

# SCIENTIFIC REPORTS



OPEN

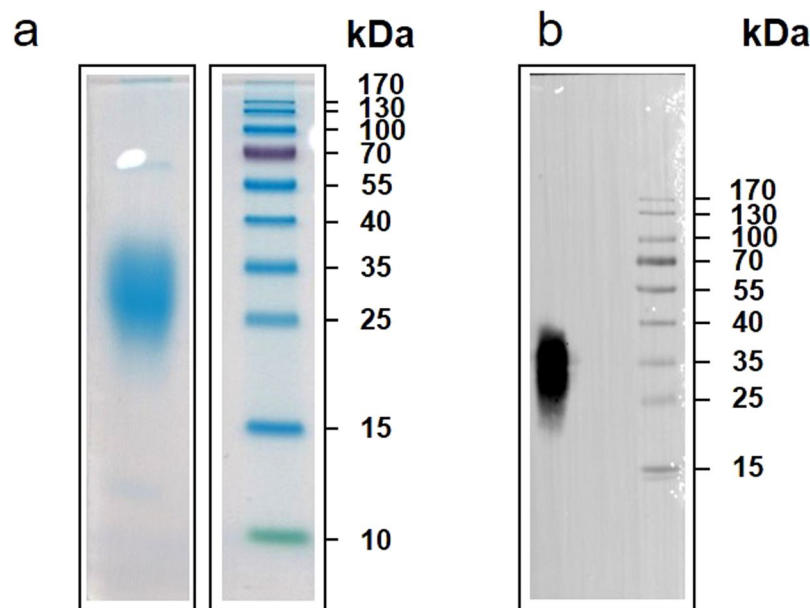
## Insights into the structure and dynamics of lysyl oxidase propeptide, a flexible protein with numerous partners

Sylvain D. Vallet<sup>1</sup>, Adriana E. Miele<sup>1</sup>, Urszula Uciechowska-Kaczmarzyk<sup>2</sup>, Adam Liwo<sup>2</sup>, Bertrand Duclos<sup>1</sup>, Sergey A. Samsonov<sup>2</sup> & Sylvie Ricard-Blum<sup>1</sup>

Lysyl oxidase (LOX) catalyzes the oxidative deamination of lysine and hydroxylysine residues in collagens and elastin, which is the first step of the cross-linking of these extracellular matrix proteins. It is secreted as a proenzyme activated by bone morphogenetic protein-1, which releases the LOX catalytic domain and its bioactive N-terminal propeptide. We characterized the recombinant human propeptide by circular dichroism, dynamic light scattering, and small-angle X-ray scattering (SAXS), and showed that it is elongated, monomeric, disordered and flexible ( $D_{\max}$ : 11.7 nm,  $R_g$ : 3.7 nm). We generated 3D models of the propeptide by coarse-grained molecular dynamics simulations restrained by SAXS data, which were used for docking experiments. Furthermore, we have identified 17 new binding partners of the propeptide by label-free assays. They include four glycosaminoglycans (hyaluronan, chondroitin, dermatan and heparan sulfate), collagen I, cross-linking and proteolytic enzymes (lysyl oxidase-like 2, transglutaminase-2, matrix metalloproteinase-2), a proteoglycan (fibromodulin), one growth factor (Epidermal Growth Factor, EGF), and one membrane protein (tumor endothelial marker-8). This suggests new roles for the propeptide in EGF signaling pathway.

Lysyl oxidase (LOX), a copper amine oxidase, catalyzes the oxidative deamination of lysine and hydroxylysine residues in collagens and elastin, which is the first step of the covalent cross-linking of these extracellular matrix (ECM) proteins<sup>1</sup>. LOX stimulates angiogenesis *in vivo*<sup>2</sup>, and in cancer, where it contributes to the formation of pre-metastatic “niche”, and to metastasis<sup>3–5</sup>. LOX is secreted as a proenzyme. It is activated by bone morphogenetic protein 1 (BMP-1), which releases the mature catalytic domain of LOX and its N-terminal propeptide. The propeptide of lysyl oxidase (LOX-PP) has various biological activities of its own<sup>6</sup>, and might thus be considered as a matricryptin, *i.e.* a bioactive ECM fragment exhibiting other functions than its parent molecule<sup>7,8</sup>. LOX-PP is required for the exit of proLOX from the endoplasmic reticulum<sup>9</sup>, and for the deposition of LOX onto elastic fibers<sup>10</sup>. It induces phenotypic reversion of ras-transformed cells<sup>11</sup>, and inhibits cell signaling and proliferation of smooth muscle cells and osteoblasts<sup>12,13</sup>. The inhibition of FGF-2 signaling by LOX-PP in NIH 3T3-L1 cells contributes to its proadipogenic effect<sup>14</sup>. LOX-PP inhibits the growth of prostate cancer cells by interacting with DNA repair proteins<sup>15</sup>. Indeed, LOX-PP has been found in the nucleus of several cell types<sup>16,17</sup>, and colocalizes with the microtubule network in osteoblastic cells<sup>17</sup>. LOX-PP interacts with intracellular or nuclear proteins, namely HSP70 and c-Raf, CIN85, UXT<sup>18–20</sup> and with the intracellular phosphatase domains of receptor-type protein tyrosine phosphatase kappa<sup>21</sup>. The major uptake pathway of recombinant LOX-PP is phosphatidylinositol-4,5-bisphosphate 3-kinase (PI3K)-dependent micropinocytosis, while a secondary pathway depends on dynamin and caveolae<sup>22</sup>. LOX-PP also interacts with two extracellular proteins, elastin<sup>10</sup> and fibronectin<sup>23</sup>, and contains cleavage sites for matrix metalloproteinases (MMPs) 2 and 10<sup>24,25</sup>. We have recently shown that LOX-PP binds to heparin<sup>26</sup> (HP).

<sup>1</sup>Univ Lyon, University Claude Bernard Lyon 1, CNRS, INSA Lyon, CPE, Institute of Molecular and Supramolecular Chemistry and Biochemistry, UMR 5246, F-69622, Villeurbanne cedex, France. <sup>2</sup>Laboratory of Molecular Modeling, Department of Theoretical Chemistry, Faculty of Chemistry, University of Gdansk, Wita Stwosza 63, 80-308, Gdańsk, Poland. Correspondence and requests for materials should be addressed to S.R.-B. (email: [sylvie.ricard-blum@univ-lyon1.fr](mailto:sylvie.ricard-blum@univ-lyon1.fr))

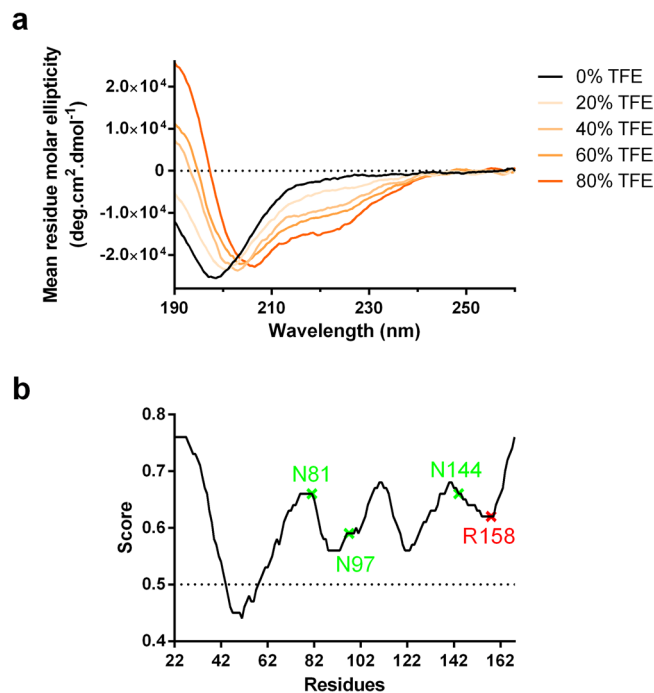


**Figure 1.** Analysis of the recombinant human propeptide of lysyl oxidase (LOX-PP) expressed in HEK293-EBNA cells. **(a)** SDS-PAGE (10% acrylamide running gel). **(b)** Western blot (anti-FLAG primary antibody<sup>51</sup>, detection by enhanced-chemiluminescence) of LOX-PP. The purification process was followed by SDS-PAGE and Western blot with a primary anti-FLAG antibody (F3165, Sigma-Aldrich<sup>51</sup>) and a secondary antibody conjugated to peroxidase (Bio-Rad, 172–1011). The gels were stained with EZBlue™ Gel Staining Reagent and were scanned with the Canon LiDE 210 scanner. For Western blots immunocomplexes were detected with SuperSignal West Pico Chemiluminescence Substrate (34080, Thermo Scientific) and visualized using the Fusion FX camera (Vilber-Lourmat) for three minutes with default settings. Pre-stained molecular weight markers were also visualized with the Fusion FX camera (Vilber-Lourmat), and superimposed on the membrane imaged by chemiluminescence with the Fusion Capt Advance software. Full-length gel and blot are presented. Lanes of the same gel were juxtaposed in **(a)**, and consecutive lanes were presented for Western blot in **(b)**.

Numerous cell studies, including those cited above, were performed with the recombinant rat propeptide, which contains both N- and O-linked carbohydrates<sup>27</sup>, and has been shown to be highly disordered by circular dichroism (CD). However, no structural data are available for the human propeptide so far. We report here the characterization of recombinant human LOX-PP by circular dichroism, dynamic light scattering (DLS), and small-angle X-ray scattering (SAXS), which showed that human LOX-PP is an elongated, monomeric, flexible, protein enriched in intrinsic disorder. We generated 3D models of LOX-PP by coarse-grained molecular dynamics (MD) simulations restrained by the distance distribution derived from SAXS measurements. LOX-PP being a heparin-binding protein<sup>26</sup>, we used these models for docking experiments with a hexasaccharide of heparin to localize the heparin-binding site. The analysis of the docked structures by all-atomic molecular dynamics and free energy showed that the hexasaccharide may interact with the propeptide *via* two regions. Furthermore, we have explored the ability of LOX-PP to interact with the ECM, and we have identified 17 new partners of LOX-PP, including four glycosaminoglycans (GAGs, chondroitin sulfate, dermatan sulfate, heparan sulfate, and hyaluronan), collagen I, cross-linking and proteolytic enzymes (lysyl oxidase-like 2, transglutaminase-2, matrix metalloproteinase-2), one proteoglycan (fibromodulin), one growth factor (Epidermal Growth Factor, EGF), and one membrane protein (Tumor Endothelial Marker-8, TEM-8, also known as anthrax receptor-1). This suggests new roles for the propeptide in ECM assembly and cross-linking, cell-matrix adhesion, and in the regulation of EGF signaling pathways.

## Results

**Expression of recombinant human LOX-PP.** Recombinant human LOX-PP expressed in Human Embryonic Kidney (HEK) 293 cells migrated with an apparent molecular weight of 30 kDa by sodium dodecyl sulfate – polyacrylamide gel electrophoresis (SDS-PAGE, Fig. 1a) although its theoretical molecular weight is 16.6 kDa. A single band was detected with an anti-FLAG antibody by Western blot (Fig. 1b). The apparent molecular weight was in agreement with the presence of glycosylation reported for the rat propeptide<sup>27</sup>. The deglycosylation of human LOX-PP by peptide N-glycosidase F (PNGase F), which removes N-linked oligosaccharides, resulted in a marked decrease in the apparent molecular weight of the human propeptide from 30 kDa to 17 kDa (Supplementary Fig. 1). This is consistent with the theoretical mass based on the amino acid sequence, and shows that N-glycans account for about 13 kDa of the molecular mass of human LOX-PP. The deglycosylated human propeptide migrated as a fuzzy and large band as previously observed for the deglycosylated rat protein<sup>27</sup>.



**Figure 2.** Secondary structure and intrinsic disorder of LOX-PP. **(a)** Averaged circular dichroism spectra ( $n = 5$ ) of LOX-PP at  $2 \mu\text{M}$  ( $60 \mu\text{g/ml}$ ) in  $10 \text{ mM}$  potassium phosphate pH 7.4. Increasing concentrations of trifluoroethanol (TFE, 20, 40, 60, and 80%) were added to assess the ability of LOX-PP to fold into  $\alpha$  helices. **(b)** Prediction with metaPrDOS of disordered residues in the sequence of LOX-PP (residues 22–168). The three glycosylated asparagine residues are represented in green and the arginine residue 158 in red, R158Q being a loss-of-function polymorphism described in breast cancer patients<sup>42</sup>.

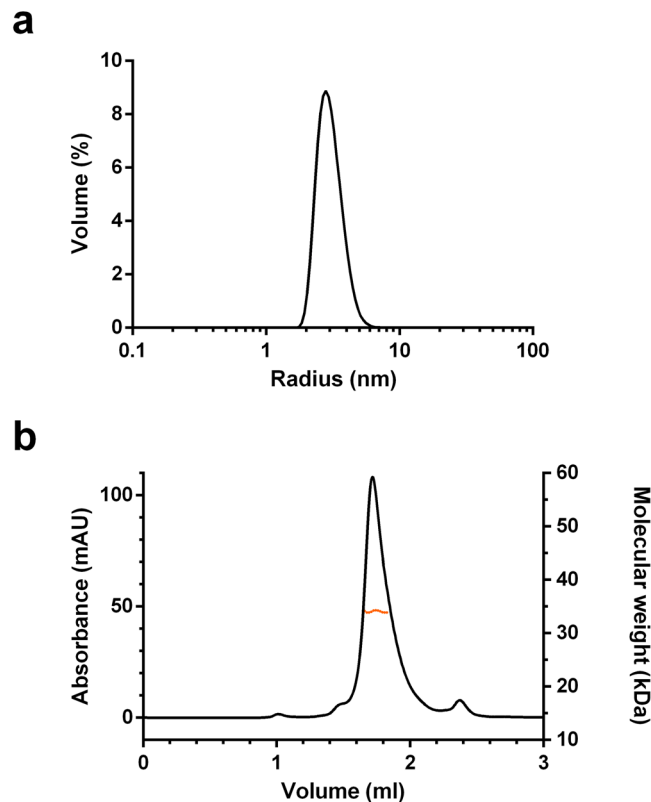
**Intrinsic disorder and secondary structure of human LOX-PP.** Circular dichroism spectra of LOX-PP showed a single minimum near 200 nm, which is characteristic of an intrinsically disordered protein (IDP) (Fig. 2a). The content in secondary structure of LOX-PP was calculated by deconvoluting the spectra with CONTIN-LL. The content in  $\alpha$ -helix and  $\beta$ -sheet was found to be 3.4% and 20.4% respectively, whereas the turn content was 11.8%, and the disorder 64.5%.

The amount of intrinsic disorder, predicted by metaPrDOS, was found to be 81% in the propeptide sequence alone and 86.4% in the context of the full-length LOX protein using the two-state prediction results with a given false positive of 5%. This suggests that the propeptide of LOX does not fold in presence of the catalytic domain of LOX. The three glycosylated asparagine residues and the arginine residue 158 (R158Q being a loss-of-function polymorphism) are in disordered regions, whereas the sequence encompassing residues 44 to 58 (QQIQWENNGQVFSLL) was predicted to be structured (Fig. 2b).

LOX-PP was analyzed in the presence of increasing concentrations of trifluoroethanol (TFE) (0–80%) to assess its ability to fold into  $\alpha$ -helices (Fig. 2a). The content in  $\alpha$ -helix of LOX-PP increased with TFE concentrations, ranging from 3.4% up to 25.2% at 80% TFE. The addition of full-length HP, a binding partner of LOX-PP, or of a HP hexasaccharide did not induce significant changes in the CD spectra of LOX-PP, either immediately after GAG addition or after a 60-min incubation (Supplementary Fig. 2).

**Analysis of LOX-PP by dynamic light scattering.** Four series of fifteen measurements were performed and gave a single peak with a polydispersity index of  $0.817 \pm 0.01$ . The hydrodynamic radius ( $R_h$ ) was calculated to be  $3.05 \pm 0.07 \text{ nm}$  (Fig. 3a) with an estimated molecular weight of  $55 \pm 3 \text{ kDa}$ , which might be due to the intrinsic disorder, flexibility and extensive glycosylation of LOX-PP.

**Analysis of LOX-PP by small-angle X-ray scattering.** The shape of human LOX-PP in solution was studied on the BioSAXS BM29 beamline (European Synchrotron Radiation Facility, ESRF, Grenoble, France). 1D SAXS profiles showed no sign of aggregation as can be seen on the logarithmic plot (Fig. 4a). LOX-PP molecular weight calculated with autoprocessing framework EDNA ranged between 16 and 22 kDa, *a priori* consistent with a monomeric species and the presence of intrinsic disorder in LOX-PP as predicted by sequence analysis and confirmed by CD. The molecular weight of LOX-PP calculated with the SAXS Molecular Weight package (SAXSMoW<sup>28</sup>, <http://saxs.ifsc.usp.br/>,  $s_{\text{max}}$  value of  $8/R_g$ ) was found to be 29.4 kDa. Size Exclusion Chromatography - Multi-Angle Light Scattering (SEC-MALS) analysis showed that glycosylated LOX-PP eluted as a single peak from a size-exclusion column with a molecular mass of 34.1 kDa (mass fraction: 97%, Fig. 3b). Both values are consistent with the apparent molecular weight calculated by SDS-PAGE ( $\sim 30 \text{ kDa}$ ) and show that glycosylated LOX-PP is a monomer.



**Figure 3.** Analysis of LOX-PP by DLS and SEC-MALS. **(a)** Size distribution analysis of LOX-PP by dynamic light scattering in HBS at 20°C (LOX-PP: 35  $\mu$ M, 1.1 mg/ml,  $4 \times 15$  curves were averaged). **(b)** LOX-PP (75  $\mu$ M) in HBS was injected on a Superdex 200 Increase 5/150 GL column (GE Healthcare) for SEC-MALS analysis including UV absorbance at 280 nm, multi-angle light scattering at 690 nm and refractometry.

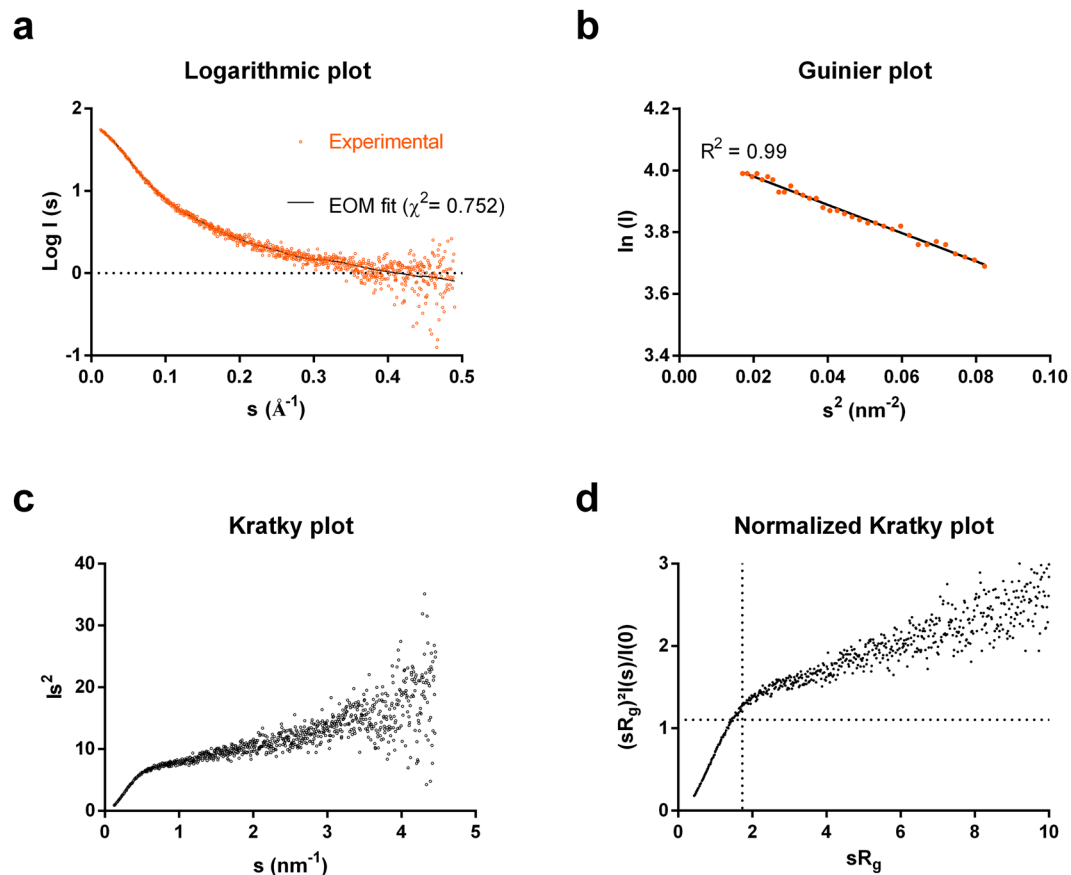
The radius of gyration ( $R_g$ ), determined with PRIMUS, was  $3.7 \pm 0.04$  nm (Fig. 4b), which is of the same order of magnitude than the theoretical  $R_g$  (3.53 nm) calculated with Flory's equation, and the hydrodynamic radius (3.05 nm).

The experimental  $R_g/R_h$  ratio (1.2) was higher than 0.7, which indicates that LOX-PP behaves as a non-globular protein, and is likely to be extended<sup>29</sup>, which is consistent with the presence of intrinsic disorder in LOX-PP. Furthermore, the Kratky plot was indicative of a non-globular natively unfolded molecule (Fig. 4c), and the normalized Kratky plot confirmed that LOX-PP is indeed flexible (Fig. 4d). Experimental maximum dimension ( $D_{max}$ ) calculated from distance distribution function was 11.7 nm.

LOX-PP was also analyzed by SEC-SAXS on the SWING beamline (Small and Wide angle X-ray scattering, French National Synchrotron Facility SOLEIL, Saint-Aubin, France) to investigate the possible influence of trace amount of aggregation on the parameters calculated by SAXS. The calculated  $R_g$  values were 3.72 nm and 3.70 nm for SEC-SAXS and SAXS respectively, and those calculated for  $D_{max}$  were 11.8 nm and 11.7 nm for SEC-SAXS and SAXS, respectively (Supplementary Fig. 3). The values of  $R_g$  and  $D_{max}$  calculated for the propeptide by both SAXS and SEC-SAXS were thus nearly identical, showing that those calculated by SAXS were not affected by aggregation.

Ensemble modelling (Ensemble Optimization Method, EOM 2.1) was used to identify the sub-ensembles of conformations of LOX-PP, which best fitted to SAXS experimental data. The best fit ( $\chi^2$ : 0.752, Fig. 4a) was obtained by a combination of six conformations of LOX-PP with an averaged  $R_g$  of 3.85 nm and  $D_{max}$  of 11.7 nm (Fig. 5a,b).  $R_{flex}$ , which allows to quantify the difference between flexible and rigid systems<sup>30</sup>, was higher for the sub-ensemble of conformers coexisting in solution fitting to SAXS data than for the random pool of conformers covering the available conformational space (89.87% and 84.98% respectively), and  $R_e$  was  $>1$  (1.43). These results predicted the co-existence of several flexible conformers of LOX-PP in solution with a 2-fold variation in the  $D_{max}$  value, from 8.6 to 18.9 nm (Fig. 5c). However, these conformers correspond to the amino acid sequence of LOX-PP only, and not to its glycosylated form because glycans cannot be included in EOM modeling.

**Coarse-grained models of the propeptide of lysyl oxidase.** Using the SAXS data-restrained UNRES/MREMD (United Residue Model of Proteins/ Multiple Replica Exchange Molecular Dynamics) modeling we obtained five three-dimensional (3D) models of LOX-PP (Fig. 6). The probability of model 1 (0.4) was significantly higher than those of the other models. Models 2, 3, and 4 had similar probabilities (0.21, 0.19, 0.16), while the probability of model 5 was very low (0.04). The N-termini of models 1 and 3 were folded and formed a globule, enriched in  $\beta$ -strands in model 3. Both N- and C-termini were folded in models 2 and 4, whereas in model 5

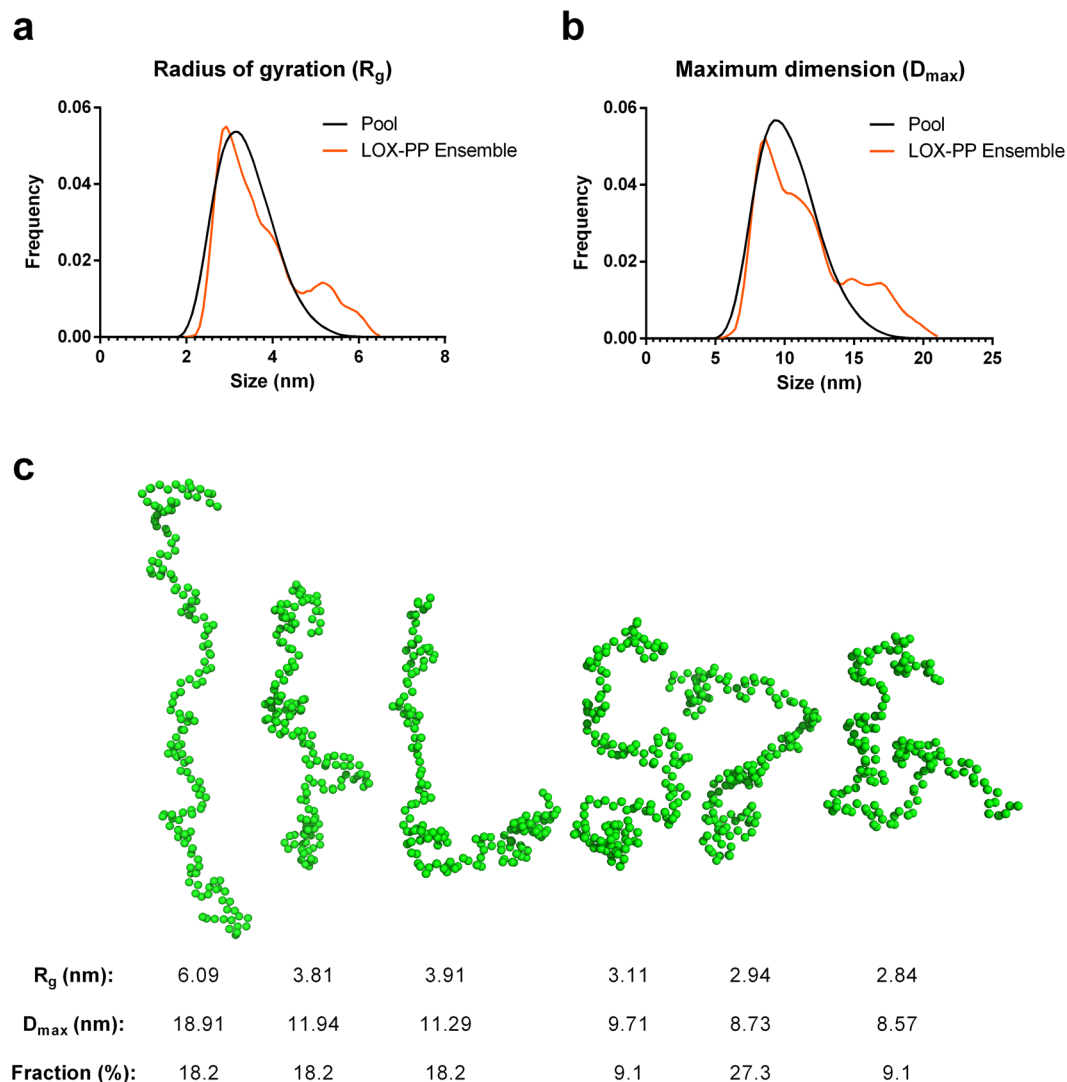


**Figure 4.** Small-angle X-ray scattering analysis of LOX-PP. (a) Logarithmic plot of experimental data. Fit of EOM 2.1 data to experimental SAXS data ( $\chi^2 = 0.752$ ). (b) Guinier plot of experimental data collected for LOX-PP in the range used to calculate the radius of gyration ( $R_g = 3.7 \pm 0.04$  nm) compared to the theoretical fit. Kratky plot (c) and Normalized Kratky plot of LOX-PP (d).

only the C-terminal part appeared to be globular. The central part of the propeptide was thus less structured than the termini, and part of it adopted an  $\alpha$ -helical conformation in models 3 and 5 (Fig. 6). The radius of gyration of model 1 was 2.93 nm, which was in good agreement with SAXS and DLS data ( $\sim 3$  nm), while the radii of gyration of models 2–5 ranged from 2.56 to 2.74 nm (2.74 nm, model 2; 2.58 nm, model 3; 2.73 nm, model 4; 2.56 nm, model 5). The  $D_{\max}$  value calculated for the five models (9.9, 8.1, 9.1, 7.9, and 8.6 nm) was lower than the experimental SAXS value (11.7 nm, Supplementary Fig. 4). It should be noted that the above models did not take into account the glycosylation of LOX-PP, whereas SAXS experiments were carried out with glycosylated LOX-PP. We generated theoretical 1D SAXS profile for each model using the  $C\alpha$ - $C\alpha$  distances and the Gaussian smoothing as described in our recent work<sup>31</sup>. The models that best fitted the experimental 1D SAXS profile were models 1 and 2 with normalized  $\chi^2$  values being 0.48 and 0.51 respectively. These values were 0.72, 0.59, and 0.88 for models 3, 4, and 5 respectively. These conformations were selected by the analysis, where the restraints were imposed on a single conformation and not on the ensemble of conformations simultaneously.

The secondary structures of all models were determined from the corresponding coordinate (.pdb) files. Models 1, 3 and 5 had the highest content of secondary structures (Table 1).  $3_{10}$ -helices and  $\beta$ -strands were found in the five models with a significant variation in  $3_{10}$ -helix content depending on the model. Topological diagrams showed that all models contained a few amount of secondary structures (4 to 9  $\beta$  strands and 1 to 4 helices, Supplementary Fig. 5).

**Heparin docking on LOX-PP and MD simulations.** We have previously shown that LOX-PP interacts with 6-kDa HP with an affinity of 190 nM<sup>26</sup> but the localization of the binding site of HP on LOX-PP is unknown. We have thus performed docking experiments with the 3D models of LOX-PP and a HP oligosaccharide in order to localize putative binding site(s) in LOX-PP sequence, and to determine if the oligosaccharide binding was able to induce LOX-PP folding. Poisson-Boltzmann surface area (PBSA) electrostatic potential calculations carried out for each model showed that arginine residues crucial for the positive electrostatic potential of LOX-PP were concentrated either at the N-terminus (models 1–3) or at the C-terminus (models 4–5). According to the computational analysis of all protein-GAG structures available in the Protein Data Bank (PDB, <https://www.rcsb.org/>) the regions with the positive electrostatic potential on the protein surface represent putative binding regions for negatively charged GAG molecules<sup>32</sup>. Therefore, we carried out docking simulations for each model of LOX-PP,



**Figure 5.** SAXS analysis by Ensemble Optimization Method (EOM). Distribution of  $R_g$  (a) and  $D_{max}$  (b) values for the generated pool and LOX-PP ensemble. (c) Beads models of LOX-PP built and selected by EOM to fit experimental data, presented by decreasing  $D_{max}$  value (Fraction %: fractions of occupancy).

sampling HP dp6 conformations in these regions. Contrasting results were obtained for LOX-PP/HP complexes depending on the model used for docking.

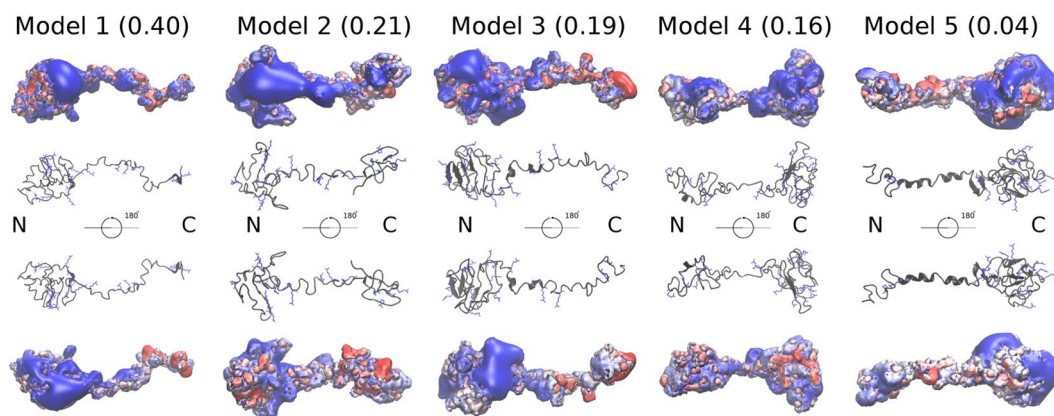
Representative clusters of docking solutions were either overlapping or localized in the same region of the propeptide (Fig. 7). A single, large, cluster, was obtained for models 2 and 5, two for models 1 and 3, and four for model 4. Both clusters were aligned but partially overlapped in model 3, whereas they were aligned end-to-end in model 1, forming an elongated binding site able to bind a longer chain of heparin. These data further indicate a high probability for HP multipose binding, which was previously computationally predicted and experimentally confirmed for other protein-GAG complexes<sup>33–35</sup>. HP predicted binding sites appeared thus to be mostly localized in globular regions, at the N-terminus for models 1–3 and at the C-terminus for models 4–5.

For each obtained cluster of docking solutions, we carried out all-atom MD simulations arbitrarily choosing three ligand poses from each cluster. In total, 30 different LOX-PP/HP complexes were analyzed. Again, our data suggest very heterogeneous nature of this interaction both for the free energy of binding and the radius of gyration changes upon GAG binding (Table 2). In most cases, the free energy of binding became more favorable during the simulation, suggesting that further contacts were established to tether LOX-PP/HP binding. The lowest energy values were obtained for models 1 and 2 complexed to HP dp6 (degree of polymerization), which suggests that they bound to the hexasaccharide with a higher affinity than the other models.

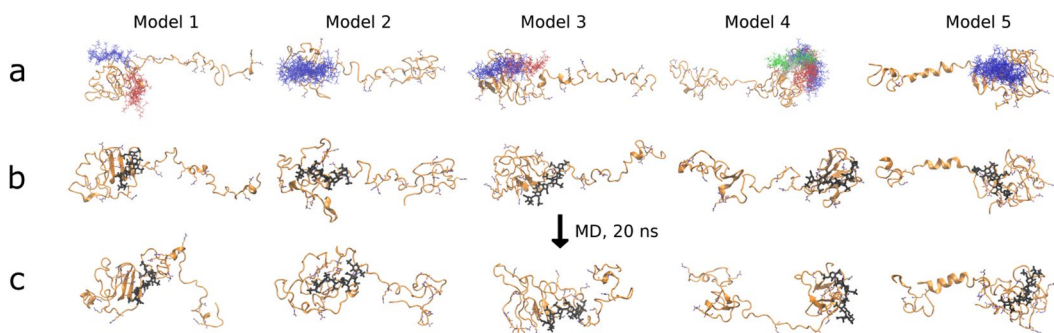
We performed 20-ns all-atom MD simulations to determine if the binding of the heparin hexasaccharide induced a conformational change in LOX-PP (Supplementary Table 1). Although a trend towards a decrease in  $R_g$  values was observed, the difference did not reach statistical significance. Furthermore, the global decrease in  $R_g$  of LOX-PP during simulation may also result from spontaneous collapsing of LOX-PP due the force field used in this study as shown by MD simulations performed in the same conditions on all the models of LOX-PP alone

LOX-PP models	$\beta$ strand (%)	$3_{10}$ helix (%)	Others (%)
1	9.6	11	79.4
2	6.8	6.2	87
3	9.6	12.3	78.1
4	11	3.4	85.6
5	13	11.6	75.3

**Table 1.** Secondary structure content of LOX-PP models determined with PDBsum from their coordinate (.pdb) files.



**Figure 6.** Models of LOX-PP obtained by coarse-grained MD simulations with UNRES. Their probability based on free energy calculations is indicated in brackets. Arginine residues are in blue sticks. The calculated electrostatic potential is represented by isosurfaces (in blue: positive electrostatic potential isosurface 3 kcal/mol·Å, in red: negative electrostatic potential isosurface  $-1$  kcal/mol·Å). Different absolute isovalues were selected for negative and positive potentials due to the high positive charge of the protein (18 arginine residues out of 147).



**Figure 7.** Docking and molecular dynamics of the heparin hexasaccharide/LOX-PP complex. **(a)** Clusters of docking solutions for heparin hexasaccharide binding to LOX-PP models obtained after the conversion from coarse-grained models to all atom representations (blue, red, green and cyan from the biggest cluster to the smallest one). The clusters were obtained from the top 50 docking solutions scored by AD3 with the DBSCAN algorithm (neighborhood search radius of 4 Å and minimal number of cluster members of 5). Structure of the LOX-PP/heparin hexasaccharide complex before **(b)** and after 20 ns all-atom MD simulations **(c)**, corresponding to the lowest free energies for each model (Supplementary Table 3). Arginine residues: thin sticks, HP hexasaccharide: black thick sticks (bottom).

(Supplementary Table 2). Consequently, the above data suggest that a 20-ns production simulation may be not long enough to allow statistically meaningful comparison of  $R_g$  values of LOX-PP upon heparin binding.

In addition, there were no significant differences (or obvious trends) in the secondary structure ( $\alpha$ -helix and  $\beta$ -strand) of LOX-PP complexed with the HP hexasaccharide between 10-ns and 20-ns MD simulations in all the models (Supplementary Table 3), which is in agreement with CD spectra.

Molecular Mechanics Generalized Born and Surface Area (MM-GBSA) free energy per residue for LOX-PP/HP complexes was calculated from 30 independent MD simulations. Binding energy contributions ranged from

Model	Cluster	$\Delta G$ (kcal/mol) at 10 ns	$\Delta G$ (kcal/mol) at 20 ns	$R_g$ (nm) at 10 ns	$R_g$ (nm) at 20 ns
1	1	$-110.4 \pm 11.7$	$-128.6 \pm 6.4$	$1.92 \pm 0.04$	$1.78 \pm 0.03$
		$-71.1 \pm 6.5$	$-116.1 \pm 7.1$	$2.62 \pm 0.05$	$2.05 \pm 0.04$
		$-80.3 \pm 9.7$	$-133.0 \pm 7.9$	$2.05 \pm 0.03$	$1.96 \pm 0.01$
	2	$-106.8 \pm 10.6$	$-141.6 \pm 11.5$	$2.44 \pm 0.04$	$2.68 \pm 0.06$
		$-138.8 \pm 9.7$	$-162.9 \pm 8.4$	$2.87 \pm 0.07$	$2.27 \pm 0.03$
		$-89.3 \pm 8.8$	$-114.1 \pm 8.3$	$3.21 \pm 0.06$	$3.11 \pm 0.05$
2	1	$-120.4 \pm 10.6$	$-141.3 \pm 7.2$	$2.25 \pm 0.05$	$2.50 \pm 0.08$
		$-98.2 \pm 8.9$	$-109.1 \pm 9.9$	$2.34 \pm 0.05$	$2.16 \pm 0.02$
		$-95.3 \pm 8.2$	$-90.8 \pm 6.6$	$2.11 \pm 0.04$	$2.31 \pm 0.04$
3	1	$-92.2 \pm 7.6$	$-70.30 \pm 7.5$	$2.22 \pm 0.08$	$2.02 \pm 0.04$
		$-55.7 \pm 5.1$	$-72.6 \pm 8.3$	$2.28 \pm 0.03$	$1.92 \pm 0.02$
		$-74.3 \pm 7.8$	$-78.2 \pm 7.1$	$1.91 \pm 0.03$	$1.77 \pm 0.03$
	2	$-74.3 \pm 8.3$	$-85.4 \pm 10.0$	$2.47 \pm 0.05$	$2.39 \pm 0.03$
		$-75.8 \pm 7.4$	$-103.9 \pm 8.1$	$2.09 \pm 0.09$	$2.15 \pm 0.05$
		$-73.3 \pm 7.4$	$-81.5 \pm 6.9$	$2.20 \pm 0.03$	$2.25 \pm 0.04$
4	1	$-43.6 \pm 4.7$	$-70.9 \pm 6.6$	$3.19 \pm 0.03$	$3.26 \pm 0.02$
		$-81.4 \pm 5.6$	$-73.0 \pm 5.8$	$2.37 \pm 0.04$	$1.90 \pm 0.03$
		$-78.4 \pm 6.5$	$-75.0 \pm 6.2$	$2.24 \pm 0.06$	$2.07 \pm 0.03$
	2	$-74.9 \pm 5.6$	$-81.0 \pm 6.2$	$2.96 \pm 0.03$	$2.94 \pm 0.04$
		$-54.4 \pm 7.8$	$-56.0 \pm 5.2$	$2.03 \pm 0.06$	$2.24 \pm 0.07$
		$-86.1 \pm 9.5$	$-71.6 \pm 6.6$	$3.05 \pm 0.05$	$3.30 \pm 1.00$
	3	$-71.3 \pm 5.8$	$-75.8 \pm 9.7$	$2.31 \pm 0.03$	$2.28 \pm 0.08$
		$-60.2 \pm 6.2$	$-85.8 \pm 6.9$	$2.16 \pm 0.05$	$2.17 \pm 0.04$
		$-77.4 \pm 6.0$	$-91.3 \pm 10.3$	$2.45 \pm 0.05$	$2.79 \pm 0.02$
		$-82.6 \pm 7.7$	$-82.1 \pm 6.7$	$3.21 \pm 0.05$	$3.38 \pm 0.08$
	4	$-45.4 \pm 3.8$	$-66.9 \pm 4.9$	$3.28 \pm 0.03$	$3.30 \pm 0.05$
		$-80.7 \pm 5.6$	$-85.0 \pm 6.8$	$3.28 \pm 0.03$	$2.84 \pm 0.03$
		$-61.1 \pm 7.4$	$-60.8 \pm 5.7$	$2.13 \pm 0.04$	$2.06 \pm 0.05$
		$-50.9 \pm 7.3$	$-71.0 \pm 5.9$	$1.99 \pm 0.06$	$1.83 \pm 0.02$
5	$-62.3 \pm 5.7$	$-71.3 \pm 5.9$	$2.25 \pm 0.04$	$2.12 \pm 0.06$	

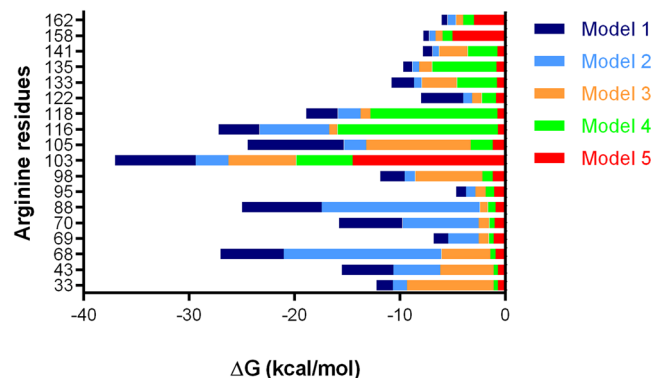
**Table 2.** MM-GBSA free energies of binding for LOX-PP/HP complexes after 10 and 20 ns of the simulation. Clusters were selected according to the procedure described in the Material and Methods section ( $n = 3$ ).

0.37 to 4.18 kcal/mol for acidic residues, from  $-0.37$  to  $-15.24$  kcal/mol for arginine residues, and from  $-0.16$  to 0.05 kcal/mol for the only histidine residue (Supplementary Table 4). These values are within the range of those calculated using this approach for protein-GAG complexes of the PDB (average total energy:  $-74.3 \pm 75.2$  kcal/mol)<sup>32</sup>.

The human propeptide does not contain lysine residues or any of the three consensus heparin-binding sites (XBBXB, XBBXXB, and XBBXXBBXXBBX, where B is a basic residue and X is a hydrophobic residue<sup>36</sup>). These data showed the crucial role of several arginine residues in the hexasaccharide binding, which were the only basic residues of LOX-PP. Two binding sites were determined, a major one encompassing the R103-R118 sequence for models 1–4 (R103 being the most frequent contributor), and another one encompassing the R68-R88 sequence mostly in models 1 and 2 (Fig. 8 and Supplementary Fig. 6). One out of the four regions of LOX-PP predicted by ANCHOR to fold upon partner binding (38–65, 74–114, 118–139, 149–168) encompassed the major part of both predicted binding sites. The regions of LOX-PP predicted to fold by MoRFchibi SYSTEM (37–61, 63–70, 84–132, and 153–158) also overlapped with the predicted binding sites.

**New binding partners: the interaction network of human LOX-PP.** To determine if LOX-PP was able to bind to other partners we first investigated if it interacted with other GAGs and to ECM proteins. Indeed, most partners of LOX-PP identified at the time we started this work were intracellular. The partners of LOX-PP reported in the literature are listed in Supplementary Table 5. Potential partners of LOX-PP were screened by Surface Plasmon Resonance imaging (SPRi), SPR, and Bio-Layer Interferometry (BLI) (Supplementary Tables 6 and 7). We have identified 17 new partners of LOX-PP including four GAGs (chondroitin sulfate, dermatan sulfate, heparan sulfate, hyaluronan), collagen I, cross-linking and proteolytic enzymes (lysyl oxidase-like 2, transglutaminase-2, and MMP-2), one proteoglycan (fibromodulin), one matricryptin (anastellin), and the ectodomain of one membrane protein (Tumor Endothelial Marker-8 also known as anthrax receptor-1<sup>37</sup>). A single representative evidence of each interaction is displayed in Supplementary Figs 7 and 8. The interaction network of LOX-PP integrating partners curated from the literature and available in MatrixDB database<sup>38</sup> (<http://matrixdb.univ-lyon1.fr/>) is displayed in Fig. 9a. All the interactions identified in this study will be publicly and freely available in MatrixDB database. Most LOX-PP partners belong to the extracellular matrix (42.5%) or are associated with it (21.2%), whereas 30.3% of them are intracellular and 6.1% are membrane proteins (Fig. 9b). Moreover,





**Figure 8.** MM-GBSA free energy for all arginine residues of LOX-PP models bound to the heparin hexasaccharide.

LOX-PP bound with very high affinity to tropoelastin and to plasminogen, and with high affinity to anastellin, a fragment of fibronectin (Table 3, Supplementary Fig. 8b).

## Discussion

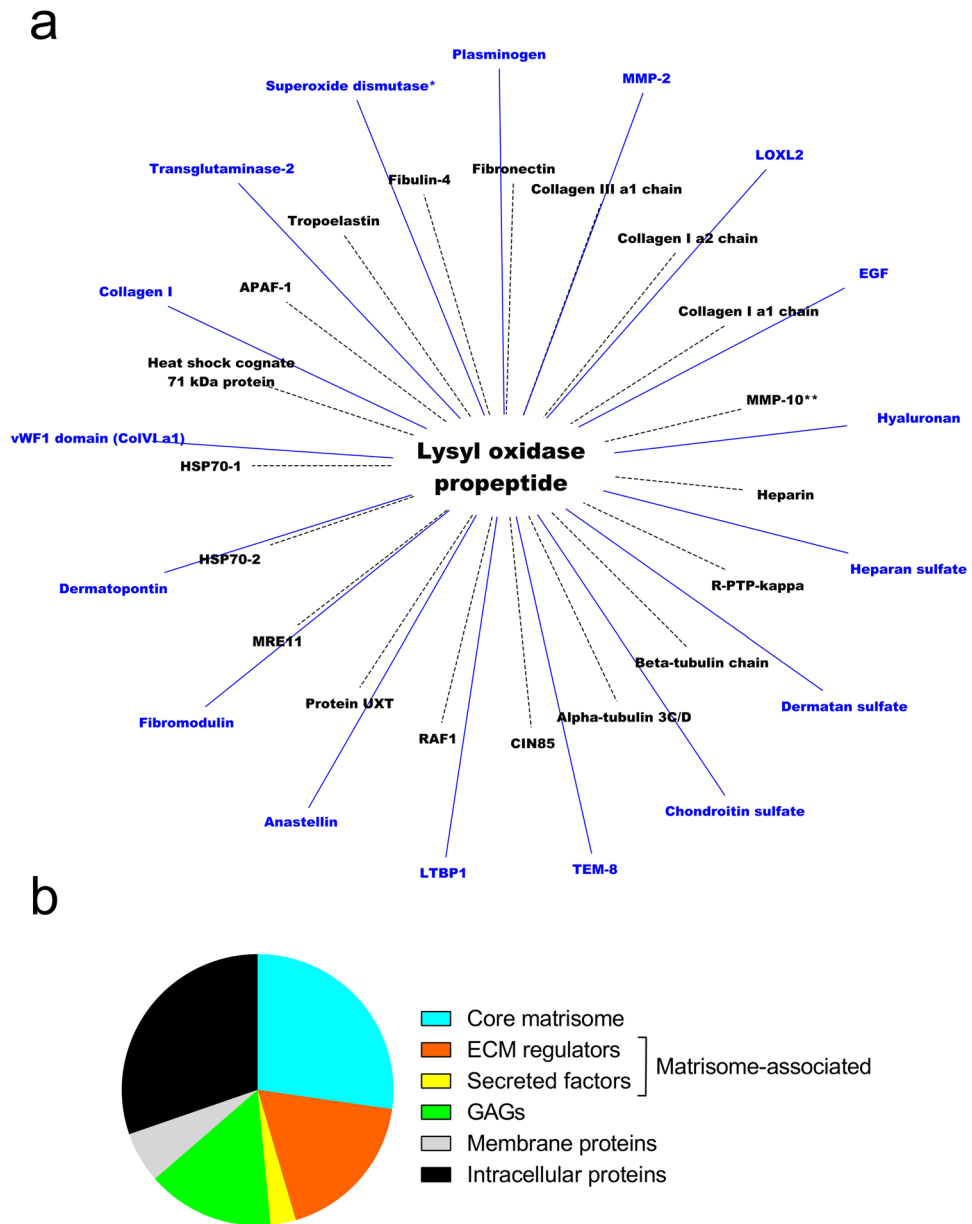
Human LOX-PP is predicted to contain more than 80% of disordered residues, which is similar to the value reported for the rat LOX-PP (67% using another predictor<sup>27</sup>). The presence of intrinsic disorder has been confirmed experimentally by CD and SAXS experiments. However, LOX-PP is partially structured as shown by deconvolution of CD spectra, which demonstrated the presence of a low amount of  $\alpha$ -helix and  $\beta$ -strands. The 3D models of the propeptide are consistent with the fact that the propeptide is not entirely disordered. Indeed, all models (except for model 5, which has the lowest probability score) contain secondary structures mostly located at the N- and C-termini and a disordered central part. The most probable models (1–3) contain a globule-like domain at the N-terminus, which is consistent with the fact that the amino acid sequence 44–58 is predicted to be structured. This sequence is located in a globular region in models 1–4, and in the N-terminal helical part in model 5. The most probable model (model 1) comprises 11% of  $3_{10}$ -helix and 9.6% of  $\beta$ -strands compared to 3.4% of  $\alpha$ -helix and 20.4% of  $\beta$ -strands determined by CD experiments. N-glycosylated<sup>39</sup> (N81, N97, and N144) and putative O-glycosylated<sup>40</sup> residues (T89, T113, T120, S130, T131, S132) are located in the disordered region of human LOX-PP, as expected for glycosylation. Indeed, the post-translational modification glycosylation has been shown to be strongly correlated with predicted disorder<sup>41</sup>. The combination of DLS and SAXS data shows that LOX-PP is an elongated, flexible, protein with a  $D_{\max}$  of 11.7 nm and a  $R_g$  of 3.7 nm, which is consistent with the  $R_g$  value calculated for model 1 (2.93 nm). The flexibility of LOX-PP is further emphasized by the predicted existence in solution of several conformers with a 2-fold variation in the  $D_{\max}$  value, and the normalized Kratky plot.

LOX-PP can adopt an  $\alpha$ -helical structure in presence of trifluoroethanol as reported for the rat propeptide<sup>27</sup>. This prompted us to determine if the propeptide was able to fold upon HP binding. No significant change was observed in CD spectra, and hence in the secondary structure of LOX-PP in presence of full-length HP or of a hexasaccharide. Docking experiments with a hexasaccharide and MD simulations with the LOX-PP models generated by SAXS-data-restrained coarse-grained simulations also failed to show any significant changes in LOX-PP secondary structure and conformation.

The models 1 and 2, which are the most probable ones according to coarse-grained simulations, bind stronger to the HP hexasaccharide. The top 10 arginine residues involved in HP binding (R116, R103, R105, R118, R68, R88, R122, R133, R135, and R43 in decreasing order of their contribution) are located in disordered regions of the propeptide, which are predicted to fold upon binding both by ANCHOR and MoRFchibi SYSTEM. The formation of  $\alpha$ -helix observed in the presence of trifluoroethanol could thus occur in the HP-binding region located between R103 and R122.

The 3D models of LOX-PP bound to the hexasaccharide show that the sequence 26–100, which is involved in the interaction of LOX-PP with c-Raf and HSP70<sup>18</sup>, is globular and contains two accessory arginine residues (R68, R88), which may contribute to heparin binding. The cleavage sites of LOX-PP by MMP-2 (Asn150-Leu151 in rat, and Asn156-Leu157 in human) and BMP-1<sup>24</sup> (Gly168-Asp169) and the arginine residue R158, crucial for the adipogenic effect of LOX-PP, are not located in the HP-binding region. This residue plays an important role in LOX-PP functions because its mutation in glutamine inhibits the adipogenic effect of LOX-PP<sup>14</sup> and impairs the ability of LOX-PP to inhibit the invasive phenotype and tumor formation of breast cancer cells<sup>42</sup>. In contrast to the above residues, the cleavage site of MMP-10 (Ser56-Leu57)<sup>25</sup>, and P111 and R116, which are required for LOX-PP binding to CIN85<sup>19</sup>, are in the HP-binding region. CIN85 being a cytoplasmic protein it is unlikely that HP could interfere with its binding to LOX-PP *in vivo*. Furthermore, LOX-PP is predicted to bind to the SUMO protein, which contributes to DNA lesion repair in mammalian cells<sup>43</sup>, via the PILL sequence (residues 90–93) with a score of 0.795 using JASSA<sup>44</sup> (<http://www.jassa.fr/index.php?m=jassa>). This potential interaction may play a role in LOX-PP binding to nuclear proteins.

We report here that LOX-PP interacts with chondroitin, dermatan and heparan sulfate and with hyaluronan. It may thus interact with cell surface proteoglycans such as syndecans and glypicans, and contribute to cell-matrix interactions. In addition, LOX-PP has intracellular and extracellular protein partners and appears



**Figure 9.** Interaction network of human LOX-PP. (a) Protein-protein and protein-GAG interaction network of LOX-PP. Previously identified partners are linked to LOX-PP by dashed edges, whereas those identified in this study are indicated by solid edges. \*Protein from *Leishmania major*. \*\*A direct interaction between MMP-10 and LOX-PP has not been demonstrated but the sequence cleaved by MMP-10 has been located in the propeptide<sup>25</sup>. APAF-1: apoptotic protease-activating factor 1, CIN85: Cbl-interacting protein of 85 kDa (SH3 domain-containing kinase-binding protein 1), MRE11: double-strand break repair protein MRE11, Protein UXT: Ubiquitously expressed transcript protein, R-PTP-kappa: Receptor-type tyrosine-protein phosphatase kappa, RAF1: proto-oncogene serine/threonine-protein kinase, TEM-8: tumor endothelial marker-8). (b) Pie chart representation of LOX-PP partners based on Naba's ECM classification<sup>84</sup>. Collagen I trimer, but not its isolated  $\alpha$  chains, was included in the chart.

to be a moonlighting protein fulfilling different roles inside and outside the cells. LOX-PP binds to ECM proteins, namely collagen I, the  $\alpha 1$  chain of collagen VI *via* its von Willebrand Factor (vWF) domain 1, the small leucine-rich proteoglycan fibromodulin, fibronectin and its anastellin fragment, and Latent-Transforming growth factor beta Binding Protein-1 (LTBP-1) *via* its EGF domains. LOX-PP likely participates in collagen fibrillogenesis in association with fibromodulin, in fibronectin supramolecular assembly and in elastic fiber formation. Several LOX-PP partners such as fibrillar collagen I<sup>45</sup>, elastin<sup>10</sup>, fibronectin<sup>23</sup>, dermatopontin<sup>46,47</sup>, and fibromodulin<sup>48</sup> bind to mature LOX, suggesting that both mature LOX and the propeptide are involved in ECM assembly and organization. LOX-PP is sufficient to promote the binding of prolyl oxidase to MMP-2, which has been previously reported<sup>24</sup>, and it interacts with the cross-linking enzymes lysyl oxidase-like 2 (LOXL-2), and transglutaminase-2. LOX-PP contains 14 glutamine residues, and may be a glutaminyl substrate of transglutaminase-2. The interaction

Immobilized partner	Analyte	ka	kd (s <sup>-1</sup> )	K <sub>D</sub> (nM)	Model
Anastellin	LOX-PP	2.93 × 10 <sup>3</sup> M <sup>-1</sup> s <sup>-1</sup>	4.45 × 10 <sup>-4</sup>	152	1:1 (BLI)
Tropoelastin	LOX-PP	3.31 × 10 <sup>2</sup> M <sup>-1</sup> s <sup>-1</sup> 3 × 10 <sup>-3</sup> s <sup>-1</sup>	4.3 × 10 <sup>-3</sup> 3.1 × 10 <sup>-7</sup>	1.39	2-state (SPR)
LOX-PP	Plasminogen	4.17 ± 0.74 × 10 <sup>4</sup> M <sup>-1</sup> s <sup>-1</sup> 2.14 ± 0.05 × 10 <sup>-3</sup> s <sup>-1</sup>	5.15 ± 0.21 × 10 <sup>-3</sup> 3.06 ± 0.16 × 10 <sup>-4</sup>	15.8 ± 1.75	2-state (SPR) n = 2

**Table 3.** Kinetic parameters, equilibrium dissociation constant (K<sub>D</sub>), and binding models of LOX-PP interactions with its partners calculated from Bio-Layer Interferometry (BLI) and Surface Plasmon Resonance (SPR) binding assays.

of the propeptide with the pro-angiogenic growth factor EGF and with TEM-8, which is overexpressed at the surface of activated endothelial cells, suggests its possible involvement in angiogenesis, and possible interference with EGF signaling pathway. Given that TEM-8 is one of the receptors of the anthrax toxin and that it binds to live *Leishmania* parasites<sup>49</sup>, LOX-PP might also contribute to the interactions between the host ECM and pathogens. Proteins containing EGF, vWF and fibronectin domains, found in LOX-PP partners, are potential partners of the propeptide, and are currently investigated to increase the coverage of the propeptide interaction network.

## Material and Methods

**Expression of recombinant LOX-PP.** LOX-PP (UniProtKB identifier of LOX: P28300, UniProtKB pro feature of LOX-PP PRO\_0000018520 encompassing residues 22–168 with the substitution P24L) was expressed with a C-terminal FLAG tag in human embryonic kidney 293 Epstein-Barr Nuclear Antigen cells (HEK293-EBNA) using the pCEP-Pu-BM40 vector<sup>50</sup>. Cells were selected with puromycin (Sigma-Aldrich, P9620, 0.5 µg/ml) and expanded in Dulbecco's Modified Eagle's medium (DMEM) high glucose medium (Sigma-Aldrich, D5796) containing 50 µg/ml of gentamicin (Sigma-Aldrich, G1272). Culture media were harvested every 48 h for fourteen days. cOmplete™ ethylenediamine tetraacetic acid (EDTA)-free Protease Inhibitor (Roche) and 0.1 mM phenylmethylsulfonyl fluoride were added to the medium which was then centrifuged at 14 000 × g for 25 min at 4 °C and stored at –80 °C until purification. The protein was purified by affinity chromatography on an anti-FLAG resin (Sigma-Aldrich, A2220) in 10 mM Hepes, 150 mM NaCl pH 7.4 (Hepes buffered saline, HBS), eluted by the FLAG peptide (200 µg/ml in HBS), and concentrated up to 6.6 mg/ml (219 µM) by centrifugation using an Amicon Ultra-0.5 device (Merck Millipore, UFC5010, 10 kDa cut-off). The purification process was followed by SDS-PAGE and Western blot with a primary anti-FLAG antibody (F3165, Sigma-Aldrich<sup>51</sup>) and a secondary antibody conjugated to peroxidase (Bio-Rad, 172–1011). The gels were stained with EZBlue™ Gel Staining Reagent and were scanned with the Canon LiDE 210 scanner. Immunocomplexes were detected in Western blot with SuperSignal West Pico Chemiluminescence Substrate (34080, Thermo Scientific) and visualized using the Fusion FX camera (Vilber-Lourmat) for three minutes with default settings. Pre-stained molecular weight markers were also visualized with the Fusion FX camera (Vilber-Lourmat), and superimposed on the membrane imaged by chemiluminescence with the Fusion Capt Advance software. The identity of the purified propeptide was controlled by Liquid Chromatography/Mass spectrometry-Mass spectrometry (UMS 3444/US8, Lyon, France).

**Prediction of intrinsic disorder in LOX-PP.** Intrinsically disordered regions in the full-length sequence of LOX-PP were predicted with the metapredictor metaPrDOS combining DISOPRED2, DISPROT, IUPred, and PrDOS predictors<sup>52</sup> (<http://prdos.hgc.jp/cgi-bin/meta/top.cgi>). Molecular recognition features, short disordered regions able to undergo disorder-to-order transition upon binding to partners, were predicted in LOX-PP with ANCHOR<sup>53</sup> (<http://anchor.enzim.hu/>) and MoRFchibi SYSTEM<sup>54</sup> ([http://www.chibi.ubc.ca/faculty/joerg-gsponer/gsponer-lab/software/morf\\_chibi/](http://www.chibi.ubc.ca/faculty/joerg-gsponer/gsponer-lab/software/morf_chibi/)).

**Circular dichroism.** The propeptide of lysyl oxidase was dialyzed against 10 mM potassium phosphate pH 7.4 (2.0 µM, 60 µg/ml) and analyzed at 20 °C in a Chirascan system (Applied Photophysics, UMS 3444/US8, Lyon, France). Spectra of LOX-PP and buffer were also recorded in presence of increasing concentrations (20, 40, 60, and 80%, v/v) of 2,2,2-trifluoroethanol (TFE, Roth, CP29). A series of experiments was performed in the same experimental conditions in the presence of full-length HP (Sigma-Aldrich, H3393, 17 kDa) and HP hexasaccharide (dp6, Iduron, HO06, 1.8 kDa). LOX-PP (2 µM) and HP (2 µM) or a HP hexasaccharide (2 µM) were mixed at a molar ratio of 1:1 immediately before the experiment. Further spectra were collected after incubating the mixture at room temperature for one hour. For all experiments five consecutive spectra were recorded in a 1-mm pathlength cell for both LOX-PP and buffer, and averaged. The buffer spectra were subtracted from LOX-PP spectra. The resulting spectra were then smoothed and deconvoluted with the CONTIN algorithm<sup>55,56</sup> and the data set 7, which contains five denatured proteins on DichroWeb<sup>57</sup> (<http://dichroweb.cryst.bbk.ac.uk/html/home.shtml>).

**Dynamic Light Scattering (DLS).** DLS measurements were performed in a Zetasizer Nano ZS (Malvern, UMR 5223, Villeurbanne, France) using a disposable plastic cuvette (Eppendorf, 0030106300). Four series of fifteen spectra of LOX-PP diluted to 1.1 mg/ml (35 µM) in HBS were acquired at a 173° angle, at 633 nm, at 20 °C. Data were analyzed with the Malvern software.

**Size-exclusion chromatography - Multi-angle light scattering (SEC-MALS).** LOX-PP diluted at 75 µM (2.3 mg/ml) in HBS was centrifuged 5 min at 14000 × g and 25 µl of the supernatant were injected on a Superdex 200 Increase 5/150 GL column (GE Healthcare) at a flow rate of 0.2 ml/min at 25 °C. The elution was

followed by UV absorbance at 280 nm, multi-angle light scattering at 690 nm (MALS, miniDAWN-TREOS detector, Wyatt Technology Corp.) and refractometry (RI, Optilab T-rex system, Wyatt Technology). ASTRA software (Wyatt Technology) was used to calculate molecular weights.

**Small-Angle X-ray Scattering (SAXS).** SAXS experiments (MX-1841, MX-1920) were performed on the BM29 BioSAXS beamline at the European Synchrotron Radiation Facility<sup>58,59</sup> (ESRF, Grenoble, France). Before analysis, LOX-PP solution was centrifuged at  $14\,000 \times g$  for 12 min to eliminate potential aggregates. LOX-PP was injected at two different concentrations (37.4 and 18.7  $\mu\text{M}$ , 1.12 and 0.56 mg/ml respectively in HBS) in the 1.8 mm-diameter quartz capillary using the automatic robotic sample changer. Ten frames per concentration were collected at 20 °C at an exposure time of 0.5 s per frame and a flow of 5  $\mu\text{l/s}$  with a 0.9919 Å wavelength on a Pilatus 1 M detector (beam center  $888 \times 91$  pixel, 172  $\mu\text{m}/\text{pixel}$ ) at 2.867 m from the sample to cover a range of momentum transfer between 0.031 and 4.94  $\text{nm}^{-1}$ . Data quality was assessed with the autoprocessing framework EDNA<sup>60</sup> before analysis. Buffer spectra were subtracted from LOX-PP spectra and data were extrapolated to infinite concentration using ATSAS suite<sup>61</sup>. Useful data range was defined by SHANUM<sup>62</sup> and PRIMUS<sup>63</sup> was used to calculate the radius of gyration of LOX-PP through Guinier's approach with a  $sR_g$  limit  $< 1.06$ . Normalized Kratky plot was generated with scÅtter<sup>64</sup> to evaluate the flexibility of LOX-PP. The theoretical  $R_g$  was calculated with Flory's equation, which associates the radius of gyration with the protein length ( $R_g = R_0 N^\nu$ ,  $N$  being the number of residues of the protein,  $R_0$  being  $2.54 \pm 0.01$  and  $\nu$   $0.522 \pm 0.01$  for intrinsically disordered proteins<sup>65</sup>).

Ensemble modelling was performed with EOM (Ensemble Optimization Method 2.1), which is adapted for ensemble modelling of flexible macromolecules<sup>30</sup>. EOM was run to analyze the behavior of LOX-PP compared to a gaussian pool of different conformations generated from the propeptide sequence. The following parameters were used: core symmetry p1, native-like structure because LOX-PP is a disordered protein, number of harmonics: 50, number of points: 100. 20 000 independent models were generated. The genetic algorithm was run 100 times to compare the averaged theoretical scattering intensity from independent ensembles of conformations against the experimental scattering data.

**Surface Plasmon Resonance imaging (SPRi) binding assays.** SPRi experiments were performed on a Biacore Flexchip system (GE Healthcare, UMS 3444/US8, Lyon, France) at 25 °C as previously described<sup>26,66,67</sup>. Briefly, 93 unique proteins, protein domains and GAGs were spotted in triplicate on a bare gold chip (Gold Affinity chip, GE Healthcare) at concentrations ranging from 0.06 to 1 mg/ml with a non-contact spotter (sciFlexarrayer S3, Scienion) (Supplementary Table 6). The chips were then blocked for  $5 \times 5$  min (Flexchip blocking buffer, GE Healthcare) and equilibrated in phosphate-buffered saline containing 0.05% Tween 20 (Sigma-Aldrich, P2287) at a flow rate of 300  $\mu\text{l}/\text{min}$  for 90 min. Purified LOX-PP was injected at 2  $\mu\text{M}$  in the same buffer and recirculated for 20 min over the chips. The dissociation of the complexes was followed in running buffer for 40 min. Kinetics and affinity of some interactions identified by SPRi were calculated either by SPR or Bio-Layer Interferometry (BLI).

**Surface Plasmon Resonance binding assays.** SPR binding assays were performed in Biacore T100 or T200 systems (GE Healthcare, UMS 3444/US8, Lyon, France) at 25 °C in HBS pH 7.4 containing 0.05% P20 (HBS-P20, GE Healthcare, BR100054). Proteins were covalently immobilized onto CM5 sensor chips (series S, GE Healthcare) *via* their amino groups as previously described<sup>68</sup>. Proteins (Supplementary Table 6) were injected over the activated sensor chips for 2 to 10 min at concentrations ranging from 50 to 500  $\mu\text{g}/\text{ml}$ . Residual activated groups were blocked by injecting 1 M ethanolamine pH 8.5 for 7 min. A control flow cell was prepared by omitting the protein during the immobilization process. Analytes (Supplementary Table 6) were injected at 30  $\mu\text{l}/\text{min}$  for 3 min on immobilized ligands and on the control flow cell to assess non-specific binding to the sensor chip surface. Regeneration was carried out by injecting 1 M NaCl, 1.5 M NaCl or 1.5 M NaCl containing 1.5 M guanidinium chloride for 15 to 30 s. Single-cycle kinetics were performed to determine affinity ( $K_D$ ) and kinetics parameters ( $k_a$ ,  $k_d$ ), which were calculated with the BiaEval software.

**Bio-Layer Interferometry (BLI) binding assays.** Interactions experiments were performed on an Octet RED96 system (Pall, ForteBIO) using amine reactive second generation (AR2G) biosensors. Sensors were hydrated in water for 10 min, then carboxylic groups were activated for 10 min with 0.1 M N-hydroxysulfosuccinimide and 0.2 M 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride. Anastellin (Sigma-Aldrich, F3542, 100  $\mu\text{g}/\text{ml}$ ) was covalently coupled to the sensor surface *via* its primary amine groups in 10 mM sodium maleate pH 6.2 for 10 min. Residual activated groups were blocked with 1 M ethanolamine pH 8.5 for 10 min. The sensors were equilibrated in HBS-P20, and then dipped into LOX-PP solutions for 3 min. Sensors were regenerated by dipping in 1 M NaCl, 1.5 M NaCl or 1.5 M NaCl containing 1.5 M guanidinium chloride. Data were analyzed with Evaluation 9.0 software.

**Building and analyzing the interaction network of LOX-PP.** The interaction network of LOX-PP was built by integrating the new partners identified in this study by SPRi, SPR and BLI, and those curated from the literature available in the interaction database MatrixDB<sup>38</sup> (<http://matrixdb.univ-lyon1.fr/>) we have created. The network was visualized using Cytoscape<sup>69</sup> (<http://www.cytoscape.org/>).

**Coarse-grained simulations of the lysyl oxidase propeptide.** 3D models of LOX-PP were generated using the UNRES coarse-grained model of proteins<sup>70,71</sup> with the recently added function of including the distance distribution from SAXS as restraints<sup>31</sup>. The full protocol is detailed in Supplementary material. The secondary structure of each model was determined from the coordinate (.pdb) files of the models with PDBsum<sup>72</sup> (<http://www.ebi.ac.uk/pdbsum>). Topological diagrams were generated with PROMOTIF software, which is based on the

Dictionary of Secondary Structure of Proteins (DSSP)<sup>73</sup>. The full-length sequence of LOX-PP (UniProtKB pro feature PRO\_0000018520) encompassing residues 22–168 was used for coarse-grained simulations.

**Molecular docking.** The five models of LOX-PP obtained by UNRES coarse-grained MD simulations were used for docking of a HP hexasaccharide (GlcNS(6S)-IdoUA(2S))<sub>3</sub> with Autodock 3<sup>74</sup> (AD3), which yields the best performance among other docking programs for GAG ligands<sup>32</sup>. Charges of LOX-PP atoms were assigned by AD3, and charges of HP dp6 were taken from the corresponding AMBER residue libraries<sup>75</sup> compatible with the GLYCAM06 force field<sup>76</sup>. IdoUA(2S) ring in <sup>1</sup>C<sub>4</sub> conformation was used. The following parameters for AD3 local docking runs were applied: a grid spacing of 0.4 Å step with 126 × 126 × 126 grid points as dimensions of the box, which contains the putative HP binding region according to PBSA electrostatic potential calculations in AMBER. Electrostatic potential isosurfaces were calculated using the PBSA program from AmberTools with a grid spacing of 1 Å. Dielectric constants for the solvent and the solute were 80 and 1 respectively, and the surface tension parameter for nonpolar solvation calculation was 0.005<sup>77</sup>. Lamarckian genetic algorithm (initial population size of 300, 10<sup>5</sup> generations, 9995 × 10<sup>5</sup> energy evaluations), 10<sup>3</sup> independent runs were used for molecular docking. All covalent bonds with the exception of N-S and O-S for N- and O-sulfate groups, were kept flexible, corresponding to 30 torsional angle degrees of freedom in total. 50 top docking solutions scored by AD3 out of 10<sup>3</sup> obtained docked structures for each docking experiment were clustered by the DBSCAN algorithm<sup>78</sup> with a neighborhood search radius of 4 Å and minimal number of cluster members of 5. The metric distance between two structures was defined as the root mean square of atomic distances (RMSD) for the nearest atoms of the same type in order to take into account GAG periodic nature<sup>79</sup>. Three arbitrarily chosen poses from each cluster were used for MD-based refinement and analysis. Prior to MD simulations, the conformations of GAG glycosidic linkages were checked to avoid distorting starting geometries that could be potentially observable in the structures produced by AD3<sup>80</sup>.

**Molecular dynamics.** MD simulations of LOX-PP/HP complexes obtained by molecular docking were carried out in AMBER16<sup>77</sup> (<http://ambermd.org/>). We used periodic boundary conditions with a truncated octahedron TIP3P water box with at least 4 Å distance from the solute to the periodic box border, which corresponded to 2–3 × 10<sup>5</sup> water molecules in the periodic box, and counterions. ff99SB force field parameters for protein and the GLYCAM06<sup>76</sup> for GAGs were used. Two energy-minimization steps were carried out: 0.5 × 10<sup>3</sup> steepest descent cycles and 10<sup>3</sup> conjugate gradient cycles with harmonic force restraints on solute atoms, 3 × 10<sup>3</sup> steepest descent cycles and 3 × 10<sup>3</sup> conjugate gradient cycles without restraints. Then, the system was heated up to 300 K for 10 ps, equilibrated for 50 ps at 300 K and 10<sup>6</sup> Pa in isothermal isobaric ensemble (NPT). Finally, a 20 ns of productive MD run was carried out in NTP ensemble. The SHAKE algorithm, 2 fs time integration, 8 Å cutoff for non-bonded interactions and the Particle Mesh Ewald method were used. All MD simulations of LOX-PP/HP complexes were repeated three times. Analysis of the trajectories was done using the cpptraj module of AMBER Tools 17<sup>77</sup> (<http://ambermd.org/>). The radius of gyration was calculated for all protein atoms using radgyr command with default parameters in cpptraj. VMD<sup>81</sup> was used for visualization of trajectories and the figures. Energetic post-processing of the trajectories and per residue energy decomposition were done using MM-GBSA with igb = 2 for protein-GAG complexes. Entropic contribution to binding was not considered<sup>82,83</sup>. Therefore, the calculated free energies should be understood as enthalpies which also implicitly partially account for the entropic component of the solvent in the implicit model.

## References

- Mäki, J. M. Lysyl oxidases in mammalian development and certain pathological conditions. *Histol. Histopathol.* **24**, 651–660 (2009).
- Baker, A.-M. *et al.* Lysyl oxidase plays a critical role in endothelial cell stimulation to drive tumor angiogenesis. *Cancer Res.* **73**, 583–594 (2013).
- Cox, T. R., Gartland, A. & Erler, J. T. Lysyl Oxidase, a Targetable Secreted Molecule Involved in Cancer Metastasis. *Cancer Res.* **76**, 188–192 (2016).
- Trackman, P. C. Enzymatic and non-enzymatic functions of the lysyl oxidase family in bone. *Matrix Biol.* **52–54**, 7–18 (2016).
- Johnston, K. A. & Lopez, K. M. Lysyl oxidase in cancer inhibition and metastasis. *Cancer Lett.* **417**, 174–181 (2018).
- Trackman, P. C. Functional importance of lysyl oxidase family propeptide regions. *J. Cell. Commun. Signal.* <https://doi.org/10.1007/s12079-017-0424-4> (2018).
- Ricard-Blum, S. & Vallet, S. D. Proteases decode the extracellular matrix cryptome. *Biochimie* **122**, 300–313 (2016).
- Ricard-Blum, S. & Vallet, S. D. Fragments generated upon extracellular matrix remodeling: Biological regulators and potential drugs. *Matrix Biol.* <https://doi.org/10.1016/j.matbio.2017.11.005> (2017).
- Grimsby, J. L., Lucero, H. A., Trackman, P. C., Ravid, K. & Kagan, H. M. Role of lysyl oxidase propeptide in secretion and enzyme activity. *J. Cell. Biochem.* **111**, 1231–1243 (2010).
- Thomassin, L. *et al.* The Pro-regions of lysyl oxidase and lysyl oxidase-like 1 are required for deposition onto elastic fibers. *J. Biol. Chem.* **280**, 42848–42855 (2005).
- Palamakumbura, A. H. *et al.* The propeptide domain of lysyl oxidase induces phenotypic reversion of ras-transformed cells. *J. Biol. Chem.* **279**, 40593–40600 (2004).
- Hurtado, P. A. *et al.* Lysyl oxidase propeptide inhibits smooth muscle cell signaling and proliferation. *Biochem. Biophys. Res. Commun.* **366**, 156–161 (2008).
- Vora, S. R. *et al.* Lysyl oxidase propeptide inhibits FGF-2-induced signaling and proliferation of osteoblasts. *J. Biol. Chem.* **285**, 7384–7393 (2010).
- Griner, J. D., Rogers, C. J., Zhu, M.-J. & Du, M. Lysyl oxidase propeptide promotes adipogenesis through inhibition of FGF-2 signaling. *Adipocyte* **6**, 12–19 (2017).
- Bais, M. V., Ozdener, G. B., Sonenshein, G. E. & Trackman, P. C. Effects of tumor-suppressor lysyl oxidase propeptide on prostate cancer xenograft growth and its direct interactions with DNA repair pathways. *Oncogene* **34**, 1928–1937 (2015).
- Li, J. *et al.* Nna1 mediates Purkinje cell dendritic development via lysyl oxidase propeptide and NF-κB signaling. *Neuron* **68**, 45–60 (2010).
- Guo, Y., Pischon, N., Palamakumbura, A. H. & Trackman, P. C. Intracellular distribution of the lysyl oxidase propeptide in osteoblastic cells. *Am. J. Physiol. Cell Physiol.* **292**, C2095–2102 (2007).

18. Sato, S. *et al.* The Ras signaling inhibitor LOX-PP interacts with Hsp70 and c-Raf to reduce Erk activation and transformed phenotype of breast cancer cells. *Mol. Cell. Biol.* **31**, 2683–2695 (2011).
19. Sato, S. *et al.* Inhibition of CIN85-mediated invasion by a novel SH3 domain binding motif in the lysyl oxidase propeptide. *PLoS ONE* **8**, e77288 (2013).
20. Sánchez-Morgan, N., Kirsch, K. H., Trackman, P. C. & Sonenshein, G. E. UXT Is a LOX-PP Interacting Protein That Modulates Estrogen Receptor Alpha Activity in Breast Cancer Cells. *J. Cell. Biochem.* **118**, 2347–2356 (2017).
21. Sánchez-Morgan, N., Kirsch, K. H., Trackman, P. C. & Sonenshein, G. E. The lysyl oxidase propeptide interacts with the receptor-type protein tyrosine phosphatase kappa and inhibits  $\beta$ -catenin transcriptional activity in lung cancer cells. *Mol. Cell. Biol.* **31**, 3286–3297 (2011).
22. Ozdener, G. B., Bais, M. V. & Trackman, P. C. Determination of cell uptake pathways for tumor inhibitor lysyl oxidase propeptide. *Mol. Oncol.* **10**, 1–23 (2016).
23. Fogelgren, B. *et al.* Cellular fibronectin binds to lysyl oxidase with high affinity and is critical for its proteolytic activation. *J. Biol. Chem.* **280**, 24690–24697 (2005).
24. Panchenko, M. V., Stetler-Stevenson, W. G., Trubetsky, O. V., Gacheru, S. N. & Kagan, H. M. Metalloproteinase activity secreted by fibrogenic cells in the processing of prolysyl oxidase. Potential role of procollagen C-proteinase. *J. Biol. Chem.* **271**, 7113–7119 (1996).
25. Schlage, P. *et al.* Time-resolved analysis of the matrix metalloproteinase 10 substrate degradome. *Mol. Cell Proteomics* **13**, 580–593 (2014).
26. Vallet, S. D. *et al.* Chapter 11: Strategies for Building Protein–Glycosaminoglycan Interaction Networks Combining SPRi, SPR, and BLI. In *Handbook of Surface Plasmon Resonance*, 398–414, <https://doi.org/10.1039/9781788010283-00398> (2017).
27. Vora, S. R. *et al.* Characterization of recombinant lysyl oxidase propeptide. *Biochemistry* **49**, 2962–2972 (2010).
28. Fischer, H. *et al.* Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale. *J. Appl. Cryst.* **43**, 101–109 (2010).
29. Burchard, W. Static and dynamic light scattering from branched polymers and biopolymers. in *Laser light scattering in biochemistry*. 3–22 (Harding, S. E., Settle, D. B. & Bloomfield, V. A. editors, 1992).
30. Tria, G., Mertens, H. D. T., Kachala, M. & Svergun, D. I. Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. *IUCr* **2**, 207–217 (2015).
31. Karczyńska, A. S. *et al.* Prediction of protein structure with the coarse-grained UNRES force field assisted by small X-ray scattering data and knowledge-based information. *Proteins* **86**, 228–239 (2018).
32. Samsonov, S. A. & Pisabarro, M. T. Computational analysis of interactions in structurally available protein-glycosaminoglycan complexes. *Glycobiology* **26**, 850–861 (2016).
33. Atkovska, K., Samsonov, S. A., Paszkowski-Rogacz, M. & Pisabarro, M. T. Multipose Binding in Molecular Docking. *Int. J. Mol. Sci.* **15**, 2622–2645 (2014).
34. Joseph, P. R. B., Mosier, P. D., Desai, U. R. & Rajarathnam, K. Solution NMR characterization of chemokine CXCL8/IL-8 monomer and dimer binding to glycosaminoglycans: structural plasticity mediates differential binding interactions. *Biochem. J.* **472**, 121–133 (2015).
35. Rother, S. *et al.* Structural and functional insights into the interaction of sulfated glycosaminoglycans with tissue inhibitor of metalloproteinase-3 - A possible regulatory role on extracellular matrix homeostasis. *Acta Biomater.* **45**, 143–154 (2016).
36. Capila, I. & Linhardt, R. J. Heparin-protein interactions. *Angew. Chem. Int. Ed. Engl.* **41**, 391–412 (2002).
37. Cryan, L. M. & Rogers, M. S. Targeting the anthrax receptors, TEM-8 and CMG-2, for anti-angiogenic therapy. *Front. Biosci.* **16**, 1574–1588 (2011).
38. Launay, G., Salza, R., Multedo, D., Thierry-Mieg, N. & Ricard-Blum, S. MatrixDB, the extracellular matrix interaction database: updated content, a new navigator and expanded functionalities. *Nucleic Acids Res.* **43**, D321–327 (2015).
39. Trackman, P. C., Bedell-Hogan, D., Tang, J. & Kagan, H. M. Post-translational glycosylation and proteolytic processing of a lysyl oxidase precursor. *J. Biol. Chem.* **267**, 8666–8671 (1992).
40. Steentoft, C. *et al.* Mining the O-glycoproteome using zinc-finger nuclease-glycoengineered SimpleCell lines. *Nat. Methods* **8**, 977–982 (2011).
41. Xie, H. *et al.* Functional anthology of intrinsic disorder. 3. Ligands, post-translational modifications, and diseases associated with intrinsically disordered proteins. *J. Proteome Res.* **6**, 1917–1932 (2007).
42. Min, C. *et al.* A loss-of-function polymorphism in the propeptide domain of the LOX gene and breast cancer. *Cancer Res.* **69**, 6685–6693 (2009).
43. Garvin, A. J. & Morris, J. R. SUMO, a small, but powerful, regulator of double-strand break repair. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **372** (2017).
44. Beauclair, G., Bridier-Nahmias, A., Zagury, J.-F., Saïb, A. & Zamborlini, A. JASSA: a comprehensive tool for prediction of SUMOylation sites and SIMs. *Bioinformatics* **31**, 3483–3491 (2015).
45. Siegel, R. Biosynthesis of collagen crosslinks: increased activity of purified lysyl oxidase with reconstituted collagen fibrils. *PNAS* **71**, 4826–4830 (1974).
46. Cronshaw, A. D. *et al.* TRAMP (tyrosine rich acidic matrix protein), a protein that co-purifies with lysyl oxidase from porcine skin. Identification of TRAMP as the dermatan sulphate proteoglycan-associated 22K extracellular matrix protein. *Matrix* **13**, 255–266 (1993).
47. Forbes, E. G., Cronshaw, A. D., MacBeath, J. R. & Hulmes, D. J. Tyrosine-rich acidic matrix protein (TRAMP) is a tyrosine-sulphated and widely distributed protein of the extracellular matrix. *FEBS Lett.* **351**, 433–436 (1994).
48. Kalamajski, S., Bihan, D., Bonna, A., Rubin, K. & Farndale, R. W. Fibromodulin Interacts with Collagen Cross-linking Sites and Activates Lysyl Oxidase. *J. Biol. Chem.* **291**, 7951–7960 (2016).
49. Fatoux-Ardore, M. *et al.* Large-scale investigation of Leishmania interaction networks with host extracellular matrix by surface plasmon resonance imaging. *Infect. Immun.* **82**, 594–606 (2014).
50. Kohfeldt, E., Maurer, P., Vannahme, C. & Timpl, R. Properties of the extracellular calcium binding module of the proteoglycan testican. *FEBS Lett.* **414**, 557–561 (1997).
51. Brizzard, B. L., Chubet, R. G. & Vizard, D. L. Immunoaffinity purification of FLAG epitope-tagged bacterial alkaline phosphatase using a novel monoclonal antibody and peptide elution. *Biotechniques* **16**, 730–735 (1994).
52. Ishida, T. & Kinoshita, K. Prediction of disordered regions in proteins based on the meta approach. *Bioinformatics* **24**, 1344–1348 (2008).
53. Dosztányi, Z., Mészáros, B. & Simon, I. ANCHOR: web server for predicting protein binding regions in disordered proteins. *Bioinformatics* **25**, 2745–2746 (2009).
54. Malhis, N., Jacobson, M. & Gsponer, J. MoRFChibi SYSTEM: software tools for the identification of MoRFs in protein sequences. *Nucleic Acids Res.* **44**, W488–493 (2016).
55. Provencher, S. W. & Glöckner, J. Estimation of globular protein secondary structure from circular dichroism. *Biochemistry* **20**, 33–37 (1981).
56. Sreerama, N. & Woody, R. W. Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal. Biochem.* **287**, 252–260 (2000).
57. Whitmore, L. & Wallace, B. A. Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers* **89**, 392–400 (2008).

58. Pernot, P. *et al.* Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *J. Synchrotron Radiat.* **20**, 660–664 (2013).
59. Round, A. *et al.* BioSAXS Sample Changer: a robotic sample changer for rapid and reliable high-throughput X-ray solution scattering experiments. *Acta Crystallogr. D Biol. Crystallogr.* **71**, 67–75 (2015).
60. Incardona, M.-F. *et al.* EDNA: a framework for plugin-based applications applied to X-ray experiment online data analysis. *J. Synchrotron Radiat.* **16**, 872–879 (2009).
61. Franke, D. *et al.* ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Cryst.* **50**, 1212–1225 (2017).
62. Konarev, P. V. & Svergun, D. I. A posteriori determination of the useful data range for small-angle scattering experiments on dilute monodisperse systems. *IUCr* **2**, 352–360 (2015).
63. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. & Svergun, D. I. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Cryst.* **36**, 1277–1282 (2003).
64. Förster, S., Apostol, L. & Bras, W. Scatter: software for the analysis of nano- and mesoscale small-angle scattering. *J. Appl. Cryst.* **43**, 639–646 (2010).
65. Kikhney, A. G. & Svergun, D. I. A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Lett.* **589**, 2570–2577 (2015).
66. Faye, C., Chautard, E., Olsen, B. R. & Ricard-Blum, S. The first draft of the endostatin interaction network. *J. Biol. Chem.* **284**, 22041–22047 (2009).
67. Salza, R. *et al.* Extended interaction network of procollagen C-proteinase enhancer-1 in the extracellular matrix. *Biochem. J.* **457**, 137–149 (2014).
68. Ricard-Blum, S. *et al.* Characterization of endostatin binding to heparin and heparan sulfate by surface plasmon resonance and molecular modeling: role of divalent cations. *J. Biol. Chem.* **279**, 2927–2936 (2004).
69. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
70. Liwo, A., Czaplowski, C., Oldziej, S. & Scheraga, H. A. Computational techniques for efficient conformational sampling of proteins. *Curr. Opin. Struct. Biol.* **18**, 134–139 (2008).
71. Liwo, A. *et al.* A unified coarse-grained model of biological macromolecules based on mean-field multipole-multipole interactions. *J. Mol. Model.* **20**, 2306 (2014).
72. Laskowski, R. A. PDBsum new things. *Nucleic Acids Res.* **37**, D355–359 (2009).
73. Hutchinson, E. G. & Thornton, J. M. HERA—a program to draw schematic diagrams of protein secondary structures. *Proteins* **8**, 203–212 (1990).
74. Morris, G. M. *et al.* Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* **19**, 1639–1662 (1998).
75. Pichert, A. *et al.* Characterization of the interaction of interleukin-8 with hyaluronan, chondroitin sulfate, dermatan sulfate and their sulfated derivatives by spectroscopy and molecular modeling. *Glycobiology* **22**, 134–145 (2012).
76. Kirschner, K. N. *et al.* GLYCAM06: a generalizable biomolecular force field. *Carbohydrates. J. Comput. Chem.* **29**, 622–655 (2008).
77. Case, D. A. *et al.* AMBER 2017. *University of California, San Francisco* (2017).
78. Ester, M., Kriegl, H., Sander, J. & Xu, X. A density-based algorithm for discovering clusters a density-based algorithm for discovering clusters in large spatial databases with noise. *KDD* **96**, 226–231 (1996).
79. Samsonov, S. A., Gehrcke, J.-P. & Pisabarro, M. T. Flexibility and explicit solvent in molecular-dynamics-based docking of protein-glycosaminoglycan systems. *J. Chem. Inf. Model.* **54**, 582–592 (2014).
80. Nivedha, A. K., Makeneni, S., Foley, B. L., Tessier, M. B. & Woods, R. J. Importance of ligand conformational energies in carbohydrate docking: Sorting the wheat from the chaff. *J. Comput. Chem.* **35**, 526–539 (2014).
81. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**(33–38), 27–28 (1996).
82. Gandhi, N. S. & Mancera, R. L. Free energy calculations of glycosaminoglycan-protein interactions. *Glycobiology* **19**, 1103–1115 (2009).
83. Homeyer, N. & Gohlke, H. Free Energy Calculations by the Molecular Mechanics Poisson-Boltzmann Surface Area Method. *Mol. Inform.* **31**, 114–122 (2012).
84. Naba, A. *et al.* The extracellular matrix: Tools and insights for the “omics” era. *Matrix Biol.* **49**, 10–24 (2016).

## Acknowledgements

This work was supported by a grant from the Fondation pour la Recherche Médicale (grant n°DBI20141231336 to SRB), Ligue Nationale contre le Cancer, Comité du Rhône (2017–01CI: UCBL to SRB) and the National Science Center of Poland (Narodowe Centrum Nauki, grant UMO-2016/21/P/ST4/03995 (SAS, UU, AL), the European Union Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 665778 to SAS). The SAXS experiments were performed on the beamline BM29 at the European Synchrotron Radiation Facility (ESRF, Grenoble, France, MX-1841 and MX-1920) and on the SWING (Small and Wide angle X-ray scattering) beamline (French National Synchrotron Facility SOLEIL, Saint-Aubin, France, 20170906). We are grateful to Romain Salza (Villeurbanne, France) for transfecting HEK293 cells with the LOX propeptide construct and for his help in performing SPRI experiments, to the local ESRF contact for providing assistance in using BM29 beamline, to Nathalie Sintès (UMR 5223, Villeurbanne, France) for providing access to the DLS instrument, to Laure Perrin-Cocon and Vincent Lotteau (International Center for Infectiology Research, Lyon, France) for their generous gift of LTBP-1 domains and to Laurent Terradot (UMR 5086, Lyon, France) for performing SEC-MALS experiment and analysis. Protein arrays were analysed in the Biacore Flexchip system of the UMS 3444/US8 Facility (Lyon, France). We thank Christophe Marquette and Loïc Blum (Institut de Chimie et Biochimie Moléculaires et Supramoléculaires, UMR 5246 CNRS, Université Lyon 1, Villeurbanne, France) for providing access to the Scienion spotter. The computational resources were provided by the Polish Grid Infrastructure (PL-GRID), the Interdisciplinary Center of Mathematical and Computer Modeling (ICM; grant GA-71-23) and our cluster at the Faculty of Chemistry, University of Gdansk (Poland).

## Author Contributions

S.D.V. performed binding assays, C.D. experiments, calculation and modelling from S.A.X.S. data, prepared Figures 1–5 and 9, drafted and formatted the manuscript, A.E.M. performed S.A.X.S. experiments, U.U.K., A.L. and S.A.S. carried out coarse-grained modeling and docking experiments and prepared Figures 6–8, B.D. designed the expression constructs and contributed to the expression of recombinant proteins, SRB designed the study, interpreted binding assays and wrote the manuscript. All authors reviewed the final version of the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-30190-6>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018