vitro* than removal of just the N-propeptide. This implies fibril thickness control by propeptide processing[24, 25].

C-propeptide - the α-chain selector

Each of the known collagen monomers is composed of three α-chains, but they are produced from about 40 different gene products. How is the selection of all these chains controlled? C-propeptides initiate the assembly of α-chains into trimers, although they are not required for the formation of the triple helix (that is stabilized by hydroxyprolines)[26, 27]. Already in early synthesis, during ER translocation, pro-α-chains associate preferentially with chaperones (Hsp47, BiP) and not with each other. The trimer formation commences when the selected C-propeptides interact and form intra- and inter-chain disulfide bonds, and the *association* is driven by highly conserved sequences. However, the trimer composition – the *selection* of the associating α-chains – is determined by C-propeptide variable sequences. For instance, the sequence of the pro-α2(I)-C-propeptide gives folding with intra-chain disulfide bond formation that creates an energetically favorable interaction site for α1(I)-chains (rather than allowing formation of homotrimers). This is the reason why collagen type I is mainly a heterotrimer composed of two α1(I)- and one α2(I)-chain[28, 29]. As a further security measure, pro-α2(I)-chain is not secreted from the ER unless it is associated with two α1(I)-chains[30].

COLLAGEN SYNTHESIS



The variable C-propeptide regions may have evolved in vertebrate genomes as a result of sequence insertion and further mutations. Conversely, invertebrates lack a major part of these chain selection sequences. Relating to the diversification of species, the close relatives of the present simpler animals like Hydra or sponge are the candidates from which the vertebrate collagens evolved. Organisms with new types of collagen acquired it partly through novel C-propeptide variants[31].

The C-propeptides must be removed in order for the collagen monomers to interact with each other and start forming fibrils. The active enzymes (cleaving at Ala-Asp or Gly-Asp), include BMP-1, mTLD (mammalian Tolloid), and mTLL-1 (mammalian Tolloid-like 1) – all being important developmental patterning proteinases, responsible for dorsal-ventral axis formation (Drosophila) and endochondral bone formation. For example, BMP-1 and mTLL-1 cleave Chordin, which releases TGF-β-like BMPs, crucial for bone formation. BMP-1 also cleaves propeptides of two SLRPs – biglycan and osteoglycin[32, 33], which could be functionally correlated with collagen fibrillogenesis.

N-propeptide – fibrillogenesis control

N-propeptide is the least conserved unit of collagens, and in some peculiar cases it is not processed, thereby influencing the final fibril diameter. For example, sea urchins have an Echinoderm-specific N-propeptide sequence, extending outwards during the fibril assembly, thus limiting the fibril diameter to about 20 nm[34]. However, in most animals this propeptide is processed, and the cleavage can start already in Golgi-to-PM carriers[20], regulating fibril formation by rendering the collagen less water-soluble. The enzymes in the process include ADAMTS-2, -3, and -14 (A Disintegrin-like And Metalloprotease domain with Thrombospondin type I motifs). In humans, absence of N-proteinase activity (e.g. ADAMTS-2 mutations in Ehlers-Danlos syndrome VIIC) leads to skin fragility and preponderance for bruises. Compensatory mechanisms (for example ADAMTS-2-deficient mice rescued by ADAMTS-14) have been demonstrated, although the compensation is gradual and varies in each tissue. In the mentioned example, the rescue is satisfactory in tendons but not in skin[35, 36]. Intriguingly, these enzymes have tissue-specific expression overlapping with the presence of particular fibrillar collagens. For instance, ADAMTS-2 is co-localized with procollagen III in the less dense, stretchable connective tissues (skin, lung, arteries, intestines). This is also where the phenotype of the ADAMTS-2-deficient mice manifests in abnormal collagen fiber formation, and where developmental expression of ADAMTS-3 and -14 is absent. Similarly, expression of ADAMTS-3 is co-localized with procollagen II in cartilage, without the other enzymes[37]. To conclude, besides being able to process procollagen I, the ADAMTS homologues are involved in tissue-specific processing of procollagens and possess a limited functional redundancy.

The significance of α2(I)-chain

It is somehow intruiging that collagen type I is a heterotrimer, since a homotrimer made of α1(I)-chains can exist in a stable form with an even higher thermal melting point (2.5 °C higher) and over 100-fold slower isothermal melting than the heterotrimer[38]. From the evolutionary point of view, α2(I)-chain is the precursor of α1(I), and its C-propeptide sequence is responsible for trimer chain selection[39]. However, a homotrimer of α2(I)-chains has never been found. Since the heterotrimer is thermally unstable at body temperature[40], the inclusion of α2(I)-chain mediates this important property; patients with missing α2(I) have a wide range of symptoms resembling osteogenesis imperfecta and Ehlers-Danlos syndrome. Possibly, the less stable monomers are necessary for molecular flexibility that allows association into thermally stable fibers (Tm ~60 °C), or for interaction

with other proteins, like MMPs that require a certain degree of unfolding[41]. Also, α2(I)-chain contains three lysine residues involved in intermolecular cross-linking of collagen, which is vital for fibril assembly. For example, in fibromodulin-deficient mice, α1(I)-chains are excessively cross-linked, have a large pool of free α2(I)-chains, and consequently, disformed collagen fibrils and weak tendons (see *Present investigations, paper IV*).

Another function of α2(I)-chain is hidden in its telopeptides (see *Telopeptides*).

Fibrillogenesis

Assembly of collagen monomers into fibrils is an endothermal entropy-driven process, where the hydration water around monomers is dislodged and rearranged during polymerization. Also, the critical concentration of collagen decreases with higher temperature, reminding of other polymerizations, like flagella or tobacco mosaic virus protein[24]. The final, millimeter-long fibrils have a characteristic 67 nm axial periodicity and are further assembled into longer and thicker fibres. Depending on tissue requirements, the fibril diameter varies from 20 nm in cornea to 500 nm in tendons. Fibrils can be arranged not only in a longitudinal manner (parallel bundles) as in tendons, but also in orthogonal networks (cornea), or in weaved basket form (bone) – the shape once again reflecting the tissue needs. Notably, the network architecture is driven by specific, yet undiscovered, cues *in vivo*, since fibroblasts passed into cell culture are unable to organize the matrix that would resemble the tissue of origin. Somehow, cell sensing and signaling through matrix receptors is involved in the process[42].

Initial extracellular fibrillogenesis of collagen and cellular sensing of the matrix

Although some of the initial fibril assembly starts in Golgi-to-PM transport carriers[20], most of the process completes in the matrix. This end stage may be controlled by the cell surface collagen-binding integrins including $\alpha_1\beta_1$, $\alpha_2\beta_1$, $\alpha_v\beta_3$, $\alpha_{10}\beta_1$, $\alpha_{11}\beta_1$ – they are all distinguished from other integrins by a collagen-binding I-domain in the α-subunit. All cells differentially express these integrins: α_1 is traditionally ascribed to the mesenchymal, and α_2 to the epithelial cells, although fibroblasts can express both subunits simultaneously[43]. During activation, $\alpha_2\beta_1$ evokes upregulation of MMP-1 and -13, while $\alpha_1\beta_1$ appears to provoke a negative feedback on these collagenases. In addition, p38 – a stress-related MAP kinase – is upregulated through $\alpha_2\beta_1$ activation.

In certain pathological conditions, expression pattern of these integrins may shift the phenotype of the cells with respect to collagen synthesis. For example, PDGF and TGF-β can upregulate $\alpha_2\beta_1$ and downregulate $\alpha_1\beta_1$[44]. One question is if these, or other, integrins recognize different structures of collagen in extracellular matrix, and if so, wouldn't this be intimately connected to the expression of other matrix molecules, like SLRPs?

$\alpha_2\beta_1$ integrin is active during collagen fibrillogenesis, as its inhibition (in cell culture) delays this process. Interestingly, inhibition of $\alpha_5\beta_1$ integrin or fibronectin mediates similar effect, which implies dependence of initial fibrillogenesis on fibronectin network. Fibronectin is a major ligand for collagen, and fibronectin-deficient fibroblasts without α_2- and α_{11}-integrin subunits cannot attach to collagen in cell culture; however, they regain this ability after transfection with the corresponding integrin expression vectors. Similarly, growth of these cells on a fibronectin coat triggers the formation of collagen network[45, 46].

Other collagen receptors include DDR-1 and DDR-2 (discoidin domain receptors) – tyrosine kinase receptors that are non-responsive to growth factors. Phosphorylated tyrosines still do interact with SH2 and PTB domains, and trigger matrix remodeling as well as cell growth, migration and differentiation. DDR1-deficient mice are small, have swollen glomerular basement membranes, non-secreting mammary gland epithelium, and

hyperactive MAPK. DDR2-deficient mice are also small, although due to diminished chondrocyte proliferation. This implies a tissue-specific usage of DDRs, which is also apparent in cancers where epithelial-derived cells express DDR1 while the stromal cells produce DDR2. Furthermore, DDR2 specificity is exclusive for fibrillar collagen[47] that is abundant in the stroma.

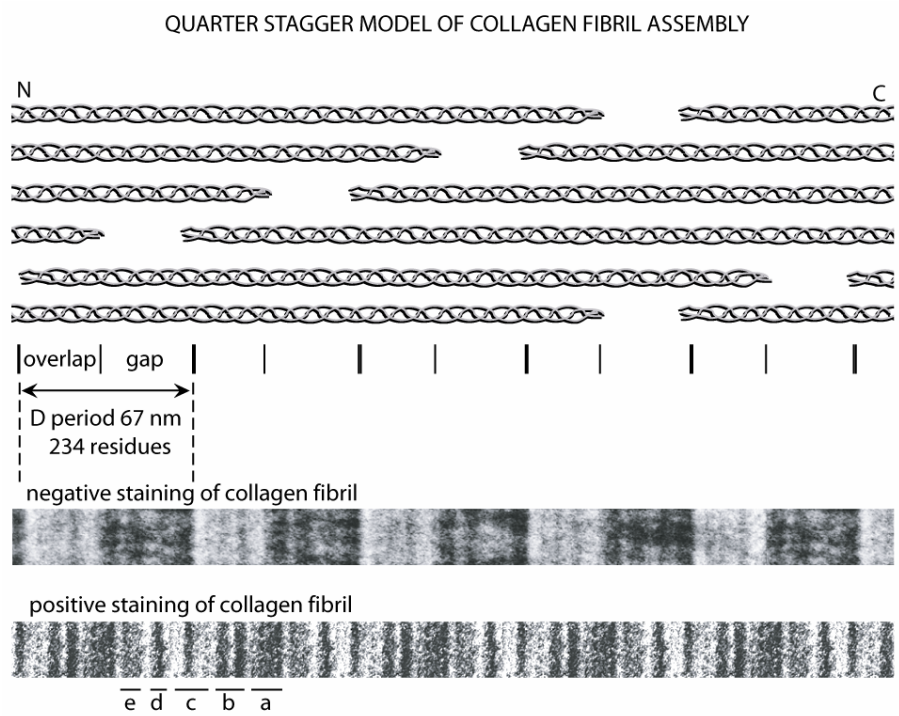
A recently found collagen-internalization receptor expressed only on mesenchymal cells is Endo180 or uPARAP – a member of mannose receptor family (constitutive endocytosis receptors). Even though fibroblasts isolated from the knockout mice are unable to efficiently internalize collagen, there is no obvious mice phenotype[48]. uPARAP preferentially collects precleaved collagen fragments rather than native collagen, which indicates a precedent, concerted action of MMPs[49]. In tumors, uPARAP is upregulated by stromal cells that internalize and degrade collagen, but lack of this receptor, maybe paradoxically, reduces the tumor burden[50].

In conclusion, cells can regulate the extracellular matrix formation by sensing and responding to the environment. Adhesion to the matrix is also relevant for their survival, migration, and differentiation. It is now recognized that these events are crucial for the outcomes of pathological conditions, where matrix remodeling is in the focus of wound healing, angiogenesis, and fibrotic responses. To carry out these processes, altered collagen synthesis, organization of matrix, and involvement of SLRPs is required.

Collagen fibril growth

After the commencing phases of collagen polymerization, the thin segments are assembled into larger fibrils, thereby increasing in both length and diameter. Growth proceeds by fusion of pre-formed thinner fibrils, both laterally – controlled by collagen- binding proteoglycans like decorin, and head-to-tail where at least one C-terminal collagen end is required[51]. This is accompanied by formation of covalent cross-links between collagens, which further enhances the mechanical strength of the connective tissue. In this context, the role of SLRPs is crucial, judging from knockout mice phenotype and experiments showing that decorin and fibromodulin can bind to and regulate collagen fibril formation *in vitro*[52-55]. Also, regarding collagen type I, other collagens are involved in the process, in particular the FACIT collagens XII and XIV, and the fibrillar collagens III and V. Some SLRPs can also interact with FACIT collagens, for example decorin with collagen XII and XIV, and fibromodulin with collagen XII[56, 57]. All these complementary proteins seem to play unique roles during fibril formation. FACIT collagens, covering the fibrils, aid in the sliding of collagen fibrils past each other during integrin-dependent collagen gel contraction[58]. This is possibly aided by decorin that may connect fibrillar collagens with FACIT. SLRPs themselves appear to regulate the fusion of thinner fibril intermediates into larger fibrils, which would influence intermolecular cross-linking of collagen. Clearly, even though collagen can self-associate into fibrils due to its inherent properties, it cannot develop the functional architecture and plasticity without the additional building blocks.

An intriguing phenomenon in a developing chicken embryo tendon is the formation of either unipolar or bipolar early collagen fibrils (that is, the molecular polarity N→C is either unchanged throughout or changed within the fibril). In the bipolar fibrils, there is a transition region where collagens join tail-to-tail, with their C-terminal ends pointing towards each other, so that the N-terminal ends face outwards. The alternative head-to-head fusion does not appear to take place, so the only bipolar fibrils during development have N-N polarity[59]. The early fibrils later associate into larger units, under condition that a lengthwise fusion connects at least one C-terminal end of another fibril unit.



Proteoglycans like SLRPs that reside on the collagen would act against preemptive lateral fusion, allowed first after formation of a sufficiently long fibril. The latter would in turn be determined by the amount of bipolar fibrils that can no longer fuse in longitudinal direction[51].

In vitro model of collagen fibrillogenesis

Collagen fibrillogenesis can be studied *in vitro* using monomers extracted from tissue with acetic acid. Optionally, pepsin is used to cleave parts of the telopeptides and thus remove the intermolecular cross-links (which increases solubility). These collagen monomers aggregate into fibrils with a characteristic 67 nm D-periodic staining pattern (as also present *in vivo*) at physiological conditions. The drawback of this system is that fibrils are usually no more than 40 nm in diameter (in tendons they can be up to 500 nm), their shape does not resemble the *in vivo* organization, and the fibril packing is less ordered. Clearly, the role of other proteins and cellular control of fibril growth is significant for these properties. However, this model of collagen fibrillogenesis has been used to study interactions with matrix molecules, and the nature of fibril assembly, but due to its limitations it is not fully reflecting the *in vivo* process.

Fibrillogenesis is traced by measuring turbidity at 400 nm, as collagen aggregates disperse the incoming light beam. A typical fibrillogenesis curve has an initial lag phase where smaller collagen polymers are formed, then a log phase where these precursors aggregate into large fibrils, and lastly a plateau phase where no more growth occurs. This

process can be manipulated by enzymatic cleavage of parts of the terminal non-helical sequences of collagen. Pepsin treatment (mentioned above) or pronase digestion of telopeptides results in delayed initial nucleation of fibrils. Removal of the N-terminal telopeptide with leucine aminopeptidase also prolongs the lag phase. On the other end, removal of the C-telopeptide by carboxypeptidase quickens the lag phase but reduces lateral growth during the later stages. In conclusion, the N-telopeptide of collagen regulates initial fibril formation, while the C-telopeptide promotes lateral growth of the precursors[60].

In vitro, using pC-collagen (with intact propeptide C) treated with pC-propeptidase, the newly formed bipolar (N-N) fibrils have finely tapered or blunted ends[61]. Their growth, examined live in a microscope, reveals accretion preferably at the pointing tip; however, on some fibrils the blunt end also grows, but with a newly developing pointing tip. Examination of the banding pattern of collagen shows that the N-terminal ends of monomers are directed toward the pointed tips, which further substantiates the theory of preferred C-to-N assembly of collagen[62]. Fibrils examined in electron microscopy resemble those found in tissues, with respect to a uniform diameter, shape, banding pattern, and the pointed tips.

Collagen molecules appear to interact by balancing the hydration-promoting forces of repulsion (at near distances) and the hydration-depleting forces of attraction (at longer distances) between charged and hydrophobic amino acids. Attraction forces can be reduced by solutes like glucose, glycerol, or 1,3-propane diol. Interestingly, 1,2-propane diol (a hydrophobic attraction disrupter) has no significant effect on fibrillogenesis, even though both propane diols also change dielectric properties of solutions. This implies that collagen fibrillogenesis is not driven by hydrophobic or electrostatic interactions but rather by special recognition sites that arrange water molecules in hydrogen bonds between the helices[63-65], a process that involves hydroxyprolines[65]. It seems that the two propane-diol isomers differently alter the structure of water, which influences collagen fibrillogenesis in two different ways.

Repulsion between collagen helices is highest at lower temperatures, when interaxial spacing is at its peak (can reach over 17Å). Conversely, attractive forces are dominating at higher temperatures, removing water molecules from between the collagens (low-hydrated fibers have 12Å spacing). In other manners, the interaxial spacing can be restrained by intermolecular cross-linking, or osmotic stress mediated by salts or non-penetrating neutral solutes (PEG, dextran). In the latter case, high osmotic pressure decreases axial spacing due to dehydration of collagens[66].

Telopeptides – determinants of cross-linking and fibrillogenesis

Telopeptides of collagen contain not only information for spatial orientation of collagens, but also the cross-linking lysine residues. The C-telopeptide is located near the triple helical five-fold repeated Gly-Pro-Hyp sequence, and contains a number of hydrophobic residues: three Phe in α1(I), and two Tyr and one Phe in α2(I). Altogether, strong hydrophobic interactions associate the α-chains in this region. In the C-terminal part of the telopeptide, the two α1(I)-chains are longer than α2(I), and they contain the Lys residue involved in cross-linking (α2(I) lacks it). According to molecular modeling, this residue bulks out – a feature that is mediated by vicinal prolines that kink the telopeptide, and by glycines that promote the chain flexibility. In such a position, the lysine interacts with the triple helical part of another collagen, creating an intermolecular cross-link.

The α1(I) C-telopeptide terminates with two Tyr residues that contribute to the hydrophobic C-terminal interactions. α2(I)-chain lacks these properties, but it still provides stabilizing N-terminal hydrophobic amino acids that actually prevent a disoriented folding of α1(I)-telopeptides and misaligned cross-links[67]. In an x-ray diffraction study, the C-telopeptides of α1(I)-chains had a hairpin conformation, presumably allowing the proper molecular register for intermolecular cross-link. This may also explain the shorter length of alpha2(I) C-telopeptide, since the need for accommodation of the bulky hairpins[68].

N-telopeptides of both α1(I)- and α2(I)-chains contain Lys residues involved in cross-linking to the triple helix. Proximal to these are hydrophilic amino acids (Asp-Glu on α1(I) and Asp on α2(I)) that could aid the interaction with the helical telopeptide receptor. According to molecular modeling studies on N-telopeptides, a type I β-turn forms when a hydrogen bond connects Asp and Ser residues, and Lys projects outwards. The N-terminal part of the α1(I)-telopeptide contains the hydrophobic Leu and two Tyr residues that project to the interior of the chain; in addition, Ile is contained within the C-terminal part of the telopeptide. Similar properties are found in the α2(I)-chain: it carries Phe, but also (in the C-terminal end) three Gly residues. Thus, a hydrophobic pocket can be formed to stabilize the β-turn structure of the telopeptide. In molecular modeling studies, the N-telopeptide lacks a regular structure in solution, but adapts the hairpin-fold conformation when adjacent to another collagen triple helix, through induced fit[69]. A question arises here: since there are three different orientations possible for the three telopeptidal lysines on a collagen monomer, does the azimuthal orientation of the formed cross-link affect the final fibril stability? And further, do the SRLPs influence this?

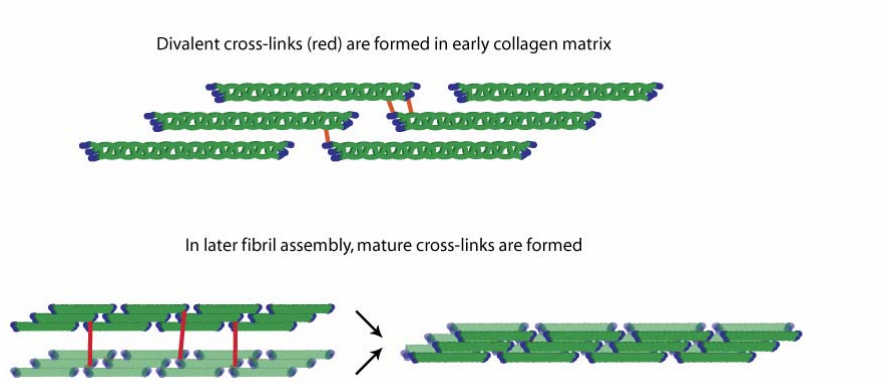
TELOPEPTIDE SEQUENCES OF HUMAN COLLAGEN TYPE I

	N-telopeptides	C-telopeptides
α1(I)	QLSYGYDEKSTGISVP	SAGFDFSFLPQPPQEKAHDGGRYYRA
α2(I)	QFDAKGGGP	GGGYDGYDGD FYRA

Peptides based on telopeptide sequences of collagen can inhibit collagen fibrillogenesis *in vitro* (reconstituted collagen fibrils), having a prominent effect on the lag growth phase, and binding via hydrophobic interactions to the 776-822 region of the collagen triple helix[70]. This part of the helix is flexible (contrary to the rigid structure of the large part of the triple helix), which makes it accessible for specific collagenase activity[71] and it may, for the very same reason, serve as an induced fit-pocket for telopeptide interaction. In conlusion, the telopeptide function is not limited to cross-linking, as collagen treated with pepsin (that cleaves telopeptides) has a delayed self-assembly rate, which can in turn be influenced by the addition of free telopeptides[60, 70, 72].

Cross-linking

Intermolecular cross-linking partially sustains the mechanical integrity of collagen, even after heat denaturation or enzymatic cleavage[73]. In general, due to the arrangement of assembled collagen molecules (quarter-stagger), the telopeptidal lysines form cross-links with the helical lysine residues. The N-terminal telopeptide of α1(I)-chain uses Lys-9N, and α2(I) – Lys-5N. They can cross-link to the C-terminal helical α1(I)Lys-930 or α2(I)Lys-933. In the other terminus, the C-terminal telopeptides of α1(I)-chain use Lys-16C, while α2(I) is devoid of any potential cross-linking site. Lys-16C can combine with the N-terminal helical α1(I)Lys-87 or α2(I)Lys-90. Interestingly, the only two histidines in the triple helix of α1(I)-chain are adjacent to the helical cross-linking lysines, possibly contributing to the process[74, 75].



Cross-link formation is initiated in extracellular matrix by the action of lysyl oxidase that converts the telopeptidal lysines or hydroxylysines into aldehydes (thus producing allysines or hydroxyallysines). In the consequent step, they connect to a helical lysine or hydroxylysine and form an immature divalent cross-link; either hydroxylysylnorleucine (HLNL) or dihydroxylysylnorleucine (DHLNL). They later mature into trivalent cross-links by reacting with another telopeptide allysine or hydroxyallysine. This gives lysyl pyridinolines or hydroxypyridinolines, respectively[75].

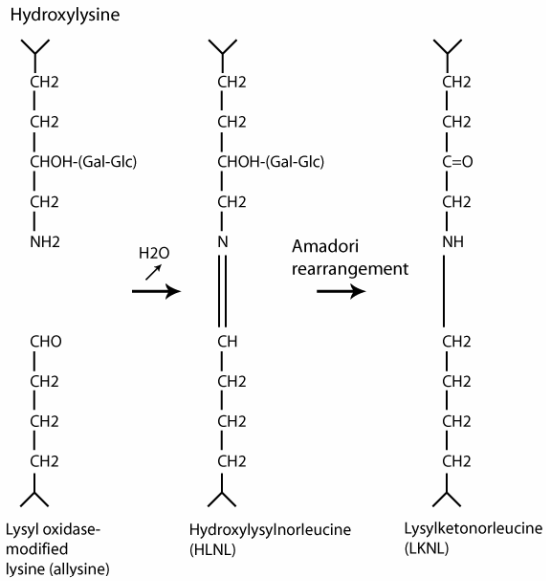
Ketoamines (in DHLNL) can react with allysine or hydroxylysine, and in the respective case a pyrrole or a pyridinoline cross-link will form. Pyrroles are made as ketoamine reacts with the aldehyde group of allysine, producing a Schiff base adduct, which then closes into a ring form. Pyridinoline is formed by reaction of ketoamine with aldehyde of hydroxyallysine, producing a Schiff base adduct that undergoes Amadori rearrangement; then, the ring closure occurs, followed by enolization.

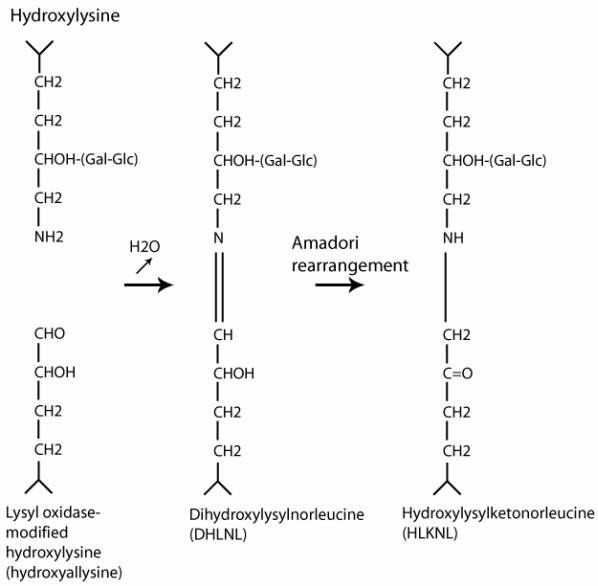
Pyrrole cross-linking involves three lysines – two in telopeptides and one in the helical part of collagen, and the major locus (in bovine tendon collagen) is α2(I)Hyl-933 x α1(I)Lys(Hyl)-9N x α2(I)Lys(Hyl)-5N. Other pyrrole cross-links include α1(I)-chain lysines or hydroxylysines at positions 87, 930, or 16C. Pyrroles also contain glucose and galactose that are attached to lysine hydroxyl groups. Similar sites of cross-linking have been found in pyridinolines[76-79] .

How are the different cross-links distributed in tissues? Hydroxylysylpyridinolines (HP) are abundant in cartilage, bone, dentin, ligament, tendon, aorta, and lung. Lysylpyridinolines (LP) are relatively high in bone and dentin. In contrast, both HP and LP comprise just a minor fraction of cross-links in skin, cornea, and basal laminae[80]. In fibrotic conditions, HP cross-links are upregulated due to overexpression of lysyl hydroxylase-2b, and this fortifies the neostromal structure[81].

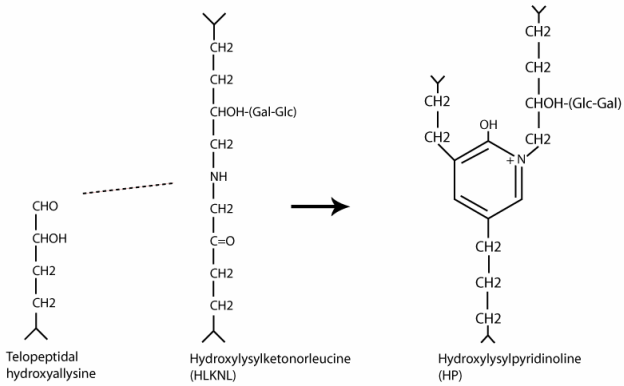
In human bone collagen, the amount of pyridinolines is lower than in non-mineralized tissues, contrary to the higher amount of divalent cross-links. This is explained by biomineralization arresting the cross-link maturation process, which is not evident in other tissues. In bones, pyridinolines are evenly distributed at both ends of collagen, with pyrrole and LP often found at the N-telopeptide site. However, the C-telopeptide site involves exclusively α1(I)-chains because α2(I) in vertebrates lacks the C-telopeptidal Lys residue. Formation of either pyrrole or pyridinoline cross-links is determined by the extent of lysine hydroxylation at telopeptide sites[78], but the physiological relevance of either process for the product that makes up our connective tissues is not known. Also, possible involvement of SLRPs in these mechanisms could prove valuable to explore.

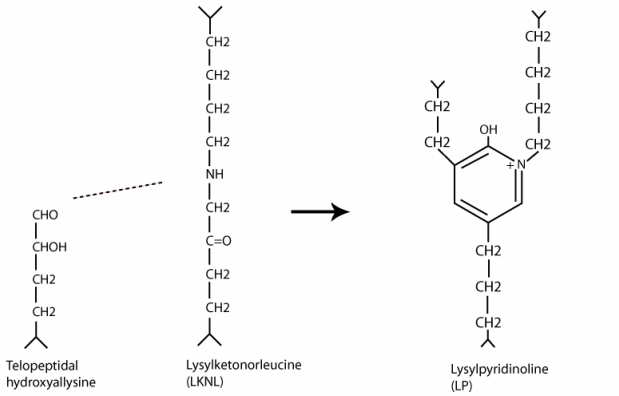
FORMATION OF DIVALENT COLLAGEN CROSS-LINKS

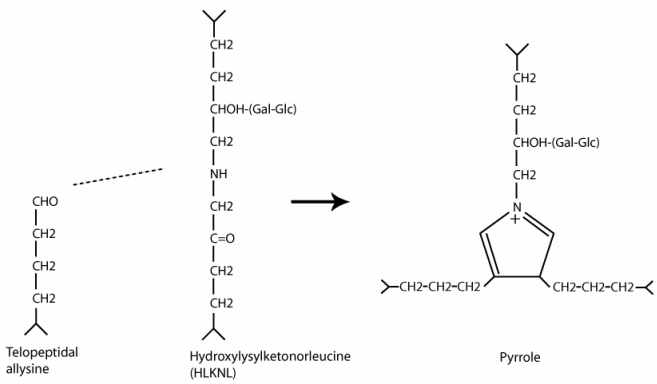




FORMATION OF TRIVALENT COLLAGEN CROSS-LINKS







Collagen-modifying enzymes

Peptidyl-prolyl cis-trans isomerase

Pre-procollagen is synthesized with about 1/10 of prolines having peptide bonds in *cis*-conformation; therefore, cis-trans isomerization is a rate-limiting step in the assembly of collagen triple helix (proline content of collagen type I is 18%). Denaturation of collagen and subsequent incubation with peptidyl-prolyl cis-trans isomerase increases the triple helical folding rate three-fold, compared with native protein[82]. From there, prolines and hydroxyprolines contribute to fold the polyproline II-like helix – that is, the preponderant secondary structure of collagen α-chains.

Prolyl-4-hydroxylase

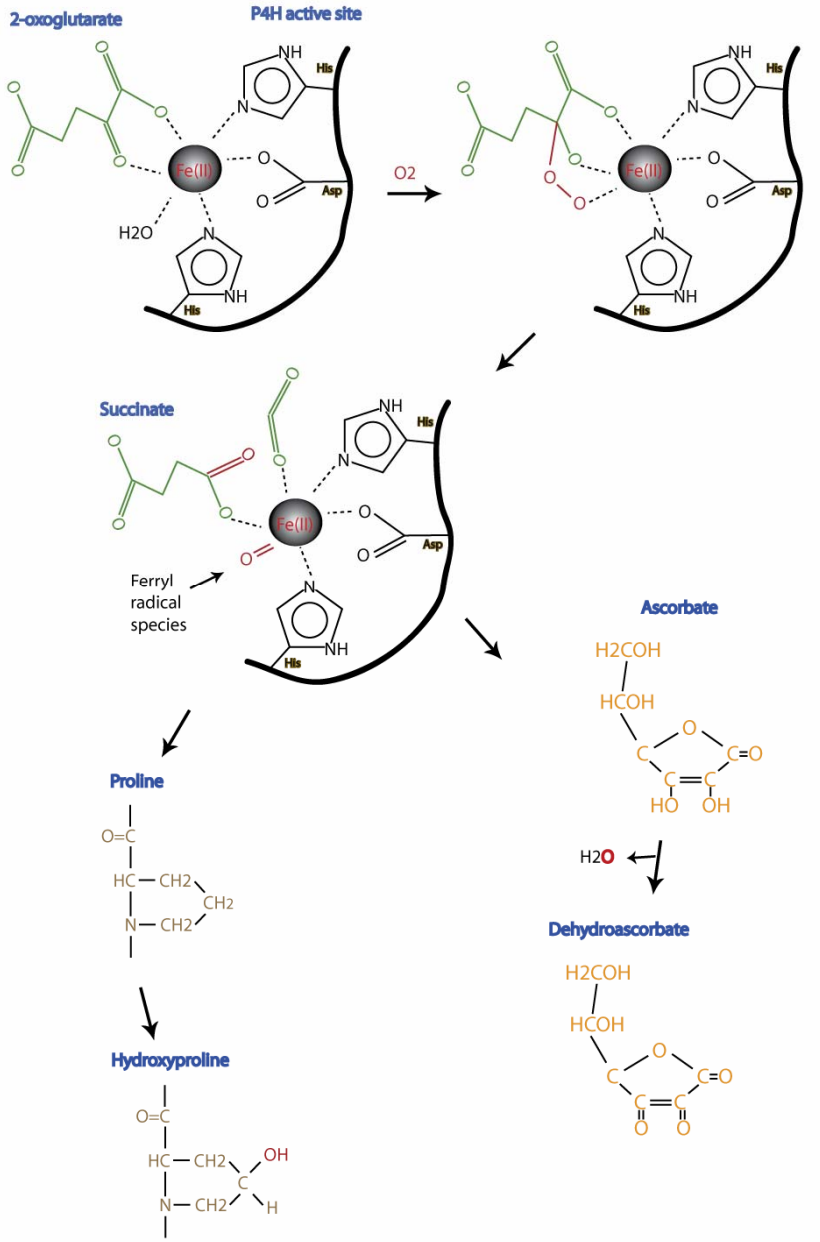
Prolyl-4-hydroxylase (P4H) is a tetramer consisting of two catalyzing α-subunits, and two structure-stabilizing β-subunits that are identical with protein disulfide isomerase (PDI)[83]. The enzyme mediates hydroxylation of prolines on α-chains, essential for formation of triple helices, and its inactivity retains procollagen in ER, which leads to scurvy.

For its function, P4H requires Fe^{2+}, ascorbate, 2-oxoglutarate, and O_2. The hydroxylation reaction is the following: 2-oxoglutarate is decarboxylated to succinate by one atom of oxygen, while the other oxygen atom creates a reactive ferryl radical species that hydroxylates the proline. Ascorbate is not consumed stoichiometrically, as opposed to 2-oxoglutarate, but is used for uncoupled (that is, without proline hydroxylation) decarboxylation of 2-oxoglutarate. There, it accepts the ferryl oxygen, thus preventing self-oxidation of P4H[84] (see figure on the next page).

Experimentally, this reaction can be inhibited by α,α'-dipyridyl (Fe^{2+}-chelator), or by depleting the ascorbate from the cell medium. This accumulates procollagen within the ER, in a non-triple helical structure and bound to P4H[85].

What is the function of hydroxylated prolines? In short, thermal stability. Measured by melting point determination, non-hydroxylated collagen denatures at 15ºC lower temperature[86], and the amount of hydroxyprolines in the triple helix is proportional to its melting point[87]. Through x-ray crystallography on collagen peptides, the conclusion is that the stability is mediated by a coat of and interchain formation of water bridges; these bind through hydrogen bonds connecting the hydroxyl and carbonyl groups of the α-chain hydroxyprolines and peptide bonds[88]. Substitution of proline for hydroxyproline in Gly-**X**-Y repeat of a collagen peptide gives lower enthalpy of denaturation (thus lower entropy), which further supports the theory that hydration and consequent water bridge formation is stabilizing collagen[89]. An alternative view points out that not the water bridges but the *trans*-conformation of hydroxyproline peptide bond fulfills this function. The conclusion is drawn from the fact that substitution of hydroxyproline with fluoroproline in a collagenous peptide increases its melting point, and fluoroproline is not readily forming hydrogen bonds with water[90]. However, considering the fact that collagen helix is stable just above the body temperature[40] it may be using the hydroxyprolines to build up a network of water bridges that gives collagen its "metastability". The substitution of fluoroproline surely gives higher thermostability but this does not need to reflect the physiological needs for a "metastable" collagen.

PROLYL 4-HYDROXYLASE REACTION MECHANISM



Protein disulfide isomerase

The propeptide domains in procollagen are covalently cross-linked, both interchain and intrachain, via disulfide bonds. If this process is inhibited, it triggers retention of procollagen in ER by PDI-mediated quality control, as PDI associates with monomeric collagen[30]. The S-S bond shuffling is then catalyzed by the ER-residing PDI, and this accelerates the trimerization of α-chains[91]. PDI, as commonly known, is not collagen-specific as it accepts other protein substrates. This enzyme also comprises a part of the P4H quarternary structure, forming the two β-subunits that restrain the enzyme from aggregation by non-specific S-S bridges[92, 93].

Lysyl hydroxylases

Specific collagen lysine residues can be hydroxylated by three isoforms of lysyl hydroxylases (LH1, LH2 and LH3), which all act on the helical X-**Lys**-Gly, but only one (LH2) on the telopeptidal lysines. The isoforms are encoded by the genes *PLOD1, PLOD2,* and *PLOD3,* and the proteins are 60% homologous. Patients with connective tissue disorders, where lysine hydroxylation of collagen is altered, have a different pattern of collagen cross-linking[94, 95], which manifests in weakened or hypertense collagenous network.

LH1 deficiency in humans (due to mutated *PLOD1* gene) causes kyphoscoliosis, or Ehlers-Danlos syndrome VIA – an autosomal recessive disorder. Lh1-deficient mice have slow locomotion probably due to dislocated joints and weak muscles, while some mice die prematurely from aorta rupture-caused hemorrhages in thoracic and abdominal cavities. This phenotype is related to irregularly shaped collagen fibrils with less pyridinoline cross-links[96]; notably, LH1 probably hydroxylates the helical lysine residues α1(I)Lys-87, α1(I)Lys-930, and α2(I)Lys-933 – that is, the lysines involved in cross-linking of collagens[97].

LH2 can be found in tissues as two splice variants. The most abundant form is LH2b that includes a proline-rich sequence encoded by an extra 63 bp exon; this is the exclusive LH2 variant in skin, lung, and aorta. Interestingly, most organs express both LH2b and LH2a, so the tissue-specific expression of LH2b may coincide with specific hydroxylation loci on collagen[98]. Missense mutations in *PLOD2* gene are correlated with Bruck syndrome, an autosomal recessive disease causing osteoporosis and fragile bones. Adding to its uniqueness, LH2 is the only hydroxylase acting on the telopeptidal lysines (X-Lys-Ala or X-Lys-Ser), but it hydroxylates helical lysines as well[99]. It is overexpressed in systemic sclerosis fibroblasts, hepatic stellate cells, and scars, where amounts of LH1, LH3, and LOX are not significantly increased. In inflammation, LH2b is triggered by profibrotic cytokines, including TGF-β, IL-4, and TNF-α[100]. Hyperactivation of LH2 consequently leads to a changed collagen cross-linking pattern with dominating pyridinoline cross-links[81, 95, 101]. LH2 is also involved in over-hydroxylation of collagen in hypertrophic tendon[102], and *in vitro* it upregulates DHLNL and pyridinoline cross-links in collagen produced by osteoblast-like cell line[103].

LH3 has both lysyl hydroxylase and lysyl glycosyltransferase activities, but the active sites are not co-localized, and glycosyltransferase activity is located in the less conserved region of LH3[104]. Lh3-deficient mice die at mid-embryonic stage, have undeveloped and thin basement membranes, relating to the fact that collagen IV (a basement membrane-specific collagen) is highly hydroxylated and glycosylated; lack of Lh3 hinders these modifications and retains collagen IV within the cells. The result is a collagen network with reduced number of pyridinoline cross-links. On the other hand, Lh3 deficiency does not appear to affect expression of other basement membrane components, including laminin. The hydroxylase activity of Lh3 may be compensated by other Lh, since mutations of just this active site (with retained glycosyltransferase activity) renders viable mice with less disorganized basement membranes[105]. Glycosylations of lysines therefore play a crucial, although yet unknown, role in collagen cross-link formation.

Hsp47

Hsp47 is a procollagen-specific chaperone sequestered to endoplasmic reticulum through its C-terminal RDEL sequence. Hsp47-deficient mice die before embryonal day 12, due to misformation of collagen triple helices (the secreted collagens are trypsin-sensitive, which is not normal). Consequently, the mice have disorganized basement membranes and irregular reticular fiber network in the mesenchyme[106]. Hsp47 is not homologous with other heat-shock proteins or chaperones, and actually belongs to the serpin family, although it is upregulated and binds to procollagen during heat shock (other chaperones included in the process are Grp78 and Grp94)[107]. For recognition, it requires triple helical Gly-X-Y sequence where Y is Arg, and it binds only weakly to single α-chains. It also prefers non-hydroxylated prolines, implying prevention of the hydroxyproline-poor regions against thermally induced unfolding[108-110].

In conclusion, Hsp47 is not involved in folding of collagen per se, since it binds the ready-made trimer and protects it against unfolding[111] (it should be restated that collagen helix melts just above the body temperature[40]). The interaction with collagen is terminated during the ER-to-Golgi transport, and Hsp47 is recycled to ER[9]. How the collagen is stabilized in later stages of transport is not known, but possibly, through early polymerization in transport vesicles, its thermal sensitivity declines.

Lysyl oxidase

Collagen cross-linking is initiated by the copper-dependent enzyme lysyl oxidase (LOX). There are five different genes encoding LOX in human and mouse genomes (*LOX, LOXL, LOXL2-4*), all positioned on different chromosomes. The most abundant transcripts found in adult tissues are those of LOX and LOXL, while the other three oxidases are only present in minute amounts. After knocking out the *Lox* gene, the mice die early after parturition due to ruptured aorta and diaphragm, probably after sudden increase in blood pressure and lung respiration that sets in after birth. This phenotype is explained by insufficient collagen and to a larger extent – elastin – cross-linking, which weakens the aorta and diaphragm connective tissues[112]. In cancer, Lox expression is boosted by HIF (hypoxia-inducible factor) in hypoxic tumors, and its presence triggers distant metastasis of breast tumors in mice[113]. This event may as well be connected to a changed collagen matrix structure.

Lox is secreted as an inactive pro-enzyme, requiring the action of bone morphogenic protein-I (procollagen C-proteinase) to remove the propeptide and activate it. Also, mammalian Tolloid-like-1 (mTLL-1) is able to cleave the propeptide, albeit at a lower efficiency, and *Bmp1-/- Tll1-/-* fibroblasts retain only 30% of Lox activity[114].

With aid of a carbonyl cofactor (lysyl tyrosyl quinone), LOX oxidizes lysine ε-amino groups in collagen, elastin, as well as in a number of other substrates, like histone protein H1. The amino group is oxidized to peptidyl α-aminoadipic-δ-semialdehyde that can later react with ε-amino groups of other lysines, to form initial cross-links between collagen molecules[115]. LOX does not bind to monomeric collagen, but requires at least a partial fibril state[116]. Therefore, the instatement of cross-links may be preceded by SLRP-driven alignment of collagens.

How does LOX recognize the proper lysines? It has a high cationic affinity, but the telopeptidal lysines are joined by acidic amino acid residues, and LOX is barely oxidizing the peptide G_4-DEK-G_4 – the N-telopeptidal sequence of α1(I)-chain. However, its k_{cat}/K_m increases dramatically if Asp in this sequence is exchanged with Gln. It seems then that acidic amino acids in the telopeptides are restraining the activity of LOX until after the initial fibril formation; then, other collagens could use their basic amino acids for neutralization of

the acidic environment, in result promoting LOX activity[117]. Perhaps the histidines near the cross-linking helical lysines contribute to the process.

Other collagen modifications

Besides P4H, proline hydroxylation can also be carried out by prolyl-3-hydroxylase, although it hydroxylates but one residue – Pro-986 on α1(I)-collagen chain (more prolines are modified by this enzyme in collagens IV and V). It can bind both folded collagen and gelatin, and is expressed prominently in tissues with fibrillar collagens[118]. Lack of prolyl 3-hydroxylase gives a recessive osteogenesis imperfecta-like disorder, with osteopenic and thinner bones, and reminds of the one caused by mutations in collagen genes *COL1A1* and *COL1A2.* Patients with non-hydroxylated Pro-986 on α1(I)-chain have overmodified collagen with 1 ºC higher melting temperature[119, 120].

Another collagen modification, with less clear function, takes place in the ER and relies on the Mn^{2+}-dependent glycosylation enzymes UDP-galactosyltransferase and UDP-glucosyltransferase. The glycosylation is placed on the hydroxylated lysines, and it could affect collagen fibrillogenesis, akin to the function of lysyl hydroxylase-3[105, 121, 122]. Collagen lysines and hydroxylysines can also be glycosylated in a non-enzymatic fashion, when free sugars react with ε-amine groups and eventually form advanced glycation end products. These can increase the stiffness of collagen in cartilage that becomes less functional and therefore less resistant to mechanical pressure, possibly being one of the mechanisms behind the development and the age correlation of osteoarthritis[123].

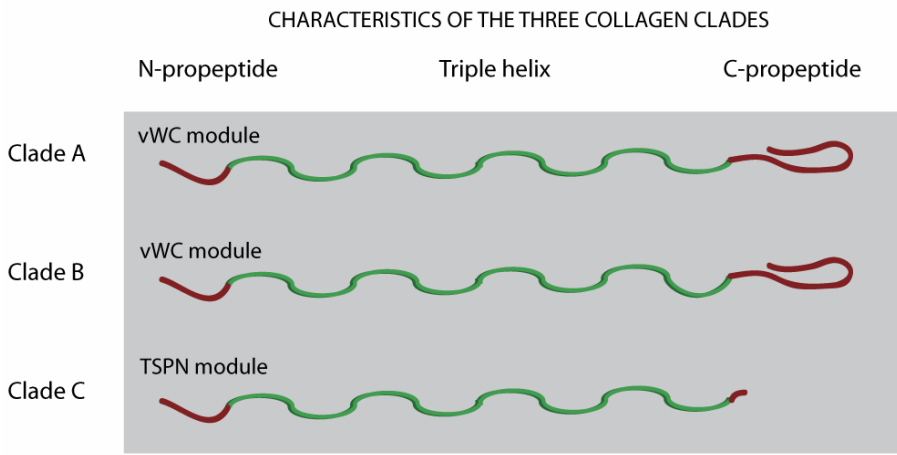
COLLAGEN AND SLRP EVOLUTION

Collagens represent some of the most archetypal proteins within the animal kingdom, found already in sponges (Parazoa). Also, collagen-like proteins are present even in bacteria (e.g. Scl-2 or hyaluronidase of *Streptococcus pyogenes*) although these may represent acquired gene material. Polymers of long and stable fibers made collagens suitable for development of specialized connective tissues, and they also constitute parts of non-matrix proteins like acetyl cholinesterase, and some innate immunity players like C1q, collectins, and macrophage scavenger receptors[34]. It is fascinating how one protein can be used in so many ways, simply by "reshuffling" the genome, while retaining its self-assembly properties and adding novel functions. As it now appears, the evolution of connective tissues has gone from the very simple, non-organized matrices of sponges, through the intermediate form of mesoglia of diploblasts, to the complex, specialized collagen networks found in our body.

There are two things in common among all animal collagens: they have the characteristic long triple-helical domain that may or may not be interrupted, and a well-conserved C-terminal non-collagenous propetide domain involved in selection of α-chains for the trimers. From studies of the most primitive animals, including sea urchins and sponges, it is obvious that the ancestor fibrillar collagens are thin and most similar to mammalian collagens V and XI[34]. Looking at the vertebrate collagen V, it can actually limit the diameter of newly forming collagen I fibrils *in vitro*[124], and collagen V-deficient mice die early in embryogenesis due to severely distorted, or total lack of, fibrillar collagen in mesenchymal tissues[125]. The role and evolutionary relationship of collagen V is perhaps not surprising, since the evolution of collagen fibrils would reasonably start with the smaller units and proceed to the wide variety of diameters and structures present in higher metazoa. Interestingly, some animals like fruit fly, or nematode worm, did not retain these collagens, but developed an alternative body support in form of exoskeleton. On the other hand, the class-related (insecta) mosquitoes still express fibrillar collagens. However, what seems to be common throughout metazoa development is the presence of basement membrane collagen – already sponges, the most primitive animals, have collagens with NC-domains homologous with nematode cuticle collagens and with mammalian collagen IV[34]. In later evolution, from these interrupted triple helix-types of collagens probably FACITs emerged, which can also be argued from the structural similarities observed in NC-domains (like cysteine bridges), as well as the modifying glycosylations that have functional significance in collagen IV and FACITs[105].

Phylogenetic analyses reveal that the vertebrate fibrillar collagens can be divided into three clades, that is "ancestor" collagen with all its descendants. Clade A includes all α-chains of collagen I, II, and III (all being quantitatively highest in expression in humans), as well as α2(V) chain. Clade B contains α1(V), α3(V), α1(XI) and α2(XI), while clade C hosts collagens XXIV and XXVII that have shorter and not as conserved triple helices. The clade grouping is based on exon structure, and N-terminal NC- domains that are nearly identical within the respective clades (B and C are identical). Other characteristics include a minor triple helix domain in the N-terminus and a C-terminal propeptide chain selection sequence in A and B, but not in C clade.

One notable exception is the lack of, or rather the existence of a much shorter, N-terminal domain in the pro-α2(I) chain. Therefore, α2(I) could represent the archetypal form of collagen chain, from which, through insertion mutations and further duplications, other collagen genes may have ensued.



A few theories describe how the diversity of interactions between fibrillar collagens, and between the different α-chains, has evolved. One explanation regards the variability of the chain selection sequences in the C-terminal propeptide of α-chains. The difference between invertebrates and vertebrates is the formers' general lack of 7-8 amino acids in this sequence. Consequently, in early vertebrate evolution a unique genomic event of insertion of such a sequence, with accompanying mutations, would allow a trial-and-error selection of multitudal collagen interactions. This is called the molecular incest theory, and assumes a common ancestral vertebrate that carried each of the three clades of collagen genes[31, 126].

A contrasting theory, essentially a "non-incest" model, of collagen clade evolution points out that already some invertebrates, including protostomes, have clade A, B and C in their genomes, all being very similar to the human collagen genes. In addition, an invertebrate like sea urchin (Echinoderm) can produce heterotypic fibrils[127]. Altogether, this pushes back the origin of the three collagen clades to the ancestors of Bilateria phylum. A parsimonious explanation is that the contemporary diversity of collagens simply arose by general rules of genomic duplications and mutations during the long evolutionary process. The formation of the C-propeptide selection sequence would therefore not have evolved by a rare genomic event (molecular incest) but rather been refined through the ages[128].

A conspicuous feature of vertebrates is the variety of connective tissues and the relatively higher amount of matrix genes (relative to total gene number) that becomes apparent in animals evolutionarily close to the invertebrates. For example, a comparison of matrix genes of *Ciona intestinalis* (sea squirt) with human genome hints at traces of genomic duplication events that occurred during evolution and inspired the formation of new connective tissues. In Ciona genome, 90% of the matrix genes are orthologues with the human matrix genes, and may represent the founders that were later refined through evolution.

A common theory is that a two-fold genome duplication event took place early during vertebrate evolution, followed by gene loss and mutations. In human genome, the presence of 400-500 paralogons of similar length and with homologous sequences supports this theory. In line with this, the 10 collagen genes in Ciona are each orthologue with 1-5 human collagen genes (31 in total). Other genes encoding matrix proteins and collagen-modifying enzymes are similarly related. In conclusion, the diversity of connective tissues in vertebrates would have emanated after the divergence of urochordates. In later stages, duplications of each of the clades of fibrillar collagens generated bone-specific collagens I,

V, XXIV and cartilage-specific II, XI, and XXVII – all three representing the respective clades A, B, and C[127, 129-131].

When did the collagen-binding integrins evolve? Mammalian cells recognize collagens through integrin αI-domains, but these have only been found in urochordates and vertebrates, not earlier. Still, the αI domain of the urochordate representative – *Ciona intestinalis* – binds only one fibrillar mammalian collagen IX, independently of Mg^{2+} or MIDAS, implying a urochordate-specific collagen recognition sequence. Mammalian αI-domains recognize the GFOGER sequence (O=hydroxyproline) on collagens, and have a more extended structure, which is found earliest in teleosts (like pufferfish). It is probable then that the collagen-specific integrins of vertebrates evolved after the divergence of ascidians. However, this branching was preceded by appearance of laminin- and RGD-recognizing integrins[132-134], which may reflect a vertebrate's pre-requirement of basal laminae and fibronectin for attachment of cells on, and formation of, collagen matrix.

As for the SLRP genes, in Ciona there are three founders identified and labeled SLRPa-c. Intriguingly, SLRPa gene is a direct orthologue with all human class I, II, and III SLRP genes (so far, twelve in total), and they are all paralogons on chromosomes 9q22, 12q23, Xq28. From this point of view, twelve out of fourteen known SLRP proteins would share one ancestor gene. This wide range of duplication events cannot be ascribed to SLRPb that only roots PODN (podocan) and FLJ23447 (hypothetical protein); similarly, SLRPc only roots CHAD (chondroadherin) and LOC15036. Some further peculiarities arise when tracing the matrix gene duplication in the Bilateria lineage. For example, Drosophila and nematodes have no collagen and SLRP genes except for the basal lamina collagen IV and XVIII, while honeybee and mosquito still carry the fibrillar collagens and the related proteoglycans[129, 135]. Perhaps a late divergence in these lineages still allowed for different solutions in formation of exoskeletons.

SMALL LEUCINE-RICH REPEAT PROTEOGLYCANS

As of now, 14 proteoglycans belong to the family featuring the characteristic tandemly repeated leucine-rich domains (LRRs) that are each 20-30 amino acids. From mammals to tomatoes to plague bacteria, LRR in itself is a widespread protein domain, covering the most basal cell mechanisms, like RNA splicing. However, the SLRP family is distinguished due to internal homologies, GAG chains (not on all SLRPs though), and the consensus cysteine-rich sequence ($CX_{2-3}CXCX_{6-9}C$) at the otherwise variable N-terminal end. The family is subdivided into five classes, sorting out the above features and the number of exons in the genes. Class I contains: biglycan, decorin, asporin; class II: lumican, fibromodulin, PRELP, keratocan, osteoadherin; class III: epiphycan, osteoglycin, opticin; class IV: chondroadherin; and class V: podocan. The five classes of SLRPs also differ in their N-terminal cysteine-rich cluster patterns. Class I has Cx_3CxCx_6C. Class II: Cx_3CxCx_9C. Class III: Cx_2CxCx_6C. Class IV: Cx_3CxCx_8C. Class V: Cx_3CxCx_7C.

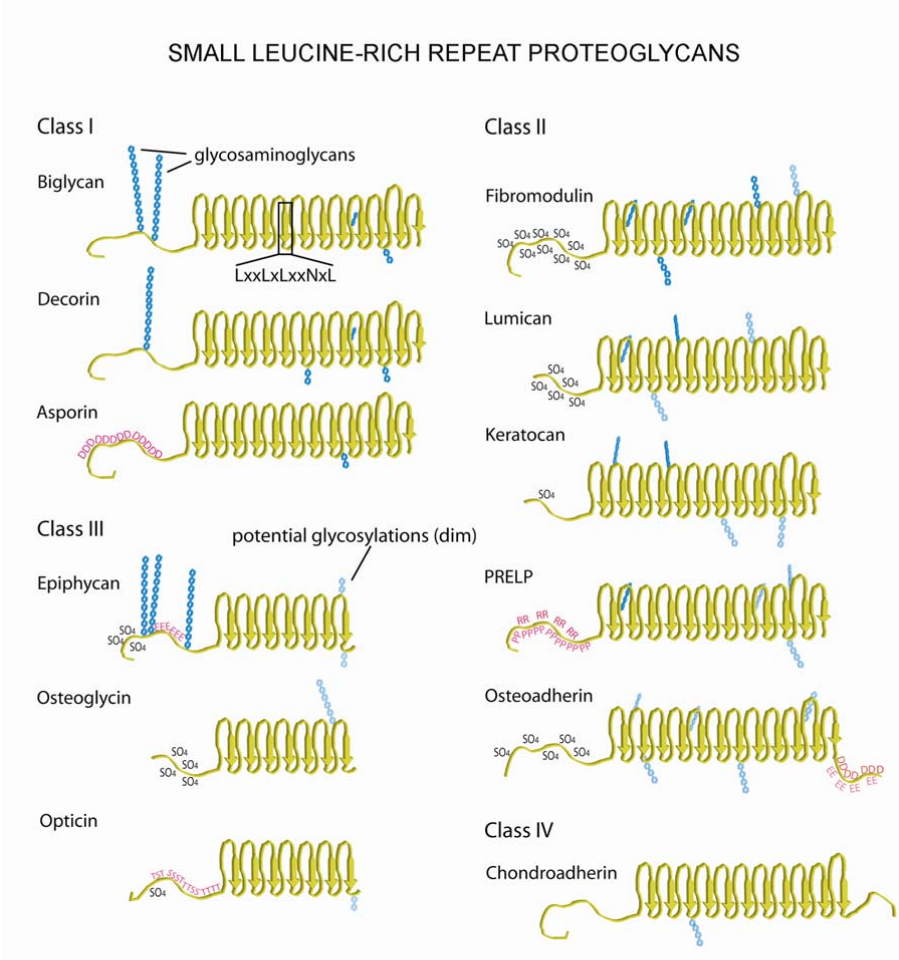
The model proposed by Matsushima N., *et al.*[136] is perhaps the most sensible to describe the variances between these different classes. It acknowledges two types of LRRs – the shorter, called *S:* xxaPzxLPxxLxxLxLxxNxI, and the longer, denoted *T:* zzxxaxxxxFxxaxxLxxLxLxxNxL (where **x** is any amino acid, **a** is Val, Leu, or Ile, and **z** is frequently a gap). Also, I (Ile) is sometimes substituted with L (Leu). Using this *S* and *T* nomenclature, the pattern of tandem LRRs in class I and II SLRPs is *STTSTTSTTSTT* (12 LRRs), and in class III: *STTSTST*. Class IV chondroadherin has 10 *T* repeats, while podocan of class V has 21 LRR repeats with a tandem *STT* pattern.

Crystal structures of decorin and biglycan show that the highly conserved LxxLxLxxNxL part of the consensus LRR repeat forms a β-strand, while the rest of the

sequence runs in antiparallel direction, forming short β-strands, 3_{10} helices, and polyproline II helices. The most variable repeats in class I and II are the two C-terminal LRRs, the first being several amino acids longer and the second being shorter than the other repeats. In contrast, in class III the most C-terminal LRR is 5-7 amino acids longer than the other *T* repeats. These sequence variations, within the otherwise well-conserved overall structure, may have some functional significance that discerns the SLRPs from each other, although this is yet to be confirmed experimentally.

The modifying sugar chains also vary between the different classes, as well as within the SLRP classes. In general, class I proteoglycans have CS/DS (except asporin), class II is associated with KS, and class III epiphycan can have three CS chains. Chondroadherin (class IV) has one potential O-glycosylation site, although it is not glycosylated as purified from tissue. Lastly, purified podocan is only a glycoprotein, with three N-glycosylations. In addition, the sugars may be more or less sulfated, depending on the spatial location, age, and health condition of the tissue. The identity of GAGs contributes to the function of collagen-bound SLRP, which is not least evident from the studies on tooth development (see *Tooth development*).

A number of tyrosine sulfations in the N-terminal part of the protein core can also be found on class II and class III SLRPs, with exception of PRELP that has a proline- and arginine-rich N-terminal domain. This, together with the sulfated GAGs, mediates an acidic property to the proteoglycans. Some SLRPs have distinct acidic qualities in their N-terminal domains: in opticin the domain is serine- and threonine-rich (and can be phosphorylated), in epiphycan it has a polyglutamate, and in asporin – a polyaspartate tail. Osteoadherin is somehow unique in having a glutamate- and aspartate-rich C-terminal domain that is not a LRR, and a similar glutamate-rich C-terminal tail is found in podocan. All these unique features of the individual SLRPs are not yet investigated for their functional relevance[136-138].


SMALL LEUCINE-RICH REPEAT PROTEOGLYCANS
Class I
glycosaminoglycans
Biglycan
LxxLxLxxNxL
Decorin
Asporin
Class II
Fibromodulin
Lumican
Keratocan
Class III
potential glycosylations (dim)
Epiphycan
Osteoglycin
Opticin
PRELP
Osteoadherin
Class IV
Chondroadherin

Structures of SLRPs and ligand interactions

The long stretch of tandemly repeated LRRs creates a bent overall shape, and the curvature degree depends on the secondary structures formed on the convex face of the protein (more α-helices give larger bend). The length of the LRR domains probably determines which secondary structures will form – the shorter *S* repeats prefer polyproline II helices, while the longer *T* repeats have short β-strands or 3_{10} helices. The concave face, on the other hand, is dominated by parallel β-strands, and these are implicated in protein-protein interactions. This is also the most conserved part of LRR proteins. Lastly, the hydrophobic residues, including conserved LRR leucines, point to the interior of the solenoid structure and stabilize the overall structure[139].

The X-ray crystal structure of decorin protein core shows that dimerization can occur through a large surface (2300 $Å^2$) of the protein's concave part, and a similar dimer structure is seen in biglycan crystals[140, 141]. Another SLRP, opticin, is also a dimer, according to light scattering studies[142]. Looking at the interacting surfaces, it is striking that about 90% of the solvent-exposed amino acids are conserved or partially conserved within the class I SLRPs. Similar homology patterns are seen in the presumed concave surfaces of class III, although not in class II proteoglycans. Dimerization could potentiate interactions with the ligands, but it is not yet known whether dimers dissociate into monomers upon binding. In the latter case, what could affect the equilibrium state between dimers and monomers? From an energetic perspective, the large surfaces that associate during dimer formation would render the structure very stable in solution. However, conformational changes induced by interactions with ligands could disturb this balance, possibly regulating the affinities of the SLRPs during different stages of collagen fibrillogenesis.

One possible function of the conserved cysteine-rich cluster in the N-terminal end of the protein core can be inferred from the crystal structure. In the N-terminus, the four cysteines form two pairs of disulfide bridges between the first LRR domain and the preceding antiparallel β-strand. This cap could protect the long LRR domain from unfolding and aggregating. A similar disulfide knot, made of two conserved cysteines (or four in chondroadherin), is found in the C-terminal end where it connects the last and the penultimate LRR repeat. The latter domain is usually extra long, and contains several non-homologous amino acids that protrude in a loop from the solenoid structure (termed ear-like domain by Scott P.G., *et al.*[140]). A different C-terminal capping motif is found in chondroadherin and podocan, the former having four cysteines, and the latter having none. These variations could convey specific functions of each SLRP, although this remains to be investigated.

The concave surface and the β-sheet-loop region of LRR proteins is often used for ligand interactions, especially with other proteins[143]. However, some LRR proteins, like TLR3, use their convex surface to bind dsRNA, and other recognition variants are reported. What sets these proteins apart are insertions of unique secondary structures that accommodate the non-protein ligands[144, 145]. To conclude, not only SLRPs but also other LRR proteins use their conserved tandemly repeated domains as protein interaction surfaces. They may however carry additional ligand-binding sites, in form of non-conserved insertions or sugar chains. All this essentially conveys a picture of a LRR protein being involved in multiple interaction mechanisms, and it would be interesting to investigate if SLRPs could be functionally regulated through binding to more than one ligand.

Surveying the literature, attention has been drawn to the fact that some SLRPs (biglycan, decorin, fibromodulin, asporin) can interact with TGF-β[146], and could thereby impose a physiological impact on cellular growth factor environment. In light of this, it must be considered that collagen – the major protein of extracellular matrix – and fibronectin, another major matrix component, also bind TGF-β[3, 147]. The physiological relevance of this event is not entirely clear, although it is proposed that matrix proteins would act as

extracellular inhibitors or depots of TGF-β. However, studies on SLRP knockout mice do not reveal any imminent effects of excessive activity of TGF-β (like excessive matrix production, suppressed immunity, etc.), which could be due to the abundant presence of this growth factor in relation to a more specialized distribution and possibly the amount of SLRPs. Naturally, TGF-β signaling is so complex that it may not be an easy task to clearly connect the function of SLRPs with the bioavailabilty of TGF-β *in vivo*. SLRPs may inhibit and function as storage molecules for TGF-β in the matrix, but direct evidence is lacking. In the upcoming sections, analysis of the individual SLRPs and their TGF-β interactions will follow.

Class I SLRPs

Asporin

Asporin is preferentially expressed in heart, liver, aorta, uterus, and in osteoarthritic cartilage. Contrary to biglycan and decorin, asporin has no consensus Ser-Gly O-glycosylation site, although it carries O-linked oligosaccharide on Ser-54, and N-linked oligosaccharide coupled to Asn-281. One uniqueness of asporin lies within the long polyaspartate tail in the N-terminal domain, polymorphic in humans, with 8-19 residues encoded by alleles D8-D19[148, 149]. This polymorphism has been linked to susceptibility to osteoarthritis, as the D14 allele is overrepresented among knee osteoarthritis patients, while D13 dominates in the control groups[150]. This correlation is not without its complexities, including ethnic differences. For example, European populations are less susceptible to the D14 allele, although according to a meta-analysis study, they are affected to a significant degree. Such differences may be due to environment, variations in study designs (post-surgery versus symptomatic patients), and other genes. In addition, knee osteoarthritis seems to be more influenced by D14 asporin, contrary to the hip variant of the disorder[151]. To explain the influence of one extra aspartate on pathobiology of the disorder one group proposed the changed inhibitory activity of D13 and D14 on TGF-β-driven chondrogenic differentiation[150]. In this study however, the experiments include comparison of transfected cells overexpressing the asporins, and internal clonal differences may influence the results. Also, the asporin-TGF-β interaction studies are based on immunoprecipitation and do not report any dose-response relations.

Since several proteins involved in biomineralization of collagen contain long sequences of polyaspartates or polyglutamates (calcium-binding), one possible function of asporin could include vertebrate bone calcification and development. In bone, hydroxyapatite crystals are deposited, at least initially, in the gap regions of collagen fibrils from which further nucleation of the crystal emanates. The orientation of the crystal in relation to the collagen fibril appears important as well, and collagen-bound acidic proteins may be implied in hydroxyapatite nucleation. Some matrix proteins, like bone sialoprotein, are capable of forming hydroxyapatite crystals in agarose gels[152], while osteopontin can inhibit this process; these effects can be diminished by chemically altering the carboxyl groups into non-acidic residues. Within this context, polyglutamates have about 1,000-fold lower activity compared to polyaspartates, which could be due to different secondary structures[153].

Asporin is one of the SLRPs expressed by fibroblasts in periodontal ligament of teeth, a tissue which is never ossified *in vivo* despite being surrounded by alveolar bone and cementum (where asporin is not present), and from which the mesenchymal stem cells differentiate to form the surrounding tissues. There, the inhibitory effect of asporin on collagen mineralization may be triggered. It was suggested that this effect is signaled through BMP-2 (mineralization inductive factor) inhibition by asporin[154], but no studies on direct role of asporin in calcification have been reported. In developing bones, asporin is expressed predominantly in perichondrium and periosteum, as well as in knee joint tendons[155]. However, it is not found in ephyseal cartilage, articular cartilage, or growth plate,

until at a later age. Considering the above information, it could be worth to examine the role of polymorphic polyaspartates in biomineralization of collagen, not least since asporin interacts with this major matrix protein (unpublished observation).

Asporin is also upregulated by TGF-β which may then be inhibited by asporin, but only at several 100-fold molar excess of this SLRP[155, 156]. Even though TGF-β may be partially regulated in a negative feedback loop by asporin, the reported ablation of TGF-β-mediated chondrogenic effects requires such excessive amounts of the protein that the finding may not relate to the *in vivo* environment.

Biglycan

Biglycan, originally called PG I, was found as a proteoglycan with two O-linked GAG chains, expressed by bone-derived cells[157], and is the only SLRP so far localized to the human X chromosome (Xq28)[158]. Biglycan is expressed in tendons during development of fibrocartilage that forms at the junction site with bone[159]. Additionally, it is present during cartilage and bone formation, during kidney differentiation [160], while much transcript can be found in lung, spleen and liver[161].

Looking at the regulatory expression mechanisms, biglycan is suppressed by retinoic acid in chondrocytes, which can influence chondrogenesis[162]. Intracellularly, it is upregulated by cAMP/protein kinase A pathway, which is important for differentiation of chondrocytes and osteoblasts[163]. During Xenopus development, biglycan serves as a negative regulator of BMP4 activity through interaction with Chordin – a BMP4 antagonist. This determines cell fate along dorsoventral axis, as inhibition by biglycan induces secondary axes in Xenopus embryos. In this pathway, the BMP4 gradient is regulated by Xolloid, and biglycan – secreted as pro-biglycan – may be itself regulated by this protease that cleaves the propeptide[164, 165]. The functional significance of this propeptide sequence is not known, however biglycan in its pro-form is most abundant in adult articular cartilage and scarce in other tissues or in young cartilage, which alludes to tissue-regulated processing of biglycan. Up to date, it is only known that removal of the propeptide induces conformational changes that affect C-terminal epitope availability for anti-biglycan antibodies. This could certainly also affect interactions with other proteins, like collagen, which is formed in the tissue and processed by the same enzyme – the Xolloid homologue, BMP-1 (procollagen C-proteinase)[165, 166]. Since collagen propeptide C removal reduces its solubility and induces nucleation, biglycan processing would correlate with this early phase of fibrillogenesis, with yet an unknown effect.

BMP4 also determines bone osteoblast differentiation, and in biglycan-deficient mice (that have an osteopenic phenotype) osteoblasts have low sensitivity to BMP4 (low SMAD-1). This results in loss of expression of bone-specific markers (like osteopontin, osteocalcin, BSP) and reduced calcium accumulation. Biglycan, being expressed pericellularly, also appears to aid in binding of BMP4 to its cell surface receptor, and could thus provide an additional cell-matrix interaction, besides collagen fibril regulation[167].

Biglycan-deficient mice grow slower during the first months of life, and have reduced bone mass with thin, less calcified trabeculae, reminding of osteopenia[168]. Furthermore, bone marrow stromal cells that differentiate to fibroblasts proliferate slower, are more apoptotic, and have a low expression of collagen I – this phenotype was explained by dysregulation of TGF-β in absence of biglycan[169], since this SLRP interacts with the cytokine[146]. In cell culture, fibroblasts overexpress biglycan when stimulated by TGF-β, and biglycan may act in a negative feedback loop for the cytokine[170]. To counter this reasoning, it has been shown that TGF-β regulates biglycan expression on a post-transcriptional level (cycloheximide-sensitive) while biglycan promoter activity is unaffected. The promoter is on the other hand supressed by TNF-α, and its activity is dependent on p38 MAP kinase and Smad4/DPC4 – this implies a cross-talk between the two signaling pathways connected by early Smad-induced GADD45β that onsets the MAPKKK signaling[171-174]. Similarly, in

decorin/biglycan-deficient mice, changed TGF-β activity and localization could affect the fate of the differentiating bone marrow stromal cells, and interestingly, mice deficient for TGF-β II receptor have high bone mass[175]. Another explanation, relating to the reported effects of biglycan on collagen[176], is that an impaired fibril formation would influence the matrix/niche in which these cells differentiate.

Other studies on biglycan regulation show its unique role in bone development. Even though the main biomineralization factor, BMP-2, has no effect on biglycan message (but upregulates decorin)[177], some other findings deserve attention. Glucocorticoids that differentiate bone marrow stromal cells into osteoblastic phenotype are simultaneously supressing biglycan and amplifying biomineralization[178]. The relative amount and distribution of class I SRLPs may therefore secure the proper calcification of the bone. To add to the story, osteoclasts in biglycan-deficient mice are overactive, reportedly due to biglycan normally safeguarding the differentiation of the osteogenic osteoblast precursors[179]. Biglycan, together with fibromodulin, can also regulate the tendon stem cell niche through interactions with bone morphogenic proteins or through influence on matrix structure[180].

Interestingly, bone metabolism is not affected in female biglycan-deficient mice and they are also resistant to ovariectomy-induced bone resorption[181], while males have reduced bone mass and strength, despite geometrically larger tibial diaphyses[182]. This would possibly connect the phenotype of these mice to estrogen activity in bone, or simply due to some uncontrolled genetic changes in the X chromosome (where the biglycan gene is located) acquired during the gene targeting procedure.

What happens if additional SLRPs are “knocked out”? Combined with fibromodulin deficiency, the mice develop premature osteoarthritis in knee joints and ectopic ossification in tendons, which is amplified with a frequent use of joints. Temporomandibular joints are also prone to accelerated osteoarthritis, although less than in articular cartilage. Tendons in these double-deficient mice are also weaker, with thinner and irregularly shaped collagen fibrils, a phenotype that is more severe than in either single knockout mouse[183, 184].

Other effects are seen in biglycan/decorin-deficient mice, where posterior-frontal suture fusion (the only fusing suture in mice) is impaired, with hypomineralized embryonal bones and upregulated BMP4 contributing to the fusion process. Since decorin is also upregulated in the cranial bones of biglycan-deficient mice[185], these are additional clues that reveal the two SLRPs’ synergic role in bone mineralization.

To contrast with the other class I SLRPs, biglycan deficiency gives weaker flexor digitorum longus tendons and normal tail tendons. The situation is different in decorin-deficient mice whose patellar tendon[186] and tail tendons are weaker, and while biglycan is upregulated in these tissues it is not able to restore the mechanical strength or a proper collagen fibril assembly[187]. Biglycan expression also rises with compressive loading applied on tendons, while decorin levels remain constant. This kind of stress triggers formation of fibrocartilage in the tissue, as cells express more TGF-β[188] – the biglycan-triggering cytokine. Shortly, biglycan and decorin act in distinct manners during connective tissue development, despite close homology and compensatory regulation. The exact nature of their functions remains to be unveiled.

In skeletal muscle, biglycan expression is low in adults, but high in newborn mice. Even in this tissue, biglycan-deficient mice have upregulated decorin, akin to some functional redundancy or co-regulation. No clear function of biglycan has been related to skeletal muscle, but its absence delays barium chloride-induced regeneration of fibers and myosin expression[189]. It is also upregulated during muscle regeneration and in dystrophic muscle. One hypothesis can be based on biglycan binding to α-sarcoglycan (in dystrophin-associated protein complex) and to collagen VI (mutated in congenital muscle dystrophy) – it could act as a bridge between cell and matrix, maintaining the integrity of muscle cell membrane and connecting the cytoskeleton to the extracellular matrix[190-192].

Biglycan binds collagen type I, appearing in the gap zone of the fibrils, as judged by immunogold labeling, and it is a binding for which decorin is able to compete[193]. Biglycan also interacts with collagens I, II, and III in an *in vitro* collagen fibrillogenesis assay, but without restricting the final fibril size, as opposed to decorin[194]. As mentioned above, biglycan also binds the microfibrillar collagen VI, close to the N-terminus (also here decorin can compete for the binding), and can organize the thin fibrils into a hexagonal-like network by connecting the network's inner junctions. In both cases the GAG chains have insignificant effect on the interaction; however, they are required to organize the collagen VI microfibrils[195, 196]. Furthermore, biglycan (or decorin) in complex with matrillins can form a bridge between collagen VI and collagen II, or aggrecan[197], which could be important for arranging the fibrillar network of the cartilage.

Decorin

Decorin is found in many connective tissues, and several phenotypes are reported on decorin-deficient mice that have a disturbed collagen matrix. An inchoate study described abnormal collagen fibril fusion in skin and tendons, with the former tissue acquiring fragility[54], and the latter being mechanically weaker[187]. Decorin deficiency also has apparent effects on the structure of kidney mesangial matrix in diabetes type I model mice. There, the glomerular morphology is altered, with consequent albuminuria and fibrosis[198]. Another phenotype is found in lungs where the airway resistance is decreased, together with heightened compliance of the respiratory system[199]; in contrary, decorin amount is decreased in asthma patients and in fibrotic lung disease model, where airway resistance is rised[200, 201]. This questions decorin's decisive role in the progression of these disorders – at least, the stromal architecture that controls the lung mechanics seems to depend more on other factors, with decorin as a contributor to the overall chain of events.

Other matrix deficiencies due to lack of decorin include slower wound healing and delayed infiltration of blood vessels[202, 203]. In wound healing, where fibrin matrix is layed down to serve as a provisional scaffold for the upcoming tissue repair, decorin interacts with fibrinogen and influences formation of the fibrin network and its contraction[204-206]. Some reports state decorin overexpression being inhibitory for scar formation in fibrotic conditions[207], but these studies are hard to evaluate due to use of controls like empty expression vectors or PBS.

How do the abnormally shaped fibrils of decorin-deficient mice relate to the role of decorin in tendon? To begin with, one study has stated the correlation between fusion of collagen fibrils and decreased amount of collagen-bound decorin[51], which sets this SLRP as a regulator of collagen fibril size. During chick tendon development, the decorin amount decreases at embryonic day 18[208], but is more abundant during postnatal development (simultaneous with reduced biglycan amount). Interestingly, expression of biglycan is higher in decorin-deficient tendons, which parallels a similar phenomenon in fibromodulin-knockout mice (where lumican is upregulated)[187]. It seems that biglycan and lumican can at least partially compensate for the lack of their other subclass SLRPs, and they may have some redundant function, although the upregulation may simply be triggered by the changes in constitution of collagen network. Presumably, SLRPs can work on collagen fibrillogenesis in a coordinated manner.

Decorin has one N-terminal O-glycosylation site, substituted with either CS or DS. The functional significance of this is not well researched, although it has been shown that removal of the (uncharacterized) chain mediates formation of larger collagen fibrils in cell culture, interestingly without overexpression of biglycan[209]. Decorin GAG can also interact with collagen XII and XIV, which would place decorin as a link between FACIT collagens, and collagen I with which it interacts with its protein core[56, 57, 210]. How the sugars, attached to collagen-bound decorin, may affect collagen fibrillogenesis is worth investigating, since the different GAGs per se grant distinct effects on collagen fibrillogenesis[211]. Conceivably,

GAGs may simply present a steric hindrance during the fusion process; another alternative is the reciprocal interaction of decorin sugars, at a force of about one-third of decorin-collagen interaction[212] – a process that could influence collagen assembly. What has been lacking in the studies so far is the identification of the decorin sugar chains, and how each of these could commit fibrillogenesis into different paths.

How much decorin is present in the tissues is determined not only by expression control but also by catabolic activity of MMPs. Decorin can actually delay MMP-1 and MMP-13 activity *in vitro*[213], and since it binds near the C-terminal part of collagen (25 nm into the chain)[214] it could possibly mask the MMP cleavage site located about 75 nm from the C-terminus. However, this effect is time- and concentration-dependent since decorin can also be cleaved by MMP-13 (at Ser240-Leu241), although not as efficiently as fibromodulin or biglycan[215]. Additionally, MMP-2, -3, and -7 are able to cleave decorin into small fragments, which probably reflects a general matrix remodeling event during inflammatory reactions[216].

The interaction of decorin at one end of the collagen monomer places it close to the lysine involved in intermolecular cross-linking of collagen[75]. Whether decorin indeed has an effect on specific cross-linking regulation should be investigated, since fibromodulin – another collagen-binding SLRP – has such a function (see *Present investigations*). It is already known that addition of decorin to cell culture of *Dcn-/-* fibroblasts does induce cross-linking between the α-chains[217], but it remains to examine the cross-linking *pattern* in collagen from decorin-deficient mice. Possibly, it is different from the one influenced by fibromodulin (or lumican), since decorin does not interfere with fibromodulin-collagen interaction[210]. Since all three SLRPs can affect collagen fibrillogenesis *in vitro*[52, 53, 218], but possess different binding sites, they could control the different cross-link formations. Decorin interaction with fibronectin[219] could further fine-tune the collagen matrix assembly, perhaps by acting as a bridge between the cells and the matrix.

Decorin can be upregulated by estrogen[219], and is somehow involved in TGF-β signaling. In fact, several studies point to a dual role of this SLRP, involved in both matrix and cell signaling regulation. The most abiding idea is developed around the fact that decorin binds TGF-β[146]. However, a clear relation between these two proteins is difficult to settle. Decorin in a myoblast cell line does not affect the amount of phosphorylated Smad-2 or translocation of Smad-4, although a decorin-dependent PI3K pathway, triggered by TGF-β, is affected through inhibition of decorin binding to a lipoprotein receptor-related protein (LRP-1). Other pathways, including ERK, IGFR-1, and p38-MAPK are not altered by the same restraint[220]. Decorin also inhibits the TGF-β-driven collagen production in tendon-derived cells[221], although the effect may be secondary depending on altered matrix structure sensed by the cells. A similar objection may explain the report that myoblasts from *Dcn-/-* mice bind more TGF-β to their responding receptors[222], or that bone marrow stromal cells become overactivated by TGF-β in absence of decorin[175]. Some comparable studies propose that decorin treatment during muscle healing improves regeneration and prevents scar healing [223], possibly through restricting a pro-fibrotic member of TGF-family – myostatin[224]. This would relate to decorin being overexpressed in endomysium and perimysium in Duchenne muscle dystrophy[191], although another report claims the opposite[225]. These studies could be followed up by investigating if SLRPs alter the bioavailability/signaling of TGF-β by simply modulating the matrix and cell response. To go one step further, one should also consider that collagen and fibronectin – much more abundant than SLRPs – also bind TGF-β, but in which conditions? Can the cytokine affinity change with an altered, decorin-driven remodeling of the tissue?

Concerning this question, a presumed inhibitory activity of other SLRPs on TGF-β has been reported. However, these conclusions have been inferred from the following: firstly, the molar ratio in which a marked inhibition is achieved is extremely high. In the first study on decorin[226], over 100-fold molar excess of decorin was required to inhibit TGF-β activity by 50%, and similar amounts were used to inhibit TGF-β-induced scarring in experimental kidney disease[227]. Contradicting these findings, in other systems decorin excess amplified the activation of TGF-β signaling[228, 229], or would be without effects[230]. These discrepancies were explained by the assertion that decorin would stimulate TGF-β

activity in low concentrations and inhibit it in higher concentrations (although this has not been proven). Another proposition was that localization of decorin, either pericellular or intercellular, would differentiate these effects. Taken together though, decorin does not appear to be an efficient TGF-β inhibitor/stimulator *in vivo*, due to the stoichiometric relation, and since TGF-β is primarily regulated by LTBP protein. The observed effects may be because SLRPs are homologous with more potent TGF-β inhibitors, just as decorin is homologous with a natural EGFR inhibitor LRIG1, in particular its LRR domain that is shed from the cell surface[231, 232]. By homology, a TGF-β binding site, although not well-conserved, could contribute to the SLRP affinity for the cytokine. Another hypothesis is simply that SLRPs influence collagen matrix in a way that allows or restricts the flow of fluid to the cells, thus indirectly affecting TGF-β availability

Class II SLRPs

Fibromodulin

Fibromodulin is associated with dense or mechanically strong connective tissues, like tendons, ligaments, and cartilage. It has five potential N-linked glycosylation sites, and was extracted from bovine articular cartilage carrying four KS chains[233, 234]. These substitutions are also in scleral and tendon fibroblast-produced fibromodulin[235]. In addition to GAGs, the proteoglycan can have up to nine tyrosine sulfations in its N-terminal part[236]; however, neither modifications have been functionally characterized. The protein core, on the other hand, interacts with collagen I and II, at other sites than decorin, but on or near the binding site of lumican[52, 210, 237]. If reduced by dithiotreitol, fibromodulin can no longer inhibit fibrillogenesis *in vitro* but can still interact with collagen[238], which alludes to the functional importance of the conserved SLRP cysteine loops. Fibromodulin can also bind collagen XII, and possibly form a bridge between the fibrillar and the FACIT collagens[56].

The binding to collagen I is near the gap zone of the assembled collagen fibrils[239]. On CNBr-derived collagen fragments, fibromodulin interacts with CB7, CB6 (together comprising the C-terminal half) and CB8 (N-terminal part) of α1(I)-chain, and with CB4 (N-terminal part) of α2(I)-chain[240]. However, the data on CNBr-fragments (chemically modified) do not pinpoint the actual binding site, which would otherwise prove valuable to examine the hypothetic role of fibromodulin in regulation of collagen cross-linking.

Fibromodulin-deficient mice have weaker tendons with abnormally shaped collagen fibrils – a large population of thinner fibril segments blends with irregularly fused big fibrils[55]. Thus, fibromodulin is required for assembly of collagen intermediates into larger units, conceivably through prevention of uncontrolled lateral accretion. During metatarsal tendon development in chicken embryos, fibromodulin increases 6-8-fold between day 14 and 19, that is when the fibril growth is dramatically induced, partly by TGF-β2 and IGF-1[241, 242].

Fibromodulin is also upregulated by TGF-β in cartilage, with less affected decorin and biglycan expression[243] – this has implications in osteoarthritis where the proteoglycan is remodeling the collagen network[241, 244]. In this tissue though, fibromodulin-deficient mice have no collagen fibril abnormities, but still develop osteoarthritis, possibly through defective joint ligaments that fail to restrict the mechanical damage imposed on the articular cartilage[245]. In addition, ligaments of these mice overproduce lumican that could have some redundant function, since both SLRPs bind within the same collagen region[237]. What further supports this view is that fibrils in both knockout mice develop to different degrees of heterogeneity, fibromodulin appearing to play a crucial role in the late fusion of thicker fibrils, and lumican in the early development[246]. Also, weak tendons are found in compound fibromodulin/lumican-deficient mice that develop osteoarthritis even earlier than single knockout animals[247]. A similar situation occurs in biglycan/fibromodulin-deficient mice[183], so

the more severe phenotype in the double-knockouts reveals the additive effect of SLRPs on collagen matrix assembly.

If there is no apparent phenotype in fibrillar collagen in cartilage, why is fibromodulin expressed there? With respect to the classical view of sulfated GAGs, fibromodulin may use its sugar chains to allow for further interactions and/or accumulate water between collagen fibrils, allowing them to slide during tension. This SLRP may also play its part in bone development, as TGF-β induces its expression by chondrocytes and osteoblasts during both endochondral and intramembraneous ossification[248, 249]. Firstly, the presence of SLRPs may also affect the resorption of collagen in the developing bone, as they can delay MMP activity on collagen *in vitro*[213]. The other issue is more complex: in mouse articular cartilage, fibromodulin expression peaks together with collagen type II up to one month of age, and is later decreased. With time, the amount and length of KS chains is decreasing, with heavily reduced sulfation in old cartilage. Distribution of fibromodulin within this tissue also varies – from pericellular and more abundant in the resting zone, to less pronounced interterritorial allocation in the proliferation and hypertrophic zones as well as in the calcified matrix[250-252]. A parallel can be drawn to decorin, whose presence in the calcified collagen and with age is also decreased[253]. Of course, the findings in these studies, based on antibody-based detection methods, could depend on the degree of proteolysis of the proteoglycans. However, to some degree, the results reflect the studies on tooth development, where SLRPs seemingly "prepare" collagen fibrils for mineralization, but are later excluded from the process.

Other studies identified fibromodulin as an age-specific marker in fibroblasts. *FMOD* is upregulated in middle-age and old-age human fibroblasts, as the only extracellular matrix protein (out of seven upregulated genes in total) in a microarray study on over 6,000 human genes. It is also peculiar that acquired cellular senescence increases expression of this proteoglycan[254, 255]. Considering that *FMOD* gene is located in the chromosome 1q32, in vicinity of *PRELP* and *OPTC* (other SLRP genes that are not upregulated), establishes a unique age-related expression pattern of fibromodulin among the SLRPs.

Relating to this fact, fibromodulin is upregulated by UV irradiation, as *FMOD* gene promoter contains DDB-1 binding sequence involved in UV-regulated gene expression. Notably, this effect is prominent in exponentially growing fibroblast primary culture, but absent in senescent fibroblasts. The lack of UV response in old individuals could be damaging for the proper regulation of extracellular matrix scaffold, and contribute to the photoaging of skin (disorganized collagen fibrils and increased collagen degradation by collagenases)[256]. This could explain the inherent fibromodulin upregulation in old fibroblasts, although the issue needs to be investigated further.

Lumican

Lumican is expressed in many connective tissues, but not in the growth plate where other SLRPs are found[251, 257]. Lumican has only two tyrosine sulfation sites[236] and four potential N-glycosylation sites, substituted with polylactosamine or KS, and is the closest homologue to fibromodulin[258, 259]. As yet another SLRP, lumican can protect collagen from MMP cleavage *in vitro*, although not as efficiently as decorin[213].

During tendon development, lumican deficiency seemingly accelerates fibril fusion by a presumed lost control of the process, as judged by electron microscopy[246]. Being the closest homologue, it can compete with fibromodulin for binding to collagen[237], and is overexpressed in tendons of fibromodulin-deficient mice[55], that is, in the tissue where it is not normally present. However, despite a similar collagen-binding site, lumican fails to rescue the fibromodulin-deficient phenotype, which might point to a precise local regulation of collagen assembly by these two SLRPs.

Lumican-deficient mice have fragile skin (low tensile strength) and opaque cornea[260]. Particularly in the posterior corneal stroma, where lumican is abundant, the collagens fuse in

an unregulated manner, producing irregularly shaped fibrils. In contrary, the anterior stroma is unaffected, which correlates with its lower expression of lumican. On a larger scale, as a consequence of lumican deficiency, the disrupted lamellar organization of the stroma results in its threefold increase in backscattering of light[261-263]. The disordered collagen network is apparent by examination of x-ray diffraction patterns, which, to compare with other corneal SLRP, is contrasting with a normal pattern observed in osteoglycin-deficient mice[264]. On the other hand, it resembles the phenotype of *Tgfb2*-/- mice that have less keratocytes and less lumican and matrix in this tissue[265]. This implicates lumican regulation by TGF-β2, its production also being stimulated by fibroblast growth factor-2 in primary keratocyte cell culture[266].

What is so unique with corneal lumican? It is the major KS-substituted proteoglycan of the cornea, and has so far been found as such only in this tissue. Removing a KS-specific sulfotransferase (N-acetylglucosamine-6-O-sulfotransferase) results in abnormal collagen fibril formation, implying that lumican (probably along with osteoglycin and keratocan) regulates fibril fusion and/or spacing between the fibrils[267]. It is also known that corneal transparency reaches 40% at day 9 postcoitus, but doesn't continue to increase until after day 15. This is also the time when polylactosamine form of lumican is being switched to KS proteoglycan, further correlating the GAG with a cornea-specific function[268]. In addition, the posterior stromal swelling that is induced at eyelid opening (8-14 days postnatal) is absent in lumican-deficient mice, and the eye develops a diffuse light scattering property[269].

In eye sclera, lumican carries less sulfated GAGs and is also in complex with aggrecan, the amount of this complex being increased with age[270]. Lumican deficiency in this tissue creates heterogenous, slightly thicker collagen fibrils[271], and what more, mice deficient in both lumican and fibromodulin have additional phenotypes, including increased axial length of the eye (10%), frequent detachment of retina, as well as thinner sclera made of thin and badly shaped big collagen fibrils – in result, a myopia-like phenotype emerges in the eye[272].

Lumican is also expressed in developing bone and cartilage, although the lumican-deficient mice have no reported phenotype in these tissues[273]. However, in combination with fibromodulin deficiency, the mice have weaker tendons and develop age-dependent osteoarthritis, possibly due to weak joint ligaments. This phenotype is more severe than in absence of only one proteoglycan[247]. Another implication in osteoarthritis could be that lumican is not cleaved by MMP-13, as opposed to fibromodulin and biglycan, and could therefore protect collagen fibrils from MMP degradation *in vivo*[213, 215].

Keratocan

Keratocan is primarily found in cornea and sclera, and in much lesser amounts in skin, tendon, ligaments, and cartilage[274, 275]. In cornea, chick keratocan is substituted with three N-linked KS chains (like lumican and osteoglycin) on sites close to aromatic amino acids[258], and the GAGs are longer and more sulfated with age[276]. Their amount also increases during the first two weeks of postnatal age, correlating with the eye opening period[277].

Some cornea plana (CNA2) patients, with the severe and recessively inherited form of this disorder, have *KERA* mutations that either substitute a conserved asparagine for serine in the consensus LRR sequence, or create a stop codon[278]. Akin to this finding, keratocan-deficient mice have a 40% thinner corneal stroma and a flatter cornea, with a 30-40% narrower cornea-iris angle, essentially distorting proper light refraction in the eye. Electron microscopic imaging revealed 10% larger collagen fibrils that appear less organized in the stroma[277]. This is a phenotype comparable to that of another KS-substituted SLRP – lumican, whose deficiency leads to thinner corneal stroma, although with a resulting corneal opacity[263]. Interestingly, keratocan deficiency does not lead to upregulation of other KS-substituted SLRPs, or decorin, which further establishes its unique role in this tissue. However, lumican deficiency downregulates keratocan, and conversely, its overexpression

also triggers keratocan synthesis[279] – this may account for the somehow overlapping phenotype of both knockout mice, and also hints at SLRPs' need to be at a certain stoichiometric ratio to fulfill their tissue function.

Finally, regulation of keratocan expression is similar to that of other SLRPs in one point – the dependence on TGF-β. Keratocytes downregulate its expression when treated with this cytokine, simultaneously producing collagens, fibronectin, and transdifferentiating to myofibroblasts that express more biglycan[280]. This effect is similar to other cell types involved in scar and fibrosis formation. Furthermore, keratocan has also been implicated as a mediator of neutrophil recruitment in corneal keratitis, binding CXCL1/KC and thus creating a chemotactic gradient[281]. However, the keratocan effect on stroma might provide as good explanation for the phenomenon, since collagen network structure affects migration of cells[282, 283].

Osteoadherin

Osteoadherin is unique among the SLRPs for its extense of acidic domains. It can carry six tyrosine sulfations in the N-terminal and two in the C-terminal part of the protein[236]. It also contains aspartate and glutamate-rich sequence near the C-terminus, and can therefore be purified on hydroxyapatite. In addition, it has six potential N-linked glycosylations, and has been purified from tissue with KS chains. Adding to the specificity, osteoadherin is expressed mainly by osteoblasts of trabeculae of primary spongiosa, distant from the growth plate, and is not expressed in cartilage. Osteoblasts bind osteoadherin through $\alpha_v\beta_3$ integrin, which can be inhibited by GRGDSL peptide, and therefore could be mediated by the RID sequence in LRR 11. Besides cell attachment, integrin binding could function as a cell signaling trigger for expression of collagen or its mineralization[284, 285]. The reasoning is that in bone, osteoadherin is concentrated in mineralized bone matrix, and since it is acidic (calcium attractant?), it may be involved in biomineralization. The expression pattern is similar to that of bone sialoprotein, and the two proteins could co-operate in the process[286]. Especially if osteadherin binds collagen, it could regulate the initial crystal nucleation.

In teeth, osteadherin is co-localized with $\alpha_v\beta_3$ integrin in odontoblasts, which is similar to osteoblasts[287]. The proteoglycan amount is increasing with more mature odontoblast cell layers in developing tooth and may also be expressed by ameloblasts. Osteoadherin is also found in mineralized matrix near collagen fibers, in mineralized dentin and in cementum. Conversely to other SLRPs though, no expression is seen in the pulp or the periodontal ligaments[288-290]. In conclusion, the presence of osteadherin in mineralized tooth or in the mineralizing front of the bone tissues, coupled with its affinity for hydroxyapatite, infer that it may be specialized for biomineralization of collagen.

Promoter analysis and *in vitro* studies on osteoblasts showed that osteoadherin is downregulated by TGF-β and upregulated by BMP-2. TGF-β can trigger expansion of committed osteoblasts but inhibit the late stage of differentiation during mineralization. In contrast, BMP-2 is inducing the osteoblastic phenotype and differentiation mainly through Runx2. It appears then that BMP-2 acts on the mineralizing cells to express osteoadherin at a proper location and differentiation stage[291].

PRELP (Proline and Arginine-rich End Leucine-rich repeat Protein)

PRELP, or prolargin, originally found in articular cartilage, has an unusual (for SLRPs) proline- and arginine-rich N-terminal, but also a short acidic segment and at least three N-linked glycosylations[292]. Its expression peaks in adult cartilage, following a scarcer distribution in neonate or fetus cartilage; also, it is not abundant in other tissues, except lung, kidney and skin[293]. PRELP can bind heparin and heparan sulfate through the basic N-terminal part, where the GAG sulfates are critical, as the interaction may be mediated by the

consensus heparin-binding sequence XBBXBX (B is a basic amino acid) or simply by the clustered basic amino acids[294]. In addition, PRELP can bind perlecan's HS-substituted domain I and V, and not the GAG-free domains, and is co-localized with this proteoglycan in basement membranes, bridging collagen-perlecan complexes. PRELP may therefore form links between the cells and the underlying basement membrane[295], and can be further involved in cell-regulated assembly of collagen fibrils.

Another tissue-specific function of PRELP may relate to its increasing amount in the adult eye sclera, regulating the mechanical properties of this connective tissue[296]. This is of interest, since mutations in PRELP, and other SRLPs, were detected in high myopia patients[297]. However, the significance of this finding is not yet researched.

Class III SLRPs

Opticin

The vitreous gel of the eye contains a collagenous scaffold that liquefies with age – essentially, the stability of collagen gel is lost and it gradually aggregates and collapses. This process is common in conditions like retinal detachment or macular hole formation. Opticin (or oculoglycan) through binding to collagen, could regulate or delay these changes, if it has a similar function as other, more characterized SLRPs. To support this view, hyaluronan (a major part of the vitreous) does not appear to be essential for upholding the collagen structure but rather is a "pad" in-between the collagen fibres[298].

Opticin has been purified from bovine eye having sialylated O-linked oligosaccharides in the N-terminal region, and its mRNA was also found in ligament and skin[299]. In eye, opticin is not only present in vitreous but also in iris, ciliary body, and retina[300], but not in sclera[301] where many other SLRPs are found. In addition, the expression of the different collagen chains[302, 303] and their organization varies between sclera and, in this case, vitreous and iris – that is, the parts responsible for transparency and refraction of light[304]. Opticin also co-precipitates with retinal growth hormone suggesting that it additively aids in ocular function as a GH-sequestering protein[305].

Opticin contains a heparin-binding sequence motif. It also binds HS-chains of collagen XVIII present in internal limiting membrane of the retina, to which vitreous collagen fibrils attach, and could therefore provide a bridge between the two collagens in the neighboring tissues[306]. Mutations in opticin gene have been detected in a number of patients with open angle glaucoma[307] and in high myopia[297], although due to the complexity of these disorders opticin involvement may only apply to individual cases. It is still probable though that an altered collagen network structure in the vitreous would contribute to the pathological progress in these patients.

Epiphycan

In epiphyseal cartilage, the most abundant proteoglycans are aggrecan, fibromodulin and decorin. Epiphycan (or PG-Lb/DSPG3), characterized for its three O-linked GAG chains, is another SLRP found in this tissue, possibly potentiating the effect of its homologous partners during collagen network assembly and biomineralization. Like the other class III SLRPs, it has only seven LRRs, the implication of which has not been investigated. It is most homologous with osteoglycin, another class III SLRP, and these two SLRPs could reciprocally complement their functions, similarly to fibromodulin and lumican, or decorin and biglycan[308, 309]. In zebrafish, epiphycan is expressed only in developing notochord (at time correlated with chondrogenesis), as well as in cranial and fin cartilage[310], which substantiates its role in chondrogenesis.

Osteoglycin

Osteoglycin (or mimecan) was initially identified as osteoinductive factor, isolated as a glycoprotein from demineralized bone, and later found substituted with KS chains in cornea (there are two potential N-linkage sites in human osteoglycin). The latter is also the tissue where osteoglycin is most abundant. In cornea, three differently sized osteoglycin fragments were found, along with differently spliced mRNAs[311] that, nevertheless, coded for the same full-length protein[312]. Possibly, a proteolytic cleavage of osteoglycin may correlate with differential control of the corneal transparency. For example, propeptide cleavage by BMP-1, converting pro-osteoglycin to its mature form, alters collagen fibrillogenesis *in vitro*, giving an early lateral assembly phase[313]. It is also known that corneal transparency can be impaired when KS chains lack sulfation or are shortened, which could affect collagen network spacing. Interestingly, during progress of chick cornea development the GAG chains of osteoglycin (and also of lumican and keratocan) are produced with increasing length and sulfation[276].

Judging from electron microscopic imaging, in one study mice deficient in osteoglycin had dysregulated collagen fibril fusion in skin and in cornea, although strangely, corneal clarity was unchanged and skin tensile strength decreased only moderately[314]. In a contrasting study, x-ray diffraction on corneas showed no disarray in collagen network organization, and not even any differences in collagen fibril diameter[315], which may be explained by measurements of different or larger areas of the cornea in the respective reports.

Lastly, to mention the few studies on osteoglycin regulation, in pathological conditions this SLRP can be suppressed by IFN-γ[316], and it is also downregulated by Runx2 isoforms[317] – the latter is important for bone development.

Class IV SLRPs

Chondroadherin

Chondroadherin is a cartilage-specific SLRP that has one extra cysteine bridge near the C-terminus (as compared with other SLRPs), as well as a free cysteine in the N-terminal part[318]. The CHAD gene differs from the other SLRPs in exon-intron structure and it resides on human chromosome 17, where no other SLRP has been found to date[319]. In growth plate cartilage, chondroadherin accumulates around the proliferative chondrocytes with increasing age[320]. It also binds to $\alpha_2\beta_1$ integrin, while chondrocytes can attach to chondroadherin-coated plates[321]. Chondroadherin can interact with collagen type II as well, as seen on cartilage-isolated complexes, and it binds near one end and near the center of a collagen monomer. In addition, it interacts with collagen VI (also recognized by $\alpha_2\beta_1$), although it does not appear that it can organize this collagen into fibrils[196]. Since chondroadherin is expressed pericellularly, unlike some other SLRPs[160], and shares a collagen receptor on the cell surface, it could modify cell signaling of chondrocytes[322].

Besides cartilage, chondroadherin has been detected in cornea, lens, and retina of the eye[323]. Some authors also used the more sensitive RT-PCR to find chondroadherin mRNA in other tissues, however correlation with *in situ* hybridization and immunohistochemistry has not been satisfactory. In conclusion, chondroadherin appears to be cartilage- and eye-specific, functioning as a cell attachment protein that could use $\alpha_2\beta_1$ integrin to mediate the chondrocyte formation of newly secreted, pericellular collagen matrix.

TENDON DEVELOPMENT

The mechanical strength of connective tissues and their plasticity depends on several factors, like collagen fibril diameter, length, orientation, and cross-linking, as well as interaction with other matrix components. For example, tail tendon from osteogenesis imperfecta mouse model (defective α2(I)-chain) has twice as weak resistance against tensile stress, when compared with wild type mice. This correlates with about 60% decreased diameter of the collagen fibres[324], and changes in collagen cross-linking pattern[325]. These mechanisms are the central players during development of the mechanical functions of tendons, and it is becoming clear that SLRPs are involved in these processes.

In 14-day chicken embryo tendon, collagen fibrils are present as approximately 20 μm long segments, growing to over 60 μm in an 18-day embryo. At both these points of development, the fibrils have short taper (1.1 μm) and long taper (4.4 μm) ends, are oriented roughly parallel with the tendon axis, with some fibrils being branched or bifurcated. For this reason, they may run at slight angles against the tendon longitudinal axis, possibly adding to the mechanical flexibility[208]. At embryonal day 18, the fibrils become more irregular, which is when many fibril segments accrue into bigger fibrils[22, 326]. Several morphological species of fibrils can be discerned at this stage. About 20% of them are short (<3 μm) with the bulk of the mass in the middle of the fibril, gradually tapering towards both ends. These fibrils are precursors for the larger fibrils, which comprise 80% of the tendon. The large fibrils are either intermediates of precursor fibrils, having thinned mass at their fusion points, or are uniform and large in diameter. Most of them, even if fused, retain the N-to-C molecular polarity, although some fibrils are bipolar as they fuse tail-to-tail with their C-terminal ends. These bipolar fibrils are also unable to fuse further, which implies that a reversed polarity mechanism controls the fibril size[59]. What determines the final polarity is not known.

Collagen fibres in mice flexor digitorum longus tendons are formed from three fibril types, all temporally regulated. Immediately after parturition, there are homogenous thin fibrils that start to accrue to heterogenous intermediates, and at day 10 this continuously growing population associates into longer and wider fibrils. At day 90, the mature tendon contains mostly the large fibrils, with some thin intermediates and a smaller fraction of growing fibrils. During this developmental process, SLRPs bind to collagens and, judged from phenotypes of knockout mice, regulate their assembly. The expression of the different SLRPs can be correlated in time with the fibril growth.

Biglycan peaks at day 1 but decreases quickly, and is at low amount in the mature tendons. Conversely, decorin peaks at day 4-10 and is then reduced to a lower steady-state level[187]. This switch of expression of homologous proteoglycans may reflect a coordinated regulation of tendon collagen fibrillogenesis; there, the initial assembly of fibril intermediates is critical for the final function. A similar situation exists for lumican and fibromodulin: lumican peaks early, and fibromodulin peaks late in fibril development, and they also compete for binding to collagen[246]. The precise role of the SLRPs is not determined, although the knockout mice all have irregularly formed collagen fibrils with ill-defined contours. Reasonably, abnormal lateral fusion of fibril intermediates in absence of a SLRP occurs, perhaps giving rise to a faster but unregulated progress of fibril formation, and resulting in abnormally sized and badly shaped mature fibrils. But there is yet another side of the story.

Tail tendon is essentially made from tendon stem cells or progenitor cells that reside in a special niche of collagen fibers, as well as biglycan and fibromodulin. These SLRPs can, at least partially, modulate BMP activity in tendon, which commits the stem cells to osteogenic, chondrogenic, or adipogenic differentiation paths. The absence of these proteoglycans increases proliferation of cells, increases BMP2 activity, as well as loss of tendon-specific markers (Scleraxis, collagen type I), which may be due to compensation for failed differentiation and matrix[180].

TOOTH DEVELOPMENT

Involvement of SLRPs with collagen matrices is apparent during development and biomineralization of teeth, although the clear mechanisms are still under veil. The advantages of studies on tooth mineralization, instead of bone, are several: one, it is a generally simpler system to study; two, there is an easy-to-follow polarized collagen secretion from odontoblasts and stepwise fibrillogenesis (from the proximal predentin to its distal part); three, there occurs rapid mineralization and no consequent remodeling. During transition from predentin to dentin, prior to mineralization of collagen, extracellular proteins are modified, and phosphoproteins are secreted to form calcium-capturing complexes. The involvement of several SLRPs in these processes has been studied to a wider extent than in bone development.

The mineralized dentin matrix is composed of 90% collagen, and 10% non-collagenous proteins. To the latter group belong phosphorylated, calcium-binding proteins like dentin sialoprotein, dentin matrix protein-1, osteocalcin, osteopontin, bone sialoprotein, and osteonectin. To date, SLRPs have been mainly identified in the tissues that precede the onset of dentin mineralization (like predentin), although they may also be present in low amounts in dentin. The sulfated GAGs are also found in non-mineralized tissues, and are scarce in dentin, correlating with SLRP expression patterns. Interestingly, CS/DS amount decreases from proximal to distal part of predentin, and proportion of CS versus DS increases towards the dentin front that itself only contains longer CS chains[327, 328]. Furthermore, inhibition of MMPs impairs mineralization and essentially formation of the hard tooth tissue, which coincides with accumulation of proteoglycans in the mineralization front[329, 330]. It must be considered that dentin has more tightly packed and structurally altered collagen fibrils[331], and removal of proteoglycans or changing their GAG content may bring about this process.

Why are such proteoglycan gradients indispensable? Since CS is a better calcium attractant than DS, DS proteoglycans in predentin may preferentially act in organization of a proper collagenous network. Later, in the distal part, the bigger CS chains would attract calcium that is necessary for mineralization. Hints for this theory come from a study where proteoglycans extracted from non-mineralized and mineralized matrices reveal changed distribution (through degradation?) at mineralization front[332]. One should consider that polyanionic proteins can inhibit or promote mineralization of collagen, as shown in studies on bone sialoprotein and osteopontin[333, 334]. Another example, regarding CS-controlled mineralization, is upregulation of hydroxyapatite growth by CS-substituted aggrecan – it accelerates the more CS chains aggrecan carries[335]. Together, these data suggest that proteoglycan modification in a gradient-wise manner, from proximal predentin to dentin, as a mode to regulate mineralization. The modifiers of proteoglycans could be MMPs (MMP-2, -3, -9, -20), especially stromelysin (MMP-3) – a proteoglycanase that is expressed more densely at the transition zone between the proximal and the distal predentin, as well as at dentin-enamel junction where another mineralization front is located[328].

To consider the role of specific SLRPs, what does fibromodulin do during tooth development? During the bell stage, it is secreted by tooth papilla-resident odontoblasts into the predentin layer. It is also expressed by the middle layer (stratum intermedium) of the enamel organ, while relatively low amount of fibromodulin is found in the distal predentin, and virtually none in the mineralized dentin or enamel. In fibromodulin-deficient mice, the enamel layer is thinner, metadentin and dentin are hypomineralized, while predentin contains thicker collagen fibrils. No compensatory mechanisms of other SLRPs – decorin, biglycan – are evident, but some proteins involved in mineralization regulation are overexpressed (bone sialoprotein, dentin matrix protein-1) or downregulated (dentin sialoprotein, osteopontin)[336]. Fibromodulin is also highly expressed as KS-substituted proteoglycan in periodontal ligament, probably securing the formation of thick and mechanically resistant collagen fibres that attach the teeth to the alveolar bone. However, fibromodulin is absent from the vicinal mineralized cementum layer and bone[290, 337, 338], and

has been purified from non-calcified cementum, along with lumican, also substituted with KS chains[339]. These SLRPs could therefore function as regulatory mineralization factors.

Concerning lumican, its amount as well as KS amount increase from the proximal to the distal predentin[328, 340]. Lumican is also present as a KS-substituted proteoglycan in the non-mineralized cementum[339], and this GAG may therefore, similarly to CS, serve as a calcium attractant. It should be noted though that the fine motif patterns of GAGs (including site-specific sulfations or length) probably determine the outcome of mineralization[341].

Biglycan is expressed in ameloblasts (enamel-forming cells), and relatively weakly in odontoblasts, but is present as a CS-proteoglycan in predentin and dentin. In biglycan-deficient mice, amelogenin expression in both cell types is upregulated, indicating that it either inhibits amelogenin, or that the changed matrix structure signals amelogenin expression. Biglycan, as well as decorin, can interfere with hydroxyapatite formation, mainly through their GAG chains – an effect that is reduced when biglycan is bound to collagen. Also, the affinity for hydroxyapatite is weaker for biglycan and decorin isolated from the distal parts of predentin, as compared with the proximal predentin fraction, and as mentioned above, CS/DS ratio is higher in the distal part. To add to the complexity, collagen fibrils in biglycan-deficient mice are thinner in the proximal and thicker in the distal dentin. Dentin mineralization in these mice is heterogenous, while enamel is seven-fold enlarged possibly due to simultaneous upregulation of amelogenin, or because of a significant loss of mineralization control. Supporting the latter hypothesis is also the four-fold thicker metadentin layer. Lastly, a number of odontogenic phosphoproteins are upregulated, stating that biglycan is one of the key players in tooth development, albeit not a critical one since adult teeth appear structurally normal[334, 342-345].

In short, it seems that biglycan has two roles in this process, and either one can be ascribed to a different part of the proteoglycan – protein core would regulate collagen fibrillogenesis, and the substituted GAGs, depending on their identity, prevent premature mineralization in one place, and stimulate it when appropriate. Similar effects may be present in bone development, considering the specific phenotype of osteoporotic bones in biglycan-deficient mice[346].

Decorin is highly expressed in odontoblasts, and weakly in ameloblasts, and it has been extracted from predentin and dentin as a CS-substituted proteoglycan. It is most abundant in predentin, associated with collagen fibers[347]. As biglycan, decorin can interact with hydroxyapatite through the GAG chains, and the inhibition of biomineralization is reduced when it binds collagen. Osteoblastic cells *in vitro* that express proportionally more decorin have delayed mineralization, compared with low-decorin-expressing cells[348]. This effect is absent in some assays where decorin efficiency is compared to biglycan, and it may be because biglycan has a thirty-fold higher affinity for hydroxyapatite (of course, these values are largely dependent on the GAG identity). Decorin content in predentin increases from the proximal to the distal part, as the total CS/DS content decreases; within this context, CS has increased sulfation and longer chains, and its amount is higher[327]. Decorin-deficient mice have more severely hypomineralized dentin than biglycan knockouts, and a five-fold enlarged metadentin.

In the other mineralization front, enamel formation is delayed, possibly due to the porous dentin that cannot support amelogenin-rich matrix; as a result, enamelin is also downregulated. As another contrast to biglycan-deficient mice, expression of other dentinogenic phosphoproteins is decreased. What is striking, the absence of decorin is not compensated by upregulation of biglycan (conversely to tail tendons), and neither the opposite is true. Also in decorin-deficient mice the adult teeth look structurally normal[334, 342-345], implying a non-critical but nevertheless significant role of this SLRP in tooth development. Taken together, it also appears that biglycan has a more pronounced role in formation of enamel, while decorin is more involved in dentin matrix construction. Both biglycan and decorin can control formation of collagen network that is deposed in dentin, and

later mineralized, and their absence leads to deficient mineralization, visible in knockout mice. Both GAG chains and protein cores contribute to the overall effect, although by different mechanisms.

The above mentioned SLRPs are also expressed in periodontal ligament, where collagen fibres attach the tooth to the alveolar bone, and where the collagen turnover is very high (this is also the tissue that is primarily destroyed by scurvy). Collagen fibrils in the respective knockout mice are more irregular in periodontal ligament than in other parts of the teeth[290]. In short, SLRPs are expressed in all connective tissues of the tooth, except the fully mineralized dentin, cementum, and enamel. Their role involves both mineralization of collagen and its structural integrity, and the differential expression reflects the differences in tissue functions.

SLRPs AND ELASTINOGENESIS

SLRPs do not just participate in establishing a working collagen matrix, but also contribute to the elastic fiber formation. Therefore, they have a dual role in forming mechanical strength, as well as resilience of connective tissues. To briefly introduce the subject, one may first consider the elastic tissue disorders and their related mutations.

Mature elastic fibers are composed of elastin core, encircled by microfibrils made of fibrillin-1 and -2 (LTBP-related proteins). Fibrillin-1 mutations give Marfan syndrome, an autosomal dominant genetic disorder that leads to aorta and heart valve abnormalities, as well as long limbs [349]. An overlapping phenotype relates to fibrillin-2 mutations that lead to long and slender fingers and toes (arachnodactyly). Notably, a TGF-β-neutralizing antibody rescues the aortal changes in Marfan syndrome model mice (*Fbn1*$^{C1039G/+}$), with reduced elastin fragmentation and improved aortic wall architecture[350]. Also, gain-of-function mutations in TGF-β receptors give aortic aneurysms in Loeys-Dietz syndrome[351]. Elastin deletion on the other hand leads to Williams syndrome (a polygenic “cocktail-party” syndrome), fragile skin, and supravalvular stenosis.

What is the relation of these phenotypes to SLRPs? To begin with, TGF-β regulates the expression of most SLRPs. Also, SLRPs have been shown to interact with elastin matrix components, and some phenotypes of the respective knockout mice could be explained by failed elastinogenesis. Decorin, whose deficiency gives fragile skin, may be involved in formation of a proper elastic skin matrix, as it can interact with elastin[352]. Furthermore, lack of biglycan leads to aortic aneurysms, and this SLRP is also expressed in the developing elastic bovine nuchal ligament, during the later elastinogenetic phase when decorin is downregulated[353, 354]. Biglycan can bind tropoelastin and microfibril-associated glycoprotein-1, and may thus regulate deposition of tropoelastin in the pre-formed elastin microfibrils, or by attaching the elastic network to the collagen matrix[355]. For example, in a study on induced tubulointerstitial injury of the kidney, biglycan and decorin were upregulated together with fibrillin-1, while absence of biglycan and fibrillin-1 impaired the elastic strength of Bowman’s capsule and tubules, with resulting haemorrhage[356]. Investigations on simultaneous binding of SLRPs to collagen and elastic fiber components could further clarify their role in elastinogenesis, and in the structural organization of connective tissues.

SLRPs AND CANCER

Extracellular matrix of tumors is subject to a constant turnover directed by the cancer cells. Degradation of collagen potentially expands the tumor mass by subverting the proliferation restriction of cancer cells, and allows a more liberated flow of fluid that supplies nutrition and oxygen. Also, cancer cell invasion, and neovasculature formation is directed by matrix remodeling.

Anticancer drugs are often based on high molecular weight molecules that need to diffuse into the tumor-built interstitium. It is now recognized that extracellular matrix forms a barrier against such diffusion, thus impairing the drug efficacy. To some degree, the content of GAGs is responsible for transport resistance, although it is not the only obstacle. The mechanical stiffness and the fine structure of collagenous networks, being dependent on the interactions with proteoglycans, present a potential encumbrance that needs to be tackled.

Involvement of SLRPs during cancer progress has been addressed in several reports. Biglycan is downregulated in skin with basal cell carcinoma[357], but is upregulated in high-grade osteosarcoma with poor therapy response, where also TGF-β is amplified[358]. This SLRP has also an anti-proliferative effect on pancreatic cancer cell line *in vitro,* at nM level[359]. These differences are probably due to biglycan tissue-specific functions and promoter usage, as implied in one study[174]. In poorly responding osteosarcomas, osteoclastic bone resorption is increased, and biglycan could provide some inhibitory effect on resorption. Also, pancreatic cancer cells are growth-inhibited by biglycan affecting either the cells or the collagen matrix – intriguingly, fibrillar collagen (regulated by biglycan?) can inhibit cell proliferation by triggering p27[283]. Therefore, the cancer cells could trade slower proliferation for a long-term growth-beneficial matrix structure. Lastly, biglycan function in the skin is unclear, although it could affect collagen fibrils and their attachment to the basement membrane. Since the basement membrane components are also downregulated in basal cell carcinoma, this would collectively aid in tissue invasion[360].

As for decorin, its transcripts are not significantly altered in relation to the abundant presence of lumican in cases of breast adenocarcinoma[361, 362]. Lumican overexpression also correlated with metastatic, more advanced, colorectal carcinomas[363]. Lumican, along with decorin (but not fibromodulin or biglycan) also appeared to be overproduced in highly progressed breast adenocarcinoma (judged by morphology), especially at the invasive edge of the tumor; however, it could not be correlated with other prognostic markers[361, 364]. Another, more defined, study on invasive node-negative cancer correlated low lumican and decorin expression with poor prognosis (including early recurrence, high EGFR, large size, poor survival). This could mean that a lumican-deficient matrix would be an easier barrier to penetrate for the tumor cells[365]. In pancreatic tumors, patients with lumican-expressing cancer cells correlate with better prognosis, while those with lumican in the stroma – with poor prognosis[366]. It needs to be reminded though that lumican is present in the stroma of normal pancreas, but a lumican-altered *tumor* stroma could be a pro-mallignant factor to the cancer cells[367].

Concerning other cancers, in neuroendocrine tumors lumican is found in both cancer cell cytoplasm and in the stroma[368]. In melanoma, lumican is not expressed in tumoral but rather in the peritumoral stroma (this could represent skin fibroblast expression), and its amount fades with increasing Clark levels, negatively correlating the amount of lumican in the surrounding healthy matrix with vertical growth of the tumor front[369].

Studies on decorin function in carcinoma have been mainly targeted toward its effect on EGFR signaling, as injections of decorin in mouse downregulate the expression of EGFR and trigger apoptosis in tumor cells[370]; also, pulmonary metastases from breast carcinoma are inhibited with the treatment[371]. However, the effect on EGFR may not be specific, and may be mimicking the action of LRIG1 (see *Decorin*). Regarding the decorin influence on collagen matrix, EGFR signaling may also change since it is involved in integrin-dependent cell adhesion[372]. In addition, studies that report decorin binding to EGFR, or dimerization and

internalization of EGFR through decorin[373], use high molar amounts of decorin compared to controls, which might bias the results. It is possible though that decorin is a matrikine precursor, and that cleavage of a part of decorin would release a potent cell-signaling peptide – however, this is yet not proved.

Decorin was also found in stroma of pancreatic cancer[374], in human gastric carcinoma[375], and in colorectal carcinoma – along with less iduronic acid-containing sugar chains[376]. Effects of adenoviral-mediated overexpression of decorin have also been investigated, with decorin delaying tumor growth and metastases, although the study included empty expression vectors or PBS as controls[377].

Not many reports on the role of fibromodulin in tumors have been published, although the few available studies are worth mentioning. Fibromodulin is upregulated by metastatic melanoma cells, along with a few other matrix proteins – biglycan, fibronectin, and α2(I) collagen – once again pointing to the collagen matrix structure as a prerequisite for metastatic potential[378]. It is also expressed by the cells of B-cell chronic lymphocytic leukemia (B-CLL), with no abnormities of the *FMOD* gene locus seen in the patients[379]. This suggests fibromodulin aiding in cell survival in bone marrow matrix, since interaction with bone marrow-derived stromal cells is also vital for B-CLL proliferation. Other authors found that fibromodulin peptides are presented on the surface of CLL cells, bound to HLA-A2 dimers, and capable of expanding T-cells[380].

From our studies, fibromodulin appears to be a dominant matrix protein in several experimental carcinoma models *in vivo*, and its presence secures the formation of a dense collagen matrix (see *Present investigations*). Fibromodulin is also a major matrix component of leiomyoma, produced by SMCs stimulated by TGF-β, and although this disorder is not malignant, it still features a fibrotic matrix where fibromodulin may play a central role[381].

SLRPs AND ATHEROSCLEROSIS

In arteries, vascular smooth muscle cells (VSMCs) play a contractile role in tunica media, and build up a matrix of collagen and elastin that enforces the vascular wall. However, in atherosclerosis, where a fibrotic neointima is formed, VSMCs change from a contractile to a synthetic phenotype and migrate into the intima to lay down a fibrotic matrix consisting of collagens, elastin, and proteoglycans. This thickening of intima, and its fragility, influences the outcome of angina pectoris, thrombosis, and myocardial infarct.

In blood vessels, SLRPs probably improve the mechanical and elastic strength of the collagen and elastin network, automatically providing support for cellular migration and proliferation. Through these mechanisms, the presence of SLRPs would affect the outcome of atherogenesis and aneurysms, and therefore make a potential target for therapy. In addition, the GAG chain quality can regulate atherogenesis through binding or retention of LDL. For example, angiotensin II (vasoconstrictor and macrophage-recruiter) upregulates production of biglycan with longer DS chains that bind LDL more avidly than control biglycan[382]. Deteriorating effects of biglycan deficiency on collagen fibril formation have been found in media and adventitia of aorta, with aortic ruptures and sudden deaths of 50% of male mice before 3 months of age. Even though elastic laminae are morphologically unaffected, the lower tensile strength of artery is suffices to trigger the phenotype. Interestingly, collagen is only altered in male mice, suggesting involvement of an estrogen-dependent or androgen-dependent mechanism[353] – androgens can in fact stimulate biglycan synthesis, as well as GAG chain length and sulfation[383]. This knockout phenotype also correlates with more than halved biglycan content in aortic aneurysms, together with unchanged decorin amount. The authors suggest the reduction of biglycan may be due to selective degradation by macrophage-induced proteolysis[384], possibly by MMP-13 that is abundant in atherosclerosis and cleaves biglycan, but not decorin or lumican[215, 385]. This could further expose collagen for degradation by MMPs, as implicated in an *in vitro* study[214].

Biglycan can be found co-localized with collagens I and III in human fibrotic restenosis tissue (lesions formed after angioplasty) and in atherosclerotic plaques. On the other hand, decorin, while present in the plaques, is not abundant in restenosis[386]. During initial intimal thickening of atherosclerotic vessels, biglycan and decorin are expressed in the outer layer of the intima, preceding the macrophage infiltration[387]. LDL can bind to biglycan in the arterial wall[388] and be retained there, which thereby involves the SLRP in the process of LDL oxidation. This would create a chemotactic gradient for VSMC migration and macrophage invasion, with an incipient foam cell formation. To contradict this, biglycan can also bind apoE, an anti-atherogenic molecule found in HDL, and could instead contribute to protection against the disease. Perhaps the binding of either ligand could prevent interaction with the other, thus regulating the outcome of the pathological process[389, 390]. Lastly, overexpression of biglycan in arteries increases VSMCs proliferation and media thickness, which contrasts with the reducing inflammatory and fibrotic effects of decorin in a similar system[391, 392]. These two SRLPs, even though being close homologues, would have competing effects on atherogenesis; however, overexpression studies can introduce uncontrolled random factors that could have influenced the results.

Fibromodulin is found in atherosclerotic plaques formed in ApoE/LDLR-deficient mice, along with decorin and PRELP[393], coinciding with upregulated synthesis of fibrotic collagen matrix. In coronary atherosclerosis, the thickened intima contains high amounts of lumican (relative to the media), and is expressed by the infiltrating VSMCs and not macrophages[394]. Since fibromodulin and lumican control the collagen fibril formation, the outcome of the plaque stability may be affected by these two SLRPs, but it is not yet clear if and how. Experiments evaluating plaque fragility and susceptibility to thrombosis in the respective knockout mice would answer the question.

SLRPs AND INFLAMMATION

Even if the findings on SLRPs in the common inflammatory disorders have been discussed already, there are other inflammatory conditions worth addressing.

To start with fibromodulin – it is implicated in development of osteoarthritis. It can be cleaved by MMP-13 (a prominent protease in this disorder upregulated by IL-1) and the tyrosine-sulfated N-terminal part is released[395]. Fibromodulin can bind complement factor C1q in its globular head, which activates C1 and the complement pathway; furthermore, fibromodulin also binds factor H, which prevents formation of membrane attack complexes[396]. Interestingly, decorin and biglycan also interact with C1q, although at the collagenous C1q domain, and are, contrary to fibromodulin, inhibitory for the complement pathway activation[397, 398]. Decorin, once again, appears to have an opposing effect on inflammation, since it has been reported to inhibit fibrotic conditions. The reasoning by the authors was that decorin action was mediated by its TGF-β- inhibitory properties, although the required several 100-fold molar ratio for IC_{50} [227] contrasts with that of decorin inhibiting C1q activation (IC_{50}) at ~20-fold molar excess. The anti-fibrotic effect could therefore be mediated primarily by amelioration of extensive tissue destruction by complement mechanisms[397]. Actually, in decorin-deficient mice, induced tubulointerstitial kidney fibrosis accelerates apoptosis of tubular epithelial cells and infiltration of mononuclear inflammatory cells[399]. The higher levels of TGF-β in these mice could therefore be secondary to complement activation and the consequent tissue destruction.

The role of biglycan in inflammation is more troublesome to define: it inhibits activation of C1q, although at about 10-fold higher relative concentration compared with decorin. However, it also acts pro-inflammatory by activating TLR-2 and -4 (that have extracellular LRR domains) on macrophages, thus inducing expression of the inflammatory cytokines TNF-α and MIP-2. The latter help in local clearing of infections, although on a major scale the overproduction of cytokines leads to sepsis – a condition that, interestingly,

is dampened in biglycan-deficient mice[400]. Innate immunity may therefore be dependent on biglycan, but the interplay between complement system and macrophages is context-dependent.

Lumican also plays a part in innate immunity, in particular the LPS-induced TLR-4 signaling. Macrophages from *Lum-/-* mice respond weakly (less TNF-α, IL-6) to LPS stimulation, and addition of recombinant lumican rescues the response. In contrast to biglycan, lumican cannot induce TNF-α signaling without LPS, as it binds the bacterial macromolecule to its protein core[401]. In a related context, corneas of *Lum-/-* mice have low p53, and keratocytes are less apoptotic through failed activation of Fas signaling. Lumican could therefore serve as an adaptor protein for Fas ligand, and be involved in limiting the inflammation by Fas+ lymphoid cells. Fas is not just a death-triggering molecule, but can also be proinflammatory for macrophages, being involved in LPS activation of TLR4, in which lumican has a role. Indeed, LPS stimulation of stromal wounds does not induce Fas in *Lum-/-* mice, and the wounded corneas have less macrophages due to low levels of the inflammatory IL-1, IL-6, and TNF-α[402, 403].

To conclude, studies on SLRPs have shown their potential for targeting a number of inflammatory conditions. Considering infections, perhaps it is not surprising that LRR proteins of extracellular matrix, where the infection agents are primarily invading, can aid in clearing the local infections by innate immunity to avoid major tissue destruction. In skin – the major physical barrier against infection – decorin, lumican and biglycan are expressed, and they could therefore be active in the initial immune response.

The LRR domain in general is not only used in innate immunity, and not just in mammals. Intriguingly, jawless fish have adaptive immune system relying on recombination of so-called VLRs (variable lymphocyte receptors), resembling the system of immunoglobulins in jawed vertebrates. These receptors contain variable number and sequences of LRR domains, which safeguard a vast range of ligand interactions[404].

Continuing with other inflammatory conditions, in bleomycin-induced lung fibrosis, fibromodulin and biglycan protein expression is elevated. They peak at 14 days after bleomycin application, and are present in the fibrotic tissue beneath the epithelial layer[405]; in this system, fibromodulin regulation seems to be controlled posttranscriptionally as mRNA expression is not altered[201]. Functionally, biglycan is proposed to affect lung tissue elastance and resistance, since its upregulation correlates with changes in these parameters[406]. Also, because biglycan influences the same factors of blood vessel walls (see *Elastinogenesis*) it may be involved in the process of tissue repair during early fibrotic progression, and actually not be pro-fibrotic. Similarly, lumican and decorin that are expressed in the smooth muscle cell layer of normal and asthmatic lungs, decrease in the grave cases of asthma[200, 407, 408], and can not contribute to the thicker subepithelial collagen layer. If the SLRPs have different pro- and anti-fibrotic properties, and if these are not general but tissue- and context-dependent, remains to be examined.

Lastly, to the process of wound healing, where complex anabolic and catabolic processes remodel the matrix and remove damaged tissue. Serine proteases like MMPs, cysteine proteases cathepsins K and L, as well as alkaline phosphatase (TRAP), help to degrade and later lay down a new provisional matrix on which the cells, if possible, can rebuild the functional tissue. Here, collagen expression correlates with expression of SLRPs, and is not always driven by inflammatory response. For example, an intrinsic fibroblast-driven matrix protein expression is found during wound healing where no large tissue destruction occurred. There, collagen is upregulated along with decorin, lumican, and fibromodulin, but interestingly, biglycan mRNA is not altered[409]. Particular granulation tissue fibroblasts can also raise their expression of collagens and SLRPs, without external TGF-β stimulation[410], although this phenotype may also be pre-established by inflammatory cells. In a study on transition from scarless to adult tissue repair, fibromodulin protein was decreased, contrary to decorin and biglycan messages, suggesting that fibromodulin could help in scarless healing[411]. Indeed, adenovirus overexpression of fibromodulin produces less scar and more tensile strength in the healed wounds, with lowered TGF-β expression[412]. Even though, as authors suggest, fibromodulin may bind TGF-β as reported[146], and

contribute to its inactivity, the observed effect could simply reflect a quickened matrix remodeling (through fibromodulin) and the consequent decrease in pro-fibrotic cytokine expression.

PRESENT INVESTIGATIONS

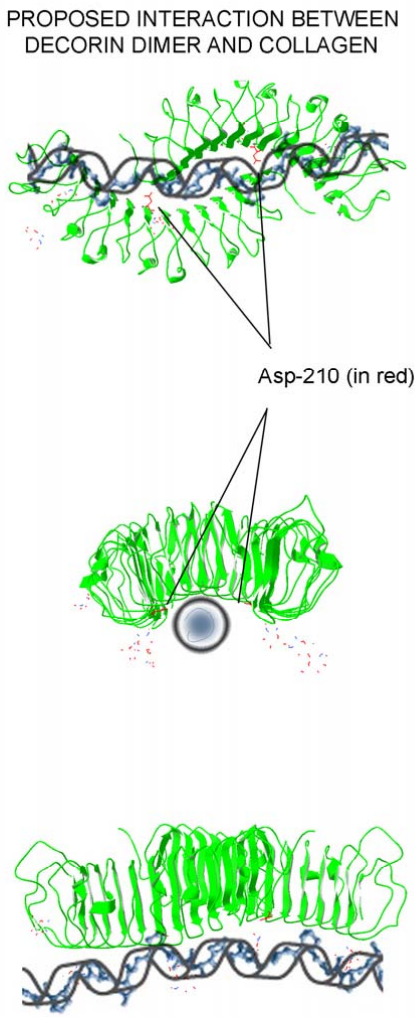
Decorin binding to collagen type I (paper I)

Decorin-collagen interaction has been the subject of several reports. The information could prove valuable for designing peptides that could selectively influence the fine-tuned collagen fibrillogenesis. Initially, two interaction sites in each half of decorin were found[413], in a study that used recombinant fragments purified with chaotropic agents, without any fusion tags. Such procedure, in my experience, renders fragments that are not water-soluble or uncorrectly folded unless they are coupled to a larger fusion tag, like GST. It appears that LRR protein solubility and correct folding requires presence of cysteine bridges in the N- and C-terminal ends of the protein core, without which the solenoid structure becomes disrupted. This problem can also be tackled by constructing full-length chimers with homologous proteins that do not have a high collagen-binding activity, as done in a study where decorin fragments were chimerized with biglycan. There, a central domain of decorin LRRs 5-6 (according to the *STT* nomenclature) interacted with collagen[414]. A later study used site-directed mutagenesis on native decorin to identify a glutamic acid residue (Glu-180, mutation E180K) in LRR 5 as essential for this interaction[415]. This mutation did not impair, or perhaps even increased, binding of decorin to collagen type VI, implying that the proteoglycan could bridge the two collagens[416]. The mechanism could be explained by the evolutionary progress in which the new collagens evolved along with the new SLRPs – for example, collagen type VI is not found in Ciona genome, and decorin is related to a founder gene SLRPa that was later duplicated to carry most of the mammalian SLRP genes.

Back to decorin Glu-180: in our present study[417] Asp-210 (mutation D210N) was identified as crucial for collagen interaction, and Glu-180 appears to rather have a stabilizing role in the β-sheet structure. Two lines of evidence would support our conclusion. One, the crystal structure of decorin[140] reveals that Glu-180 can form potential hydrogen bonds with three neighboring amino acids in the β-sheet; and two, computation of the protein stability change that E180K mutation would impose (based on changes of atom pair potentials and torsion angle potentials), indicated an unfavorable energy state. In these computations, mutation D210N is not affecting the structural stability, and Asp-210 is also more solvent-exposed as compared with Glu-180. To confirm our findings we also used a synthetic peptide (SYIRIADTNIT) that includes Asp-210 and its neighboring amino acids on LRR 6 to successfully inhibit the decorin-collagen interaction.

One unresolved issue however still remains – the probable existence of a cooperative binding site on other LRR. This is since neither the peptide nor the D210N mutation fully impaired the interaction, and also because an earlier study[414] implicated a possible existence of a weaker binding site in the C-terminal part of decorin. Possibly, the collagen-binding site on decorin, and maybe even on other SLRPs, involves two “main” amino acids from two distant LRRs, as well as other weak interactions (including the omnipresent van der Waal forces).

How does decorin bind to collagen? Considering the notion that decorin can form dimers in solution[140] and that Asp-210 side chain projects the carboxy group away from the dimerization surface could mean that collagen is bound on the ”loop” surface of the protein core, connected by bonds from two Asp-210, one from each decorin monomer (see figure on the next page). Of course, decorin may bind collagen in its monomeric form, although that would be energetically unfavorable[140]. An alternative that reconciles the two hypotheses would be that decorin function is regulated through its dimerization, and its quarternary structure would differentially regulate collagen fibrillogenesis or interactions with matrix proteins.



Fibromodulin binding to collagen type I (paper II)

Lack of fibromodulin in tendons and ligaments gives mechanical fragility and osteoarthritis, and this SLRP is also expressed in fibrotic conditions in carcinoma. Being another class of SLRP than decorin, its interaction with collagen may differentially regulate the fibril formation

process, as the interaction site itself differs from that of decorin[210]. The knowledge of how the different SLRPs bind to collagen will likely contribute to the general knowledge of their concerted effects on collagen fibrillogenesis, and possibly the intermolecular cross-linking. From a clinical perspective, targeting the SLRP-collagen interaction could manipulate the processes of fibrosis that are common in many disorders.

To study the fibromodulin-collagen interaction, we expressed a number of fibromodulin fragments in bacteria. These fragments were together covering the entire sequence of fibromodulin, and tagged with GST to ensure solubility and easy antibody detection. Measuring the interaction between the fragments and collagen was done by isothermal calorimetry and solid-phase assays. In the first method, a pre-determined amount of the fragment (ligand) is injected into the calorimeter cell that contains collagen (macromolecule) that is continuously stirred in a physiological buffer. Injections are sequential, so for each new injection the available ligands are interacting with the free binding sites on the macromolecule, during which heat is produced (in case of an exothermal reaction). The injections continue and eventually, when the macromolecule binding sites become saturated, the heat of interaction is no longer produced by the system, so the final injections emanate the same negligible amount of heat as a blank injection.

When the fibromodulin fragments were tested for collagen interaction in calorimetry, the two fragments that contained LRRs 11-12 were most active, LRR 5-7 had a lower binding enthalpy, and other parts of fibromodulin were inactive. A similar pattern was seen on a solid-phase assay, where collagen-coated plastic wells were incubated with the fragments, and the interaction was detected with an anti-GST antibody. Furthermore, the fragments that bound to collagen could also inhibit the binding of full-length fibromodulin at a reasonable ratio (800 ng of the fragment inhibited 100 ng of the full-length protein).

Comparing the two methods for studying protein interactions – calorimetry vs. solid-phase assay – the former generates dissociation constants of larger quantity, and this needs to be commented upon. There is one major difference between the two methods, in that calorimetry is run in solution at 37ºC, which is a condition where collagen is assembling into bigger fibrils. Naturally, the only available interaction sites will be present on the surface of the collagen fibrils, while collagens within the fibril will be unavailable for interaction. This is not corrected for when calculating dissociation constants that consider the concentration of macromolecule with assumed availability of *all* binding sites. In solid-phase assay we used acid-solubilized collagen, which in this condition is monomeric, and of course, presents a higher number of available binding sites in relation to concentration. Consequently, the same amount of ligand will bind to more sites in a solid-phase assay than in the calorimetric cell, even if using the same concentration of collagen.

To identify the amino acids involved in collagen binding, we made five mutations in fibromodulin cDNA, and then expressed the full-length proteins to test their collagen affinity. The mutations were placed in LRR 11, in its conserved β-sheet and the less conserved loop region that is immediately adjacent to the β-sheet structure. In effect, two amino acids were critical for collagen interaction – glutamate and lysine – located around the last aspargine in the consensus LRR sequence (LxxLxLxxNxL). Also, one other mutation (D350N), in the β-sheet, had an impairing effect on the interaction – however, this mutated protein had a different CD spectrum than the others, implying a change in the conserved LRR structure. This, we interpreted as having an indirect influence on collagen affinity. In this context, one may mention the study on decorin where a glutamate mutation in a similar position within the β-sheet also impaired collagen binding[415] (discussed above) – the reason for the loss of function could have been a structural alteration that indirectly ablated collagen affinity.

Lastly, to conclude our study on fibromodulin-collagen interaction, we designed a synthetic peptide based on the primary sequence of fibromodulin LRR 11, and covering the amino acids that were critical for interaction according to mutagenesis data. This peptide (RLDGNEIKR) inhibited the interaction of full-length fibromodulin with collagen with inhibition constant at μM range. The reasons for this relatively high amont of peptide needed for inhibition may be due to the weaker binding site of fibromodulin (LRR 5-7) contributing to the interaction. Also, the other forces (van der Waal, hydrophobic etc.) from the other LRR

domains also strengthen the interaction between full-length protein, so the short synthetic peptide needs to be more concentrated to effectively break them.

In conclusion, we found a high-affinity collagen-binding site on fibromodulin LRR 11, and a lower affinity site in the central part of the protein core – LRR 5-7. This information will be used to further study the SLRPs differential interactions with collagens, and their function in matrix turnover.

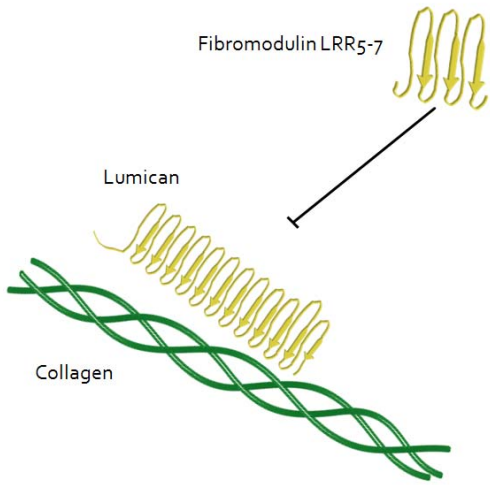
Mutations of Glu-353 and Lys-355 in LRR 11 impair collagen affinity of fibromodulin

Lumican binding to collagen type I and relation to fibromodulin (paper III)

Since fibromodulin and lumican are class II SLRP homologues, inhibit each other's binding to collagen, and are expressed in a compensatory manner in the knockout mice[55, 237], it is curious how they differ in function during collagen matrix assembly. It should be mentioned that lumican is widely expressed, while fibromodulin is found mainly in dense, strong connective tissues. One study therefore proposed, based on collagen fibril morphological changes during development of the knockout mice, that lumican acts early in fibrillogenesis, while fibromodulin is required for accretion of larger collagen fibrils[246]. To find the mechanism of SLRP-driven regulation of this process we reasoned that we first needed to find the collagen-binding sites of both SLRPs, and then use mutated proteoglycans or synthetic peptides in fibrillogenesis assay to test their different effects on collagen. Possibly, the knowledge could also be used for *in vivo* studies of the SLRP effect on fibrotic tissue formation that is common in many human disorders.

We found the major collagen-binding site of fibromodulin in its LRR 11, but also a weaker affinity site in LRR 5-7. Since the system we designed for fibromodulin study was satisfactory for the project, we used similar experimental methods to uncover the collagen-binding site of lumican.

The GST-tagged lumican fragments had lower affinity than the fibromodulin fragments, which confirmed the previously reported higher collagen affinity of fibromodulin[237]. In addition, only one lumican fragment bound to collagen – the one covering LRRs 5-7. What was immediately obvious was the corresponding span of LRRs also having collagen affinity in fibromodulin. Naturally, we analyzed the inhibitory effect of the fibromodulin LRR 5-7 on full-length lumican binding to collagen, and the results were positive – apparently, the homology between these two SLRPs conserves one collagen-binding site, but fibromodulin also carries an additional unique site in LRR 11. Perhaps this is the reason for the relatively higher collagen affinity of this SLRP versus lumican?



To further analyze the lumican active site, we chose to mutate an aspartic acid in LRR 7, and this mutation impaired collagen affinity of the protein. Why just one codon mutation? It was selected from the criteria that are becoming more obvious when it comes to SLRP's, and other LRR proteins', interaction with their protein ligands. The critical amino acids are often conserved among different species, they have so far always been present in the β-sheet–loop region of a conserved leucine-rich repeat, and they are often basic or acidic. These criteria, looking at lumican LRR 5-7, were only met by Asp-213, hence the choice for mutagenesis. One may also mention that some LRR proteins actually have ligand-binding sites on the convex face of the solenoid structure, but the ligands in these cases include nucleic acids and lipopolysaccharide. Considering that SLRPs also carry GAG chains that can interact with other ligands, it appears that SLRPs, and possibly other LRR proteins, can use their β-sheet–loop structures to attach to one protein and their other parts to connect to another ligand. This is worth keeping in mind in later studies on effects of lumican and fibromodulin on collagen fibrillogenesis.

Future perspectives:

How do the different SLRPs regulate collagen fibrillogenesis, are they dependent on fibril thickness, why do they inhibit each other's binding to collagen?

Do the differently placed collagen-binding sites, that is different parts of LRR domain, have implications for specific regulation of collagen fibril formation?

How do the GAGs play their role in the process?

Can the synthetic peptides that inhibit SLRP-collagen interactions be used *in vivo* to manipulate fibrotic tissues in disorders like cancer, atherosclerosis, etc.?

Fibromodulin function in collagen matrix assembly (paper IV)

The mechanical properties of tendons not only depend on cross-linking of collagen, but also on collagen fibril diameter, organization, length, and their interaction with other matrix components.

It is conceivable that SLRPs, through interactions with collagen, can orient the rotation of collagen monomers, and consequently determine which of the available lysine residues (eleven in each collagen monomer) become cross-linked to the other monomer. In turn, this will limit the selection of the lysines that cross-link into the growing fibril. If this has any effect on the final product of collagen fibril formation is not known, although the different packing of collagens within the fibril could well mediate an altered function. Interestingly, since lysyl oxidase requires a pre-formed thin collagen polymer to interact with[418], a preceding action of SLRPs could orient the lysines, anticipating the aldehyde formation by lysyl oxidase.

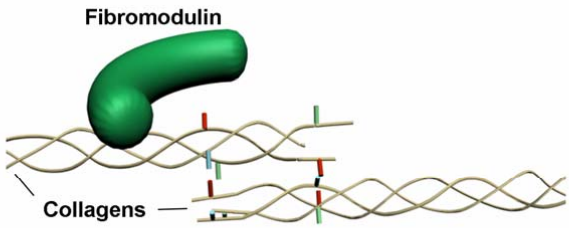
The above reasoning is drawn from a series of experiments on fibromodulin-deficient mouse tail tendons. This tissue is easy to work with, it contains mostly collagen type I, and it was earlier shown to carry a phenotype in the knockout mice[55]. Initially, we extracted tendon collagen in three steps, using gradually harsher solubilizing agents: PBS, acetic acid, and lastly, pepsin (dissolved in acetic acid). The rationale behind this method is this: PBS extracts mostly the monomeric, newly secreted collagen; then, acetic acid solubilizes the less cross-linked collagens, and to dissolve the more cross-linked collagens one needs to cleave their telopeptides with pepsin (or pronase B). Following this procedure, collagen amount in each extracted fraction can be determined and analyzed for differences between wild type and fibromodulin-deficient mice. We mainly used acetic acid-extracted collagen since it is extracted in high amounts, is soluble in acetic acid, and has been processed *in vivo*. It is also to some degree cross-linked, and not treated with pepsin. In some cases, total collagen is necessary to use, in particular for cross-link analyses.

During the sequential extraction procedure, a different distribution of collagen between the fractions was evident. In the wild type mice, most collagen was solubilized by acetic acid, and a smaller fraction needed to be treated by pepsin to release it into the solvent. However, in the fibromodulin-deficient tendon the situation was inversed, as most collagen required pepsin solubilization. This was the first indication that cross-linking of collagen could be regulated by fibromodulin.

When the acetic acid-extracted fractions were run on SDS-PAGE, changed patterns of monomeric and dimeric α-chains were apparent – the knockout mice had more free α2(I)-chains and less dimers. They also had more trimeric γ-components, which altogether pointed to a disturbed cross-linking of α2(I)-chain and an excessive cross-linking of α1(I)-chains. Fibromodulin appeared then to be a specific regulator of α2(I)-chain azimuthal orientation during collagen fibrillogenesis, since this action would define the accessibility of the cross-linking lysine residues. Alternatively, fibromodulin could shield the “inappropriate” lysines from treatment by the cross-link-initiating lysyl oxidase. This is not improbable since fibromodulin has an acidic domain that would repel the alkalophilic enzyme from interaction with lysines. Whichever is true, the effect of fibromodulin is on intermolecular cross-linking of collagen.

What kind of cross-linking is regulated by fibromodulin? Is it only the quantity of the involved α2(I)-chains that changes in the knockout mice or is there a specific cross-linking site that this SLRP can influence? Analysis of the total collagen for quantity of pyrrole cross-links (the predominant cross-link in tendons) showed that they are 30% more abundant in fibromodulin-deficient mice. This finding is somehow confusing, since we observed that they have mechanically weaker tendons (unpublished data), and intuitively, one may correlate higher cross-linking amount with higher tensile strength. As mentioned though, the quality, that is, which lysines become incorporated in the process, may contribute to the mechanical properties of tendon. It is not a new concept: examples of excessive, non-specific cross-linking of collagen that lead to deteriorated physical properties of tissue include UV-damage

of skin or diabetic glycation end products. At the time of writing, the analyses on the cross-link site specificity are underway and will hopefully reveal the function of fibromodulin in collagen fibrillogenesis.



When fibromodulin binds to collagen, it may shield some lysines (colored) from cross-linking to each other. The two collagen monomers therefore connect through a specific intermolecular cross-linking pattern.

Lack of fibromodulin has some dramatic consequences on the physical properties of collagen. Not only are the tendons weaker, but the tendon fibres melt at a lower temperature, and the extracted collagen has a contrasting melting thermogram profile. The latter can be explained by imperfect cross-linking leading to collagen polymers that dissociate into monomers to a higher extent, at a certain temperature. The triple helical stability does not appear to be affected, although the total enthalpy of denaturation is lower in the fibromodulin-deficient collagen, implying that the acetic acid-extracted fraction contains already denatured collagen. This collagen is quantified during hydroxyproline determination, but of course does not contribute to the melting enthalpy, and therefore equimolar amounts of collagens have different total enthalpy.

Furthermore, the cellular response to the failed collagen fibrillogenesis consists of an overall expansion of collagen turnover – the upregulated synthesis of collagen *ex vivo* is reflected in higher incorporation of radiolabelled glycine into the protein, and also by higher messages of collagen-modifying enzymes. Also present are the larger amounts of MMP-2, a gelatinase that degrades denatured collagen, that is, the fraction that would not cross-link correctly and impaired the fibrillogenesis. We do not yet know how this regulation of cellular collagen production is signalled, but several collagen receptors may be involved. How the fibril structure is recognized by cells, if this is indeed the case, could be the subject of further investigation.

In conclusion, this study shows the function of fibromodulin during collagen fibrillogenesis. This SLRP acts as a specific regulator of α2(I)-chain cross-linking in the growing collagen polymers.

Future perspectives:

If fibromodulin is a specific regulator of α2(I)-chain cross-linking, how does lumican, and for that matter – the other collagen-binding SLRPs – function in the cross-linking process? What implications does it have for the overall collagen matrix physiological properties?

Can these be differentially manipulated, e.g. in pathological conditions, by targeting the specific SLRPs by synthetic peptides or in other ways?

Fibromodulin in carcinoma (paper V)

Many tumors can be described as sites of a prolonged chronic inflammation, where the inflammatory cells and cancer cells stimulate stromal fibroblasts to produce collagenous matrix. This matrix plays several roles in the development of tumors. It can serve as a scaffold for cancer cell migration and invasion, for angioneogenesis required for nutritional replenishment of the cells, and lastly – from the pharmaceutical point of view – it can present a challenge to drug delivery since its constitution will influence the diffusion of drugs into the tumor interstitium[419]. It is this last factor that was the subject of our study, in particular, we wanted to examine the role of fibromodulin in neoformation of tumor stroma and how it would affect the interstitial fluid pressure (IFP) – the influential factor in drug delivery.

IFP in tumors can be lowered by admistration of a number of systemic drugs, like TNF-α, or antagonists to VEGF, PDGF, or TGF-β. The mechanism through which they work relates to stromal changes, although of differing natures: for example, countering TGF-β decreases matrix deposition, while PDGF inhibitors release the contracting force of fibroblasts on the matrix[420]. Consequently, the fluid flow in the stroma becomes more open as IFP recedes to a lower level, allowing for efficient diffusion of drugs. What is interesting when using TGF-β inhibitor (Fc:TβRII) is the downregulation of fibromodulin message, and none of the other matrix genes. Of course, the amounts of matrix proteins may still be affected by the treatment, but the finding in itself justified the examination of tumors grafted in the fibromodulin-deficient mice, and comparison with mice treated with Fc:TβRII. We also observed a conspicuous presence of fibromodulin in the stroma of several carcinoma models, which further encouraged the study. Of special interest were the effects on tumor stroma and in turn, the interstitial fluid pressure. In a long-term perspective, drug efficacy in the respective mice would also be tested.

In this study, we used tumors grown in nude (immunodeficient) mice that were either fibromodulin-deficient or wild type. The procedure included a subcutaneous injection of KAT-4 (thyroidal carcinoma) or PROb (colorectal carcinoma) cancer cells and after three weeks the tumors were excised from the animal and examined for IFP and stromal content. IFP was measured by the so-called wick-in needle method, where a perforated, hollow needle is inserted into the tumor stroma, and the pressure is transferred through a connected cord to a sensitive membrane that amplifies the signal that then can be measured. The stroma of the tumors was analyzed visually in electron microscope, and with hydroxyproline quantification assay (for collagen content).

In the microscope, we observed thinning of collagen fibrils in tumors grown in both fibromodulin-deficient and Fc:TβRII-treated mice, and interestingly, the trends were similar in both mice, with respect to fibril diameter and shape. Likewise, the three-dimensional imaging of tumor stroma (made on whole-mount samples after lysis of cells) in a scanning electron microscope, showed a more disperse collagenous matrix in these same mice. Furthermore, the total amount of collagen in tumors, determined by hydroxyproline levels, was lowered in the knockout mice, which corroborated the microscopy data. Somewhat surprising was the same amount of blood vessels in fibromodulin-deficient and wild type mice, implying that fibromodulin did not influence angiogenesis in this system.

What does it mean for the physiological properties of the tumor stroma? The changed amount and organization of matrix would reasonably affect the flow of fluid within the stroma, and indeed, after measuring the extracellular fluid volume in the tumors, it was clear that lack of fibromodulin allows a higher perfusion of liquid in a scarcer fibrotic tissue. This complies with the classical law of physics – Boyle's law – that states the proportional relationship between volume and pressure. However, the IFP is a more complex phenomenon, also related to osmotic pressure and hydraulic conductivity of the tissue, in addition to the mentioned factors. It must be stressed though how important collagenous

matrix is in controlling the IFP since all the listed parameters can be manipulated, more or less directly, by the matrix constitution.

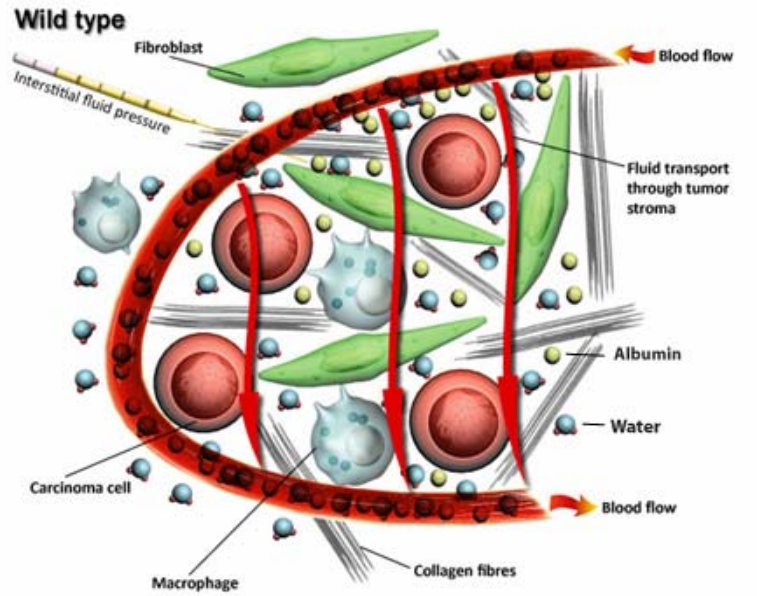
In this context we discovered a critical role of the SLRP that is normally sequestered to mechanically strong connective tissues. Fibromodulin could therefore be targeted in tumors in combination with the already established anti-cancer drugs. This could increase their delivery into the tumors with high IFP that forms the barrier against infusion of the drug from the blood into the interstitium. It needs to be mentioned that patients with high IFP tumors often have poor prognosis, which could make our, and similar studies, worth further investigation. Since we already characterized fibromodulin-collagen interaction site, we could use this information for opening a new venue for tumor treatment, possibly without severe side-effects of systemic drugs.

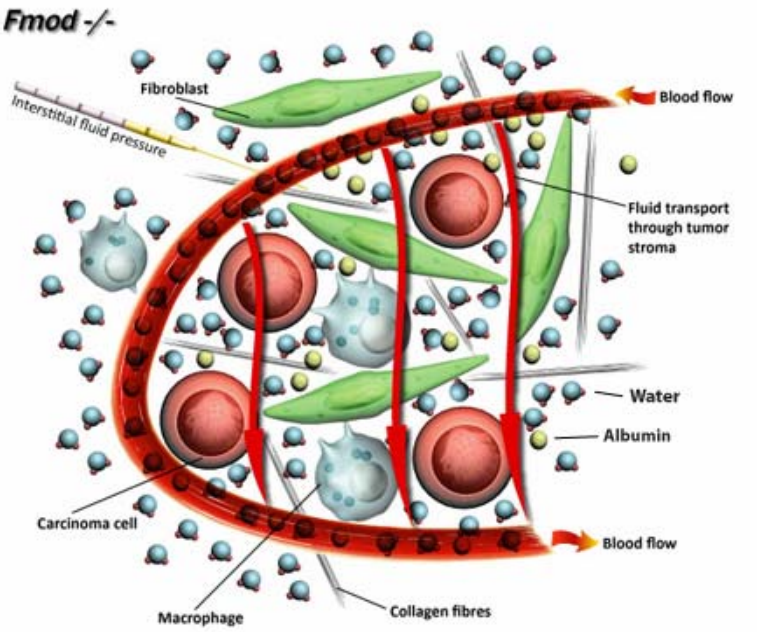
Future perspectives:

Collagen cross-linking patterns in tumors – do they influence tumor progression? Are SLRPs intimately connected to this process?

How do other SLRPs affect tumor stroma and IFP? Can a combined targeting of their collagen interaction dramatically improve drug delivery?

Next page: tumors grown in wild type mice have a dense collagen network with impaired fluid flow, due to a compacted interstitium. In Fmod-/- mice, the tumors have a scarcer collagen network, which creates more space for fluid flow, and the IFP consequently decreases.


Wild type
Fibroblast
Interstitial fluid pressure
Blood flow
Fluid transport through tumor stroma
Albumin
Water
Carcinoma cell
Blood flow
Macrophage
Collagen fibres


Fmod -/-
Fibroblast
Interstitial fluid pressure
Blood flow
Fluid transport through tumor stroma
Water
Albumin
Carcinoma cell
Blood flow
Macrophage
Collagen fibres

Fibromodulin in atherosclerosis (paper VI)

During atherosclerotic process, vascular smooth muscle cells proliferate into the newly formed neointima and lay down collagen matrix that makes up the fibrotic cap of the plaque. This matrix, depending on its composition, can be fragile or can actually stabilize the plaque from mechanical damage. The process has therefore a major influence on the outcome of atherosclerosis, including the direct complications like thrombosis. The latter takes place in instable plaques that are easily ruptured by the turbid flow of blood, and thereby – through exposure of collagen – initiate a coagulation reaction. Therefore, the studies on plaque stability in relation to the matrix are necessary to develop better strategies for prevention of thrombosis and treatment of atherosclerotic disorder.

One of the matrix proteins deposited in the fibrotic neointima is fibromodulin – normaly expressed in the media of arteries where it probably contributes to the formation of the reticular network. Since this proteoglycan has a major influence on collagen fibrillogenesis we wished to examine its role during the plaque formation process, and in particular to get an overview of the injury response in fibromodulin-deficient mice. The experiments focused on microscopic analysis of plaques, examining collagen matrix, and cellular proliferation and apoptosis. Furthermore, we made *in vitro* studies on cell proliferation, in addition to immunohistochemistry on lumican – in case any compensatory mechanism of this fibromodulin homologue would be evident.

The plaques were induced by collar injury on carotid arteries of wild type and fibromodulin-deficient mice. Upon analysis of protein expression by immunohistochemistry, both fibromodulin and lumican were present in neointimae of the plaques, and their message peaked three days after the collar injury. The number of cells per area unit in the plaques was higher, although the total plaque size tended to be lower, in fibromodulin-deficient mice. This could be explained by a more disperse matrix, and through electron microscopic examination we concluded that the collagen fibrils in neointima had a broader range of diameters, compared with wild type mice. Apparently, the dysregulated fibril formation in fibromodulin-knockout mice contributed to the reduction of neointimal area.

We reasoned that the altered matrix formation could directly influence VSMC proliferation and apoptosis. Notably, the cells in fibromodulin-deficient neointimae stained more avidly for TUNEL marker, and their message of Flip – an anti-apoptotic protein – was decreased. How does it comply with the higher proliferation? It could simply represent a compensatory response of the cells that cannot find sufficient survival support due to matrix alteration. However, these events may also be linked in time of injury, which would require further analyses.

We also observed that rat smooth muscle cells seeded *in vitro* migrated faster on collagen isolated from fibromodulin-deficient mice than on wild type collagen. Additionally, the migration of smooth muscle cells isolated from the knockout mice is faster on a plastic surface, compared with the wild type cells. These observations find support in earlier studies that have shown the influence of collagen structure – monomeric versus fibrillar – on cellular migration, as well as their proliferation and survival. Even though we do not know how the signaling of such process is regulated, one possibility is that DDR2 (collagen receptors) may be involved, since the message for this protein is increased in fibromodulin-deficient plaques.

In summary, fibromodulin affects collagen matrices formed during neointimal response to collar injury, and this in turn regulates proliferation and migration of VSMCs. As a whole, the plaques made with or without fibromodulin can have a different morphological status. This opens up for studies on atherosclerotic relevance of fibromodulin, including the stability of the plaques and their predisposition for thrombosis.

Future perspectives:

Crossing apoE-deficient and fibromodulin-deficient mice to generate a model for studies on fibromodulin influence on plaque stability. This model will have spontaneously formed atherosclerotic plaques (due to lack of apoE) when fed with a fat-rich diet. The plaque stability can be asserted by morphological analyses, like fibrotic tissue damages or thrombus formation.

Analysis of other SLRPs involvement in neointima formation – can they work in synergy, and can this synergy be broken by administration of drugs? What happens to the stability or other physiological properties of the plaque?

How can the trade-off between the total size of the plaque and its stability be controlled by SLRPs' influence on collagen fibrils? Naturally, collagen is more abundant in larger plaques, but the bigger the size the higher risk for occlusion of blood flow. On the other hand, more collagen makes the plaque more stable against thrombosis. Can SLRPs, through alteration of collagen matrix, bring out the best of both worlds? That is, can we manipulate the plaque growth so it becomes small as well as stable?

ACKNOWLEDGEMENTS

Since I don't want to forget anyone here, and I'm not the one for long speeches, I feel it is best to first of all thank everyone who made my stay at BMC a pleasurable event, and to all the people outside the lab and my family back in Poland who all bring joy to life on this planet (you know who you are!). Big thanks also go to all the addicted and addictive music people I had privilege to meet during the last few years, including Henrik Gade, my old and wise guitar teachers Stefan and Dana, and the fun (and sometimes crazy) people I had in my band! Those were the days! There are of course certain persons that have also been central in my professional scientific career (hey, I can finally say that!) and thus I thank my supervisors Åke Oldberg and Anders Aspberg for everything scientific and not-so-scientific, as well as the unofficial supervisor Kristofer Rubin for not least getting me interested in the field of cancer research, for the fine and funny collaborations, and talks on how to make the world's most expensive apple wine (colored black!). Speaking of cancer, Alexei Salnikov has bravely attempted to guide me into the confusing concept of interstitial fluid pressure, connected with discussions on Russian "drinks" and the inferiority of communism (how did these things sneaked in there?!). Also, I need to thank Erik and Lina for teaching me the fine art of electron microscopy, without which we wouldn't get the PNAS paper! Finally, all of you whom I bothered with my questions, or who in other ways kept my projects and spirit running high during the years at BMC – THANK YOU!!!

POPULARIZED SCIENTIFIC SUMMARIES

In Polish – Streszczenie Popularnonaukowe po Polsku

Niniejsza praca poświęcona jest biologii tkanki łącznej i próby jej manipulacji w celu udoskonaleniu metod leczenia wielu chorób. Ponieważ tkanka łączna pełni ważne funkcje w naszym organiźmie, jej struktura kontrolowana jest poprzez różne mechanizmy, które są głównym tematem naszych studiów. Struktura tkanki łącznej jest zmienna i wynika z potrzeb różnych organów, jak np. kości, ścięgna, czy skóry. Choroby takie jak rak, miażdżyca, czy artretyzm, powodują przemiany tej tkanki, co przekłada się na progresję choroby. Możliwośc kontroli nad tymi przemianami ma duże szanse okazać się nową, komplementarną metodą leczenia tych, i wielu innych chorób.

Tkanka łączna składa się w większości z kolagenu, który dziś używany jest także jako popularny środek kosmetyczny. Kolagen jest specjalnym białkiem które może się łączyć w różnomierne włókna, a proces ten jest regulowany przez komórki i przez inne białka, między innymi tzw. SLRPs (small leucine-rich proteoglycans). Nieznany jest wpływ tych białek na budowę włókien kolagenowych, i z tego właśnie powodu potrzebne są badania opisane w tej pracy.

Nasz organizm produkuje wiele różnych białek SLRP, i każde z nich ma swoją rolę w uporządkowaniu kolagenów w specjalne struktury. Ich odmienne możliwości leżą w różnych częściach białka, które składają się z aminokwasów, poukładanych według unikatowych wzorów. Nawet jeśli ogólna struktura danego białka SLRP przypomina inne SLRP, to właśnie te zmienności w aminokwasowych łańcuchach rozstrzygają zróżnicowane funkcje każdego SLRP. Te z kolei różnie oddziałują na budowę włókien z kolagenu, co pokazujemy w jednej z naszych publikacji (artykuł IV). Oczywiście, to ma wpływ na ostateczną strukturę i funkcję tkanki łącznej w skórze, chrząstce, czy w ścięgnach. W naszych studiach (artykuł I-III) zidentyfikowaliśmy na trzech różnych cząsteczkach SLRP miejsca wiązania z kolagenem, co dalej może sprawdzic się w efektywnym manipulowaniu tkanki łącznej w ciężkich chorobach, jak w przypadku raka czy chorób naczyń krwionośnych. Demonstrujemy tu między innymi (artykuł V), że jeden SLRP tak wpływa na budowę tkanki łącznej w raku, że może uniemożliwić efektywną chemioterapię. Tu należy podreślić, że wielu pacjentów z małym przepływem cieczy w nowotworach ma mniejsze szanse na wyleczenie choroby. Jeśli więc umożliwiona będzie zmiana budowy nowotwora poprzez regulowanie SLRP, wtedy będzie można poprawić skuteczność chemioterapii i w ten sposób zwiększyć efektywność leczenia nowotworów.

Jak działa SLRP na budowę stwardnionych i zwapniałych tkanek w chorobach naczyń krwionośnych? Według naszych obserwacji (artykuł VI), wyeliminowanie jednego SLRP podczas procesu twardnienia, prowadzi do zredukowania tej patologicznej tkanki, i z kolei do lepszego przepływu krwi. To redukuje ryzyko zawału serca, i może nawet da kontrolę nad stabilnością tkanki przeciw formacji zakrzepów krwi. Te zwykle mają tendencję do powstawania gdy turbulentny przepływ krwi rozbija niestabilne, zwapniałe tkanki. Takie zakrzepy mogą ”utknąć” w naczyniach które prowadzą krew do ważnych organów, i wtedy doprowadzają do niebezpiecznego dla życia stanu.

W podsumowaniu, nasze badania nad SLRP, ich kontrolą nad tworzeniem włókien kolagenowych, a także jak tę kontrolę można przejąć, i jaki to ma wpływ na progresje groźnych chorób, może się okazać ważne w przyszłej, nowatorskiej medycynie.

In Swedish - Populärvetenskaplig Sammanfattning på Svenska

Denna avhandling kretsar runt bindväv och våra försök att manipulera den, så att vi i framtiden kan förändra sjukdomarnas förlopp. Eftersom bindväv har flera olika funktioner i kroppen, måste dess uppbyggnad kontrolleras genom olika mekanismer. Hur dessa inverkar på produktionen av bindvävens struktur har varit det övergripande syftet med mina studier.

Bindväv kan se olika ut beroende på var den finns – ben, senor eller hud utnyttjar denna vävnad till olika ändamål, för att inte glömma tillstånd som cancer, ateroskleros eller artros, där bindväven är central för hur grav sjukdomen blir. Vår förmåga att utifrån kontrollera bindvävens struktur kan visa sig vara oumbärligt för behandling av dessa, och många andra, sjukdomar.

Bindväv består till större del av kollagen, som idag är en populär ”skönhetsmedel” och ingår i flera kosmetiska produkter. Detta protein har förmåga att byggas upp till större, väldefinierade fiberstrukturer, vilket bestäms av våra celler och andra proteiner, bland annat så kallade SLRPs (small leucine-rich repeat proteoglycans). Hur dessa proteiner fungerar under kollagenets fiberbildning är inte helt känt, vilket förmådde oss att utforska just detta fält.

Det finns många, till synes väldigt lika SLRPs, som verkar fintrimma kollagenets fiberstruktur på olika sätt. Deras skilda förmågor ligger i proteinernas olika delar bestående av byggstenar (aminosyror) som är ordnade på olika sätt i varje protein. Även om den övergripande tredimensionella strukturen av en SLRP påminner om en annan SLRP, är det de små skillnaderna som bestämmer deras enskilda funktion under kollagenets fiberbildning, vilket vi visar i en av våra studier (paper IV). Detta har givetvis konsekvenser för det slutliga utseendet och funktionen av bindväv i t.ex. hud, knäbrosk, eller senor. Vi har lyckats indentifiera bindningsställen mellan tre av dessa SLRPs och kollagen (papers I-III), vilket på lång sikt kan leda till en effektiv manipulering av bindvävsbildning i bl.a. de stora folksjukdomarna som cancer eller hjärt- och kärlsjukdomar. I en av våra studier (paper V) visar vi att bindvävens struktur i cancersvulster (reglerad av en SLRP) skulle kunna påverka cancercellernas känslighet för behandling med cellgifter, genom att tillåta en större mängd läkemedel att flöda in i den sjuka vävnaden. Detta är en viktig upptäckt eftersom många cancerpatienter med lågt vätskeflöde i sina tumörer har en sämre överlevnadsstatistik. Om vi lyckas att aktivt inverka på bildning av kollagenfibrer i cancersvulster, genom att utnyttja vår kunskap om SLRPs, skulle fler cancerbehandlingar vara mer effektiva och därmed bidra till fler räddade liv.

Hur påverkar SLRP bildning av blodkärlsplack i hjärt- och kärlsjukdomar? Enligt våra studier, om man slår ut en SLRP under tiden att placket bildas blir placket mindre och inte lika obstruktiv för blodflödet. Detta kan minska risken eller gravheten av bl.a. hjärtinfarkt, och kan även inverka på plackets stabilitet. I så fall, kan det inte lika lätt slås sönder av det turbulenta blodflödet och sedan fastna i en mindre blodkärl där den skulle stoppa blodflödet till någon viktig organ.

Sammanfattningsvis, vår forskning om hur SLRP påverkar kollagenets fiberstruktur, hur samspel mellan dessa proteiner kan manipuleras, samt hur detta yttrar sig i olika sjukdomsförlopp kan visa sig vara viktig för behandling av flera av de stora folksjukdomarna.

REFERENCES

1. Cooper, G.M. & Hausman, R.E. *The Cell: A Molecular Approach* (Sinauer, Washington, 2004).
2. Schuppan, D.*, et al.* Collagens in the liver extracellular matrix bind hepatocyte growth factor. *Gastroenterology* **114**, 139-152 (1998).
3. Shibuya, H., Okamoto, O. & Fujiwara, S. The bioactivity of transforming growth factor-beta1 can be regulated via binding to dermal collagens in mink lung epithelial cells. *J Dermatol Sci* **41**, 187-195 (2006).
4. Shimabukuro, Y.*, et al.* Basic fibroblast growth factor regulates expression of heparan sulfate in human periodontal ligament cells. *Matrix Biol* **26**, 26 (2007).
5. Somasundaram, R. & Schuppan, D. Type I, II, III, IV, V, and VI collagens serve as extracellular ligands for the isoforms of platelet-derived growth factor (AA, BB, and AB). *J Biol Chem* **271**, 26884-26891 (1996).
6. Myllyharju, J. & Kivirikko, K.I. Collagens, modifying enzymes and their mutations in humans, flies and worms. *Trends Genet* **20**, 33-43 (2004).
7. Gelse, K., Poschl, E. & Aigner, T. Collagens--structure, function, and biosynthesis. *Adv Drug Deliv Rev* **55**, 1531-1546 (2003).
8. Smith, T., Ferreira, L.R., Hebert, C., Norris, K. & Sauk, J.J. Hsp47 and cyclophilin B traverse the endoplasmic reticulum with procollagen into pre-Golgi intermediate vesicles. A role for Hsp47 and cyclophilin B in the export of procollagen from the endoplasmic reticulum. *J Biol Chem* **270**, 18323-18328 (1995).
9. Satoh, M., Hirayoshi, K., Yokota, S., Hosokawa, N. & Nagata, K. Intracellular interaction of collagen-specific stress protein HSP47 with newly synthesized procollagen. *J Cell Biol* **133**, 469-483 (1996).
10. Stephens, D.J. & Pepperkok, R. Imaging of procollagen transport reveals COPI-dependent cargo sorting during ER-to-Golgi transport in mammalian cells. *J Cell Sci* **115**, 1149-1160 (2002).
11. Mironov, A.A.*, et al.* ER-to-Golgi carriers arise through direct en bloc protrusion and multistage maturation of specialized ER exit domains. *Dev Cell* **5**, 583-594 (2003).
12. Bonfanti, L.*, et al.* Procollagen traverses the Golgi stack without leaving the lumen of cisternae: evidence for cisternal maturation. *Cell* **95**, 993-1003 (1998).
13. Mironov, A.A.*, et al.* Small cargo proteins and large aggregates can traverse the Golgi by a common mechanism without leaving the lumen of cisternae. *J Cell Biol* **155**, 1225-1238 (2001).
14. Matsuura-Tokita, K., Takeuchi, M., Ichihara, A., Mikuriya, K. & Nakano, A. Live imaging of yeast Golgi cisternal maturation. *Nature* **441**, 1007-1010 (2006).
15. Rabouille, C. & Klumperman, J. Opinion: The maturing role of COPI vesicles in intra-Golgi transport. *Nat Rev Mol Cell Biol* **6**, 812-817 (2005).
16. Hulmes, D.J., Bruns, R.R. & Gross, J. On the state of aggregation of newly secreted procollagen. *Proc Natl Acad Sci U S A* **80**, 388-392 (1983).
17. Trelstad, R.L. & Coulombre, A.J. Morphogenesis of the collagenous stroma in the chick cornea. *J Cell Biol* **50**, 840-858 (1971).
18. Polishchuk, E.V., Di Pentima, A., Luini, A. & Polishchuk, R.S. Mechanism of constitutive export from the golgi: bulk flow via the formation, protrusion, and en bloc cleavage of large trans-golgi network tubular domains. *Mol Biol Cell* **14**, 4470-4485 (2003).
19. Polishchuk, R.S.*, et al.* Correlative light-electron microscopy reveals the tubular-saccular ultrastructure of carriers operating between Golgi apparatus and plasma membrane. *J Cell Biol* **148**, 45-58 (2000).
20. Canty, E.G.*, et al.* Coalignment of plasma membrane channels and protrusions (fibripositors) specifies the parallelism of tendon. *J Cell Biol* **165**, 553-563 (2004).
21. Canty, E.G.*, et al.* Actin filaments are required for fibripositor-mediated collagen fibril alignment in tendon. *J Biol Chem* **281**, 38592-38598 (2006).
22. Provenzano, P.P. & Vanderby, R., Jr. Collagen fibril morphology and organization: implications for force transmission in ligament and tendon. *Matrix Biol* **25**, 71-84 (2006).
23. Provenzano, P.P., Hurschler, C. & Vanderby, R., Jr. Microstructural morphology in the transition region between scar and intact residual segments of a healing rat medial collateral ligament. *Connect Tissue Res* **42**, 123-133 (2001).
24. Kadler, K.E., Hojima, Y. & Prockop, D.J. Assembly of collagen fibrils de novo by cleavage of the type I pC-collagen with procollagen C-proteinase. Assay of critical concentration demonstrates that collagen self-assembly is a classical example of an entropy-driven process. *J Biol Chem* **262**, 15696-15701 (1987).
25. Miyahara, M.*, et al.* Formation of collagen fibrils by enzymic cleavage of precursors of type I collagen in vitro. *J Biol Chem* **259**, 9891-9898 (1984).
26. Doege, K.J. & Fessler, J.H. Folding of carboxyl domain and assembly of procollagen I. *J Biol Chem* **261**, 8924-8935 (1986).
27. Bulleid, N.J., Dalley, J.A. & Lees, J.F. The C-propeptide domain of procollagen can be replaced with a transmembrane domain without affecting trimer formation or collagen triple helix folding during biosynthesis. *Embo J* **16**, 6694-6701 (1997).
28. Alvares, K., Siddiqui, F., Malone, J. & Veis, A. Assembly of the type 1 procollagen molecule: selectivity of the interactions between the alpha 1(I)- and alpha 2(I)-carboxyl propeptides. *Biochemistry* **38**, 5401-5411 (1999).
29. Lees, J.F., Tasab, M. & Bulleid, N.J. Identification of the molecular recognition sequence which determines the type-specific assembly of procollagen. *Embo J* **16**, 908-916 (1997).
30. Bottomley, M.J., Batten, M.R., Lumb, R.A. & Bulleid, N.J. Quality control in the endoplasmic reticulum: PDI mediates the ER retention of unassembled procollagen C-propeptides. *Curr Biol* **11**, 1114-1118 (2001).
31. Boot-Handford, R.P., Tuckwell, D.S., Plumb, D.A., Rock, C.F. & Poulsom, R. A novel and highly conserved collagen (pro(alpha)1(XXVII)) with a unique expression pattern and unusual molecular characteristics establishes a new clade within the vertebrate fibrillar collagen family. *J Biol Chem* **278**, 31067-31077 (2003).
32. Li, S.W.*, et al.* The C-proteinase that processes procollagens to fibrillar collagens is identical to the protein previously identified as bone morphogenic protein-1. *Proc Natl Acad Sci U S A* **93**, 5127-5130 (1996).
33. Scott, I.C.*, et al.* Mammalian BMP-1/Tolloid-related metalloproteinases, including novel family member mammalian Tolloid-like 2, have differential enzymatic activities and distributions of expression relevant to patterning and skeletogenesis. *Dev Biol* **213**, 283-300 (1999).

34. Exposito, J.Y., Cluzel, C., Garrone, R. & Lethias, C. Evolution of collagens. *Anat Rec* **268**, 302-316 (2002).
35. Colige, A.*, et al.* Cloning and characterization of ADAMTS-14, a novel ADAMTS displaying high homology with ADAMTS-2 and ADAMTS-3. *J Biol Chem* **277**, 5756-5766 (2002).
36. Colige, A.*, et al.* cDNA cloning and expression of bovine procollagen I N-proteinase: a new member of the superfamily of zinc-metalloproteinases with binding sites for cells and other matrix components. *Proc Natl Acad Sci U S A* **94**, 2374-2379 (1997).
37. Le Goff, C.*, et al.* Regulation of procollagen amino-propeptide processing during mouse embryogenesis by specialization of homologous ADAMTS proteases: insights on collagen biosynthesis and dermatosparaxis. *Development* **133**, 1587-1596 (2006).
38. Kuznetsova, N.V., McBride, D.J. & Leikin, S. Changes in thermal stability and microunfolding pattern of collagen helix resulting from the loss of alpha2(I) chain in osteogenesis imperfecta murine. *J Mol Biol* **331**, 191-200 (2003).
39. Malone, J.P., Alvares, K. & Veis, A. Structure and assembly of the heterotrimeric and homotrimeric C-propeptides of type I collagen: significance of the alpha2(I) chain. *Biochemistry* **44**, 15269-15279 (2005).
40. Leikina, E., Mertts, M.V., Kuznetsova, N. & Leikin, S. Type I collagen is thermally unstable at body temperature. *Proc Natl Acad Sci U S A* **99**, 1314-1318 (2002).
41. Minond, D.*, et al.* The roles of substrate thermal stability and P2 and P1' subsite identity on matrix metalloproteinase triple-helical peptidase activity and collagen specificity. *J Biol Chem* **281**, 38302-38313 (2006).
42. Canty, E.G. & Kadler, K.E. Procollagen trafficking, processing and fibrillogenesis. *J Cell Sci* **118**, 1341-1353 (2005).
43. Leitinger, B. & Hohenester, E. Mammalian collagen receptors. *Matrix Biol* **26**, 146-155 (2007).
44. Heino, J. The collagen receptor integrins have distinct ligand recognition and signaling functions. *Matrix Biol* **19**, 319-323 (2000).
45. Li, S., Van Den Diepstraten, C., D'Souza, S.J., Chan, B.M. & Pickering, J.G. Vascular smooth muscle cells orchestrate the assembly of type I collagen via alpha2beta1 integrin, RhoA, and fibronectin polymerization. *Am J Pathol* **163**, 1045-1056 (2003).
46. Velling, T., Risteli, J., Wennerberg, K., Mosher, D.F. & Johansson, S. Polymerization of type I and III collagens is dependent on fibronectin and enhanced by integrins alpha 11beta 1 and alpha 2beta 1. *J Biol Chem* **277**, 37377-37381 (2002).
47. Vogel, W.F., Abdulhussein, R. & Ford, C.E. Sensing extracellular matrix: an update on discoidin domain receptor function. *Cell Signal* **18**, 1108-1116 (2006).
48. Engelholm, L.H.*, et al.* uPARAP/Endo180 is essential for cellular uptake of collagen and promotes fibroblast collagen adhesion. *J Cell Biol* **160**, 1009-1015 (2003).
49. Madsen, D.H.*, et al.* Extracellular collagenases and the endocytic receptor, urokinase plasminogen activator receptor-associated protein/Endo180, cooperate in fibroblast-mediated collagen degradation. *J Biol Chem* **282**, 27037-27045 (2007).
50. Curino, A.C.*, et al.* Intracellular collagen degradation mediated by uPARAP/Endo180 is a major pathway of extracellular matrix turnover during malignancy. *J Cell Biol* **169**, 977-985 (2005).
51. Graham, H.K., Holmes, D.F., Watson, R.B. & Kadler, K.E. Identification of collagen fibril fusion during vertebrate tendon morphogenesis. The process relies on unipolar fibrils and is regulated by collagen-proteoglycan interaction. *J Mol Biol* **295**, 891-902 (2000).
52. Hedbom, E. & Heinegard, D. Interaction of a 59-kDa connective tissue matrix protein with collagen I and collagen II. *J Biol Chem* **264**, 6898-6905 (1989).
53. Vogel, K.G., Paulsson, M. & Heinegard, D. Specific inhibition of type I and type II collagen fibrillogenesis by the small proteoglycan of tendon. *Biochem J* **223**, 587-597 (1984).
54. Danielson, K.G.*, et al.* Targeted disruption of decorin leads to abnormal collagen fibril morphology and skin fragility. *J Cell Biol* **136**, 729-743 (1997).
55. Svensson, L.*, et al.* Fibromodulin-null mice have abnormal collagen fibrils, tissue organization, and altered lumican deposition in tendon. *J Biol Chem* **274**, 9636-9647 (1999).
56. Font, B., Eichenberger, D., Rosenberg, L.M. & van der Rest, M. Characterization of the interactions of type XII collagen with two small proteoglycans from fetal bovine tendon, decorin and fibromodulin. *Matrix Biol* **15**, 341-348 (1996).
57. Font, B., Aubert-Foucher, E., Goldschmidt, D., Eichenberger, D. & van der Rest, M. Binding of collagen XIV with the dermatan sulfate side chain of decorin. *J Biol Chem* **268**, 25015-25018 (1993).
58. Nishiyama, T., McDonough, A.M., Bruns, R.R. & Burgeson, R.E. Type XII and XIV collagens mediate interactions between banded collagen fibers in vitro and may modulate extracellular matrix deformability. *J Biol Chem* **269**, 28193-28199 (1994).
59. Holmes, D.F., Lowe, M.P. & Chapman, J.A. Vertebrate (chick) collagen fibrils formed in vivo can exhibit a reversal in molecular polarity. *J Mol Biol* **235**, 80-83 (1994).
60. Helseth, D.L., Jr. & Veis, A. Collagen self-assembly in vitro. Differentiating specific telopeptide-dependent interactions using selective enzyme modification and the addition of free amino telopeptide. *J Biol Chem* **256**, 7118-7128 (1981).
61. Holmes, D.F., Chapman, J.A., Prockop, D.J. & Kadler, K.E. Growing tips of type I collagen fibrils formed in vitro are near-paraboloidal in shape, implying a reciprocal relationship between accretion and diameter. *Proc Natl Acad Sci U S A* **89**, 9855-9859 (1992).
62. Kadler, K.E., Hojima, Y. & Prockop, D.J. Collagen fibrils in vitro grow from pointed tips in the C- to N-terminal direction. *Biochem J* **268**, 339-343 (1990).
63. Kuznetsova, N., Chi, S.L. & Leikin, S. Sugars and polyols inhibit fibrillogenesis of type I collagen by disrupting hydrogen-bonded water bridges between the helices. *Biochemistry* **37**, 11888-11895 (1998).
64. Leikin, S., Rau, D.C. & Parsegian, V.A. Temperature-favoured assembly of collagen is driven by hydrophilic not hydrophobic interactions. *Nat Struct Biol* **2**, 205-210 (1995).
65. Bella, J., Eaton, M., Brodsky, B. & Berman, H.M. Crystal and molecular structure of a collagen-like peptide at 1.9 A resolution. *Science* **266**, 75-81 (1994).
66. Leikin, S., Rau, D.C. & Parsegian, V.A. Direct measurement of forces between self-assembled proteins: temperature-dependent exponential forces between collagen triple helices. *Proc Natl Acad Sci U S A* **91**, 276-280 (1994).
67. Malone, J.P. & Veis, A. Heterotrimeric type I collagen C-telopeptide conformation as docked to its helix receptor. *Biochemistry* **43**, 15358-15366 (2004).

68. Orgel, J.P., Wess, T.J. & Miller, A. The in situ conformation and axial location of the intermolecular cross-linked non-helical telopeptides of type I collagen. *Structure* **8**, 137-142 (2000).
69. Malone, J.P., George, A. & Veis, A. Type I collagen N-telopeptides adopt an ordered structure when docked to their helix receptor during fibrillogenesis. *Proteins* **54**, 206-215 (2004).
70. Prockop, D.J. & Fertala, A. Inhibition of the self-assembly of collagen I into fibrils with synthetic peptides. Demonstration that assembly is driven by specific binding sites on the monomers. *J Biol Chem* **273**, 15598-15604 (1998).
71. Brown, R.A., Hukins, D.W., Weiss, J.B. & Twose, T.M. Do mammalian collagenases and DNA restriction endonucleases share a similar mechanism for cleavage site recognition? *Biochem Biophys Res Commun* **74**, 1102-1108 (1977).
72. Kuznetsova, N. & Leikin, S. Does the triple helical domain of type I collagen encode molecular recognition and fiber assembly while telopeptides serve as catalytic domains? Effect of proteolytic cleavage on fibrillogenesis and on collagen-collagen interaction in fibers. *J Biol Chem* **274**, 36083-36088 (1999).
73. Wang, X., Li, X., Bank, R.A. & Agrawal, C.M. Effects of collagen unwinding and cleavage on the mechanical integrity of the collagen network in bone. *Calcif Tissue Int* **71**, 186-192 (2002).
74. Bailey, A.J., Robins, S.P. & Balian, G. Biological significance of the intermolecular crosslinks of collagen. *Nature* **251**, 105-109 (1974).
75. Eyre, D.R., Paz, M.A. & Gallop, P.M. Cross-linking in collagen and elastin. *Annu Rev Biochem* **53**, 717-748 (1984).
76. Kuypers, R., Tyler, M., Kurth, L.B., Jenkins, I.D. & Horgan, D.J. Identification of the loci of the collagen-associated Ehrlich chromogen in type I collagen confirms its role as a trivalent cross-link. *Biochem J* **283 (Pt 1)**, 129-136 (1992).
77. Kemp, P.D. & Scott, J.E. Ehrlich chromogens, probable cross-links in elastin and collagen. *Biochem J* **252**, 387-393 (1988).
78. Hanson, D.A. & Eyre, D.R. Molecular site specificity of pyridinoline and pyrrole cross-links in type I collagen of human bone. *J Biol Chem* **271**, 26508-26516 (1996).
79. Reiser, K., McCormick, R.J. & Rucker, R.B. Enzymatic and nonenzymatic cross-linking of collagen and elastin. *Faseb J* **6**, 2439-2449 (1992).
80. Eyre, D.R., Koob, T.J. & Van Ness, K.P. Quantitation of hydroxypyridinium crosslinks in collagen by high-performance liquid chromatography. *Anal Biochem* **137**, 380-388 (1984).
81. van der Slot, A.J.*, et al.* Increased formation of pyridinoline cross-links due to higher telopeptide lysyl hydroxylase levels is a general fibrotic phenomenon. *Matrix Biol* **23**, 251-257 (2004).
82. Bachinger, H.P. The influence of peptidyl-prolyl cis-trans isomerase on the in vitro folding of type III collagen. *J Biol Chem* **262**, 17144-17148 (1987).
83. Pihlajaniemi, T.*, et al.* Molecular cloning of the beta-subunit of human prolyl 4-hydroxylase. This subunit and protein disulphide isomerase are products of the same gene. *Embo J* **6**, 643-649 (1987).
84. Kivirikko, K.I. & Myllyharju, J. Prolyl 4-hydroxylases and their protein disulfide isomerase subunit. *Matrix Biol* **16**, 357-368 (1998).
85. Walmsley, A.R., Batten, M.R., Lad, U. & Bulleid, N.J. Intracellular retention of procollagen within the endoplasmic reticulum is mediated by prolyl 4-hydroxylase. *J Biol Chem* **274**, 14884-14892 (1999).
86. Berg, R.A. & Prockop, D.J. The thermal transition of a non-hydroxylated form of collagen. Evidence for a role for hydroxyproline in stabilizing the triple-helix of collagen. *Biochem Biophys Res Commun* **52**, 115-120 (1973).
87. Berisio, R., Granata, V., Vitagliano, L. & Zagari, A. Characterization of collagen-like heterotrimers: implications for triple-helix stability. *Biopolymers* **73**, 682-688 (2004).
88. Bella, J., Brodsky, B. & Berman, H.M. Hydration structure of a collagen peptide. *Structure* **3**, 893-906 (1995).
89. Kawahara, K.*, et al.* Effect of hydration on the stability of the collagen-like triple-helical structure of [4(R)-hydroxyprolyl-4(R)-hydroxyprolylglycine]10. *Biochemistry* **44**, 15812-15822 (2005).
90. Holmgren, S.K., Taylor, K.M., Bretscher, L.E. & Raines, R.T. Code for collagen's stability deciphered. *Nature* **392**, 666-667 (1998).
91. Koivu, J. & Myllyla, R. Interchain disulfide bond formation in types I and II procollagen. Evidence for a protein disulfide isomerase catalyzing bond formation. *J Biol Chem* **262**, 6159-6164 (1987).
92. Koivu, J. & Myllyla, R. Protein disulfide-isomerase retains procollagen prolyl 4-hydroxylase structure in its native conformation. *Biochemistry* **25**, 5982-5986 (1986).
93. John, D.C., Grant, M.E. & Bulleid, N.J. Cell-free synthesis and assembly of prolyl 4-hydroxylase: the role of the beta-subunit (PDI) in preventing misfolding and aggregation of the alpha-subunit. *Embo J* **12**, 1587-1595 (1993).
94. Eyre, D.R. & Glimcher, M.J. Reducible crosslinks in hydroxylysine-deficient collagens of a heritable disorder of connective tissue. *Proc Natl Acad Sci U S A* **69**, 2594-2598 (1972).
95. Brinckmann, J.*, et al.* Overhydroxylation of lysyl residues is the initial step for altered collagen cross-links and fibril architecture in fibrotic skin. *J Invest Dermatol* **113**, 617-621 (1999).
96. Takaluoma, K.*, et al.* Tissue-specific changes in the hydroxylysine content and cross-links of collagens and alterations in fibril morphology in lysyl hydroxylase 1 knock-out mice. *J Biol Chem* **282**, 6588-6596 (2007).
97. Eyre, D., Shao, P., Weis, M.A. & Steinmann, B. The kyphoscoliotic type of Ehlers-Danlos syndrome (type VI): differential effects on the hydroxylation of lysine in collagens I and II revealed by analysis of cross-linked telopeptides from urine. *Mol Genet Metab* **76**, 211-216 (2002).
98. Yeowell, H.N. & Walker, L.C. Tissue specificity of a new splice form of the human lysyl hydroxylase 2 gene. *Matrix Biol* **18**, 179-187 (1999).
99. Takaluoma, K., Lantto, J. & Myllyharju, J. Lysyl hydroxylase 2 is a specific telopeptide hydroxylase, while all three isoenzymes hydroxylate collagenous sequences. *Matrix Biol* **26**, 396-403 (2007).
100. van der Slot, A.J.*, et al.* Elevated formation of pyridinoline cross-links by profibrotic cytokines is associated with enhanced lysyl hydroxylase 2b levels. *Biochim Biophys Acta* **1741**, 95-102 (2005).
101. van der Slot, A.J.*, et al.* Identification of PLOD2 as telopeptide lysyl hydroxylase, an important enzyme in fibrosis. *J Biol Chem* **278**, 40967-40972 (2003).
102. Gerriets, J.E., Curwin, S.L. & Last, J.A. Tendon hypertrophy is associated with increased hydroxylation of nonhelical lysine residues at two specific cross-linking sites in type I collagen. *J Biol Chem* **268**, 25553-25560 (1993).
103. Pornprasertsuk, S., Duarte, W.R., Mochida, Y. & Yamauchi, M. Lysyl hydroxylase-2b directs collagen cross-linking pathways in MC3T3-E1 cells. *J Bone Miner Res* **19**, 1349-1355 (2004).

104. Heikkinen, J., *et al.* Lysyl hydroxylase 3 is a multifunctional protein possessing collagen glucosyltransferase activity. *J Biol Chem* **275**, 36158-36163 (2000).
105. Ruotsalainen, H., *et al.* Glycosylation catalyzed by lysyl hydroxylase 3 is essential for basement membranes. *J Cell Sci* **119**, 625-635 (2006).
106. Nagai, N., *et al.* Embryonic lethality of molecular chaperone hsp47 knockout mice is associated with defects in collagen biosynthesis. *J Cell Biol* **150**, 1499-1506 (2000).
107. Nakai, A., Satoh, M., Hirayoshi, K. & Nagata, K. Involvement of the stress protein HSP47 in procollagen processing in the endoplasmic reticulum. *J Cell Biol* **117**, 903-914 (1992).
108. Koide, T., Takahara, Y., Asada, S. & Nagata, K. Xaa-Arg-Gly triplets in the collagen triple helix are dominant binding sites for the molecular chaperone HSP47. *J Biol Chem* **277**, 6178-6182 (2002).
109. Tasab, M., Batten, M.R. & Bulleid, N.J. Hsp47: a molecular chaperone that interacts with and stabilizes correctly-folded procollagen. *Embo J* **19**, 2204-2211 (2000).
110. Koide, T., *et al.* Specific recognition of the collagen triple helix by chaperone HSP47: minimal structural requirement and spatial molecular orientation. *J Biol Chem* **281**, 3432-3438 (2006).
111. Makareeva, E. & Leikin, S. Procollagen triple helix assembly: an unconventional chaperone-assisted folding paradigm. *PLoS ONE* **2**, e1029 (2007).
112. Hornstra, I.K., *et al.* Lysyl oxidase is required for vascular and diaphragmatic development in mice. *J Biol Chem* **278**, 14387-14393 (2003).
113. Erler, J.T., *et al.* Lysyl oxidase is essential for hypoxia-induced metastasis. *Nature* **440**, 1222-1226 (2006).
114. Uzel, M.I., *et al.* Multiple bone morphogenetic protein 1-related mammalian metalloproteinases process pro-lysyl oxidase at the correct physiological site and control lysyl oxidase activation in mouse embryo fibroblast cultures. *J Biol Chem* **276**, 22537-22543 (2001).
115. Kagan, H.M. & Li, W. Lysyl oxidase: properties, specificity, and biological roles inside and outside of the cell. *J Cell Biochem* **88**, 660-672 (2003).
116. Siegel, R.C. Biosynthesis of collagen crosslinks: increased activity of purified lysyl oxidase with reconstituted collagen fibrils. *Proc Natl Acad Sci U S A* **71**, 4826-4830 (1974).
117. Nagan, N. & Kagan, H.M. Modulation of lysyl oxidase activity toward peptidyl lysine by vicinal dicarboxylic amino acid residues. Implications for collagen cross-linking. *J Biol Chem* **269**, 22366-22371 (1994).
118. Vranka, J.A., Sakai, L.Y. & Bachinger, H.P. Prolyl 3-hydroxylase 1, enzyme characterization and identification of a novel family of enzymes. *J Biol Chem* **279**, 23615-23621 (2004).
119. Cabral, W.A., *et al.* Prolyl 3-hydroxylase 1 deficiency causes a recessive metabolic bone disorder resembling lethal/severe osteogenesis imperfecta. *Nat Genet* **39**, 359-365 (2007).
120. Morello, R., *et al.* CRTAP is required for prolyl 3- hydroxylation and mutations cause recessive osteogenesis imperfecta. *Cell* **127**, 291-304 (2006).
121. Harwood, R., Grant, M.E. & Jackson, D.S. Studies on the glycosylation of hydroxylysine residues during collagen biosynthesis and the subcellular localization of collagen galactosyltransferase and collagen glucosyltransferase in tendon and cartilage cells. *Biochem J* **152**, 291-302 (1975).
122. Brownell, A.G. & Veis, A. The intracellular location of the glycosylation of hydroxylysine of collagen. *Biochem Biophys Res Commun* **63**, 371-377 (1975).
123. Verzijl, N., *et al.* Crosslinking by advanced glycation end products increases the stiffness of the collagen network in human articular cartilage: a possible mechanism through which age is a risk factor for osteoarthritis. *Arthritis Rheum* **46**, 114-123 (2002).
124. Birk, D.E., Fitch, J.M., Babiarz, J.P., Doane, K.J. & Linsenmayer, T.F. Collagen fibrillogenesis in vitro: interaction of types I and V collagen regulates fibril diameter. *J Cell Sci* **95 (Pt 4)**, 649-657 (1990).
125. Wenstrup, R.J., *et al.* Type V collagen controls the initiation of collagen fibril assembly. *J Biol Chem* **279**, 53331-53337 (2004).
126. Boot-Handford, R.P. & Tuckwell, D.S. Fibrillar collagen: the key to vertebrate evolution? A tale of molecular incest. *Bioessays* **25**, 142-151 (2003).
127. Cluzel, C., Lethias, C., Garrone, R. & Exposito, J.Y. Distinct maturations of N-propeptide domains in fibrillar procollagen molecules involved in the formation of heterotypic fibrils in adult sea urchin collagenous tissues. *J Biol Chem* **279**, 9811-9817 (2004).
128. Aouacheria, A., *et al.* Invertebrate data predict an early emergence of vertebrate fibrillar collagen clades and an anti-incest model. *J Biol Chem* **279**, 47711-47719 (2004).
129. Huxley-Jones, J., Robertson, D.L. & Boot-Handford, R.P. On the origins of the extracellular matrix in vertebrates. *Matrix Biol* **26**, 2-11 (2007).
130. Wada, H., Okuyama, M., Satoh, N. & Zhang, S. Molecular evolution of fibrillar collagen in chordates, with implications for the evolution of vertebrate skeletons and chordate phylogeny. *Evol Dev* **8**, 370-377 (2006).
131. Zhang, X., *et al.* The collagens of hydra provide insight into the evolution of metazoan extracellular matrices. *J Biol Chem* **282**, 6792-6802 (2007).
132. Tulla, M., *et al.* Analysis of an ascidian integrin provides new insight into early evolution of collagen recognition. *FEBS Lett* **581**, 2434-2440 (2007).
133. Ewan, R., *et al.* The integrins of the urochordate Ciona intestinalis provide novel insights into the molecular evolution of the vertebrate integrin family. *BMC Evol Biol* **5**, 31 (2005).
134. Huhtala, M., Heino, J., Casciari, D., de Luise, A. & Johnson, M.S. Integrin evolution: insights from ascidian and teleost fish genomes. *Matrix Biol* **24**, 83-95 (2005).
135. Panopoulou, G. & Poustka, A.J. Timing and mechanism of ancient vertebrate genome duplications -- the adventure of a hypothesis. *Trends Genet* **21**, 559-567 (2005).
136. Matsushima, N., Ohyanagi, T., Tanaka, T. & Kretsinger, R.H. Super-motifs and evolution of tandem leucine-rich repeats within the small proteoglycans--biglycan, decorin, lumican, fibromodulin, PRELP, keratocan, osteoadherin, epiphycan, and osteoglycin. *Proteins* **38**, 210-225 (2000).
137. Iozzo, R.V. The family of the small leucine-rich proteoglycans: key regulators of matrix assembly and cellular growth. *Crit Rev Biochem Mol Biol* **32**, 141-174 (1997).
138. McEwan, P.A., Scott, P.G., Bishop, P.N. & Bella, J. Structural correlations in the family of small leucine-rich repeat proteins and proteoglycans. *J Struct Biol* **155**, 294-305 (2006).
139. Kajava, A.V. Structural diversity of leucine-rich repeat proteins. *J Mol Biol* **277**, 519-527 (1998).

140. Scott, P.G., *et al.* Crystal structure of the dimeric protein core of decorin, the archetypal small leucine-rich repeat proteoglycan. *Proc Natl Acad Sci U S A* **101**, 15633-15638 (2004).
141. Scott, P.G., Dodd, C.M., Bergmann, E.M., Sheehan, J.K. & Bishop, P.N. Crystal structure of the biglycan dimer and evidence that dimerization is essential for folding and stability of class I small leucine-rich repeat proteoglycans. *J Biol Chem* **281**, 13324-13332 (2006).
142. Le Goff, M.M., Hindson, V.J., Jowitt, T.A., Scott, P.G. & Bishop, P.N. Characterization of opticin and evidence of stable dimerization in solution. *J Biol Chem* **278**, 45280-45287 (2003).
143. Kobe, B. & Kajava, A.V. The leucine-rich repeat as a protein recognition motif. *Curr Opin Struct Biol* **11**, 725-732 (2001).
144. Bell, J.K., Askins, J., Hall, P.R., Davies, D.R. & Segal, D.M. The dsRNA binding site of human Toll-like receptor 3. *Proc Natl Acad Sci U S A* **103**, 8792-8797 (2006).
145. Bell, J.K., *et al.* Leucine-rich repeats and pathogen recognition in Toll-like receptors. *Trends Immunol* **24**, 528-533 (2003).
146. Hildebrand, A., *et al.* Interaction of the small interstitial proteoglycans biglycan, decorin and fibromodulin with transforming growth factor beta. *Biochem J* **302 (Pt 2)**, 527-534 (1994).
147. Fava, R.A. & McClure, D.B. Fibronectin-associated transforming growth factor. *J Cell Physiol* **131**, 184-189 (1987).
148. Lorenzo, P., *et al.* Identification and characterization of asporin. a novel member of the leucine-rich repeat protein family closely related to decorin and biglycan. *J Biol Chem* **276**, 12201-12211 (2001).
149. Henry, S.P., *et al.* Expression pattern and gene characterization of asporin. a newly discovered member of the leucine-rich repeat protein family. *J Biol Chem* **276**, 12212-12221 (2001).
150. Kizawa, H., *et al.* An aspartic acid repeat polymorphism in asporin inhibits chondrogenesis and increases susceptibility to osteoarthritis. *Nat Genet* **37**, 138-144 (2005).
151. Nakamura, T., *et al.* Meta-analysis of association between the ASPN D-repeat and osteoarthritis. *Hum Mol Genet* **16**, 1676-1681 (2007).
152. Hunter, G.K. & Goldberg, H.A. Nucleation of hydroxyapatite by bone sialoprotein. *Proc Natl Acad Sci U S A* **90**, 8562-8565 (1993).
153. Hunter, G.K., Kyle, C.L. & Goldberg, H.A. Modulation of crystal formation by bone phosphoproteins: structural specificity of the osteopontin-mediated inhibition of hydroxyapatite formation. *Biochem J* **300 (Pt 3)**, 723-728 (1994).
154. Yamada, S., *et al.* PLAP-1/asporin, a novel negative regulator of periodontal ligament mineralization. *J Biol Chem* **282**, 23070-23080 (2007).
155. Kou, I., Nakajima, M. & Ikegawa, S. Expression and regulation of the osteoarthritis-associated protein asporin. *J Biol Chem* **282**, 32193-32199 (2007).
156. Nakajima, M., *et al.* Mechanisms for asporin function and regulation in articular cartilage. *J Biol Chem* **282**, 32185-32192 (2007).
157. Fisher, L.W., Termine, J.D. & Young, M.F. Deduced protein sequence of bone small proteoglycan I (biglycan) shows homology with proteoglycan II (decorin) and several nonconnective tissue proteins in a variety of species. *J Biol Chem* **264**, 4571-4576 (1989).
158. Fisher, L.W., *et al.* Human biglycan gene. Putative promoter, intron-exon junctions, and chromosomal localization. *J Biol Chem* **266**, 14371-14377 (1991).
159. Perez-Castro, A.V. & Vogel, K.G. In situ expression of collagen and proteoglycan genes during development of fibrocartilage in bovine deep flexor tendon. *J Orthop Res* **17**, 139-148 (1999).
160. Wilda, M., *et al.* A comparison of the expression pattern of five genes of the family of small leucine-rich proteoglycans during mouse development. *J Bone Miner Res* **15**, 2187-2196 (2000).
161. Wegrowski, Y., Pillarisetti, J., Danielson, K.G., Suzuki, S. & Iozzo, R.V. The murine biglycan: complete cDNA cloning, genomic organization, promoter function, and expression. *Genomics* **30**, 8-17 (1995).
162. Pearson, D. & Sasse, J. Differential regulation of biglycan and decorin by retinoic acid in bovine chondrocytes. *J Biol Chem* **267**, 25364-25370 (1992).
163. Ungefroren, H., Cikos, T., Krull, N.B. & Kalthoff, H. Biglycan gene promoter activity in osteosarcoma cells is regulated by cyclic AMP. *Biochem Biophys Res Commun* **235**, 413-417 (1997).
164. Moreno, M., *et al.* Biglycan is a new extracellular component of the Chordin-BMP4 signaling pathway. *EMBO J* **24**, 1397-1405 (2005).
165. Scott, I.C., *et al.* Bone morphogenetic protein-1 processes probiglycan. *J Biol Chem* **275**, 30504-30511 (2000).
166. Roughley, P.J., White, R.J. & Mort, J.S. Presence of pro-forms of decorin and biglycan in human articular cartilage. *Biochem J* **318 (Pt 3)**, 779-784 (1996).
167. Chen, X.D., Fisher, L.W., Robey, P.G. & Young, M.F. The small leucine-rich proteoglycan biglycan modulates BMP-4-induced osteoblast differentiation. *FASEB J* **18**, 948-958 (2004).
168. Xu, T., *et al.* Targeted disruption of the biglycan gene leads to an osteoporosis-like phenotype in mice. *Nat Genet* **20**, 78-82 (1998).
169. Chen, X.D., Shi, S., Xu, T., Robey, P.G. & Young, M.F. Age-related osteoporosis in biglycan-deficient mice is related to defects in bone marrow stromal cells. *J Bone Miner Res* **17**, 331-340 (2002).
170. Kahari, V.M., Larjava, H. & Uitto, J. Differential regulation of extracellular matrix proteoglycan (PG) gene expression. Transforming growth factor-beta 1 up-regulates biglycan (PGI), and versican (large fibroblast PG) but down-regulates decorin (PGII) mRNA levels in human fibroblasts in culture. *J Biol Chem* **266**, 10608-10615 (1991).
171. Ungefroren, H. & Krull, N.B. Transcriptional regulation of the human biglycan gene. *J Biol Chem* **271**, 15787-15795 (1996).
172. Ungefroren, H., Lenschow, W., Chen, W.B., Faendrich, F. & Kalthoff, H. Regulation of biglycan gene expression by transforming growth factor-beta requires MKK6-p38 mitogen-activated protein Kinase signaling downstream of Smad signaling. *J Biol Chem* **278**, 11041-11049 (2003).
173. Chen, W.B., *et al.* Smad4/DPC4-dependent regulation of biglycan gene expression by transforming growth factor-beta in pancreatic tumor cells. *J Biol Chem* **277**, 36118-36128 (2002).
174. Ungefroren, H., Groth, S., Ruhnke, M., Kalthoff, H. & Fandrich, F. Transforming growth factor-beta (TGF-beta) type I receptor/ALK5-dependent activation of the GADD45beta gene mediates the induction of biglycan expression by TGF-beta. *J Biol Chem* **280**, 2644-2652 (2005).
175. Bi, Y., *et al.* Extracellular matrix proteoglycans control the fate of bone marrow stromal cells. *J Biol Chem* **280**, 30481-30489 (2005).

176. Corsi, A.*, et al.* Phenotypic effects of biglycan deficiency are linked to collagen fibril abnormalities, are synergized by decorin deficiency, and mimic Ehlers-Danlos-like changes in bone and other connective tissues. *J Bone Miner Res* **17**, 1180-1189 (2002).
177. Lecanda, F., Avioli, L.V. & Cheng, S.L. Regulation of bone matrix protein expression and induction of differentiation of human osteoblasts and human bone marrow stromal cells by bone morphogenetic protein-2. *J Cell Biochem* **67**, 386-396 (1997).
178. Kimoto, S., Cheng, S.L., Zhang, S.F. & Avioli, L.V. The effect of glucocorticoid on the synthesis of biglycan and decorin in human osteoblasts and bone marrow stromal cells. *Endocrinology* **135**, 2423-2431 (1994).
179. Bi, Y.*, et al.* Biglycan deficiency increases osteoclast differentiation and activity due to defective osteoblasts. *Bone* **38**, 778-786 (2006).
180. Bi, Y.*, et al.* Identification of tendon stem/progenitor cells and the role of the extracellular matrix in their niche. *Nat Med* **13**, 1219-1227 (2007).
181. Nielsen, K.L.*, et al.* Biglycan deficiency interferes with ovariectomy-induced bone loss. *J Bone Miner Res* **18**, 2152-2158 (2003).
182. Wallace, J.M.*, et al.* The mechanical phenotype of biglycan-deficient mice is bone- and gender-specific. *Bone* **39**, 106-116 (2006).
183. Ameye, L.*, et al.* Abnormal collagen fibrils in tendons of biglycan/fibromodulin-deficient mice lead to gait impairment, ectopic ossification, and osteoarthritis. *FASEB J* **16**, 673-680 (2002).
184. Wadhwa, S., Embree, M.C., Kilts, T., Young, M.F. & Ameye, L.G. Accelerated osteoarthritis in the temporomandibular joint of biglycan/fibromodulin double-deficient mice. *Osteoarthritis Cartilage* **13**, 817-827 (2005).
185. Wadhwa, S.*, et al.* Impaired posterior frontal sutural fusion in the biglycan/decorin double deficient mice. *Bone* **40**, 861-866 (2007).
186. Robinson, P.S.*, et al.* Influence of decorin and biglycan on mechanical properties of multiple tendons in knockout mice. *J Biomech Eng* **127**, 181-185 (2005).
187. Zhang, G.*, et al.* Decorin regulates assembly of collagen fibrils and acquisition of biomechanical properties during tendon development. *J Cell Biochem* **98**, 1436-1449 (2006).
188. Robbins, J.R., Evanko, S.P. & Vogel, K.G. Mechanical loading and TGF-beta regulate proteoglycan synthesis in tendon. *Arch Biochem Biophys* **342**, 203-211 (1997).
189. Casar, J.C., McKechnie, B.A., Fallon, J.R., Young, M.F. & Brandan, E. Transient up-regulation of biglycan during skeletal muscle regeneration: delayed fiber growth along with decorin increase in biglycan-deficient mice. *Dev Biol* **268**, 358-371 (2004).
190. Bowe, M.A., Mendis, D.B. & Fallon, J.R. The small leucine-rich repeat proteoglycan biglycan binds to alpha-dystroglycan and is upregulated in dystrophic muscle. *J Cell Biol* **148**, 801-810 (2000).
191. Fadic, R.*, et al.* Increase in decorin and biglycan in Duchenne Muscular Dystrophy: role of fibroblasts as cell source of these proteoglycans in the disease. *J Cell Mol Med* **10**, 758-769 (2006).
192. Lechner, B.E., Lim, J.H., Mercado, M.L. & Fallon, J.R. Developmental regulation of biglycan expression in muscle and tendon. *Muscle Nerve* **34**, 347-355 (2006).
193. Schonherr, E.*, et al.* Interaction of biglycan with type I collagen. *J Biol Chem* **270**, 2776-2783 (1995).
194. Douglas, T., Heinemann, S., Bierbaum, S., Scharnweber, D. & Worch, H. Fibrillogenesis of collagen types I, II, and III with small leucine-rich proteoglycans decorin and biglycan. *Biomacromolecules* **7**, 2388-2393 (2006).
195. Wiberg, C.*, et al.* Biglycan and decorin bind close to the n-terminal region of the collagen VI triple helix. *J Biol Chem* **276**, 18947-18952 (2001).
196. Wiberg, C., Heinegard, D., Wenglen, C., Timpl, R. & Morgelin, M. Biglycan organizes collagen VI into hexagonal-like networks resembling tissue structures. *J Biol Chem* **277**, 49120-49126 (2002).
197. Wiberg, C.*, et al.* Complexes of matrilin-1 and biglycan or decorin connect collagen VI microfibrils to both collagen II and aggrecan. *J Biol Chem* **278**, 37698-37704 (2003).
198. Williams, K.J.*, et al.* Decorin deficiency enhances progressive nephropathy in diabetic mice. *Am J Pathol* **171**, 1441-1450 (2007).
199. Fust, A., LeBellego, F., Iozzo, R.V., Roughley, P.J. & Ludwig, M.S. Alterations in lung mechanics in decorin-deficient mice. *Am J Physiol Lung Cell Mol Physiol* **288**, L159-166 (2005).
200. de Medeiros Matsushita, M.*, et al.* Airway proteoglycans are differentially altered in fatal asthma. *J Pathol* **207**, 102-110 (2005).
201. Westergren-Thorsson, G.*, et al.* Altered expression of small proteoglycans, collagen, and transforming growth factor-beta 1 in developing bleomycin-induced pulmonary fibrosis in rats. *J Clin Invest* **92**, 632-637 (1993).
202. Jarvelainen, H.*, et al.* A role for decorin in cutaneous wound healing and angiogenesis. *Wound Repair Regen* **14**, 443-452 (2006).
203. Schonherr, E.*, et al.* Decorin deficiency leads to impaired angiogenesis in injured mouse cornea. *J Vasc Res* **41**, 499-508 (2004).
204. Johnson, P.Y., Potter-Perigo, S., Gooden, M.D., Vernon, R.B. & Wight, T.N. Decorin synthesized by arterial smooth muscle cells is retained in fibrin gels and modulates fibrin contraction. *J Cell Biochem* **101**, 281-294 (2007).
205. Dugan, T.A., Yang, V.W., McQuillan, D.J. & Hook, M. Decorin modulates fibrin assembly and structure. *J Biol Chem* **281**, 38208-38216 (2006).
206. Dugan, T.A., Yang, V.W., McQuillan, D.J. & Hook, M. Decorin binds fibrinogen in a Zn2+-dependent interaction. *J Biol Chem* **278**, 13655-13662 (2003).
207. Huijun, W., Long, C., Zhigang, Z., Feng, J. & Muyi, G. Ex vivo transfer of the decorin gene into rat glomerulus via a mesangial cell vector suppressed extracellular matrix accumulation in experimental glomerulonephritis. *Exp Mol Pathol* **78**, 17-24 (2005).
208. Birk, D.E., Nurminskaya, M.V. & Zycband, E.I. Collagen fibrillogenesis in situ: fibril segments undergo post-depositional modifications resulting in linear and lateral growth during matrix development. *Dev Dyn* **202**, 229-243 (1995).
209. Ruhland, C.*, et al.* The glycosaminoglycan chain of decorin plays an important role in collagen fibril formation at the early stages of fibrillogenesis. *FEBS J* **274**, 4246-4255 (2007).
210. Hedbom, E. & Heinegard, D. Binding of fibromodulin and decorin to separate sites on fibrillar collagens. *J Biol Chem* **268**, 27307-27312 (1993).

211. Kvist, A.J., *et al.* Chondroitin sulfate perlecan enhances collagen fibril formation. Implications for perlecan chondrodysplasias. *J Biol Chem* **281**, 33127-33139 (2006).
212. Liu, X., Yeh, M.L., Lewis, J.L. & Luo, Z.P. Direct measurement of the rupture force of single pair of decorin interactions. *Biochem Biophys Res Commun* **338**, 1342-1345 (2005).
213. Geng, Y., McQuillan, D. & Roughley, P.J. SLRP interaction can protect collagen fibrils from cleavage by collagenases. *Matrix Biol* **25**, 484-491 (2006).
214. Keene, D.R., *et al.* Decorin binds near the C terminus of type I collagen. *J Biol Chem* **275**, 21801-21804 (2000).
215. Monfort, J., *et al.* Degradation of small leucine-rich repeat proteoglycans by matrix metalloprotease-13: identification of a new biglycan cleavage site. *Arthritis Res Ther* **8**, R26 (2006).
216. Imai, K., Hiramatsu, A., Fukushima, D., Pierschbacher, M.D. & Okada, Y. Degradation of decorin by matrix metalloproteinases: identification of the cleavage sites, kinetic analyses and transforming growth factor-beta1 release. *Biochem J* **322 (Pt 3)**, 809-814 (1997).
217. Seidler, D.G., *et al.* A physiologic three-dimensional cell culture system to investigate the role of decorin in matrix organisation and cell survival. *Biochem Biophys Res Commun* **332**, 1162-1170 (2005).
218. Rada, J.A., Cornuet, P.K. & Hassell, J.R. Regulation of corneal collagen fibrillogenesis in vitro by corneal proteoglycan (lumican and decorin) core proteins. *Exp Eye Res* **56**, 635-648 (1993).
219. Hong, S.H., *et al.* Analysis of estrogen-regulated genes in mouse uterus using cDNA microarray and laser capture microdissection. *J Endocrinol* **181**, 157-167 (2004).
220. Cabello-Verrugio, C. & Brandan, E. A novel modulatory mechanism of transforming growth factor-beta signaling through decorin and LRP-1. *J Biol Chem* **282**, 18842-18850 (2007).
221. Bates, S.J., *et al.* Mannose-6-phosphate, an inhibitor of transforming growth factor-beta, improves range of motion after flexor tendon repair. *J Bone Joint Surg Am* **88**, 2465-2472 (2006).
222. Droguett, R., Cabello-Verrugio, C., Riquelme, C. & Brandan, E. Extracellular proteoglycans modify TGF-beta bioavailability attenuating its signaling during skeletal muscle differentiation. *Matrix Biol* **25**, 332-341 (2006).
223. Li, Y., *et al.* Decorin gene transfer promotes muscle cell differentiation and muscle regeneration. *Mol Ther* **15**, 1616-1622 (2007).
224. Zhu, J., *et al.* Relationships between transforming growth factor-beta1, myostatin, and decorin: implications for skeletal muscle fibrosis. *J Biol Chem* **282**, 25852-25863 (2007).
225. Zanotti, S., *et al.* Decorin and biglycan expression is differentially altered in several muscular dystrophies. *Brain* **128**, 2546-2555 (2005).
226. Yamaguchi, Y., Mann, D.M. & Ruoslahti, E. Negative regulation of transforming growth factor-[beta] by the proteoglycan decorin. *Nature* **346**, 281-284 (1990).
227. Border, W.A., *et al.* Natural inhibitor of transforming growth factor-beta protects against scarring in experimental kidney disease. *Nature* **360**, 361-364 (1992).
228. Riquelme, C., *et al.* Antisense inhibition of decorin expression in myoblasts decreases cell responsiveness to transforming growth factor beta and accelerates skeletal muscle differentiation. *J Biol Chem* **276**, 3589-3596 (2001).
229. Takeuchi, Y., Kodama, Y. & Matsumoto, T. Bone matrix decorin binds transforming growth factor-beta and enhances its bioactivity. *J Biol Chem* **269**, 32634-32638 (1994).
230. Hausser, H., Groning, A., Hasilik, A., Schonherr, E. & Kresse, H. Selective inactivity of TGF-beta/decorin complexes. *FEBS Lett* **353**, 243-245 (1994).
231. Laederich, M.B., *et al.* The leucine-rich repeat protein LRIG1 is a negative regulator of ErbB family receptor tyrosine kinases. *J Biol Chem* **279**, 47050-47056 (2004).
232. Goldoni, S., *et al.* A soluble ectodomain of LRIG1 inhibits cancer cell growth by attenuating basal and ligand-dependent EGFR activity. *Oncogene* **26**, 368-381 (2007).
233. Plaas, A.H., Neame, P.J., Nivens, C.M. & Reiss, L. Identification of the keratan sulfate attachment sites on bovine fibromodulin. *J Biol Chem* **265**, 20634-20640 (1990).
234. Oldberg, A., Antonsson, P., Lindblom, K. & Heinegard, D. A collagen-binding 59-kd protein (fibromodulin) is structurally related to the small interstitial proteoglycans PG-S1 and PG-S2 (decorin). *EMBO J* **8**, 2601-2604 (1989).
235. Antonsson, P., Heinegard, D. & Oldberg, A. Posttranslational modifications of fibromodulin. *J Biol Chem* **266**, 16859-16861 (1991).
236. Onnerfjord, P., Heathfield, T.F. & Heinegard, D. Identification of tyrosine sulfation in extracellular leucine-rich repeat proteins using mass spectrometry. *J Biol Chem* **279**, 26-33 (2004).
237. Svensson, L., Narlid, I. & Oldberg, A. Fibromodulin and lumican bind to the same region on collagen type I fibrils. *FEBS Lett* **470**, 178-182 (2000).
238. Font, B., Eichenberger, D., Goldschmidt, D., Boutillon, M.M. & Hulmes, D.J. Structural requirements for fibromodulin binding to collagen and the control of type I collagen fibrillogenesis--critical roles for disulphide bonding and the C-terminal region. *Eur J Biochem* **254**, 580-587 (1998).
239. Hedlund, H., Mengarelli-Widholm, S., Heinegard, D., Reinholt, F.P. & Svensson, O. Fibromodulin distribution and association with collagen. *Matrix Biol* **14**, 227-232 (1994).
240. Viola, M., *et al.* Fibromodulin interactions with type I and II collagens. *Connect Tissue Res* **48**, 141-148 (2007).
241. Nurminskaya, M.V. & Birk, D.E. Differential expression of genes associated with collagen fibril growth in the chicken tendon: identification of structural and regulatory genes by subtractive hybridization. *Arch Biochem Biophys* **350**, 1-9 (1998).
242. Nurminskaya, M.V. & Birk, D.E. Differential expression of fibromodulin mRNA associated with tendon fibril growth: isolation and characterization of a chicken fibromodulin cDNA. *Biochem J* **317 (Pt 3)**, 785-789 (1996).
243. Burton-Wurster, N., *et al.* TGF beta 1 and biglycan, decorin, and fibromodulin metabolism in canine cartilage. *Osteoarthritis Cartilage* **11**, 167-176 (2003).
244. Dourado, G.S., Adams, M.E., Matyas, J.R. & Huang, D. Expression of biglycan, decorin and fibromodulin in the hypertrophic phase of experimental osteoarthritis. *Osteoarthritis Cartilage* **4**, 187-196 (1996).
245. Gill, M.R., Oldberg, A. & Reinholt, F.P. Fibromodulin-null murine knee joints display increased incidences of osteoarthritis and alterations in tissue biochemistry. *Osteoarthritis Cartilage* **10**, 751-757 (2002).
246. Ezura, Y., Chakravarti, S., Oldberg, A., Chervoneva, I. & Birk, D.E. Differential expression of lumican and fibromodulin regulate collagen fibrillogenesis in developing mouse tendons. *J Cell Biol* **151**, 779-788 (2000).
247. Jepsen, K.J., *et al.* A syndrome of joint laxity and impaired tendon integrity in lumican- and fibromodulin-deficient mice. *J Biol Chem* **277**, 35532-35540 (2002).

248. Gori, F., Schipani, E. & Demay, M.B. Fibromodulin is expressed by both chondrocytes and osteoblasts during fetal bone development. *J Cell Biochem* **82**, 46-57 (2001).
249. Baffi, M.O., Moran, M.A. & Serra, R. Tgfbr2 regulates the maintenance of boundaries in the axial skeleton. *Dev Biol* **296**, 363-374 (2006).
250. Saamanen, A.M., Salminen, H.J., Rantakokko, A.J., Heinegard, D. & Vuorio, E.I. Murine fibromodulin: cDNA and genomic structure, and age-related expression and distribution in the knee joint. *Biochem J* **355**, 577-585 (2001).
251. Alini, M. & Roughley, P.J. Changes in leucine-rich repeat proteoglycans during maturation of the bovine growth plate. *Matrix Biol* **19**, 805-813 (2001).
252. Roughley, P.J., White, R.J., Cs-Szabo, G. & Mort, J.S. Changes with age in the structure of fibromodulin in human articular cartilage. *Osteoarthritis Cartilage* **4**, 153-161 (1996).
253. Hoshi, K., Kemmotsu, S., Takeuchi, Y., Amizuka, N. & Ozawa, H. The primary calcification in bones follows removal of decorin and fusion of collagen fibrils. *J Bone Miner Res* **14**, 273-280 (1999).
254. Ly, D.H., Lockhart, D.J., Lerner, R.A. & Schultz, P.G. Mitotic misregulation and human aging. *Science* **287**, 2486-2492 (2000).
255. Faraonio, R., Pane, F., Intrieri, M., Russo, T. & Cimino, F. In vitro acquired cellular senescence and aging-specific phenotype can be distinguished on the basis of specific mRNA expression. *Cell Death Differ* **9**, 862-864 (2002).
256. Bevilacqua, M.A.*, et al.* Fibromodulin gene transcription is induced by ultraviolet irradiation, and its regulation is impaired in senescent human fibroblasts. *J Biol Chem* **280**, 31809-31817 (2005).
257. Ying, S.*, et al.* Characterization and expression of the mouse lumican gene. *J Biol Chem* **272**, 30306-30313 (1997).
258. Dunlevy, J.R., Neame, P.J., Vergnes, J.P. & Hassell, J.R. Identification of the N-linked oligosaccharide sites in chick corneal lumican and keratocan that receive keratan sulfate. *J Biol Chem* **273**, 9615-9621 (1998).
259. Chakravarti, S., Stallings, R.L., SundarRaj, N., Cornuet, P.K. & Hassell, J.R. Primary structure of human lumican (keratan sulfate proteoglycan) and localization of the gene (LUM) to chromosome 12q21.3-q22. *Genomics* **27**, 481-488 (1995).
260. Chakravarti, S.*, et al.* Lumican regulates collagen fibril assembly: skin fragility and corneal opacity in the absence of lumican. *J Cell Biol* **141**, 1277-1286 (1998).
261. Chakravarti, S., Zhang, G., Chervoneva, I., Roberts, L. & Birk, D.E. Collagen fibril assembly during postnatal development and dysfunctional regulation in the lumican-deficient murine cornea. *Dev Dyn* **235**, 2493-2506 (2006).
262. Beecher, N., Chakravarti, S., Joyce, S., Meek, K.M. & Quantock, A.J. Neonatal development of the corneal stroma in wild-type and lumican-null mice. *Invest Ophthalmol Vis Sci* **47**, 146-150 (2006).
263. Chakravarti, S.*, et al.* Corneal opacity in lumican-null mice: defects in collagen fibril structure and packing in the posterior stroma. *Invest Ophthalmol Vis Sci* **41**, 3365-3373 (2000).
264. Quantock, A.J., Meek, K.M. & Chakravarti, S. An x-ray diffraction investigation of corneal structure in lumican-deficient mice. *Invest Ophthalmol Vis Sci* **42**, 1750-1756 (2001).
265. Saika, S.*, et al.* TGFbeta2 in corneal morphogenesis during mouse embryonic development. *Dev Biol* **240**, 419-432 (2001).
266. Long, C.J.*, et al.* Fibroblast growth factor-2 promotes keratan sulfate proteoglycan expression by keratocytes in vitro. *J Biol Chem* **275**, 13918-13923 (2000).
267. Hayashida, Y.*, et al.* Matrix morphogenesis in cornea is mediated by the modification of keratan sulfate by GlcNAc 6-O-sulfotransferase. *Proc Natl Acad Sci U S A* **103**, 13333-13338 (2006).
268. Cornuet, P.K., Blochberger, T.C. & Hassell, J.R. Molecular polymorphism of lumican during corneal development. *Invest Ophthalmol Vis Sci* **35**, 870-877 (1994).
269. Song, J.*, et al.* Neonatal corneal stromal development in the normal and lumican-deficient mouse. *Invest Ophthalmol Vis Sci* **44**, 548-557 (2003).
270. Dunlevy, J.R. & Rada, J.A. Interaction of lumican with aggrecan in the aging human sclera. *Invest Ophthalmol Vis Sci* **45**, 3849-3856 (2004).
271. Austin, B.A., Coulon, C., Liu, C.Y., Kao, W.W. & Rada, J.A. Altered collagen fibril formation in the sclera of lumican-deficient mice. *Invest Ophthalmol Vis Sci* **43**, 1695-1701 (2002).
272. Chakravarti, S.*, et al.* Ocular and scleral alterations in gene-targeted lumican-fibromodulin double-null mice. *Invest Ophthalmol Vis Sci* **44**, 2422-2432 (2003).
273. Raouf, A.*, et al.* Lumican is a major proteoglycan component of the bone matrix. *Matrix Biol* **21**, 361-367 (2002).
274. Corpuz, L.M.*, et al.* Molecular cloning and tissue distribution of keratocan. Bovine corneal keratan sulfate proteoglycan 37A. *J Biol Chem* **271**, 9759-9763 (1996).
275. Tasheva, E.S., Funderburgh, J.L., Corpuz, L.M. & Conrad, G.W. Cloning, characterization and tissue-specific expression of the gene encoding bovine keratocan, a corneal keratan sulfate proteoglycan. *Gene* **218**, 63-68 (1998).
276. Dunlevy, J.R., Beales, M.P., Berryhill, B.L., Cornuet, P.K. & Hassell, J.R. Expression of the keratan sulfate proteoglycans lumican, keratocan and osteoglycin/mimecan during chick corneal development. *Exp Eye Res* **70**, 349-362 (2000).
277. Liu, C.Y., Birk, D.E., Hassell, J.R., Kane, B. & Kao, W.W. Keratocan-deficient mice display alterations in corneal structure. *J Biol Chem* **278**, 21672-21677 (2003).
278. Pellegata, N.S.*, et al.* Mutations in KERA, encoding keratocan, cause cornea plana. *Nat Genet* **25**, 91-95 (2000).
279. Carlson, E.C.*, et al.* Keratocan, a cornea-specific keratan sulfate proteoglycan, is regulated by lumican. *J Biol Chem* **280**, 25541-25547 (2005).
280. Funderburgh, J.L., Funderburgh, M.L., Mann, M.M., Corpuz, L. & Roth, M.R. Proteoglycan expression during transforming growth factor beta -induced keratocyte-myofibroblast transdifferentiation. *J Biol Chem* **276**, 44173-44178 (2001).
281. Carlson, E.C.*, et al.* Keratocan and lumican regulate neutrophil infiltration and corneal clarity in LPS-induced keratitis by direct interaction with KC/CXCL1. *J Biol Chem* (2007).
282. Sixt, M., Hallmann, R., Wendler, O., Scharffetter-Kochanek, K. & Sorokin, L.M. Cell adhesion and migration properties of beta 2-integrin negative polymorphonuclear granulocytes on defined extracellular matrix molecules. Relevance for leukocyte extravasation. *J Biol Chem* **276**, 18878-18887 (2001).
283. Koyama, H., Raines, E.W., Bornfeldt, K.E., Roberts, J.M. & Ross, R. Fibrillar collagen inhibits arterial smooth muscle proliferation through regulation of Cdk2 inhibitors. *Cell* **87**, 1069-1078 (1996).

284. Sommarin, Y., Wendel, M., Shen, Z., Hellman, U. & Heinegard, D. Osteoadherin, a cell-binding keratan sulfate proteoglycan in bone, belongs to the family of leucine-rich repeat proteins of the extracellular matrix. *J Biol Chem* **273**, 16723-16729 (1998).
285. Wendel, M., Sommarin, Y. & Heinegard, D. Bone matrix proteins: isolation and characterization of a novel cell-binding keratan sulfate proteoglycan (osteoadherin) from bovine bone. *J Cell Biol* **141**, 839-847 (1998).
286. Ramstad, V.E., Franzen, A., Heinegard, D., Wendel, M. & Reinholt, F.P. Ultrastructural distribution of osteoadherin in rat bone shows a pattern similar to that of bone sialoprotein. *Calcif Tissue Int* **72**, 57-64 (2003).
287. Lucchini, M.*, et al.* Alpha v beta 3 integrin expression in human odontoblasts and co-localization with osteoadherin. *J Dent Res* **83**, 552-556 (2004).
288. Petersson, U., Hultenby, K. & Wendel, M. Identification, distribution and expression of osteoadherin during tooth formation. *Eur J Oral Sci* **111**, 128-136 (2003).
289. Buchaille, R., Couble, M.L., Magloire, H. & Bleicher, F. Expression of the small leucine-rich proteoglycan osteoadherin/osteomodulin in human dental pulp and developing rat teeth. *Bone* **27**, 265-270 (2000).
290. Matheson, S., Larjava, H. & Hakkinen, L. Distinctive localization and function for lumican, fibromodulin and decorin to regulate collagen fibril organization in periodontal tissues. *J Periodontal Res* **40**, 312-324 (2005).
291. Rehn, A.P., Chalk, A.M. & Wendel, M. Differential regulation of osteoadherin (OSAD) by TGF-beta1 and BMP-2. *Biochem Biophys Res Commun* **349**, 1057-1064 (2006).
292. Bengtsson, E., Neame, P.J., Heinegard, D. & Sommarin, Y. The primary structure of a basic leucine-rich repeat protein, PRELP, found in connective tissues. *J Biol Chem* **270**, 25639-25644 (1995).
293. Grover, J., Chen, X.N., Korenberg, J.R., Recklies, A.D. & Roughley, P.J. The gene organization, chromosome location, and expression of a 55-kDa matrix protein (PRELP) of human articular cartilage. *Genomics* **38**, 109-117 (1996).
294. Bengtsson, E., Aspberg, A., Heinegard, D., Sommarin, Y. & Spillmann, D. The amino-terminal part of PRELP binds to heparin and heparan sulfate. *J Biol Chem* **275**, 40695-40702 (2000).
295. Bengtsson, E.*, et al.* The leucine-rich repeat protein PRELP binds perlecan and collagens and may function as a basement membrane anchor. *J Biol Chem* **277**, 15061-15068 (2002).
296. Johnson, J.M., Young, T.L. & Rada, J.A. Small leucine rich repeat proteoglycans (SLRPs) in the human sclera: identification of abundant levels of PRELP. *Mol Vis* **12**, 1057-1066 (2006).
297. Majava, M.*, et al.* Novel mutations in the small leucine-rich repeat protein/proteoglycan (SLRP) genes in high myopia. *Hum Mutat* **28**, 336-344 (2007).
298. Bishop, P.N. Structural macromolecules and supramolecular organisation of the vitreous gel. *Prog Retin Eye Res* **19**, 323-344 (2000).
299. Reardon, A.J.*, et al.* Identification in vitreous and molecular cloning of opticin, a novel member of the family of leucine-rich repeat proteins of the extracellular matrix. *J Biol Chem* **275**, 2123-2129 (2000).
300. Hobby, P.*, et al.* Cloning, modeling, and chromosomal localization for a small leucine-rich repeat proteoglycan (SLRP) family member expressed in human eye. *Mol Vis* **6**, 72-78 (2000).
301. Friedman, J.S.*, et al.* Protein localization in the human eye and genetic screen of opticin. *Hum Mol Genet* **11**, 1333-1342 (2002).
302. Schmut, O., Reich, M.E. & Zirm, M. [The proof of different types of collagen in the bovine eye (author's transl)]. *Albrecht Von Graefes Arch Klin Exp Ophthalmol* **196**, 71-77 (1975).
303. Bishop, P.N., Crossman, M.V., McLeod, D. & Ayad, S. Extraction and characterization of the tissue forms of collagen types II and IX from bovine vitreous. *Biochem J* **299 (Pt 2)**, 497-505 (1994).
304. Sebag, J. & Balazs, E.A. Morphology and ultrastructure of human vitreous fibers. *Invest Ophthalmol Vis Sci* **30**, 1867-1871 (1989).
305. Sanders, E.J., Walter, M.A., Parker, E., Aramburo, C. & Harvey, S. Opticin binds retinal growth hormone in the embryonic vitreous. *Invest Ophthalmol Vis Sci* **44**, 5404-5409 (2003).
306. Hindson, V.J., Gallagher, J.T., Halfter, W. & Bishop, P.N. Opticin binds to heparan and chondroitin sulfate proteoglycans. *Invest Ophthalmol Vis Sci* **46**, 4417-4423 (2005).
307. Acharya, M.*, et al.* Evaluation of the OPTC gene in primary open angle glaucoma: functional significance of a silent change. *BMC Mol Biol* **8**, 21 (2007).
308. Johnson, H.J.*, et al.* Characterization of epiphycan, a small proteoglycan with a leucine-rich repeat core protein. *J Biol Chem* **272**, 18709-18717 (1997).
309. Johnson, J.*, et al.* Expression and localization of PG-Lb/epiphycan during mouse development. *Dev Dyn* **216**, 499-510 (1999).
310. Zhou, W., Shirabe, K. & Kuwada, J.Y. Molecular cloning and expression of two small leucine-rich proteoglycan (SLRP) genes, dspg3l and optcl, in zebrafish. *Gene Expr Patterns* **6**, 482-488 (2006).
311. Funderburgh, J.L.*, et al.* Mimecan, the 25-kDa corneal keratan sulfate proteoglycan, is a product of the gene producing osteoglycin. *J Biol Chem* **272**, 28089-28095 (1997).
312. Tasheva, E.S., Funderburgh, M.L., McReynolds, J., Funderburgh, J.L. & Conrad, G.W. The bovine mimecan gene. Molecular cloning and characterization of two major RNA transcripts generated by alternative use of two splice acceptor sites in the third exon. *J Biol Chem* **274**, 18693-18701 (1999).
313. Ge, G.*, et al.* Bone morphogenetic protein-1/tolloid-related metalloproteinases process osteoglycin and enhance its ability to regulate collagen fibrillogenesis. *J Biol Chem* **279**, 41626-41633 (2004).
314. Tasheva, E.S.*, et al.* Mimecan/osteoglycin-deficient mice have collagen fibril abnormalities. *Mol Vis* **8**, 407-415 (2002).
315. Beecher, N.*, et al.* An x-ray diffraction study of corneal structure in mimecan-deficient mice. *Invest Ophthalmol Vis Sci* **46**, 4046-4049 (2005).
316. Tasheva, E.S. & Conrad, G.W. Interferon-gamma regulation of the human mimecan promoter. *Mol Vis* **9**, 277-287 (2003).
317. Tasheva, E.S., Klocke, B. & Conrad, G.W. Analysis of transcriptional regulation of the small leucine rich proteoglycans. *Mol Vis* **10**, 758-772 (2004).
318. Neame, P.J., Sommarin, Y., Boynton, R.E. & Heinegard, D. The structure of a 38-kDa leucine-rich protein (chondroadherin) isolated from bovine cartilage. *J Biol Chem* **269**, 21547-21554 (1994).
319. Grover, J., Chen, X.N., Korenberg, J.R. & Roughley, P.J. The structure and chromosome location of the human chondroadherin gene (CHAD). *Genomics* **45**, 379-385 (1997).

320. Shen, Z., Gantcheva, S., Mansson, B., Heinegard, D. & Sommarin, Y. Chondroadherin expression changes in skeletal development. *Biochem J* **330 (Pt 1)**, 549-557 (1998).
321. Camper, L., Heinegard, D. & Lundgren-Akerlund, E. Integrin alpha2beta1 is a receptor for the cartilage matrix protein chondroadherin. *J Cell Biol* **138**, 1159-1167 (1997).
322. Mansson, B., Wenglen, C., Morgelin, M., Saxne, T. & Heinegard, D. Association of chondroadherin with collagen type II. *J Biol Chem* **276**, 32883-32888 (2001).
323. Tasheva, E.S., Ke, A. & Conrad, G.W. Analysis of the expression of chondroadherin in mouse ocular and non-ocular tissues. *Mol Vis* **10**, 544-554 (2004).
324. Misof, K., Landis, W.J., Klaushofer, K. & Fratzl, P. Collagen from the osteogenesis imperfecta mouse model (oim) shows reduced resistance against tensile stress. *J Clin Invest* **100**, 40-45 (1997).
325. Sims, T.J., Miles, C.A., Bailey, A.J. & Camacho, N.P. Properties of collagen in OIM mouse tissues. *Connect Tissue Res* **44 Suppl 1**, 202-205 (2003).
326. Birk, D.E., Zycband, E.I., Woodruff, S., Winkelmann, D.A. & Trelstad, R.L. Collagen fibrillogenesis in situ: fibril segments become long fibrils as the developing tendon matures. *Dev Dyn* **208**, 291-298 (1997).
327. Waddington, R.J., Hall, R.C., Embery, G. & Lloyd, D.M. Changing profiles of proteoglycans in the transition of predentine to dentine. *Matrix Biol* **22**, 153-161 (2003).
328. Hall, R., Septier, D., Embery, G. & Goldberg, M. Stromelysin-1 (MMP-3) in forming enamel and predentine in rat incisor-coordinated distribution with proteoglycans suggests a functional role. *Histochem J* **31**, 761-770 (1999).
329. Fanchon, S.*, et al.* Involvement of matrix metalloproteinases in the onset of dentin mineralization. *Eur J Oral Sci* **112**, 171-176 (2004).
330. Bourd-Boittin, K.*, et al.* Matrix metalloproteinase inhibition impairs the processing, formation and mineralization of dental tissues during mouse molar development. *Exp Cell Res* **304**, 493-505 (2005).
331. Beniash, E., Traub, W., Veis, A. & Weiner, S. A transmission electron microscope study using vitrified ice sections of predentin: structural changes in the dentin collagenous matrix prior to mineralization. *J Struct Biol* **132**, 212-225 (2000).
332. Takeuchi, Y., Matsumoto, T., Ogata, E. & Shishiba, Y. Isolation and characterization of proteoglycans synthesized by mouse osteoblastic cells in culture during the mineralization process. *Biochem J* **266**, 15-24 (1990).
333. Boskey, A.L.*, et al.* Osteopontin-hydroxyapatite interactions in vitro: inhibition of hydroxyapatite formation and growth in a gelatin-gel. *Bone Miner* **22**, 147-159 (1993).
334. Boskey, A.L., Spevak, L., Doty, S.B. & Rosenberg, L. Effects of bone CS-proteoglycans, DS-decorin, and DS-biglycan on hydroxyapatite formation in a gelatin gel. *Calcif Tissue Int* **61**, 298-305 (1997).
335. Chen, C.C., Boskey, A.L. & Rosenberg, L.C. The inhibitory effect of cartilage proteoglycans on hydroxyapatite growth. *Calcif Tissue Int* **36**, 285-290 (1984).
336. Goldberg, M., Septier, D., Oldberg, A., Young, M.F. & Ameye, L.G. Fibromodulin-deficient mice display impaired collagen fibrillogenesis in predentin as well as altered dentin mineralization and enamel formation. *J Histochem Cytochem* **54**, 525-537 (2006).
337. Matias, M.A., Li, H., Young, W.G. & Bartold, P.M. Immunohistochemical localization of fibromodulin in the periodontium during cementogenesis and root formation in the rat molar. *J Periodontal Res* **38**, 502-507 (2003).
338. Watanabe, T. & Kubota, T. Characterization of fibromodulin isolated from bovine periodontal ligament. *J Periodontal Res* **33**, 1-7 (1998).
339. Cheng, H., Caterson, B., Neame, P.J., Lester, G.E. & Yamauchi, M. Differential distribution of lumican and fibromodulin in tooth cementum. *Connect Tissue Res* **34**, 87-96 (1996).
340. Hall, R.C., Embery, G. & Lloyd, D. Immunochemical localization of the small leucine-rich proteoglycan lumican in human predentine and dentine. *Arch Oral Biol* **42**, 783-786 (1997).
341. Chen, C.C. & Boskey, A.L. Mechanisms of proteoglycan inhibition of hydroxyapatite growth. *Calcif Tissue Int* **37**, 395-400 (1985).
342. Goldberg, M.*, et al.* Proteoglycans in predentin: the last 15 micrometers before mineralization. *Connect Tissue Res* **44 Suppl 1**, 184-188 (2003).
343. Goldberg, M.*, et al.* Targeted disruption of two small leucine-rich proteoglycans, biglycan and decorin, excerpts divergent effects on enamel and dentin formation. *Calcif Tissue Int* **77**, 297-310 (2005).
344. Milan, A.M., Sugars, R.V., Embery, G. & Waddington, R.J. Dentinal proteoglycans demonstrate an increasing order of affinity for hydroxyapatite crystals during the transition of predentine to dentine. *Calcif Tissue Int* **75**, 197-204 (2004).
345. Sugars, R.V.*, et al.* Molecular interaction of recombinant decorin and biglycan with type I collagen influences crystal growth. *Connect Tissue Res* **44 Suppl 1**, 189-195 (2003).
346. Xu, T.S.*, et al.* Targeted disruption of the biglycan gene leads to an osteoporosis-like phenotype in mice. *Nature Genetics* **20**, 78-82 (1998).
347. Orsini, G.*, et al.* Immunohistochemical identification of decorin and biglycan in human dentin: a correlative field emission scanning electron microscopy/transmission electron microscopy study. *Calcif Tissue Int* **81**, 39-45 (2007).
348. Mochida, Y., Duarte, W.R., Tanzawa, H., Paschalis, E.P. & Yamauchi, M. Decorin modulates matrix mineralization in vitro. *Biochem Biophys Res Commun* **305**, 6-9 (2003).
349. Dietz, H.C.*, et al.* Marfan syndrome caused by a recurrent de novo missense mutation in the fibrillin gene. *Nature* **352**, 337-339 (1991).
350. Habashi, J.P.*, et al.* Losartan, an AT1 antagonist, prevents aortic aneurysm in a mouse model of Marfan syndrome. *Science* **312**, 117-121 (2006).
351. Loeys, B.L.*, et al.* Aneurysm syndromes caused by mutations in the TGF-beta receptor. *N Engl J Med* **355**, 788-798 (2006).
352. Itabashi, T.*, et al.* Interaction between proteoglycans and alpha-elastin in construction of extracellular matrix of human yellow ligament. *Connect Tissue Res* **46**, 67-73 (2005).
353. Heegaard, A.M.*, et al.* Biglycan deficiency causes spontaneous aortic dissection and rupture in mice. *Circulation* **115**, 2731-2738 (2007).
354. Reinboth, B.J., Finnis, M.L., Gibson, M.A., Sandberg, L.B. & Cleary, E.G. Developmental expression of dermatan sulfate proteoglycans in the elastic bovine nuchal ligament. *Matrix Biol* **19**, 149-162 (2000).

355. Reinboth, B., Hanssen, E., Cleary, E.G. & Gibson, M.A. Molecular interactions of biglycan and decorin with elastic fiber components: biglycan forms a ternary complex with tropoelastin and microfibril-associated glycoprotein 1. J Biol Chem 277, 3950-3957 (2002).
356. Schaefer, L., et al. Regulation of fibrillin-1 by biglycan and decorin is important for tissue preservation in the kidney during pressure-induced injury. Am J Pathol 165, 383-396 (2004).
357. Hunzelmann, N., et al. Altered immunohistochemical expression of small proteoglycans in the tumor tissue and stroma of basal cell carcinoma. J Invest Dermatol 104, 509-513 (1995).
358. Mintz, M.B., et al. An expression signature classifies chemotherapy-resistant pediatric osteosarcoma. Cancer Res 65, 1748-1754 (2005).
359. Weber, C.K., et al. Biglycan is overexpressed in pancreatic cancer and induces G1-arrest in pancreatic cancer cell lines. Gastroenterology 121, 657-667 (2001).
360. Chopra, A., Maitra, B. & Korman, N.J. Decreased mRNA expression of several basement membrane components in basal cell carcinoma. J Invest Dermatol 110, 52-56 (1998).
361. Leygue, E., et al. Lumican and decorin are differentially expressed in human breast carcinoma. J Pathol 192, 313-320 (2000).
362. Eshchenko, T.Y., Rykova, V.I., Chernakov, A.E., Sidorov, S.V. & Grigorieva, E.V. Expression of different proteoglycans in human breast tumors. Biochemistry (Mosc) 72, 1016-1020 (2007).
363. Seya, T., et al. Lumican expression in advanced colorectal cancer with nodal metastasis correlates with poor prognosis. Oncol Rep 16, 1225-1230 (2006).
364. Leygue, E., et al. Expression of lumican in human breast carcinoma. Cancer Res 58, 1348-1352 (1998).
365. Troup, S., et al. Reduced expression of the small leucine-rich proteoglycans, lumican, and decorin is associated with poor outcome in node-negative invasive breast cancer. Clin Cancer Res 9, 207-214 (2003).
366. Ishiwata, T., et al. Role of lumican in cancer cells and adjacent stromal tissues in human pancreatic cancer. Oncol Rep 18, 537-543 (2007).
367. Ping Lu, Y., Ishiwata, T. & Asano, G. Lumican expression in alpha cells of islets in pancreas and pancreatic cancer cells. J Pathol 196, 324-330 (2002).
368. Shinji, S., et al. Different expression levels of lumican in human carcinoid tumor and neuroendocrine cell carcinoma. Int J Oncol 26, 873-880 (2005).
369. Brezillon, S., et al. Expression of lumican, a small leucine-rich proteoglycan with antitumour activity, in human malignant melanoma. Clin Exp Dermatol 32, 405-416 (2007).
370. Seidler, D.G., et al. Decorin protein core inhibits in vivo cancer growth and metabolism by hindering epidermal growth factor receptor function and triggering apoptosis via caspase-3 activation. J Biol Chem 281, 26408-26418 (2006).
371. Reed, C.C., et al. Decorin prevents metastatic spreading of breast cancer. Oncogene 24, 1104-1110 (2005).
372. Genersch, E., Schuppan, D. & Lichtner, R.B. Signaling by epidermal growth factor differentially affects integrin-mediated adhesion of tumor cells to extracellular matrix proteins. J Mol Med 74, 609-616 (1996).
373. Zhu, J.X., et al. Decorin evokes protracted internalization and degradation of the epidermal growth factor receptor via caveolar endocytosis. J Biol Chem 280, 32468-32479 (2005).
374. Koninger, J., et al. Overexpressed decorin in pancreatic cancer: potential tumor growth inhibition and attenuation of chemotherapeutic action. Clin Cancer Res 10, 4776-4783 (2004).
375. Theocharis, A.D., Vynios, D.H., Papageorgakopoulou, N., Skandalis, S.S. & Theocharis, D.A. Altered content composition and structure of glycosaminoglycans and proteoglycans in gastric carcinoma. Int J Biochem Cell Biol 35, 376-390 (2003).
376. Tsara, M.E., Theocharis, A.D. & Theocharis, D.A. Compositional and structural alterations of proteoglycans in human rectum carcinoma with special reference to versican and decorin. Anticancer Res 22, 2893-2898 (2002).
377. Tralhao, J.G., et al. In vivo selective and distant killing of cancer cells using adenovirus-mediated decorin gene transfer. FASEB J 17, 464-466 (2003).
378. Clark, E.A., Golub, T.R., Lander, E.S. & Hynes, R.O. Genomic analysis of metastasis reveals an essential role for RhoC. Nature 406, 532-535 (2000).
379. Mikaelsson, E., et al. Fibromodulin, an extracellular matrix protein: characterization of its unique gene and protein expression in B-cell chronic lymphocytic leukemia and mantle cell lymphoma. Blood 105, 4828-4835 (2005).
380. Mayr, C., et al. Fibromodulin as a novel tumor-associated antigen (TAA) in chronic lymphocytic leukemia (CLL), which allows expansion of specific CD8+ autologous T lymphocytes. Blood 105, 1566-1573 (2005).
381. Levens, E., Luo, X., Ding, L., Williams, R.S. & Chegini, N. Fibromodulin is expressed in leiomyoma and myometrium and regulated by gonadotropin-releasing hormone analogue therapy and TGF-beta through Smad and MAPK-mediated signalling. Mol Hum Reprod 11, 489-494 (2005).
382. Figueroa, J.E. & Vijayagopal, P. Angiotensin II stimulates synthesis of vascular smooth muscle cell proteoglycans with enhanced low density lipoprotein binding properties. Atherosclerosis 162, 261-268 (2002).
383. Hashimura, K., et al. Androgens stimulate human vascular smooth muscle cell proteoglycan biosynthesis and increase lipoprotein binding. Endocrinology 146, 2085-2090 (2005).
384. Theocharis, A.D. & Karamanos, N.K. Decreased biglycan expression and differential decorin localization in human abdominal aortic aneurysms. Atherosclerosis 165, 221-230 (2002).
385. Mao, D., Lee, J.K., VanVickle, S.J. & Thompson, R.W. Expression of collagenase-3 (MMP-13) in human abdominal aortic aneurysms and vascular smooth muscle cells in culture. Biochem Biophys Res Commun 261, 904-910 (1999).
386. Riessen, R., et al. Regional differences in the distribution of the proteoglycans biglycan and decorin in the extracellular matrix of atherosclerotic and restenotic human coronary arteries. Am J Pathol 144, 962-974 (1994).
387. Nakashima, Y., Fujii, H., Sumiyoshi, S., Wight, T.N. & Sueishi, K. Early human atherosclerosis: accumulation of lipid and proteoglycans in intimal thickenings followed by macrophage infiltration. Arterioscler Thromb Vasc Biol 27, 1159-1165 (2007).
388. O'Brien, K.D., et al. Comparison of apolipoprotein and proteoglycan deposits in human coronary atherosclerotic plaques: colocalization of biglycan with apolipoproteins. Circulation 98, 519-527 (1998).
389. Klezovitch, O. & Scanu, A.M. Domains of apolipoprotein E involved in the binding to the protein core of biglycan of the vascular extracellular matrix: potential relationship between retention and anti-atherogenic properties of this apolipoprotein. Trends Cardiovasc Med 11, 263-268 (2001).
390. Olin, K.L., Potter-Perigo, S., Barrett, P.H., Wight, T.N. & Chait, A. Biglycan, a vascular proteoglycan, binds differently to HDL2 and HDL3: role of apoE. Arterioscler Thromb Vasc Biol 21, 129-135 (2001).

391. Shimizu-Hirota, R., et al. Extracellular matrix glycoprotein biglycan enhances vascular smooth muscle cell proliferation and migration. Circ Res 94, 1067-1074 (2004).
392. Al Haj Zen, A., et al. Decorin overexpression reduces atherosclerosis development in apolipoprotein E-deficient mice. Atherosclerosis 187, 31-39 (2006).
393. Strom, A., Ahlqvist, E., Franzen, A., Heinegard, D. & Hultgardh-Nilsson, A. Extracellular matrix components in atherosclerotic arteries of Apo E/LDL receptor deficient mice: an immunohistochemical study. Histol Histopathol 19, 337-347 (2004).
394. Onda, M., et al. Expression of lumican in thickened intima and smooth muscle cells in human coronary atherosclerosis. Exp Mol Pathol 72, 142-149 (2002).
395. Heathfield, T.F., Onnerfjord, P., Dahlberg, L. & Heinegard, D. 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 (2004).
396. Sjoberg, A., Onnerfjord, P., Morgelin, M., Heinegard, D. & Blom, A.M. The extracellular matrix and inflammation: fibromodulin activates the classical pathway of complement by directly binding C1q. J Biol Chem 280, 32301-32308 (2005).
397. Groeneveld, T.W., et al. Interactions of the extracellular matrix proteoglycans decorin and biglycan with C1q and collectins. J Immunol 175, 4715-4723 (2005).
398. Krumdieck, R., Hook, M., Rosenberg, L.C. & Volanakis, J.E. The proteoglycan decorin binds C1q and inhibits the activity of the C1 complex. J Immunol 149, 3695-3701 (1992).
399. Schaefer, L., et al. Absence of decorin adversely influences tubulointerstitial fibrosis of the obstructed kidney by enhanced apoptosis and increased inflammatory reaction. Am J Pathol 160, 1181-1191 (2002).
400. Schaefer, L., et al. The matrix component biglycan is proinflammatory and signals through Toll-like receptors 4 and 2 in macrophages. J Clin Invest 115, 2223-2233 (2005).
401. Wu, F., et al. A novel role of the lumican core protein in bacterial lipopolysaccharide-induced innate immune response. J Biol Chem 282, 26409-26417 (2007).
402. Vij, N., Roberts, L., Joyce, S. & Chakravarti, S. Lumican suppresses cell proliferation and aids Fas-Fas ligand mediated apoptosis: implications in the cornea. Exp Eye Res 78, 957-971 (2004).
403. Vij, N., Roberts, L., Joyce, S. & Chakravarti, S. Lumican regulates corneal inflammatory responses by modulating Fas-Fas ligand signaling. Invest Ophthalmol Vis Sci 46, 88-95 (2005).
404. Pancer, Z., et al. Somatic diversification of variable lymphocyte receptors in the agnathan sea lamprey. Nature 430, 174-180 (2004).
405. Venkatesan, N., Ebihara, T., Roughley, P.J. & Ludwig, M.S. Alterations in large and small proteoglycans in bleomycin-induced pulmonary fibrosis in rats. Am J Respir Crit Care Med 161, 2066-2073 (2000).
406. Ebihara, T., Venkatesan, N., Tanaka, R. & Ludwig, M.S. Changes in extracellular matrix and tissue viscoelasticity in bleomycin-induced lung fibrosis. Temporal aspects. Am J Respir Crit Care Med 162, 1569-1576 (2000).
407. Pini, L., et al. Differences in proteoglycan deposition in the airways of moderate and severe asthmatics. Eur Respir J 29, 71-77 (2007).
408. Dolhnikoff, M., Morin, J., Roughley, P.J. & Ludwig, M.S. Expression of lumican in human lungs. Am J Respir Cell Mol Biol 19, 582-587 (1998).
409. Provenzano, P.P., Alejandro-Osorio, A.L., Valhmu, W.B., Jensen, K.T. & Vanderby, R., Jr. Intrinsic fibroblast-mediated remodeling of damaged collagenous matrices in vivo. Matrix Biol 23, 543-555 (2005).
410. Hakkinen, L., Westermarck, J., Kahari, V.M. & Larjava, H. Human granulation-tissue fibroblasts show enhanced proteoglycan gene expression and altered response to TGF-beta 1. J Dent Res 75, 1767-1778 (1996).
411. Soo, C., et al. Differential expression of fibromodulin, a transforming growth factor-beta modulator, in fetal skin development and scarless repair. Am J Pathol 157, 423-433 (2000).
412. Stoff, A., et al. Effect of adenoviral mediated overexpression of fibromodulin on human dermal fibroblasts and scar formation in full-thickness incisional wounds. J Mol Med 85, 481-496 (2007).
413. Schonherr, E., Hausser, H., Beavan, L. & Kresse, H. Decorin-type I collagen interaction. Presence of separate core protein-binding domains. J Biol Chem 270, 8877-8883 (1995).
414. Svensson, L., Heinegard, D. & Oldberg, A. Decorin-binding sites for collagen type I are mainly located in leucine-rich repeats 4-5. J Biol Chem 270, 20712-20716 (1995).
415. Kresse, H., Liszio, C., Schonherr, E. & Fisher, L.W. Critical role of glutamate in a central leucine-rich repeat of decorin for interaction with type I collagen. J Biol Chem 272, 18404-18410 (1997).
416. Nareyeck, G., et al. Differential interactions of decorin and decorin mutants with type I and type VI collagens. Eur J Biochem 271, 3389-3398 (2004).
417. Kalamajski, S., Aspberg, A. & Oldberg, A. The decorin sequence SYIRIADTNIT binds collagen type I. J Biol Chem 282, 16062-16067 (2007).
418. Cronlund, A.L., Smith, B.D. & Kagan, H.M. Binding of lysyl oxidase to fibrils of type I collagen. Connect Tissue Res 14, 109-119 (1985).
419. Mueller, M.M. & Fusenig, N.E. Friends or foes - bipolar effects of the tumour stroma in cancer. Nat Rev Cancer 4, 839-849 (2004).
420. Heldin, C.H., Rubin, K., Pietras, K. & Ostman, A. High interstitial fluid pressure - an obstacle in cancer therapy. Nat Rev Cancer 4, 806-813 (2004).

APPENDICES (PAPERS I-VI)