Differential MMP-14 Targeting by Biglycan, Decorin, Fibromodulin and Lumican Unraveled by In Silico Approach

Romain Rivet, Rajas Mallenahalli Rao, Pierre Nizet, Nicolas Belloy, Louise Huber, Manuel Dauchez, Laurent Ramont, Stéphanie Baud, Stéphane Brézillon

american physiological society
AMERICAN JOURNAL OF PHYSIOLOGY
CELL PHYSIOLOGY

Am J Physiol Cell Physiol 324: C353–C365, 2023.
First published December 19, 2022; doi:10.1152/ajpcell.00429.2022

RESEARCH ARTICLE

Deciphering the Role of Proteoglycans and Glycosaminoglycans in Health and Disease

Differential MMP-14 targeting by biglycan, decorin, fibromodulin, and lumican unraveled by in silico approach

Romain Rivet,[1] Rajas Mallenahalli Rao,[1,2] Pierre Nizet,[1] Nicolas Belloy,[1,2] Louise Huber,[1] Manuel Dauchez,[1,2] Laurent Ramont,[1,3] Stéphanie Baud,[1,2] and Stéphane Brézillon[1]

[1]*CNRS UMR 7369, Matrice Extracellulaire et Dynamique Cellulaire (MEDyC), Université de Reims Champagne Ardenne, Reims, France;* [2]*P3M, Multi-Scale-Molecular Modeling Platform, Université de Reims Champagne Ardenne, Reims, France; and* [3]*CHU Reims, Service Biochimie Pharmacologie-Toxicologie, Reims, France*

Abstract

Small leucine-rich proteoglycans (SLRPs) are major regulators of extracellular matrix assembly and cell signaling. Lumican, a member of the SLRPs family, and its derived peptides were shown to possess antitumor activity by interacting directly with the catalytic domain of MMP-14 leading to the inhibition of its activity. The aim of the present report was to characterize by in silico three-dimensional (3D) modeling the structure and the dynamics of four SLRPs including their core protein and their specific polysaccharide chains to assess their capacity to bind to MMP-14 and to regulate its activity. Molecular docking experiments were performed to identify the specific amino acids of MMP-14 interacting with each of the four SLRPs. The inhibition of each SLRP (100 nM) on MMP-14 activity was measured and the constants of inhibition (K_i) were evaluated. The impact of the number of glycan chains, structures, and dynamics of lumican on the interaction with MMP-14 was assessed by molecular dynamics simulations. Molecular docking analysis showed that all SLRPs bind to MMP-14 through their concave face, but in different regions of the catalytic domain of MMP-14. Each SLRPs inhibited significantly the MMP-14 activity. Finally, molecular dynamics showed the role of glycan chains in interaction with MMP-14 and shielding effect of SLRPs. Altogether, the results demonstrated that each SLRP exhibited inhibition of MMP-14 activity. However, the differential targeting of MMP-14 by the SLRPs was shown to be related not only to the core protein conformation but also to the glycan chain structures and dynamics.

MMP-14; small leucine-rich proteoglycans (SLRPs); glycosylation; structure and molecular modeling

INTRODUCTION

Among the many constituents of the extracellular matrix (ECM), small leucine-rich proteoglycans (SLRPs) are a family of 18 proteoglycans that are important regulators of ECM assembly and cell signaling. They have been implicated in the regulation of cancer growth and progression.

The 18 members of the SLRP family are classified into five classes based on such criteria as conservation and homology at the protein and genomic levels, the presence of characteristic N-terminal cysteine-rich clusters, and chromosomal organization (1–3). Regardless of the classification used, however, some SLRPs share common functionality (4).

From the structural point of view, SLRPs are proteoglycans sharing multiple leucine-rich repeats (LRRs). A LRR domain consists of tandem repeats rich in leucine and other small hydrophobic residues and its length varies from 20 to 39 amino acids depending on the SLRP; LRR's characteristic pattern of 11 amino acids is LxxLxLxxNxL [where x is any amino acid) and the consensus leucine can be substituted by isoleucine, valine, or less often by other hydrophobic amino acids (1)]. Since the resolution of the crystal structure of pancreatic ribonuclease inhibitor (RNI; 5), which was the first X-ray structure obtained for a leucine-rich protein, SLRP three-dimensional (3D) structures have been solved using X-ray, notably the ones of biglycan (BGN; 6), decorin (DCN; 7), and fibromodulin (FMOD; 8). These crystallographic structures display an arch shape solenoid-like morphology, where each turn of the solenoid corresponds to a single LRR. The LRRs are composed of elements of secondary structures such as α-helix, polyproline type II helix, 3–10 helix, β-turn, and β-sheets.

From a functional point of view, the primary role of SLRPs is the assembly as well as the maintenance of the structure of the ECM and more particularly, they are involved in the process of fibrillogenesis of collagens (9). Thus, in various connective tissues, lumican (LUM) has been shown to play a crucial role in regulating the assembly of collagen into fibrils

*R. Rivet and R. M. Rao contributed equally to this work. S. Baud and S. Brézillon contributed equally to this work.
Correspondence: S. Brézillon (stephane.brezillon@univ-reims.fr).
Submitted 23 September 2022 / Revised 9 December 2022 / Accepted 9 December 2022

(10). Fibromodulin is also involved in this same process by binding to collagen (11). In addition to their impact in maintaining the structure of the ECM, other roles have been identified and associated with SLRPs, particularly in the context of tumor progression (3, 12). Finally, it has been demonstrated that certain SLRPs such as decorin and lumican regulate the activity of matrix metalloproteinase (MMP)-14 and belong to its numerous substrates (13–16).

MMP-14 was the first characterized membrane-type matrix metalloproteinase. It plays a significant role in cell migration, not only by regulating the activity or expression of downstream MMPs but also by processing and activating migration-associated molecules. During migration and invasion, MMP-14 localizes at lamellipodia, the migration front of the cells (17). This localization is achieved by interaction of MMP-14 with CD44. It was also shown that MMP-14 can be colocalized with β1 integrin (18). Studies have reported that MMP-14 modulates pro-MMP-2 activation through the formation of MMP-14 dimer/tissue inhibitor of metalloproteinase-2 (TIMP-2) and proMMP-2 complex (19). The inhibition of the active form presented at the cell surface is one of the critical steps to regulate its activity. MMP-14 is inhibited by endogenous inhibitors TIMP-2, -3, and 4, but not by TIMP-1 (20).

Like decorin and lumican, biglycan and fibromodulin are involved in cancer-like pathologies (21–24). On the other hand, to date, there is no data to discuss a potential regulatory role of BGN and FMOD with respect to MMP-14. However, given the high degree of similarity between BGN, DCN, FMOD, and LUM core protein amino acid sequences, and their role in collagen fibrillogenesis, it would seem consistent to predict and investigate through combined in silico and biochemical approaches the potential interaction between MMP-14 and biglycan, and between MMP-14 and fibromodulin and thus potential regulation pathway of MMP-14 activity through these ones.

In the present work, the comparative study of the interaction of the catalytic domain of MMP-14 with BGN, DCN, FMOD, and LUM SLRPs is presented. From the 3D structures of the core proteins (extracted from the PDB database or constructed by homology modeling), it was first possible to consider the structural and hindrance impact of N-glycosylations on the structure and accessibility of SLRPs. Then, rigid molecular docking experiments allowed the comparison of the different MMP-14/SLRP complexes and could be correlated with biochemical tests giving access to the inhibition of each SLRP on the MMP-14 activity as well as inhibition constants. Finally, for the MMP-14/LUM complex, shielding effects of N-glycosylations on the dynamics of the molecular system were thoroughly evaluated using molecular dynamics (MD) simulations.

METHODS

Design of the In Silico Protocol

Among the four SLRPs our study focuses on, only one got its 3D structure solved (human fibromodulin). For the three other ones, a 3D structure was built using homology modeling and the Modeler software. The PDB code, name, and length of the chains of the templates that were used are specified in Table 1, as well as the percentage of identity between the whole sequences of the core protein of the SLRPs in amino acids or the resolved structures.

Rigid Protein-Protein Docking Experiments and Their Analyses

The Hex software (26) was used to perform the docking experiments. Hex was submitted to the Critical Assessment of PRediction of Interactions (CAPRI) test, which is used to validate a docking method on known/characterized molecular complexes (27). The configurable parameters of Hex were tuned based on the redocking experiment: TIMP-2 was docked on MMP-14 and the solutions were compared with the crystallographic reference [PDB code = 1BQQ (28)] through the evaluation of the root mean square deviation (RMSD) between the crystallographic structure and the best solution provided by the docking experiments. All docking results presented in this paper were obtained with the surface correlation method, the Optimized Potentials for Liquid Simulations (OPLS) electrostatic potential, and the (bumps + OPLS) postprocessing protocol.

From the Hex program, the 100 best solutions (ranked according to an energy criterion) were collected. The interactions between MMP-14 and the different human SLRPs were studied with the help of the CONTACT program [which is part of the CCP4 suite (29)]. When a distance between two atoms of residues belonging to different proteins was lower than 4 Å, then a contact was identified. Focusing either on the MMP-14 or on the SLRP moieties of the formed complexes, normalized contact frequencies could be evaluated along the sequences of the considered partners.

Molecular Dynamics Simulations of Glycosylated LUM

The constructed model was glycosylated with polylactosamine repeats using doglycans program (30), to the residues

Table 1. *Structural characteristics of the four investigated human SLRPs*

Human SLRP	UniProtKB AC	Length (Chain)	PDB ID Associated to the SLRP	% of Identity (Template)	PDB ID Associated to the Protein Template
Biglycan	P21810	368 (38–368)		94.58 (biglycan bos taurus)	2FT3
Decorin	P07585	359 (31–359)		89.44 (decorin bos taurus)	1XKU
Fibromodulin	Q06828	376 (19–376)	5MX0		
Lumican	P51884	338 (19–338)		39.58 (fibromodulin human)	5MX0

Structural data information for the four small leucine-rich proteoglycans (SLRPs) of interest along with their associated 3D structures can be retrieved from the UniProtKB/Swiss-Prot database (25). Accession numbers are given in the second column. The third column gives the number of amino acids contained in the protein and specifies the positions of the main chain (without the signal peptide) in parentheses. The PDB identifiers of the experimental or template structure used for homology reconstruction are specified in the fourth and sixth columns, respectively. When the SLRP structure has been obtained using homology reconstruction, the percentage of identity between the amino acid sequence of the protein and the potential resolved template is specified in the fifth column.

N88, N127, N160, and N252. The number of repeats were set to ensure that the net molecular weight of the glycoprotein is 48 kDa, as observed from experimental studies. Molecular weight was measured and orientation of glycans adjusted to remove stereochemical clashes using UCSF Chimera program (31). The protein was described using AMBER99SBILDN force field (32) and the carbohydrates were described using GLYCAM_06h force field (33). The glycosylated structure was solvated with TIP3P water molecules (34) in a cubical box of $15 \times 15 \times 15$ nm^3 dimensions, and four neutralizing Na^+ counterions were added. The system was energy minimized for 100,000 steps using the steepest-descent method. Upon minimization, initial equilibration was performed for 4 ns with 0.002 ps time-step, where the system was heated to 310 K. Further equilibration was performed at 310 K temperature and 1 atm pressure in five stages, with each stage of 2 ns with 0.002 ps time-step. During each step, position restraints were gradually reduced from 1,000 kJ/(mol·nm^2), 500 kJ/(mol·nm^2), 100 kJ/(mol·nm^2), 10 kJ/(mol·nm^2), and 0 kJ/(mol·nm^2), with the objective of gradual relaxation of the system. The temperature and pressure during equilibration were maintained using V-rescale thermostat and Parinello–Rahman barostat, respectively. Following equilibration, production simulation was performed for 100 ns with 0.002 ps time-step, with the same parameters as mentioned earlier (without the positional restraints). Three more replicates of 100 ns each were performed by randomizing the initial velocity seed during equilibration, bringing the net sampling to 400 ns. All the aforementioned MD protocols were performed using GROMACS 2020.4 program (35, 36).

Reagents

Recombinant human biglycan (2667-CM-050), decorin (143-DE-100), fibromodulin (9840-FM-050), and recombinant human furin protein (1503-SE) were supplied by R&D Systems (Minneapolis, MN). Human MMP-14 recombinant protein (RP-77533) was provided by Invitrogen (Waltham, MA). GM6001 MMP Inhibitor (Ilomastat, CC1000) and pro-MMP2 (PF037-10UG) were supplied by Merck (Darmstadt, Germany). APMA (4-aminophenylmercuric acetate) was furnished by Anaspec (Fremont, CA). Recombinant human lumican (57 kDa) was produced as previously described (37).

Proteolytic Activities of Matrix Metalloproteinases

Before all activity assays, 1 μL of proMMP-14 (RP-77533, Invitrogen) was activated using 1 μL of furin convertase (R&D Systems) for 30 min at 37°C in 8 μL of buffer [Tris 50 mM, $CaCl_2$ 3 mM, Brij-35 0.5% (vol/vol), pH 9]. To determine the direct effect of biglycan, decorin, fibromodulin, and lumican (37) on MMP-14 activity in vitro, the activated MMP-14 (8.625 nM) was preincubated in the assay buffer (Tris 50 mM, $CaCl_2$ 3 mM, $ZnCl_2$ 1 μM, pH 8.5) at 37°C overnight before assay with increasing concentrations (2.5, 25, and 250 nM) of each recombinant SLRP compared with negative control (without SLRP) and positive control (in the presence of 62.5 μM Ilomastat, Merck). The MMP-14 activity was measured in 96-well plates using a 2.5 μM substrate: 520 MMP FRET Substrate 15 (AS-60582-01, Eurogentec, Seraing, Belgium) and 3.45 nM of MMP-14. Increasing concentrations (1, 10, and 100 nM) of SLRPs were tested whereas the concentration of Ilomastat remained constant (25 μM). Fluorescence was measured in the FAM channel with a thermocycler (CFX96, BioRad, Hercules, CA) every minute for 30 min at 37°C. The assays were carried out in duplicate. Results were obtained from four independent experiments. The K_i were calculated by fitting the MMP-14 activity to the Morrison's equation.

Similarly, the MMP-2 activity was measured in 96-well plates using human MMP-2 recombinant protein. Only the activation differs, it was obtained by mixing 1 μL of proMMP-2 (6.68 μM) and 1 μL of APMA (10 mM) in 8 μL of buffer [Tris 50 mM, $CaCl_2$ 3 mM, Brij-35 0.5% (vol/vol), pH 9] and incubated for 60 min at 37°C. The activated MMP-2 (8.35 nM) was preincubated in the assay buffer (Tris 50 mM, $CaCl_2$ 3 mM, $ZnCl_2$ 1 μM, pH 8.5) at 37°C overnight before assay. The MMP-2 activity was measured using 2.5 μM of substrate (520 MMP FRET Substrate 15, Eurogentec) and 3.34 nM of MMP-2.

RESULTS

SLRP Core Proteins: From Sequences to Structural Aspects and Posttranslational Modifications

Starting from the information collected on the UniProt server and considering the primary sequences of the four human SLRPs investigated in this study, a comparison between the global organization of the primary sequence was carried out and made it possible to illustrate the general arrangement of the leucine-rich repeats (LRRs; Fig. 1*A*). The comparison was facilitated by schematically aligning the position of LRR1 for the four SLRPs. The positions (evidenced or putative) of O-glycosylations and N-glycosylations were also reported in the scheme.

Representations displayed in Fig. 1, *B* and *C* allow to bring focus on the sequence-structure-function triad. Whereas Fig. 1*B* allows comparison between sequence alignment and local secondary structure alignment, Fig. 1*C* highlights the great similarity in 3D shape and organization of the four considered human SLRPs. Despite the fact that sequence alignment and overall LLRs alignment between class-I and class-II SLRPs are not fully correlated, the striking conclusion arising from Fig. 1*B* is that sequence alignment also drives a very good local secondary structure alignment; most of the β-strands (blue arrows) are aligned and once more, a very good correspondence is evidenced within a given SLRP-class: on three locations within their N-terminal region and one location within their C-terminal region BGN and DCN align strands whereas no such local secondary structure corresponds in FMOD and LUM. From the structural alignment of the four proteins (Fig. 1*C*), it can be observed that the β-strands forming the β-sheets situated on the concave face of the SLRPs are the driving structural elements in this alignment process.

Glimpse at the Glycoprotein Structures

The glycoproteins presented in Fig. 2 are modeled considering the inner core protein and by grafting on their surface the classical biantennary structure such as the one described by Guillot et al. (38). In the case of biglycan and decorin, the number of N-glycosylations is equal to three

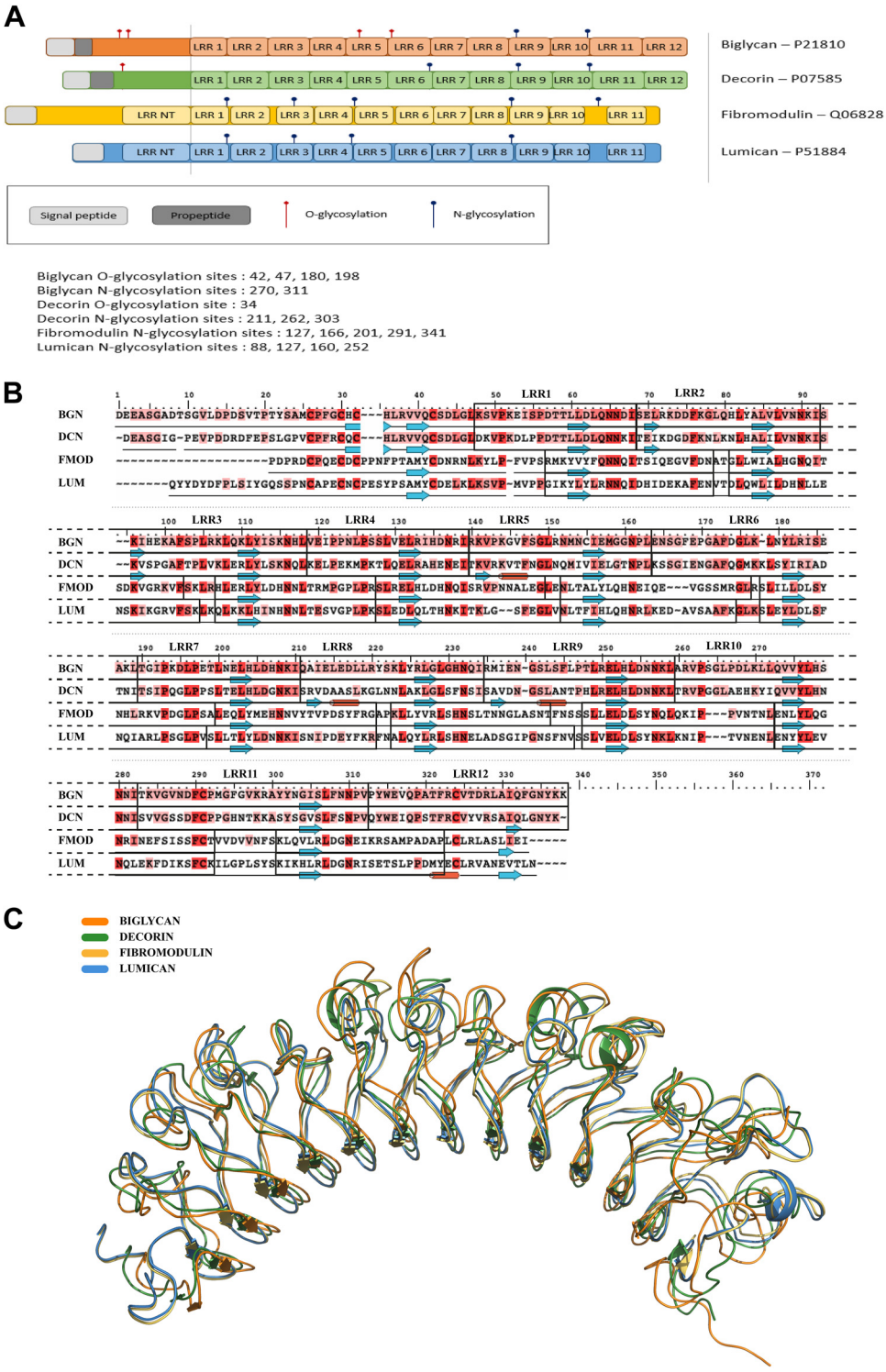


Figure 1. Comparisons of the human biglycan (BGN), decorin (DCN), fibromodulin (FMOD), and lumican (LUM) core protein structures and posttranslational modification (PTM) positions. *A*: schematic comparison of the leucine-rich repeat (LRR) sequences of BGN (orange), DCN (green), FMOD (yellow), and LUM (blue) from LRR1 to LRR12 and positions of their O- and N-glycosylation sites. The locations of the LRR and glycosylation sites were extracted from the UniProt server using the sequence references specified on the right. Signal peptide and propeptide are depicted in light gray and dark gray, respectively. *B*: dual presentation of the sequence alignment and the local secondary structure alignment. Sequence conservation is highlighted by colored letters ranging from pink (identity for two out of four sequences) to dark red (identity for all four sequences). Elements of the local secondary structure are depicted using blue arrows (β-sheets) and red cylinders (α-helices). Each LRR position is indicated with rectangular boxes. *C*: structural alignment of the four small leucine-rich proteoglycan (SLRP) structures. Each SLRP is represented with the same color as the one used in *A*.

at the most and the sugar chains are located on the convex face, near the C-terminal part of the core proteins. This distribution and orientations of the glycosylations confer asymmetric conformation to the molecules and impact their properties and interactions. Both in fibromodulin and lumican, the number of N-glycosylations is at least equal to four, and the N-glycosylation sites are distributed along the length of the glycoproteins on the convex or on the lateral faces.

Characterization of the Interactions between MMP-14 and Human SLRPs Core Proteins

In addition to a specific transmembrane domain, MMP-14 (Fig. 3*A*) is characterized by multidomain structure consisting of a signal peptide, a pro-peptide, a catalytic domain, a hinge region, and a hemopexin domain.

The 3D structure of the catalytic domain of MMP-14 is also illustrated in Fig. 3*A*. On this surface representation,

AJP-Cell Physiol • doi:10.1152/ajpcell.00429.2022 • www.ajpcell.org

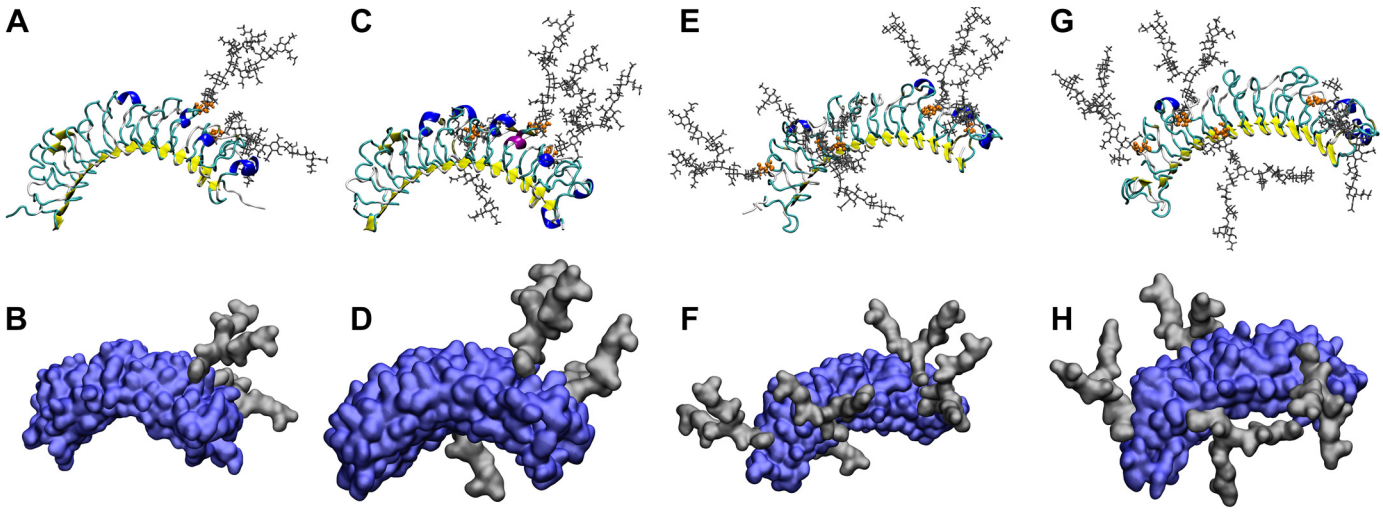


Figure 2. Secondary structures, surface representations, and N-glycosylation positions on human biglycan, decorin, fibromodulin, and lumican. Secondary structures, surface representations, and N-glycosylation positions on human biglycan (*A* and *B*), decorin (*C* and *D*), fibromodulin (*E* and *F*), and lumican (*G* and *H*). For each human small leucine-rich proteoglycan (SLRP), two types of representations are shown: a cartoon representation of the backbone (*A*, *C*, *E*, and *G*) and a surface representation that also considers the occupancy of the side chains (*B*, *D*, *F*, and *H*). The cartoon representations are colored according to the secondary structure of the core proteins and the residues bearing the N-glycosylations are displayed using orange Van der Waals motifs. Bi-antennary glycosylated chains are modeled with gray licorice (*A*, *C*, *E*, and *G*) or gray surfaces (*B*, *D*, *F*, and *H*).

three zones are evidenced: *1*) the catalytic zone (in green) containing residues known to interact with TIMP-2, *2*) the MT-LOOP area (in red), and *3*) a potential N-glycosylation site (in orange).

The docking of human SLRPs on the catalytic domain of MMP-14 is illustrated in Fig. 3*B*. Most of the traditional software developed to decipher interactions between macromolecules such as proteins do not integrate post-translational modifications (PTMs) and thus the docking simulations are performed without any glycan chains. The complexes associated with the best favorable energies of interaction were selected and the snapshots of Fig. 3*B* show two different modes of interaction. On one hand, in the case of human decorin and fibromodulin, even though SLRPs interact with MMP-14 through their concave face, the catalytic zone of MMP-14 (in green) is not involved. On the other hand, the catalytic zone of MMP-14 is clearly implicated in the interaction with the concave face of human biglycan and lumican.

Quantification of the normalized contact frequency (Ncf) over each set of 100 results generated by the docking experiments allows the graphical representations displayed in Fig. 3, *C* and *D*. The Ncf was calculated for each residue of each protein partner and the catalytic zone/MT-LOOP are highlighted in green and red, respectively (Fig. 3*C*), whereas LRRs are highlighted in orange and blue alternately (Fig. 3*D*). Focusing specifically on the two regions of interest of MMP-14 (MT-LOOP and catalytic zone) makes it possible to compare more precisely the way in which each SLRP interacts with MMP-14: thus, the values presented in Table 2 underline the fact that the lumican presents a higher averaged Ncf with the MT-LOOP than with the catalytic zone. On the contrary, decorin creates a greater averaged Ncf with the catalytic zone. Biglycan and fibromodulin show a comparable propensity to interact with each of the two regions.

Proteolytic Activities of MMPs

In this study, the direct effect of biglycan, decorin, fibromodulin, or lumican on MMP-14 activity was evaluated (Fig. 4). Each SLRPs directly inhibited the activity of MMP-14 in a dose-dependent manner. The inhibition of MMP-14 activity was significant when each SLRP was used at 100 nM in contrast to 1 and 10 nM.

More precisely, the mean percentage of inhibition of the MMP-14 activity by individual SLRP (100 nM) and their respective K_i were calculated (n = 4): BGN: 92%, K_i: 19.0 nM; DCN: 76%, K_i: 30.9 nM; FMOD: 83%, K_i: 27.1 nM; LUM: 86%, K_i: 29.3 nM. The presence of biglycan nearly totally (92%) blocked MMP-14 activity. Decorin, fibromodulin, and lumican exhibited also strong MMP-14 inhibition capacity (more than 75% inhibition). The K_i was calculated with increasing concentrations of SLRPs (1, 10, and 100 nM) by fitting to the Morrison equation (Supplemental Fig. S1; see https://doi.org/10.6084/m9.figshare.21171814).

In contrast to MMP-14 activity, the MMP-2 activity was not altered in the presence of increasing concentrations of each SLRP (1, 10, and 100 nM) as shown in Supplemental Fig. S2 (see https://doi.org/10.6084/m9.figshare.21171823).

Altogether, the results demonstrated that each SLRP exhibited a similar inhibition of MMP-14 activity with only minor difference of efficiency. To better understand the targeting of MMP-14 by the SLRPs, the oligosaccharide chain structures and dynamics of lumican were investigated to evaluate the impact of glycan shielding on the accessibility of the LRR of lumican core protein and of the four sites of cleavage by MMP-14 in lumican core protein (Fig. 5).

The Presence of N-Glycosylations Modifies the Accessibility to the LRRs of LUM

From the observation of the lumican trajectory, the overall sampling of the lumican glycans was mapped. This

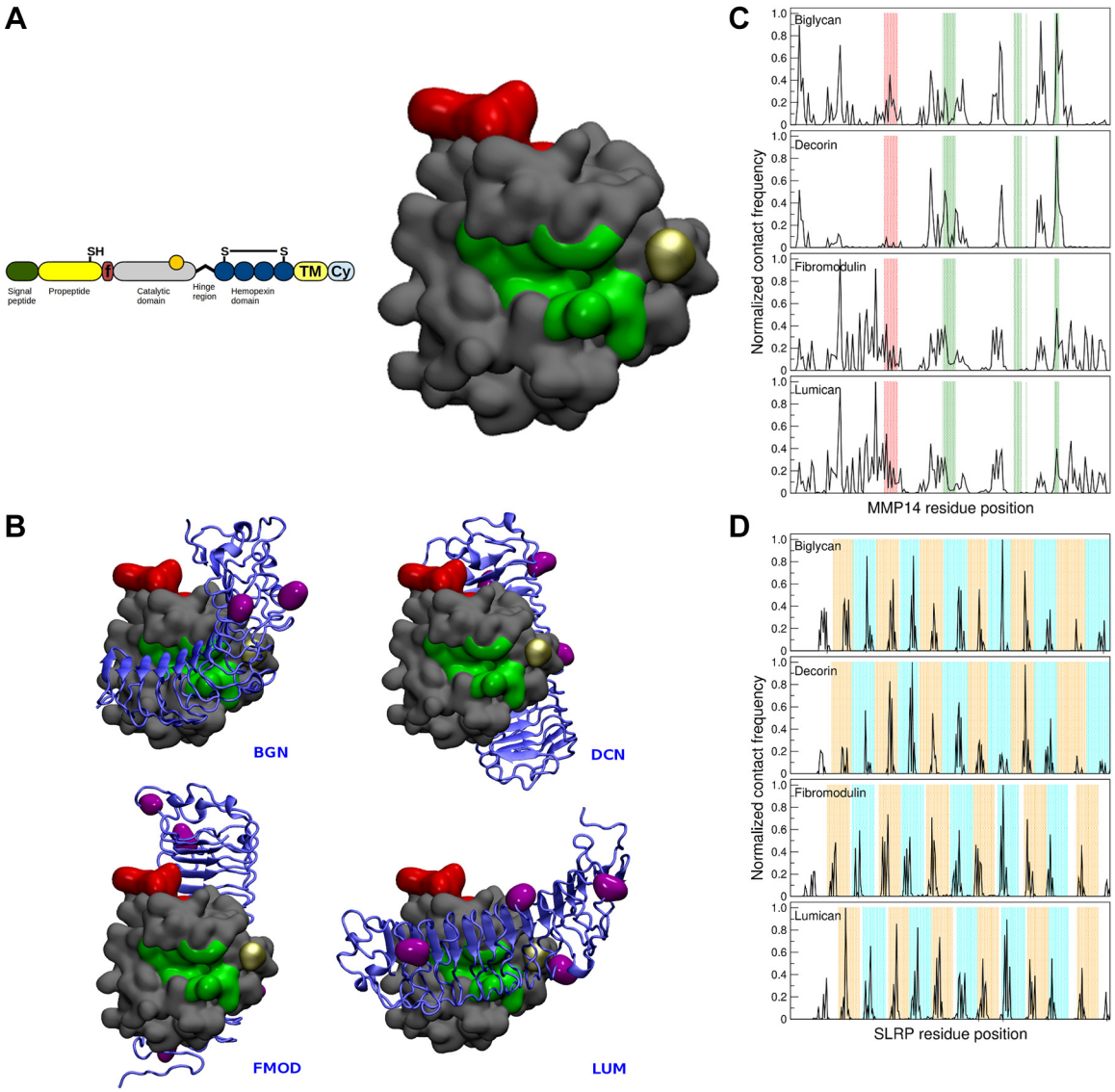


Figure 3. Rigid protein-protein docking experiments. *A*: domain structure of matrix metalloproteinase (MMP)-14 and surface representation of MMP-14 catalytic domain. The coordinates extracted from the pdb structure 1BQQ present a catalytic site (green) and the MT-LOOP (red). The ASN229 is highlighted in orange as a possible N-glycosylation site. *B*: complexes formed with human small leucine-rich proteoglycans (SLRPs). The results presented in this panel were obtained with the HEX software and correspond to the best binding energy. MMP-14 is pictured using the surface mode and the color code defined in *A*. Human SLRPs are drawn using the cartoon representation. Residues bearing the N-glycosylations are displayed using purple surfaces. *C*: MMP-14 residue positions interacting with SLRPs in the MT-LOOP (in red) and in the catalytic pocket (in green) of MMP-14. *D*: SLRPs residue positions interacting with MMP-14. The leucine-rich repeats (LRRs) in biglycan (BGN), decorin (DCN), fibromodulin (FMOD), and lumican (LUM) are indicated in orange and blue, alternatively.

enables the visualization of the lumican surface residues that are accessible and those that are buried, which would be otherwise accessible in the unglycosylated lumican. Much of the N-terminal half of lumican was found to be buried, including the disordered N-terminal loop. Although the N-terminal surface was largely buried, certain regions of the C-terminal half were well accessible and were not subject to any obstruction from glycans during the simulations (Fig. 5, *A*–*C*). Interestingly, these regions are involved in protein-protein interactions with various protein partners, including MMP-14. These include LRR9, which is observed to bind to the catalytic domain of MMP-14 from earlier experimental studies (40, 41). In addition, LRR7 and LRR11, which bind to collagen for lumican as well as for other SLRPs such as fibromodulin, were also accessible in the simulations (42, 43).

Although visualization of the glycan coverage provides a partial picture about how glycans impact the accessibility of surface residues, it is also necessary to access quantifiable observables. Thus, the absolute as well as the relative solvent accessible surface areas were calculated for different residues over the course of the simulation. The same calculation was also performed for the unglycosylated lumican model, which was compared with the average accessible area of the glycosylated lumican from the simulations. Residues whose relative accessibility is >20% are considered to be exposed (44); this threshold was thus used to classify residues as

Table 2. *Normalized contact frequency between two specific regions of MMP-14 and the different SLRPs*

	BGN	DCN	FMOD	LUM
MT-LOOP				
P163	0.03	0.01	0.07	0.12
Y164	0.23	0.09	0.42	0.53
A165	0.03	0.00	0.00	0.04
Y166	0.45	0.01	0.18	0.29
I167	0.17	0.00	0.00	0.04
R168	0.23	0.04	0.22	0.23
E169	0.13	0.00	0.04	0.08
G170	0.04	0.01	0.06	0.09
	0.16	**0.02**	**0.12**	**0.18**
Catalytic zone				
G196	0.05	0.23	0.23	0.18
G197	0.31	0.51	0.38	0.30
F198	0.21	0.39	0.24	0.18
L199	0.00	0.05	0.07	0.04
A200	0.03	0.01	0.05	0.02
H201	0.06	0.11	0.06	0.03
A202	0.04	0.03	0.06	0.03
Y203	0.17	0.34	0.11	0.08
V236	0.00	0.00	0.00	0.00
H239	0.00	0.00	0.00	0.00
E240	0.00	0.00	0.01	0.01
H243	0.05	0.01	0.02	0.01
P259	0.15	0.28	0.17	0.14
F260	1.00	1.00	0.56	0.40
Y261	0.48	0.40	0.20	0.16
	0.17	**0.22**	**0.14**	**0.10**

The 100 solutions gathered from the docking experiments were analyzed for each matrix metalloproteinase (MMP)-14/small leucine-rich proteoglycan (SLRP) complex and a normalized contact frequency could be evaluated for each protein residue. The data presented in this table are focused on the MT-LOOP and catalytic zone. For each residue defining the region, the normalized contact frequency is presented and for each SLRP, the last column (gray filling and bold characters) gives the region averaged normalized contact frequency. BGN, biglycan; DCN, decorin; FMOD, fibromodulin; LUM, lumican.

accessible or buried. Calculations were performed on two sets of residues: on the C-terminal LRR regions, which are observed to participate in protein-protein interactions with various partners, including MMP-14, and on MMP-14 cleavage sites, which have been characterized from previous experimental studies (16).

Considering the β-strand residues of LRR regions 7, 9, and 11, it could be observed that, while the LRR9 and 11 regions were accessible in both states (unglycosylated and glycosylated), LRR7 was found to be buried (Fig. 5*D*). Interestingly, glycosylation increased the accessibility of LRR7, though the residues T209 and Y211 were still considered buried. Similarly, accessibility of the residues of LRR9 reduced upon glycosylation, while an increase was observed for the residue H308 of LRR11 (Fig. 5*D*). Of the four MMP-14 cleavage sites, two are located at the N-terminal (in LRR1), and the other two at the C-terminal (in LRR9 and 10). We observed that the accessibility for the N-terminal cleavage sites residues decreased when the lumican is glycosylated. Particularly, upon glycosylation, relative accessibility of the residues Y70 (first cleavage site) and A85 (second cleavage site) reduced from 23% to 17% and from 21% to 13%, respectively (Fig. 5*E*). These residues, which were accessible in the unglycosylated lumican, became buried in the presence of glycosylation. Though the residues of the C-terminal cleavage sites remained accessible in both glycosylated as well as the unglycosylated states, a similar reduction of accessibilities was also observed for these residues as well (from 54% to 41%, 46% to 42% for sites 3 and 4, respectively, Fig. 5*E*).

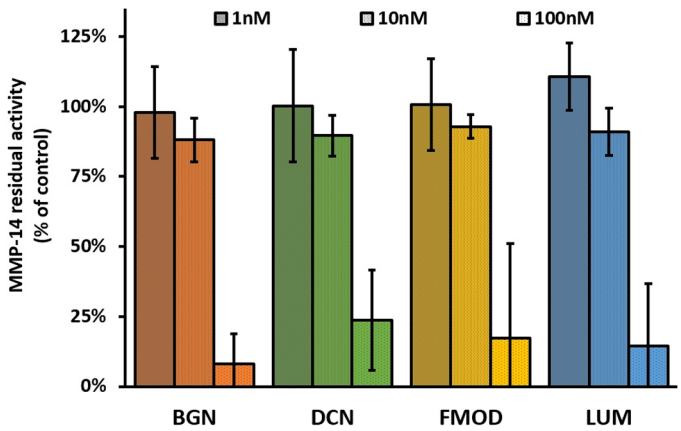


Figure 4. Matrix metalloproteinase (MMP)-14 activity assay. The effect of biglycan (BGN), decorin (DCN), fibromodulin (FMOD), and lumican (LUM) on recombinant MMP-14 activity was measured with increasing concentrations of SLRPs (1, 10, 100 nM) as described in METHODS. Data are presented as mean values ± SD from four independent experiments.

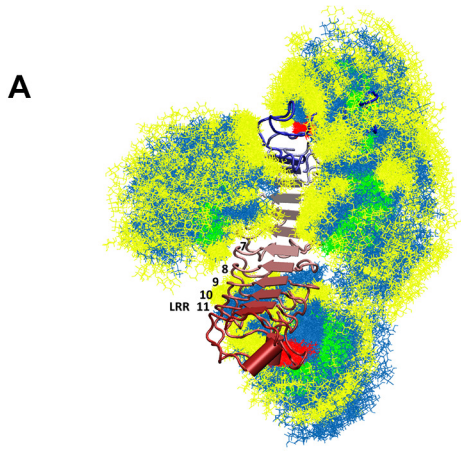
DISCUSSION

From Comparisons of Primary and Secondary Structures to Deciphering the Potential Impact of the Distribution of N-Glycosylations on the Surface of SLRPs

It is worth highlighting the particularly good identity rates obtained through the sequence alignments that allowed reconstructions by homology of the BGN and DCN molecules (94.58% and 89.44%, respectively). Even if the sequence identity rate between the LUM and FMOD is approximately twice as small, it is equal to 39.58% and therefore above the threshold of 30% which is the confidence limit above which it is generally accepted to be found to obtain a reconstruction of decent quality. Moreover, among the three experimental structures obtained by X-rays and used as templates in the construction of the models, two present a correct resolution since it is lower than or equal to 2.5 Å (2.21 Å and 2.5 Å for 5MX0-FMOD and 1XKU-DCN, respectively). The third structure, associated to the bovine BGN is certainly of lower quality (3.4 Å), but presents the advantage of being able to reconstruct the human BGN from an alignment of better quality compared with the reconstruction carried out by Vlachakis et al. (45), which was carried out from a sequence alignment between BGN and DCN (PDB ID 1XKU) presenting a similarity rate equal to 56.7%.

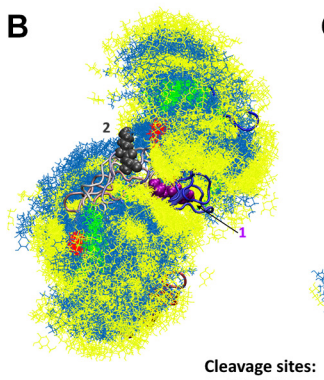
All the elements mentioned earlier allow great confidence in the comparison and the exploitation that will be carried out using these 3D structures, both from a static and dynamic point of view.

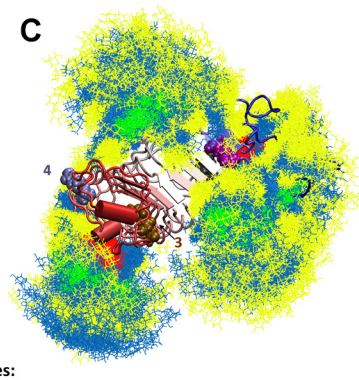
Moreover, the cross-comparison of the alignments of sequences, secondary structures, and three-dimensional structures presented in Fig. 1 not only makes it possible to underline, as mentioned previously, the great similarity of



D

LRR	Residue	AccAr from MD (Å²)	RelAcc (%)	Accessibility	AccAr from starting structure (Å²)	RelAcc (%)	Accessibility
7	T209	11.9 ± 5.1	8.2	Buried	3.3	2.2	Buried
	L210	0.3 ± 1.0	0.0	Buried	0.0	0.0	Buried
	Y211	35.3 ± 8.5	15.4	Buried	31.3	13.6	Buried
9	E258	47.4 ± 11.1	25.9	Accessible	63.1	34.4	Accessible
	L259	0.1 ± 0.6	0.0	Buried	0.0	0.0	Buried
	D260	17.2 ± 4.8	11.2	Buried	22.1	14.5	Buried
11	H308	41.4 ± 15.1	21.1	Accessible	39.4	20.1	Accessible
	L309	0.5 ± 1.4	0.0	Buried	0.0	0.0	Buried
	R310	75.0 ± 12.0	31.1	Accessible	92.1	38.1	Accessible





Cleavage sites:
1: 70-YL-71
2: 84-KA-85
3: 275-NL-276
4: 285-QL-286

E

Cleavage site	Residue	AccAr from MD (Å²)	RelAcc (%)	Accessibility	AccAr from starting structure (Å²)	RelAcc (%)	Accessibility
1 (in LRR 1)	Y70	38.7 ± 18.7	16.8	Buried	53.8	23.4	Accessible
	L71	0.4 ± 1.3	0.0	Buried	0.0	0.0	Buried
2 (in LRR 1)	K84	160.5 ± 2.4	75.0	Accessible	141.3	66.0	Accessible
	A85	14.9 ± 13.5	13.1	Buried	23.3	20.6	Accessible
3 (in LRR 9)	N275	65.4 ± 12.9	41.3	Accessible	84.9	53.7	Accessible
	L276	0.5 ± 1.2	0.0	Buried	0.0	0.0	Buried
4 (in LRR 10)	O285	80.0 ± 18.3	42.3	Accessible	71.9	45.5	Accessible
	L286	0.3 ± 1.0	0.0	Buried	0.0	0.0	Buried

Figure 5. Carbohydrate shielding on lumican leucine-rich repeat (LRR) accessibility. Distribution of glycan conformations over the lumican (*A*–*C*). *A*: front view showing the solvent inaccessible as well as solvent accessible LRRs. The C-terminal LRRs (LRR7 to LRR11) are labeled adjacent to the respective β-strands. *B*: top-view showing the N-terminal half of lumican, with the residues of cleavage sites 1 and 2 represented as Van der Waals spheres. *C*: bottom-view showing the C-terminal half of lumican, with the residues of cleavage sites 3 and 4 represented as Van der Waals spheres. Protein is represented as cartoon, colored according to blue-white-red scheme (N-terminal to C-terminal), and carbohydrate residues are represented as sticks and colored according to the SNFG scheme (39). The cleavage sites represented here are taken from experimental studies on lumican proteolysis by MMP-14 (16). *D*: solvent accessible surface areas (AccAr) and relative accessibility (RelAcc) of selected lumican residues that are part of LRR7, LRR9, and LRR11. Cases where AccAr is improved upon glycosylation are highlighted in gray. *E*: solvent accessible surface areas (AccAr) and relative accessibility (RelAcc) of lumican residues situated in the cleavage sites. Cases where AccAr is decreased upon glycosylation are highlighted in gray. A residue is considered buried when its RelAcc is <20, and it is considered accessible when its RelAcc is >20.

the four SLRPs considered in this study but also to find the subdivision into two classes of these.

To our knowledge, this is the first time that a study proposes to consider the structural and functional aspects of four SLRPs in parallel. Admittedly, earlier studies have been able to focus on and discuss the structure-function relationship of SLRPs, but only through the prism of a single SLRP: Weber et al. (46) studied DCN, Zeltz et al. (41) characterized LUM, Vlachakis et al. (45), as previously mentioned, were interested in the 3D structure of BGN, and Paracuellos et al. (8) the structure of FMOD.

Finally, since the sites of N-glycosylations are known for each SLRPs, bi-antennary chains were grafted on the core proteins. Depending on the number and positions of the glycosidic sites, it can be observed from the static accessible surface representations in Fig. 2 that the steric hindrance of the sugar chains might strongly impact the structures and the accessibility of the glycoproteins. Thus, one can infer that the number of glycans and their position on the SLRPs surface might affect the dynamical properties and behaviors of the macromolecules and have a crucial role on the overall structures and function of the SLRPs. Ultimately, the interactions between SLRPs and other molecules of the ECM may be affected by the presence or not of the glycans. Despite the complexity of the systems, it is then relevant to try to better understand the potential interactions between different human SLRPs and the catalytic domain of matrix metalloproteinase MMP-14 using in silico methodologies and more particularly docking simulations.

Modulation of Interactions between SLRP Core Proteins and MMP-14

From the 3D structures of two molecules, molecular docking makes it possible to predict, based on energetic criteria, the formation of ligand/protein, peptide/protein, or even protein/protein complexes. In the context of the study

presented here, the docking method considered is a so-called rigid method because, given the large sizes of the molecules, their internal flexibility was not taken into consideration. However, the protein/protein docking method used makes it possible to predict the formation of the SLRP/MMP-14 complex, to compare the observed docking modes, and to relate these observations to experimental data. Thus, observation of the first pose of the MMP-14/LUM complex (Fig. 3*B*) shows that the molecular interface is constituted by the cleft of the active site of MMP-14 and the region corresponding to the LRR fragments located at C-terminal. This observation is consistent with the fact that Lumican was able to inhibit MMP-14 activity. In addition, the complexes shown in Fig. 3*B* separate SLRPs into two groups: (BGN, LUM) on one side and (DCN, FMOD) on the other side. This distribution seems consistent with previous results obtained from our group showing that peptides derived from decorin or fibromodulin were unable to inhibit MMP-14 activity (40). However, the fact that DCN and FMOD are able to inhibit MMP-14 activity without a direct interaction with the MMP-14 catalytic pocket would suggest that they may interfere indirectly in the regulation of MMP-14 activity by eventually impairing the dimerization of MMP-14 leading to a decreased activity of the enzyme. This would imply that recombinant MMP-14 molecules are able to dimerize in solution as it has been described in cell membrane by Itoh (47).

The statistical analysis of the set of 100 generated docking solutions modulates this distribution. Indeed, the evaluation of the normalized contact frequency, assigned to each residue, then averaged over each region of interest of the MMP-14, makes it possible to constitute no longer two groups of proteins but three: LUM for which Ncf is more important on average on the MT-LOOP, DCN for which Ncf is more important on average on the catalytic zone, and finally BGN and FMOD for which the average values of Ncf are equivalent on the two regions of interest of the MMP-14.

A finer understanding of the interactions and a better description of the contact regions should be possible by complementing these static results using extensive MD simulations and by elaborating more complex systems through the consideration of the PTMs.

SLRP Effect on Matrix Metalloproteinases Activity

BGN, DCN, FMOD, and LUM did not show any inhibitory effect on human recombinant MMP-2 activity. Lumican did not change the MMP-2 activity as previously shown in mesenchymal stem cells (MSC) treated with recombinant lumican (48).

Our previous study showed that lumican (core protein and glycosylated forms) and partly decorin (glycosylated form) were able to inhibit human recombinant MMP-14 activity (14).

In the present study, the direct effect of lumican (glycosylated form only, 57 kDa), fibromodulin (core protein and glycosylated form, 41 and 45 kDa, respectively), biglycan (core protein and glycosylated form, 40 and 42 kDa, respectively), and decorin (core protein and glycosylated forms, 39 and 42 kDa, respectively) on MMP-14 activity was determined and compared. All of them inhibited strongly (more than 75% of inhibition) the MMP-14 activity with slightly less efficiency for decorin. Thus, while the glycosylated form of decorin was a poor inhibitor (−15%) of MMP-14 activity (14), the core protein of decorin used in the present study was a strong inhibitor (−76%) demonstrating the impact of glycosylation in the regulation of MMP-14 activity suggesting an inaccessibility to the catalytic pocket of MMP-14 due to the presence of the glycan chains. Further studies are required to analyze the glycan shielding of DCN. In contrast, upon glycosylation, lumican was shown to be a stronger inhibitor of MMP-14 activity than the unglycosylated lumican core protein suggesting an increased accessibility of the lumican residue interacting directly with the MMP-14 catalytic pocket.

In the docking studies, two sets of possible MMP14-lumican interfaces were observed: the top-scoring MMP14-lumican complex had the region around the catalytic site as the interface. However, in numerous complexes, the regions surrounding the MT-LOOP of the MMP-14 catalytic domain had high incidence of contacts with lumican. This allows for the possibility of dual binding modes of lumican with MMP-14, which needs to be further explored in future modeling and simulation studies on lumican and MMP-14. Although lumican that is localized in corneal tissue has negatively charged glycosaminoglycans (GAGs) with sulfate groups, lumican in other extracellular matrices lacks such charged GAGs (49, 50). Thus, in such tissues, there is less possibility of GAGs interacting with positively charged MMP-14 residues via salt bridges. However, GAGs could stabilize MMP-14-lumican complex through the formation of hydrogen bonds between carbohydrate residues and MMP-14 residues.

GAG-protein binding is mostly through electrostatic interactions between negatively charged sulfate groups and positively charged amino acids in the protein (51). Interestingly, electrostatics-driven interactions between enzyme and GAGs were recently reported (52). Indeed, Bojarski et al. (52) demonstrated that cathepsin-GAG interactions regulate the enzyme activity in subtle molecular mechanisms.

The lumican glycan shielding was therefore analyzed in depth by MD investigation of the polysaccharide chains as shown in Fig. 5 The in silico results revealed an increased accessibility of the LRRs (LRR7 to LRR11) of the C-terminal part of the core protein of lumican upon glycosylation. Interestingly, several synthetic peptides derived from these LRRs of lumican core protein were identified and demonstrated to be as significant inhibitors of MMP-14 activity as the entire molecule of glycosylated lumican (40). The specific residues of lumican core protein buried by glycan chains are detailed and discussed later with regards to the LRR accessibility. Further works are necessary to investigate the effect of the extended glycosylated forms of FMOD and BGN (not used in the present work) on MMP-14 activity.

In addition, lumican and decorin were described as substrates for MMP-14 (13, 15, 16). However, the core protein and the glycosylated forms of decorin were not cleaved by MMP-14 in our hands as shown in our previous work (14). Decorin degradation by MMP-14 required a pH below 6.5 and the presence of Zn^{2+} (13). More recently, MMP-14 was shown to cleave DCN to generate nonglycanated DCN. DCN deglycanation resulted in reduced intracellular DCN-collagen binding and increased the production of truncated COL6A chains. Thus, GAG modification of DCN was demonstrated to depend on MMP-14 activity and to regulate collagen assembly (15). Lumican core protein was shown to be a

substrate of MMP-14 in contrast to lumican glycosylated form (14). These results suggested that the glycosylation of lumican could protect from its degradation by MMP-14. Thus, the most efficient form of lumican (glycosylated form) able to inhibit MMP-14 activity would be preserved by the glycan chains. Indeed, the MD analysis of the glycan shielding demonstrated also that the MMP-14 cleavage sites 1 and 2 (13) in the N-terminal part of the core protein of lumican were buried by glycan chain space occupancy that would prevent lumican proteolysis by MMP-14. Moreover, the MMP-14 cleavage sites 3 and 4 of the C-terminal part of the lumican core protein were found to be relatively accessible. The specific residues of lumican core protein buried by glycan chains are detailed and discussed later with regards to the MMP-14 cleavage sites.

Lumican (glycosylated form, or core protein form or its fragments produced after MMP-14 cleavage of the core protein) could inhibit the activity of MMP-14 but differentially. During proteolysis of lumican by MMP-14, lumican or its fragments may bind to the active site of the enzyme that could lead to changes in the conformation of MMP-14 and loss of its activity. Further studies are necessary to analyze whether FMOD and BGN core proteins are cleaved by MMP-14 and to investigate the impact of their glycosylation on their potential protective role against the digestion by MMP-14.

All the four SLRPs substantially inhibited MMP14 activity in a dose-dependent manner. However, among the four SLRPs studied, decorin had the lowest inhibition activity (76%) against MMP-14 (Fig. 4). From the contact frequency plots, it is observed that decorin forms fewer contacts with the catalytic site and there are no interactions with the MT-LOOP. Similarly, in decorin, LRR8 is significantly less involved in the interface compared with the LRR8 regions of the other three SLRPs (Fig. 3*D*). Thus, binding to the MT-LOOP and catalytic site with LRR8 could be a prerequisite for effective inhibition of MMP-14 activity.

GAGs bind to many different classes of proteins such as FGF2, CCL5, antithrombin (51) and in the same manner SLRPs exhibit also many different partners in addition to MMP-14. Lumican is able to interact with collagen type I fibers (10, 53), α2 integrin subunit (37), TGFR2, TLR2, FASL and CD14 (12, 54, 55). Thus, the impact of the molecular conformation and dynamic of the keratan sulfate (KS) GAG chains of lumican should be taken into account during its binding to its different targets. Fibromodulin was demonstrated to interact directly with collagen type I (8, 11, 56) but the glycanated forms of FMOD and their KS GAG chain role would require further investigations.

Similarly, the chondroitin sulfate (CS) and the dermatan sulfate (DS) GAG chains of decorin may interfere in the DCN binding to Met (57, 58) and other DCN receptors such as EGFR, VEGFR2, Toll-like receptor (TLR) 2, TLR4, and TGFβ (59, 60). Biglycan was recently shown to interact with type I insulin-like receptor (IGF-IR) (61) and is also a ligand for TLR2, TLR4, LRP6, and major histocompatibility complex (MHC) class 1 receptors (60, 62). Thus, the glycanated forms of BGN and DCN and their CS and DS GAG chain role would require further investigations. Both DCN and BGN are necessary for maintaining collagen structure (63). It has been demonstrated that GAGs modification of DCN depends on MMP-14 activity and regulates collagen assembly (15).

Influence of GAGs on Lumican

From the visualization of the total glycan sampling in the simulations, it is observed that much of the lumican N-terminal region, including the disordered region, is well shielded by the N-glycosylations (Fig. 5, *A* and *B*). This observation is interesting, since this region is observed to have numerous modified tyrosine sulfate residues that are involved in protein-protein interactions in lumican as well as in other SLRPs such as fibromodulin (56, 64). However, it must be noted that in the current simulations, the tyrosine residues were unmodified; thus, the accessibility patterns may change in the presence of tyrosine sulfate modifications. It would be interesting to explore this in future studies. At the same time, the C-terminal half of lumican is accessible despite the presence of N-glycosylations (Fig. 5, *A* and *C*). The fact that these regions constitute LRR regions that are involved in protein-protein interactions with various partners (40–43), suggests that glycosylations shield lumican against the actions of proteolytic enzymes such as MMP-14, while at the same time, regions of lumican implicated in protein/protein interactions remain accessible to partners such as collagen and MMP-14.

This was evident from the solvent accessible surface area calculations, where MMP-14 cleavage sites at N-terminal were efficiently shielded by N-glycosylations. Though the cleavage sites toward the C-terminal half remain exposed both in the glycosylated and unglycosylated state, N-glycosylation substantially reduced their accessibility (Fig. 5*D*). The differences in accessibility of the two sites stem from the fact that the sites located at the N-terminal half are also close to the glycosylated asparagine residues, which allows better shielding with respect to these residues (Fig. 5*B*). On the other hand, the C-terminal cleavage sites are located farther from the glycosylated asparagine residues, resulting in less shielding of these sites (Fig. 5*C*).

Although the N-terminal half of lumican is strongly shielded by N-glycosylations, the C-terminal LRRs, LRR9 and 11, which are known to be involved in protein-protein interactions remain accessible regardless of the presence/absence of N-glycosylations. However, subtle changes in accessibility patterns of LRR regions were also noticed. Residues toward the C-terminal region tend to be more accessible compared with ones toward the N-terminal region. This could also stem from closer proximity of LRR7 and 9 to N-terminal as compared with the LRR11 (Fig. 5*D*). Though LRR7 is buried in both glycosylated and unglycosylated states, the presence of N-glycosylation increased the value of the accessible surface area. Since LRR7 in lumican is involved in interaction with collagen (42), the analyses of MD simulations suggest that N-glycosylations could also facilitate lumican-collagen interactions.

Though lumican consists of sulfated GAGs in corneal tissues, in other tissues, lumican GAGs remain unsulfated. However, the extent of GAG chain length variation in lumican is not well characterized and explored through experimental studies. Nevertheless, the effects of GAG chain length can be explored in the case of fibromodulin and lumican.

Though both SLRPs have relatively high sequence identity (~60%), MMP-14 binds to lumican with higher affinity than fibromodulin. This could be because fibromodulin has a longer GAG chain length compared with lumican, and has

numerous branched fucose residues (65, 66). This could affect the binding affinity of MMP-14, due to the putative obstructions created by the GAG chains over the C-terminal LRR regions. On the other hand, the C-terminal LRRs of lumican remain accessible despite wide sampling of GAG chains in our simulations (Fig. 5), primarily due to shorter length of the GAG chains.

An important caveat associated to the accessibility calculations in the present work is that many residues have large standard deviation associated with the surface area values (Fig. 5, *D* and *E*). However, this is expected because the glycans surrounding these residues are themselves very dynamic, which could be the cause of a strong variation in the solvent accessibility of these residues.

In conclusion, the glycosylations significantly impact the accessibility of residues, resulting in better shielding of residues from cleavage by MMP-14. These results are consistent with the observations made from earlier experimental studies that demonstrated that the unglycosylated lumican is more susceptible to digestion by MMP-14 and the glycosylated lumican is immune from the same. Our simulations provide a possible molecular mechanism by which glycosylations protect lumican from the proteolytic digestion by MMP-14, while at the same time enabling protein-protein interactions with other protein partners.

DATA AVAILABILITY

Data will be made available upon reasonable request.

SUPPLEMENTAL DATA

Supplemental Fig. S1: https://doi.org/10.6084/m9.figshare.21171814.

Supplemental Fig. S2: https://doi.org/10.6084/m9.figshare.21171823.

ACKNOWLEDGMENTS

The authors thank the supercomputer center ROMEO for providing CPU-GPU times and the Multiscale Molecular Modeling Platform for providing facilities in the analyses of simulations.

This article is part of the special collection “Deciphering the Role of Proteoglycans and Glycosaminoglycans in Health and Disease.” Liliana Schaefer, MD, served as Guest Editor of this collection.

GRANTS

The authors thank financial and infrastructural support from Chair Modélisation moléculaire et Agroressources: Ingrédients, Cosmétique et Santé, Université de Reims Champagne-Ardenne (MAgICS).

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

R.M.R., N.B., M.D., L.R., S.Baud, and S.Brézillon conceived and designed research; R.R., R.M.R., P.N., N.B., L.H., and S.Brézillon performed experiments; R.R., R.M.R., P.N., N.B., L.H., M.D., L.R., S.Baud, and S.Brézillon analyzed data; R.R., R.M.R., P.N., N.B., M.D., L.R., S.Baud, and S.Brézillon interpreted results of experiments; R.R., R.M.R., N.B., S.Baud, and S.Brézillon prepared figures; R.M.R., N.B., S.Baud, and S.Brézillon drafted manuscript; R.R., R.M.R., P.N., N.B., M.D., L.R., S.Baud, and S.Brézillon edited and revised manuscript; R.R., R.M.R., P.N., N.B., L.H., M.D., L.R., S.Baud, and S.Brézillon approved final version of manuscript.

REFERENCES

1. **McEwan PA**, **Scott PG**, **Bishop PN**, **Bella J.** Structural correlations in the family of small leucine-rich repeat proteins and proteoglycans. *J Struct Biol* 155: 294–305, 2006. doi:10.1016/j.jsb.2006.01.016.
2. **Schaefer L**, **Iozzo RV.** Biological functions of the small leucine-rich proteoglycans: from genetics to signal transduction. *J Biol Chem* 283: 21305–21309, 2008. doi:10.1074/jbc.R800020200.
3. **Pietraszek-Gremplewicz K**, **Karamanou K**, **Niang A**, **Dauchez M**, **Belloy N**, **Maquart F-X**, **Baud S**, **Brézillon S.** Small leucine-rich proteoglycans and matrix metalloproteinase-14: key partners? *Matrix Biol* 75–76: 271–285, 2019. doi:10.1016/j.matbio.2017.12.006.
4. **Ameye L**, **Young MF.** Mice deficient in small leucine-rich proteoglycans: novel in vivo models for osteoporosis, osteoarthritis, Ehlers-Danlos syndrome, muscular dystrophy, and corneal diseases. *Glycobiology* 12: 107R–116R, 2002. doi:10.1093/glycob/cwf065.
5. **Kobe B**, **Deisenhofer J.** Crystallization and preliminary X-ray analysis of porcine ribonuclease inhibitor, a protein with leucine-rich repeats. *J Mol Biol* 231: 137–140, 1993. doi:10.1006/jmbi.1993.1263.
6. **Scott PG**, **Dodd CM**, **Bergmann EM**, **Sheehan JK**, **Bishop PN.** 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. doi:10.1074/jbc.M513470200.
7. **Scott PG**, **McEwan PA**, **Dodd CM**, **Bergmann EM**, **Bishop PN**, **Bella J.** Crystal structure of the dimeric protein core of decorin, the archetypal small leucine-rich repeat proteoglycan. *Proc Natl Acad Sci USA* 101: 15633–15638, 2004. doi:10.1073/pnas.0402976101.
8. **Paracuellos P**, **Kalamajski S**, **Bonna A**, **Bihan D**, **Farndale RW**, **Hohenester E.** Structural and functional analysis of two small leucine-rich repeat proteoglycans, fibromodulin and chondroadherin. *Matrix Biol* 63: 106–116, 2017. doi:10.1016/j.matbio.2017.02.002.
9. **Kalamajski S**, **Oldberg Å.** The role of small leucine-rich proteoglycans in collagen fibrillogenesis. *Matrix Biol* 29: 248–253, 2010. doi:10.1016/j.matbio.2010.01.001.
10. **Chakravarti S**, **Magnuson T**, **Lass JH**, **Jepsen KJ**, **LaMantia C**, **Carroll H.** Lumican regulates collagen fibril assembly: skin fragility and corneal opacity in the absence of lumican. *J Cell Biol* 141: 1277–1286, 1998. doi:10.1083/jcb.141.5.1277.
11. **Oldberg Å**, **Kalamajski S**, **Salnikov AV**, **Stuhr L**, **Mörgelin M**, **Reed RK**, **Heldin N-E**, **Rubin K.** Collagen-binding proteoglycan fibromodulin can determine stroma matrix structure and fluid balance in experimental carcinoma. *Proc Natl Acad Sci USA* 104: 13966–13971, 2007. doi:10.1073/pnas.0702014104.
12. **Nikitovic D**, **Papoutsidakis A**, **Karamanos NK**, **Tzanakakis GN.** Lumican affects tumor cell functions, tumor–ECM interactions, angiogenesis and inflammatory response. *Matrix Biol* 35: 206–214, 2014. doi:10.1016/j.matbio.2013.09.003.
13. **Mimura T**, **Han KY**, **Onguchi T**, **Chang J-H**, **Kim T**, **Kojima T**, **Zhou Z**, **Azar DT.** MT1-MMP-mediated cleavage of decorin in corneal angiogenesis. *J Vasc Res* 46: 541–550, 2009. doi:10.1159/000226222.
14. **Pietraszek K**, **Chatron-Colliet A**, **Brézillon S**, **Perreau C**, **Jakubiak-Augustyn A**, **Krotkiewski H**, **Maquart F-X**, **Wegrowski Y.** Lumican: a new inhibitor of matrix metalloproteinase-14 activity. *FEBS Lett* 588: 4319–4324, 2014. doi:10.1016/j.febslet.2014.09.040.
15. **Daquinag AC**, **Gao Z**, **Fussell C**, **Sun K**, **Kolonin MG.** Glycosaminoglycan modification of decorin depends on MMP14 activity and regulates collagen assembly. *Cells* 9: 2646, 2020. doi:10.3390/cells9122646.
16. **Li Y**, **Aoki T**, **Mori Y**, **Ahmad M**, **Miyamori H**, **Takino T**, **Sato H.** Cleavage of lumican by membrane-type matrix metalloproteinase-1 abrogates this proteoglycan-mediated suppression of tumor cell colony formation in soft agar. *Cancer Res* 64: 7058–7064, 2004. doi:10.1158/0008-5472.CAN-04-1038.

17. **Sato H**, **Takino T**, **Miyamori H.** Roles of membrane-type matrix metalloproteinase-1 in tumor invasion and metastasis. *Cancer Sci* 96: 212–217, 2005. doi:10.1111/j.1349-7006.2005.00039.x.
18. **Wolf K**, **Müller R**, **Borgmann S**, **Bröcker E-B**, **Friedl P.** Amoeboid shape change and contact guidance: T-lymphocyte crawling through fibrillar collagen is independent of matrix remodeling by MMPs and other proteases. *Blood* 102: 3262–3269, 2003. doi:10.1182/blood-2002-12-3791.
19. **Quintero-Fabián S**, **Arreola R**, **Becerril-Villanueva E**, **Torres-Romero JC**, **Arana-Argáez V**, **Lara-Riegos J**, **Ramírez-Camacho MA**, **Alvarez-Sánchez ME.** Role of matrix metalloproteinases in angiogenesis and cancer. *Front Oncol* 9: 1370, 2019. doi:10.3389/fonc.2019.01370.
20. **Bigg HF**, **Morrison CJ**, **Butler GS**, **Bogoyevitch MA**, **Wang Z**, **Soloway PD**, **Overall CM.** Tissue inhibitor of metalloproteinases-4 inhibits but does not support the activation of gelatinase A via efficient inhibition of membrane type 1-matrix metalloproteinase. *Cancer Res* 61: 3610–3618, 2001.
21. **Pinto F**, **Santos-Ferreira L**, **Pinto M**, **Gomes C**, **Reis C.** The extracellular small leucine-rich proteoglycan biglycan is a key player in gastric cancer aggressiveness. *Cancers (Basel)* 13: 1330, 2021. doi:10.3390/cancers13061330.
22. **Khan FU**, **Owusu-Tieku NYG**, **Dai X**, **Liu K**, **Wu Y**, **Tsai H-I**, **Chen H**, **Sun C**, **Huang L.** Wnt/β-catenin pathway-regulated fibromodulin expression is crucial for breast cancer metastasis and inhibited by aspirin. *Front Pharmacol* 10: 1308, 2019. doi:10.3389/fphar.2019.01308.
23. **Shu C**, **Flannery C**, **Little C**, **Melrose J.** Catabolism of fibromodulin in developmental rudiment and pathologic articular cartilage demonstrates novel roles for MMP-13 and ADAMTS-4 in C-terminal processing of SLRPs. *Int J Mol Sci* 20: 579, 2019. doi:10.3390/ijms20030579.
24. **Hocking AM**, **Strugnell RA**, **Ramamurthy P**, **McQuillan DJ.** Eukaryotic expression of recombinant biglycan. *J Biol Chem* 271: 19571–19577, 1996. doi:10.1074/jbc.271.32.19571.
25. **The UniProt Consortium.** UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res* 49: D480–D489, 2021. doi:10.1093/nar/gkaa1100.
26. **Ritchie D.** Recent progress and future directions in protein-protein docking. *Curr Protein Pept Sci* 9: 1–15, 2008. doi:10.2174/138920308783565741.
27. **Ritchie DW.** Evaluation of protein docking predictions using Hex 3.1 in CAPRI rounds 1 and 2. *Proteins* 52: 98–106, 2003. doi:10.1002/prot.10379.
28. **Fernandez-Catalan C**, **Bode W**, **Huber R**, **Turk D**, **Calvete JJ**, **Lichte A**, **Tschesche H**, **Maskos K.** Crystal structure of the complex formed by the membrane type 1-matrix metalloproteinase with the tissue inhibitor of metalloproteinases-2, the soluble progelatinase A receptor. *EMBO J* 17: 5238–5248, 1998. doi:10.1093/emboj/17.17.5238.
29. **Winn MD**, **Ballard CC**, **Cowtan KD**, **Dodson EJ**, **Emsley P**, **Evans PR**, **Keegan RM**, **Krissinel EB**, **Leslie AGW**, **McCoy A**, **McNicholas SJ**, **Murshudov GN**, **Pannu NS**, **Potterton EA**, **Powell HR**, **Read RJ**, **Vagin A**, **Wilson KS.** Overview of the *CCP* 4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67: 235–242. 2011. doi:10.1107/S0907444910045749.
30. **Danne R**, **Poojari C**, **Martinez-Seara H**, **Rissanen S**, **Lolicato F**, **Róg T**, **Vattulainen I.** *DoGlycans*—tools for preparing carbohydrate structures for atomistic simulations of glycoproteins, glycolipids, and carbohydrate polymers for GROMACS. *J Chem Inf Model* 57: 2401–2406, 2017. doi:10.1021/acs.jcim.7b00237.
31. **Pettersen EF**, **Goddard TD**, **Huang CC**, **Couch GS**, **Greenblatt DM**, **Meng EC**, **Ferrin TE.** UCSF chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 25: 1605–1612, 2004. doi:10.1002/jcc.20084.
32. **Lindorff-Larsen K**, **Piana S**, **Palmo K**, **Maragakis P**, **Klepeis JL**, **Dror RO**, **Shaw DE.** Improved side-chain torsion potentials for the Amber Ff99SB protein force field: improved protein side-chain potentials. *Proteins* 78: 1950–1958, 2010. doi:10.1002/prot.22711.
33. **Kirschner KN**, **Yongye AB**, **Tschampel SM**, **González-Outeiriño J**, **Daniels CR**, **Foley BL**, **Woods RJ.** GLYCAM06: a generalizable biomolecular force field. Carbohydrates. *J Comput Chem* 29: 622–655, 2008. doi:10.1002/jcc.20820.
34. **Price DJ**, **Brooks CL.** A modified TIP3P water potential for simulation with ewald summation. *J Chem Phys* 121: 10096–10103, 2004. doi:10.1063/1.1808117.
35. **Berendsen HJC**, **van der Spoel D**, **van Drunen R.** GROMACS: a message-passing parallel molecular dynamics implementation. *Comput Phys Commun* 91: 43–56, 1995. doi:10.1016/0010-4655(95)00042-E.
36. **Páll S**, **Zhmurov A**, **Bauer P**, **Abraham M**, **Lundborg M**, **Gray A**, **Hess B**, **Lindahl E.** Heterogeneous parallelization and acceleration of molecular dynamics simulations in GROMACS. *J Chem Phys* 153: 134110, 2020. doi:10.1063/5.0018516.
37. **Zeltz C**, **Brézillon S**, **Käpylä J**, **Eble JA**, **Bobichon H**, **Terryn C**, **Perreau C**, **Franz CM**, **Heino J**, **Maquart F-X**, **Wegrowski Y.** Lumican inhibits cell migration through A2β1 integrin. *Exp Cell Res* 316: 2922–2931, 2010. doi:10.1016/j.yexcr.2010.08.002.
38. **Guillot A**, **Dauchez M**, **Belloy N**, **Jonquet J**, **Duca L**, **Romier B**, **Maurice P**, **Debelle L**, **Martiny L**, **Durlach V**, **Baud S**, **Blaise S.** Impact of sialic acids on the molecular dynamic of bi-antennary and tri-antennary glycans. *Sci Rep* 6: 35666, 2016. doi:10.1038/srep35666.
39. **Varki A**, **Cummings RD**, **Esko JD**, **Freeze HH**, **Stanley P**, **Marth JD**, **Bertozzi CR**, **Hart GW**, **Etzler ME.** Symbol nomenclature for glycan representation. *Proteomics* 9: 5398–5399, 2009. doi:10.1002/pmic.200900708.
40. **Pietraszek K**, **Brézillon S**, **Perreau C**, **Malicka-Błaszkiewicz M**, **Maquart F-X**, **Wegrowski Y.** Lumican—derived peptides inhibit melanoma cell growth and migration. *PLoS One* 8: e76232, 2013. doi:10.1371/journal.pone.0076232.
41. **Zeltz C**, **Brézillon S**, **Perreau C**, **Ramont L**, **Maquart F-X**, **Wegrowski Y.** Lumcorin: a leucine-rich repeat 9-derived peptide from human lumican inhibiting melanoma cell migration. *FEBS Lett* 583: 3027–3032, 2009. doi:10.1016/j.febslet.2009.08.012.
42. **Kalamajski S**, **Oldberg Å.** Homologous sequence in lumican and fibromodulin leucine-rich repeat 5-7 competes for collagen binding. *J Biol Chem* 284: 534–539, 2009. doi:10.1074/jbc.M805721200.
43. **Kalamajski S**, **Oldberg Å.** Fibromodulin binds collagen type I via Glu-353 and Lys-355 in leucine-rich repeat 11. *J Biol Chem* 282: 26740–26745, 2007. doi:10.1074/jbc.M704026200.
44. **Li X**, **Pan X-M.** New method for accurate prediction of solvent accessibility from protein sequence. *Proteins* 42: 1–5, 2001. doi:10.1002/1097-0134(20010101)42:1<1::AID-PROT10>3.0.CO;2-N.
45. **Vlachakis D**, **Tsaniras SC**, **Feidakis C**, **Kossida S.** Molecular modelling study of the 3D structure of the biglycan core protein, using homology modelling techniques. *J Mol Biochem* 2: 85–93, 2013. https://www.jmolbiochem.com/index.php/JmolBiochem/article/view/95/53.
46. **Weber IT**, **Harrison RW**, **Iozzo RV.** Model structure of decorin and implications for collagen fibrillogenesis. *J Biol Chem* 271: 31767–31770, 1996. doi:10.1074/jbc.271.50.31767.
47. **Itoh Y.** Membrane-type matrix metalloproteinases: their functions and regulations. *Matrix Biol* 44-46: 207–223, 2015. doi:10.1016/j.matbio.2015.03.004.
48. **Malinowski M**, **Pietraszek K**, **Perreau C**, **Boguslawski M**, **Decot V**, **Stoltz J-F**, **Vallar L**, **Niewiarowska J**, **Cierniewski C**, **Maquart F-X**, **Wegrowski Y**, **Brézillon S.** Effect of lumican on the migration of human mesenchymal stem cells and endothelial progenitor cells: involvement of matrix metalloproteinase-14. *PLoS One* 7: e50709, 2012. doi:10.1371/journal.pone.0050709.
49. **Karamanou K**, **Perrot G**, **Maquart F-X**, **Brézillon S.** Lumican as a multivalent effector in wound healing. *Adv Drug Deliv Rev* 129: 344–351, 2018. doi:10.1016/j.addr.2018.02.011.
50. **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. doi:10.1002/path.1037.
51. **Xu D**, **Prestegard JH**, **Linhardt RJ**, **Esko JD.** Chapter 38: Proteins that bind sulfated glycosaminoglycans. In: *Essentials of Glycobiology* (4th ed.), edited by Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, Mohnen D, Kinoshita T, Packer NH, Prestegard JH, Schnaar RL, Seeberger PH. Cold Spring Harbor, NY: Cold Spring Laboratory Press, 2022.
52. **Bojarski KK**, **Sage J**, **Lalmanach G**, **Lecaille F**, **Samsonov SA.** In silico and in vitro mapping of specificity patterns of glycosaminoglycans towards cysteine cathepsins B, L, K, S and V. *J Mol Graph Model* 113: 108153, 2022. doi:10.1016/j.jmgm.2022.108153.
53. **Stamov DR**, **Müller A**, **Wegrowski Y**, **Brezillon S**, **Franz CM.** Quantitative analysis of type I collagen fibril regulation by lumican and decorin using AFM. *J Struct Biol* 183: 394–403, 2013. doi:10.1016/j.jsb.2013.05.022.

54. **Brézillon S**, **Pietraszek K**, **Maquart F-X**, **Wegrowski Y.** Lumican effects in the control of tumour progression and their links with metalloproteinases and integrins. *FEBS J* 280: 2369–2381, 2013. doi:10.1111/febs.12210.
55. **Maiti G**, **Frikeche J**, **Lam CY-M**, **Biswas A**, **Shinde V**, **Samanovic M**, **Kagan JC**, **Mulligan MJ**, **Chakravarti S.** Matrix lumican endocytosed by immune cells controls receptor ligand trafficking to promote TLR4 and restrict TLR9 in sepsis. *Proc Natl Acad Sci USA* 118: e2100999118, 2021. doi:10.1073/pnas.2100999118.
56. **Kalamajski S**, **Bihan D**, **Bonna A**, **Rubin K**, **Farndale RW.** Fibromodulin interacts with collagen cross-linking sites and activates lysyl oxidase. *J Biol Chem* 291: 7951–7960, 2016. doi:10.1074/jbc.M115.693408.
57. **Neill T**, **Iozzo RV.** The role of decorin proteoglycan in mitophagy. *Cancers (Basel)* 14: 804, 2022. doi:10.3390/cancers14030804.
58. **Buraschi S**, **Neill T**, **Iozzo RV.** Decorin is a devouring proteoglycan: remodeling of intracellular catabolism via autophagy and mitophagy. *Matrix Biol* 75–76: 260–270, 2019. doi:10.1016/j.matbio.2017.10.005.
59. **Gubbiotti MA**, **Vallet SD**, **Ricard-Blum S**, **Iozzo RV.** Decorin interacting network: a comprehensive analysis of decorin-binding partners and their versatile functions. *Matrix Biol* 55: 7–21, 2016. doi:10.1016/j.matbio.2016.09.009.
60. **Diehl V**, **Huber LS**, **Trebicka J**, **Wygrecka M**, **Iozzo RV**, **Schaefer L.** The role of decorin and biglycan signaling in tumorigenesis. *Front Oncol* 11: 801801, 2021. doi:10.3389/fonc.2021.801801.
61. **Giatagana E-M**, **Berdiaki A**, **Gaardløs M**, **Samsonov SA**, **Tzanakakis GN**, **Nikitovic D.** Biglycan interacts with type I insulin-like receptor (IGF-IR) signaling pathway to regulate osteosarcoma cell growth and response to chemotherapy. *Cancers (Basel)* 14: 1196, 2022. doi:10.3390/cancers14051196.
62. **Moreth K**, **Frey H**, **Hubo M**, **Zeng-Brouwers J**, **Nastase M-V**, **Hsieh LT-H**, **Haceni R**, **Pfeilschifter J**, **Iozzo RV**, **Schaefer L.** Biglycan-triggered TLR-2- and TLR-4-signaling exacerbates the pathophysiology of ischemic acute kidney injury. *Matrix Biol* 35: 143–151, 2014. doi:10.1016/j.matbio.2014.01.010.
63. **Robinson KA**, **Sun M**, **Barnum CE**, **Weiss SN**, **Huegel J**, **Shetye SS**, **Lin L**, **Saez D**, **Adams SM**, **Iozzo RV**, **Soslowsky LJ**, **Birk DE.** Decorin and biglycan are necessary for maintaining collagen fibril structure, fiber realignment, and mechanical properties of mature tendons. *Matrix Biol* 64: 81–93, 2017. doi:10.1016/j.matbio.2017.08.004.
64. **Shao H**, **Lee S**, **Gae-Scott S**, **Nakata C**, **Chen S**, **Hamad AR**, **Chakravarti S.** Extracellular matrix lumican promotes bacterial phagocytosis, and Lum−/− mice show increased pseudomonas aeruginosa lung infection severity. *J Biol Chem* 287: 35860–35872, 2012 [Erratum in *J Biol Chem* 289: 33874, 2014]. doi:10.1074/jbc.M112.380550.
65. **Lauder MR**, **Huckerby NT**, **Nieduszynski AI**, **Plaas HKA.** Age-related changes in the structure of the keratan sulphate chains attached to fibromodulin isolated from articular cartilage. *Biochem J* 330: 753–757, 1998. doi:10.1042/bj3300753.
66. **Lauder RM**, **Huckerby TN**, **Nieduszynski IA.** The structure of the keratan sulphate chains attached to fibromodulin from human articular cartilage. *Glycoconj J* 14: 651–660, 1997. doi:10.1023/a:1018552913584.