



HAL
open science

The glycosaminoglycan interactome 2.0

Sylvain Vallet, Coline Berthollier, Sylvie Ricard-Blum

► **To cite this version:**

Sylvain Vallet, Coline Berthollier, Sylvie Ricard-Blum. The glycosaminoglycan interactome 2.0. American Journal of Physiology - Cell Physiology, 2022, Deciphering the Role of Proteoglycans and Glycosaminoglycans in Health and Disease, 322 (6), pp.C1271-C1278. 10.1152/ajpcell.00095.2022 . hal-04871745

HAL Id: hal-04871745

<https://hal.science/hal-04871745v1>

Submitted on 7 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

The glycosaminoglycan interactome 2.0

Sylvain D. Vallet^{1,2}, Coline Berthollier¹ and Sylvie Ricard-Blum¹

¹Univ. Lyon, Université Lyon 1, ICBMS, UMR 5246, 69622 Villeurbanne, France

²Current address: Univ. Grenoble Alpes, CNRS, CEA, IBS, F-38000 Grenoble, France

ORCID numbers

Sylvie RICARD-BLUM 0000-0001-9263-1851

Sylvain D. VALLET 0000-0001-9885-7499

Corresponding author:

Professor Sylvie RICARD-BLUM

E-mail: sylvie.ricard-blum@univ-lyon1.fr

Institut de Chimie et Biochimie Moléculaires et Supramoléculaires

ICBMS, UMR 5246

Université Lyon 1, Bâtiment Lederer

43 Boulevard du 11 novembre 1918

F-69622 Villeurbanne, France

Phone: +3(0)472448232, no fax

Running title: GAG interactome 2.0

Keywords: Glycosaminoglycans, Interaction networks, Extracellular matrix

Abbreviations: AP-MS: affinity purification-mass spectrometry, BP: binding protein, CS: chondroitin sulfate, DS: dermatan sulfate, ECM: extracellular matrix, GAG: glycosaminoglycan, GAGBP: GAG-binding protein, HA: hyaluronan, HP: heparin, HS: heparan sulfate, IdoA: iduronic acid, KS: keratan sulfate, MS: mass spectrometry, sHA: sulfated hyaluronan.

Abstract

Glycosaminoglycans (GAGs) are complex linear polysaccharides, which are covalently attached to core proteins (except for hyaluronan) to form proteoglycans. They play key roles in the organization of the extracellular matrix, and at the cell surface where they contribute to the regulation of cell signaling and of cell adhesion. To explore the mechanisms and pathways underlying their functions, we have generated an expanded dataset of 4290 interactions corresponding to 3464 unique GAG-binding proteins, four times more than the first version of the GAG interactome (*Vallet and Ricard-Blum, 2021 J Histochem Cytochem 69:93-104*). The increased size of the GAG network is mostly due to the addition of GAG-binding proteins captured from cell lysates and biological fluids by affinity chromatography and identified by mass spectrometry. We review here the interaction repertoire of natural GAGs and of synthetic sulfated hyaluronan, the specificity and molecular functions of GAG-binding proteins, and the biological processes and pathways they are involved in. This dataset is also used to investigate the differences between proteins binding to iduronic acid-containing GAGs (dermatan sulfate and heparin/heparan sulfate) and those interacting with GAGs lacking iduronic acid (chondroitin sulfate, hyaluronan, and keratan sulfate).

Introduction

Sulfated glycosaminoglycans (GAGs), namely chondroitin sulfate (CS), dermatan sulfate (DS) (1, 2), keratan sulfate (KS) (3, 4), heparin (HP), and heparan sulfate (HS) (5) are linear complex polysaccharides which are covalently attached to core proteins to form proteoglycans (6, 7). They are mostly present at the cell surface and in the extracellular matrix (ECM), but also in the nucleus as shown for HS (8), and in the intracellular granules of certain mast cells for CS, DS and HP (9). The disaccharide units of these GAGs are made of a uronic acid, and an

hexosamine except for KS containing a galactose instead of a uronic acid. Hyaluronan, a large non-sulfated GAG, does not bind covalently to proteins but links several proteoglycan molecules to form high-molecular weight aggregates in the ECM. It has also been detected in the cytoplasm and nucleus (10, 11). GAGs play key roles in ECM assembly, the regulation of cell behavior and cell signaling (4, 12–15), development (16), and diseases (17) including cancer (12, 13, 18, 19), infectious diseases (20) and neurodegenerative diseases (21) via their binding to a variety of proteins. Numerous interactions have been reported for CS (22, 23), HP/HS (24–26), and KS (27). We drafted the first comprehensive version of the interactome of the six GAGs listed above by compiling GAG-protein interactions mostly identified by biophysical techniques and protein arrays (28). The availability of datasets of hundreds of GAG-binding proteins (GAGBPs) identified by affinity purification followed by proteomic analysis (AP-MS, (29) for review) drastically increased the number of GAGBPs. These proteins identified by affinity proteomics in cell lysates or biological fluids might not all establish direct interactions with GAGs, and the biological relevance of these interactions remains to be determined, but these proteomic datasets provide information on intracellular GAG-binding partners, which are often under-estimated. Furthermore, several hundreds of DS-binding proteins which have been identified by proteomics, have not been included so far in global GAG networks. A number of them are autoantigens, suggesting a role for DS in autoimmunity, which remains under-investigated. This prompted us to expand the first version of the GAG interactome by integrating those GAGBPs. This comprehensive network will be a useful dataset of GAGBPs, which is described in the first part of this review, and will provide new insights in the binding specificities of GAGs in association with their roles in the regulation of biological processes and pathways, which are analyzed using computational tools and reported in the second part of the review.

Glycosaminoglycan-binding protein dataset

We added to the first draft of the GAG interactome animal GAG-binding proteins identified by affinity proteomics using GAGs as ligands and mass spectrometry to identify GAGBPs (30–40). Few individual interactions were also added to the dataset (41–44). All the identifiers were mapped to the human reviewed ones using InParanoid (45) or OrthoDB (46). BLAST and sequence alignments were performed when needed to identify human orthologs. Human orthologs of large murine datasets (32) were identified using UniProtKB retrieving tool by first generating a list of the primary gene name of each murine entry, and then a list of reviewed human proteins UniProtKB accession numbers. When no human ortholog was identified using this approach, manual retrieval of human orthologs was carried out using the tools mentioned above. Proteins that were not unambiguously identified were not included to the dataset. We also updated GAGBPs annotated by the following Gene Ontology (GO) terms (GO:0008201, heparin binding; GO:0035374, chondroitin sulfate binding; GO:1904399, heparan sulfate binding; GO:0005540, hyaluronic acid binding) and SwissProt keywords (KW-0358, heparin-binding). The GAGBP dataset designed as GAG interactome 2.0 is available in **Supplemental Table S1** (<https://figshare.com/s/523df8467338eecd4a4>), which includes data sources. It comprises 4290 GAG-mediated interactions involving 3464 unique GAG-binding proteins (**Supplemental Figure S1**, <https://figshare.com/s/523df8467338eecd4a4>). The GAG interactome reported in this review contains about four times more GAGBPs than the first GAG interactome, with a marked increase in the number of dermatan sulfate-binding proteins (DSBPs, mostly autoantigens) (31, 35) and of heparin/heparan sulfate-binding proteins (HP/HSBPs) (32, 33) (**Table 1**). We have also included in this version of the GAG interactome the proteins binding to unnatural sulfated HA, which has pro-osteogenic effect (34). The

interaction network of the six natural GAGs is shown in **Figure 1**. HP/HSBPs comprise the largest subnetwork of GAGBPs (74.5%) followed by DSBPs (15.3%). CSBPs (3.4%), HABPs (1.1%) and KSBPs (5.7%) represent only ~10% of the total GAGBPs (**Table 1**).

The number of GAGBPs binding specifically to one GAG and those shared by two, three, four or five GAGs are shown in **Figure 2** and **Supplemental Figure S2** (<https://figshare.com/s/523df8467338eecd4a4>), and listed in **Supplemental Table S2** (<https://figshare.com/s/523df8467338eecd4a4>). HP/HS and KS have the highest binding specificity with 62 and 67% respectively of their protein partners binding specifically to them. In contrast, only one third of the partners of CS and DS (33% and 35% respectively) interact specifically with them whereas HA has an intermediate specificity (48%). For DS this is mostly due to the addition of autoantigen DSBPs to the network because DS did not have any specific binding partners in the first draft of the interactome. Sulfated HA shares only two partners with HA, HADH and COL6A3 encoding mitochondrial hydroxyacyl-coenzyme A dehydrogenase and the α 3 chain of collagen VI respectively, and 267 with HP/HS, which suggests that sulfation rather than the disaccharide composition plays a role in the ability of sulfated HA to interact with these proteins. A single protein (insulin-like growth factor-binding protein 7, IGFBP7) interacts with all the natural GAGs in the network tested (**Figure 2**, **Supplemental Table S2**, <https://figshare.com/s/523df8467338eecd4a4>). The limited number of proteins (142) binding to both iduronic acid-containing GAGs (IdoA-GAGs: DS and HP/HS) and to non-iduronic acid-containing GAGs (non-IdoA-GAGs: CS, HA, and KS) (**Figure 3**) suggests that iduronic acid contributes to the specificity of protein binding.

The protein domains, molecular functions, biological processes and biological pathways significantly overrepresented in the GAGBP dataset were determined by performing enrichment analyses using bioinformatic tools. Bioactive ECM protein fragments were annotated according to the gene encoding the full-length protein they are derived from. For multimeric ECM proteins, the genes encoding their different polypeptide chains were included in the analyses. The protein domains overrepresented in GAGBPs were determined with the Functional Enrichment analysis tool FunRich (47). There is a significant enrichment of signal peptide ($p < 0.001$) in all GAGs except DSBPs, KSBPs and sulfated HABPs (**Supplemental Figure S3**, <https://figshare.com/s/523df8467338eecd4a4>). 78% of HABPs and 56% CSBPs contain a signal peptide and are thus secreted or membrane proteins. In contrast, only one third of HP/HSBPs and IdoA-GAGBPs have a signal peptide, most of their partners are thus intracellular. The most overrepresented domain in HABPs is the LINK domain. Tyrosine kinase catalytic domain (TyrKc) and Serine/Threonine protein kinases, catalytic domain (S_TKc) are overrepresented in KSBPs (17% and 19% respectively, **Supplemental Figure S3**, <https://figshare.com/s/523df8467338eecd4a4>). The most significantly enriched domain in DSBPs is the RNA recognition motif (RRM) but it is restricted to ~8% of the proteins. This motif is found in nuclear autoantigens, which is consistent with the fact that most DSBPs of the dataset are autoantigens (31, 35). These results are in agreement with those obtained using the Cytoscape App BiNGO, used to assess overrepresentation of Gene Ontology (GO) categories in Biological Networks (BiNGO) (48). The Cellular Component GO terms "Extracellular region", "Extracellular region part", "Extracellular space", and/or "Extracellular matrix" were significantly enriched in CSBPs, HABPs and HP/HSBPs. The GO terms "Membrane raft", "Plasma membrane part", and "Cytosol" were enriched in KSBPs, and "Ribonuclear complex", "Macromolecular complex", "Organelle lumen" in DSBPs. Although the results should be

interpreted with caution due to large difference in the number of their partners, sulfated HABPs are enriched in GO terms "Cytoplasm", "Cytoplasm part", "Intracellular part", in contrast to HABPs which are mostly extracellular.

Molecular functions, biological processes and pathways of GAGBPs

Pathways overrepresented in GAGBPs were identified using Reactome (49). The 25 most relevant pathways, sorted by p-value, of each GAGBP datasets are available in **Supplemental Table S3** (<https://figshare.com/s/523df8467338eecd4a4>). "Extracellular matrix organization" was a pathway enriched in all GAGBPs except for KSBPs. "Signal transduction" was enriched in KSBPs and HP/HSBPs, "Immune system" and "Metabolism of proteins" in DSBPs, HP/HSBPs and sulfated HABPs (**Figure 4**). "Immune system", "Disease", "Metabolism of proteins", "Cellular response to stimuli", "Metabolism of RNA", "Vesicle-mediated transport", and "Autophagy", and to a lesser extent "Transport of small molecules" are significantly overrepresented in IdoA-GAGBPs but not in non-IdoA-GAGBPs (**Figure 5**) suggesting that both subsets of GAGBPs have different molecular functions, and that these differences could be explained in part by the presence of iduronic acid in the GAGs they bind to. Iduronic acid alone does not account for these differences. DSBPs and HP/HSBPs, which both interact with IdoA-GAGs, are involved in several common pathways but they are also associated with specific pathways such as "Cell cycle", "Organelle biosynthesis and maintenance", and "Chromatin organization" and "Signal transduction" to a smaller extent for CSBPs, and with "Signal transduction" and "Transport of small molecules" for HP/HSBPs.

The overrepresented GO terms "Molecular Function" and "Biological Process" in GAGBPs were assessed with BiNGO and visualized as networks. "Protein binding" was overrepresented in IdoA-GAGBPs (*i.e.*, DSBPs and HP/HSBPs), and sulfated HABPs. "GAG binding" and

"Polysaccharide binding" are overrepresented in CSBPs, HABPs, HP/HSBPs, and IdoA-GAGBPs in contrast to non-IdoA-GAGBPs and KSBPs which are significantly enriched in the molecular functions "Protein kinase activity", and "Phosphotransferase activity" (**Supplemental Figure S4**, <https://figshare.com/s/523df8467338eecd4a4>). The GO terms "RNA binding", "Protein binding" and "Structural molecule activity" are enriched in DSBPs and HSBPs and "Catalytic activity", "Oxidoreductase activity" in sulfated HABPs. The major overrepresented biological processes are "Nervous system development" for CSBPs, "Gene expression" for DSBPs, "Metabolic process", "Vesicle-mediated transport" for sulfated HABPs and HP/HSBPs and IdoA-GAGBPs, "Localization" and "Establishment of localization" for HP/HSBPs, IdoA-GAGBPs, and sulfated HABPs, "Cell adhesion" for HABPs, and "Protein amino acid phosphorylation for KSBPs" (**Supplemental Figure S5**, <https://figshare.com/s/523df8467338eecd4a4>).

Further investigation of the proteins binding to both GAGs containing IdoA and those lacking IdoA were carried out using Reactome. These 142 proteins (**Figure 3**) have the same enrichment profile than non-IdoA-GAGBPs except that they are enriched in pathways associated with diseases and not in those involved in cell cycle (**Figure 5**). Their analysis using Metascape, a gene annotation and analysis resource (<https://metascape.org>, (50)) shows that they are enriched in matrisome-associated proteins, and in proteins involved in axon guidance, signaling by tyrosine kinase receptors, regulation of proteolysis and hyaluronic acid binding (**Supplemental Figure S6**, <https://figshare.com/s/523df8467338eecd4a4>). Indeed, the proteins interacting with both IdoA-GAGs and non-IdoA-GAGs include 42% of HABPs. They are also enriched in proteins associated with glypican-1 pathway, and diseases of signal transduction by growth factor receptors and second messengers (**Supplemental Figure S6**, <https://figshare.com/s/523df8467338eecd4a4>). The proteins binding only to IdoA-GAGs are

enriched in proteins encoded by housekeeping genes, translation, focal adhesion and VEGFA-VEGFR2 signaling pathway, whereas MAPK signaling pathway, and regulation of cell migration are enriched in proteins binding only to non-IdoA-GAGs (**Supplemental Figure S6**, <https://figshare.com/s/523df8467338eecd4a4>).

Conclusions and Future Prospects

This new version of the GAG interactome includes GAGBPs identified in the extracellular matrix, at the cell surface, and also within the cells (e.g., in the nucleus) thanks to the affinity proteomics approach developed to identify GAGBPs in cell lysates and serum. This expanded version of the GAG interactome includes 4,290 interactions and 3,464 unique GAGBPs, mostly DSBPs and new HP/HSBPs. Although GAGBPs identified by affinity proteomics contribute to the biological functions of GAGs, their direct binding to GAGs and the biological relevance of these interactions remain to be demonstrated. The addition to the network of DSBPs identified by proteomics has markedly increased the binding specificity of DSBPs up to 35%. Indeed, all the DSBPs included in the first draft of the GAG interactome were able to bind to at least one other GAG (28), which illustrates the interest of large-scale proteomic approaches to increase the coverage of GAG interactomes in order to draw meaningful conclusions from their analyses. The limited number of proteins binding to both iduronic acid-containing GAGs and to non-iduronic acid-containing GAGs suggests that iduronic acid plays a role in the specificity of protein binding. Signal transduction, developmental biology, extracellular matrix organization, hemostasis and programmed cell death are overrepresented in both IdoA-GAGBPs and non-IdoA-GAGBPs. Other biological pathways such as immune system, disease, metabolism of proteins, metabolism of RNA, DNA replication, transport of small molecules and autophagy are overrepresented in IdoA-GAGBPs but not to non-IdoA-GAGBPs. IdoA-GAGBPs,

and sulfated HABPs to a lesser extent, are enriched in proteins associated with a large number of biological pathways, DSBPs being involved in a higher number of pathways than HP/HSBPs. The characterization of GAG sequences containing iduronic acid binding to proteins, and the conformations they adopt upon binding to proteins will be crucial to determine if they share common features accounting for the additional functions overrepresented in IdoA-GAGBPs. The comparison of HABPs and sulfated HABPs is limited due to the large difference (~1:10 ratio) in the number of proteins in both datasets. A single Reactome pathway "Extracellular matrix organization" is overrepresented in both HABPs and sulfated HABPs but only two overrepresented Reactome pathways have been identified in HABPs, likely because of the small number of HABPs included in the GAG interactome. Sulfated HABPs share some features with HP/HSBPs including immune system, cellular response to stimuli, programmed cell death, metabolism of proteins, and extracellular matrix organization. Interestingly, several molecular functions and biological processes are overrepresented in IdoA-containing GAGs and in sulfated HABPs, suggesting that sulfation plays a role in selecting several functional categories of GAGBPs. Synthetic chemically-modified GAGs can be used to decipher the molecular mechanisms of biological functions and pathways, and could be of interest to target specific pathways for therapeutic purpose. The global GAG interactome 2.0 is a scaffold to build specific GAG subnetworks by integrating transcriptomic or proteomic data, and to prioritize interactions within the network using parameters related to the strength (equilibrium dissociation constant), kinetics (association and dissociation rates) and binding sites as discussed in our recent review (29). Part of the interaction data comprising the GAG interactome 2.0 is available in MatrixDB and IntAct databases (51, 52) and the whole dataset will be made available in these databases after curation according to the rules of the International Molecular Exchange consortium (53) and to the FAIR principles (54).

Most GAG interactions, except for hyaluronan, take place *in vivo* when GAGs are covalently attached to core proteins to form proteoglycans. GAG interactions should thus be contextualized by integrating them either in global proteoglycan interactomes such as the interactomes of decorin (55) and of the four syndecans (56) or in specific interactomes to focus on GAG interactions in various biological contexts (e.g., the interactome of nuclear syndecan-1 in mesothelioma cells (57)). The number of proteoglycan interactomes should increase in the future due to the identification of new proteoglycans using glycoproteomics including fifteen additional chondroitin sulfate proteoglycans in *Caenorhabditis elegans* (58), and thirteen human CSPGs (59, 60).

The detailed characterization of binding sites at the molecular level, both in GAGs and their protein partners, is needed to decipher the molecular mechanisms of GAG functions. This will be achieved experimentally using for example NMR (61), and vibrational spectroscopy to study GAG-cell membrane interactions (62). This approach has been used to generate the first vibrational sum-frequency generation spectra of CS interacting with dipalmitoyl phosphatidylcholine at air-liquid interface (62, 63). In addition, techniques based on NMR and MS-hybridized technologies including ion mobility spectrometry and infrared spectroscopy (64) will ease the sequencing of GAG oligosaccharides interacting with proteins, and the sequencing of GAGs bound to different core proteins in various tissues and various physiological and pathological situations.

REFERENCES

1. **Mizumoto S, Yamada S.** An Overview of *in vivo* Functions of Chondroitin Sulfate and Dermatan Sulfate Revealed by Their Deficient Mice. *Front Cell Dev Biol* 9: 764781, 2021. doi: 10.3389/fcell.2021.764781.
2. **Zhang B, Chi L.** Chondroitin Sulfate/Dermatan Sulfate-Protein Interactions and Their Biological Functions in Human Diseases: Implications and Analytical Tools. *Front Cell Dev Biol* 9: 693563, 2021. doi: 10.3389/fcell.2021.693563.

3. **Pomin VH.** Keratan sulfate: an up-to-date review. *Int J Biol Macromol* 72: 282–289, 2015. doi: 10.1016/j.ijbiomac.2014.08.029.
4. **Caterson B, Melrose J.** Keratan sulfate, a complex glycosaminoglycan with unique functional capability. *Glycobiology* 28: 182–206, 2018. doi: 10.1093/glycob/cwy003.
5. **Gallagher J.** Fell-Muir Lecture: Heparan sulphate and the art of cell regulation: a polymer chain conducts the protein orchestra. *Int J Exp Pathol* 96: 203–231, 2015. doi: 10.1111/iep.12135.
6. **Iozzo RV, Schaefer L.** Proteoglycan form and function: A comprehensive nomenclature of proteoglycans. *Matrix Biol* 42: 11–55, 2015. doi: 10.1016/j.matbio.2015.02.003.
7. **Karamanos NK, Piperigkou Z, Theocharis AD, Watanabe H, Franchi M, Baud S, Brézillon S, Götte M, Passi A, Vigetti D, Ricard-Blum S, Sanderson RD, Neill T, Iozzo RV.** Proteoglycan Chemical Diversity Drives Multifunctional Cell Regulation and Therapeutics. *Chem Rev* 118: 9152–9232, 2018. doi: 10.1021/acs.chemrev.8b00354.
8. **Stewart MD, Sanderson RD.** Heparan sulfate in the nucleus and its control of cellular functions. *Matrix Biol* 35: 56–59, 2014. doi: 10.1016/j.matbio.2013.10.009.
9. **Mulloy B, Lever R, Page CP.** Mast cell glycosaminoglycans. *Glycoconj J* 34: 351–361, 2017. doi: 10.1007/s10719-016-9749-0.
10. **Evanko SP, Wight TN.** Intracellular localization of hyaluronan in proliferating cells. *J Histochem Cytochem* 47: 1331–1342, 1999. doi: 10.1177/002215549904701013.
11. **Skandalis SS, Karalis T, Heldin P.** Intracellular hyaluronan: Importance for cellular functions. *Semin Cancer Biol* 62: 20–30, 2020. doi: 10.1016/j.semcancer.2019.07.002.
12. **Knelson EH, Nee JC, Blobe GC.** Heparan sulfate signaling in cancer. *Trends Biochem Sci* 39: 277–288, 2014. doi: 10.1016/j.tibs.2014.03.001.
13. **Marques C, Reis CA, Vivès RR, Magalhães A.** Heparan Sulfate Biosynthesis and Sulfation Profiles as Modulators of Cancer Signalling and Progression. *Front Oncol* 11: 778752, 2021. doi: 10.3389/fonc.2021.778752.
14. **Vigetti D, Karousou E, Viola M, Deleonibus S, De Luca G, Passi A.** Hyaluronan: biosynthesis and signaling. *Biochim Biophys Acta* 1840: 2452–2459, 2014. doi: 10.1016/j.bbagen.2014.02.001.
15. **Garantziotis S.** Modulation of hyaluronan signaling as a therapeutic target in human disease. *Pharmacol Ther* 232: 107993, 2022. doi: 10.1016/j.pharmthera.2021.107993.
16. **Townley RA, Bülow HE.** Deciphering functional glycosaminoglycan motifs in development. *Curr Opin Struct Biol* 50: 144–154, 2018. doi: 10.1016/j.sbi.2018.03.011.
17. **Shi D, Sheng A, Chi L.** Glycosaminoglycan-Protein Interactions and Their Roles in Human Disease. *Front Mol Biosci* 8: 639666, 2021. doi: 10.3389/fmolb.2021.639666.
18. **Vitale D, Kumar Katakam S, Greve B, Jang B, Oh E-S, Alaniz L, Götte M.** Proteoglycans and glycosaminoglycans as regulators of cancer stem cell function and therapeutic resistance. *FEBS J* 286: 2870–2882, 2019. doi: 10.1111/febs.14967.

19. **Bartolini B, Caravà E, Caon I, Parnigoni A, Moretto P, Passi A, Vigetti D, Viola M, Karousou E.** Heparan Sulfate in the Tumor Microenvironment. *Adv Exp Med Biol* 1245: 147–161, 2020. doi: 10.1007/978-3-030-40146-7_7.
20. **Aquino RS, Park PW.** Glycosaminoglycans and infection. *Front Biosci (Landmark Ed)* 21: 1260–1277, 2016. doi: 10.2741/4455.
21. **Jin W, Zhang F, Linhardt RJ.** Glycosaminoglycans in Neurodegenerative Diseases. *Adv Exp Med Biol* 1325: 189–204, 2021. doi: 10.1007/978-3-030-70115-4_9.
22. **Mizumoto S, Yamada S, Sugahara K.** Molecular interactions between chondroitin-dermatan sulfate and growth factors/receptors/matrix proteins. *Curr Opin Struct Biol* 34: 35–42, 2015. doi: 10.1016/j.sbi.2015.06.004.
23. **Djrbal L, Lortat-Jacob H, Kwok J.** Chondroitin sulfates and their binding molecules in the central nervous system. *Glycoconj J* 34: 363–376, 2017. doi: 10.1007/s10719-017-9761-z.
24. **Ori A, Wilkinson MC, Fernig DG.** A systems biology approach for the investigation of the heparin/heparan sulfate interactome. *J Biol Chem* 286: 19892–19904, 2011. doi: 10.1074/jbc.M111.228114.
25. **Xu D, Esko JD.** Demystifying heparan sulfate-protein interactions. *Annu Rev Biochem* 83: 129–157, 2014. doi: 10.1146/annurev-biochem-060713-035314.
26. **Gómez Toledo A, Sorrentino JT, Sandoval DR, Malmström J, Lewis NE, Esko JD.** A Systems View of the Heparan Sulfate Interactome. *J Histochem Cytochem* 69: 105–119, 2021. doi: 10.1369/0022155420988661.
27. **Conrad AH, Zhang Y, Tasheva ES, Conrad GW.** Proteomic analysis of potential keratan sulfate, chondroitin sulfate A, and hyaluronic acid molecular interactions. *Invest Ophthalmol Vis Sci* 51: 4500–4515, 2010. doi: 10.1167/iovs.09-4914.
28. **Vallet SD, Clerc O, Ricard-Blum S.** Glycosaminoglycan-Protein Interactions: The First Draft of the Glycosaminoglycan Interactome. *J Histochem Cytochem* 69: 93–104, 2021. doi: 10.1369/0022155420946403.
29. **Ricard-Blum S, Perez S.** Glycosaminoglycan interaction networks and databases. *Curr Opin Struct Biol* 74: 102355, 2022. doi: 10.1016/j.sbi.2022.102355.
30. **Gesslbauer B, Derler R, Handwerker C, Seles E, Kungl AJ.** Exploring the glycosaminoglycan-protein interaction network by glycan-mediated pull-down proteomics. *Electrophoresis* 37: 1437–1447, 2016. doi: 10.1002/elps.201600043.
31. **Wang JY, Zhang W, Rho J-H, Roehrl MW, Roehrl MH.** A proteomic repertoire of autoantigens identified from the classic autoantibody clinical test substrate HEp-2 cells. *Clin Proteomics* 17: 35, 2020. doi: 10.1186/s12014-020-09298-3.
32. **Nunes QM, Su D, Brownridge PJ, Simpson DM, Sun C, Li Y, Bui TP, Zhang X, Huang W, Rigden DJ, Beynon RJ, Sutton R, Fernig DG.** The heparin-binding proteome in normal pancreas and murine experimental acute pancreatitis. *PLoS One* 14: e0217633, 2019. doi: 10.1371/journal.pone.0217633.

33. **Sandoval DR, Gomez Toledo A, Painter CD, Tota EM, Sheikh MO, West AMV, Frank MM, Wells L, Xu D, Bicknell R, Corbett KD, Esko JD.** Proteomics-based screening of the endothelial heparan sulfate interactome reveals that C-type lectin 14a (CLEC14A) is a heparin-binding protein. *J Biol Chem* 295: 2804–2821, 2020. doi: 10.1074/jbc.RA119.011639.
34. **Großkopf H, Vogel S, Müller CD, Köhling S, Dürig J-N, Möller S, Schnabelrauch M, Rademann J, Hempel U, von Bergen M, Schubert K.** Identification of intracellular glycosaminoglycan-interacting proteins by affinity purification mass spectrometry. *Biol Chem* 402: 1427–1440, 2021. doi: 10.1515/hsz-2021-0167.
35. **Wang JY, Zhang W, Roehrl MW, Roehrl VB, Roehrl MH.** An autoantigen profile of human A549 lung cells reveals viral and host etiologic molecular attributes of autoimmunity in COVID-19. *J Autoimmun* 120: 102644, 2021. doi: 10.1016/j.jaut.2021.102644.
36. **Joffrin AM, Hsieh-Wilson LC.** Photoaffinity Probes for the Identification of Sequence-Specific Glycosaminoglycan-Binding Proteins. *J Am Chem Soc* 142: 13672–13676, 2020. doi: 10.1021/jacs.0c06046.
37. **Guyot N, Labas V, Harichaux G, Chessé M, Poirier J-C, Nys Y, Réhault-Godbert S.** Proteomic analysis of egg white heparin-binding proteins: towards the identification of natural antibacterial molecules. *Sci Rep* 6: 27974, 2016. doi: 10.1038/srep27974.
38. **Kumar V, Hassan MI, Tomar AK, Kashav T, Nautiyal J, Singh S, Singh TP, Yadav S.** Proteomic analysis of heparin-binding proteins from human seminal plasma: a step towards identification of molecular markers of male fertility. *J Biosci* 34: 899–908, 2009. doi: 10.1007/s12038-009-0104-5.
39. **Rho J, Zhang W, Murali M, Roehrl MHA, Wang JY.** Human proteins with affinity for dermatan sulfate have the propensity to become autoantigens. *Am J Pathol* 178: 2177–2190, 2011. doi: 10.1016/j.ajpath.2011.01.031.
40. **Lee J, Rho J-H, Roehrl MH, Wang JY.** Dermatan Sulfate Is a Potential Regulator of IgH via Interactions With Pre-BCR, GTF2I, and BiP ER Complex in Pre-B Lymphoblasts. *Front Immunol* 12: 680212, 2021. doi: 10.3389/fimmu.2021.680212.
41. **Thielens NM, Gout E, Lacroix M, Reiser J-B, Gaboriaud C.** Analysis of the Ligand Recognition Specificities of Human Ficolins Using Surface Plasmon Resonance. *Methods Mol Biol* 2227: 205–226, 2021. doi: 10.1007/978-1-0716-1016-9_19.
42. **Talsma DT, Poppelaars F, Dam W, Meter-Arkema AH, Vivès RR, Gál P, Boons G-J, Chopra P, Naggi A, Seelen MA, Berger SP, Daha MR, Stegeman CA, van den Born J, COMBAT Consortium.** MASP-2 Is a Heparin-Binding Protease; Identification of Blocking Oligosaccharides. *Front Immunol* 11: 732, 2020. doi: 10.3389/fimmu.2020.00732.
43. **Zhang F, Zhao J, Liu X, Linhardt RJ.** Interactions between Sclerostin and Glycosaminoglycans. *Glycoconj J* 37: 119–128, 2020. doi: 10.1007/s10719-019-09900-3.
44. **Vallet SD, Berthollier C, Salza R, Muller L, Ricard-Blum S.** The Interactome of Cancer-Related Lysyl Oxidase and Lysyl Oxidase-Like Proteins. *Cancers (Basel)* 13: E71, 2020. doi: 10.3390/cancers13010071.
45. **Sonnhammer ELL, Östlund G.** InParanoid 8: orthology analysis between 273 proteomes, mostly eukaryotic. *Nucleic Acids Res* 43: D234–239, 2015. doi: 10.1093/nar/gku1203.

46. **Kriventseva EV, Kuznetsov D, Tegenfeldt F, Manni M, Dias R, Simão FA, Zdobnov EM.** OrthoDB v10: sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs. *Nucleic Acids Res* 47: D807–D811, 2019. doi: 10.1093/nar/gky1053.
47. **Pathan M, Keerthikumar S, Ang C-S, Gangoda L, Quek CYJ, Williamson NA, Mouradov D, Sieber OM, Simpson RJ, Salim A, Bacic A, Hill AF, Stroud DA, Ryan MT, Agbinya JI, Mariadason JM, Burgess AW, Mathivanan S.** FunRich: An open access standalone functional enrichment and interaction network analysis tool. *Proteomics* 15: 2597–2601, 2015. doi: 10.1002/pmic.201400515.
48. **Maere S, Heymans K, Kuiper M.** BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics* 21: 3448–3449, 2005. doi: 10.1093/bioinformatics/bti551.
49. **Gillespie M, Jassal B, Stephan R, Milacic M, Rothfels K, Senff-Ribeiro A, Griss J, Sevilla C, Matthews L, Gong C, Deng C, Varusai T, Ragueneau E, Haider Y, May B, Shamovsky V, Weiser J, Brunson T, Sanati N, Beckman L, Shao X, Fabregat A, Sidiropoulos K, Murillo J, Viteri G, Cook J, Shorsler S, Bader G, Demir E, Sander C, Haw R, Wu G, Stein L, Hermjakob H, D’Eustachio P.** The reactome pathway knowledgebase 2022. *Nucleic Acids Res* 50: D687–D692, 2022. doi: 10.1093/nar/gkab1028.
50. **Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, Chanda SK.** Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* 10: 1523, 2019. doi: 10.1038/s41467-019-09234-6.
51. **Clerc O, Deniaud M, Vallet SD, Naba A, Rivet A, Perez S, Thierry-Mieg N, Ricard-Blum S.** MatrixDB: integration of new data with a focus on glycosaminoglycan interactions. *Nucleic Acids Res* 47: D376–D381, 2019. doi: 10.1093/nar/gky1035.
52. **Del Toro N, Shrivastava A, Ragueneau E, Meldal B, Combe C, Barrera E, Perfetto L, How K, Ratan P, Shirodkar G, Lu O, Mészáros B, Watkins X, Pundir S, Licata L, Iannuccelli M, Pellegrini M, Martin MJ, Panni S, Duesbury M, Vallet SD, Rappsilber J, Ricard-Blum S, Cesareni G, Salwinski L, Orchard S, Porras P, Panneerselvam K, Hermjakob H.** The IntAct database: efficient access to fine-grained molecular interaction data. *Nucleic Acids Res* 50: D648–D653, 2022. doi: 10.1093/nar/gkab1006.
53. **Porras P, Barrera E, Bridge A, Del-Toro N, Cesareni G, Duesbury M, Hermjakob H, Iannuccelli M, Jurisica I, Kotlyar M, Licata L, Lovering RC, Lynn DJ, Meldal B, Nanduri B, Paneerselvam K, Panni S, Pastrello C, Pellegrini M, Perfetto L, Rahimzadeh N, Ratan P, Ricard-Blum S, Salwinski L, Shirodkar G, Shrivastava A, Orchard S.** Towards a unified open access dataset of molecular interactions. *Nat Commun* 11: 6144, 2020. doi: 10.1038/s41467-020-19942-z.
54. **Wilkinson MD, Dumontier M, Aalbersberg IJJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J-W, da Silva Santos LB, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, Gray AJG, Groth P, Goble C, Grethe JS, Heringa J, ’t Hoen PAC, Hooft R, Kuhn T, Kok R, Kok J, Lusher SJ, Martone ME, Mons A, Packer AL, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone S-A, Schultes E, Sengstag T, Slater T, Strawn G, Swertz MA, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B.** The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3: 160018, 2016. doi: 10.1038/sdata.2016.18.

55. **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.
56. **Gondelaud F, Ricard-Blum S.** Structures and interactions of syndecans. *FEBS J* 286: 2994–3007, 2019. doi: 10.1111/febs.14828.
57. **Kumar-Singh A, Shrinet J, Parniewska MM, Fuxe J, Dobra K, Hjerpe A.** Mapping the Interactome of the Nuclear Heparan Sulfate Proteoglycan Syndecan-1 in Mesothelioma Cells. *Biomolecules* 10: E1034, 2020. doi: 10.3390/biom10071034.
58. **Noborn F, Gomez Toledo A, Nasir W, Nilsson J, Dierker T, Kjellén L, Larson G.** Expanding the chondroitin glycoproteome of *Caenorhabditis elegans*. *J Biol Chem* 293: 379–389, 2018. doi: 10.1074/jbc.M117.807800.
59. **Noborn F, Gomez Toledo A, Sihlbom C, Lengqvist J, Fries E, Kjellén L, Nilsson J, Larson G.** Identification of chondroitin sulfate linkage region glycopeptides reveals prohormones as a novel class of proteoglycans. *Mol Cell Proteomics* 14: 41–49, 2015. doi: 10.1074/mcp.M114.043703.
60. **Noborn F, Nikpour M, Persson A, Nilsson J, Larson G.** Expanding the Chondroitin Sulfate Glycoproteome - But How Far? *Front Cell Dev Biol* 9: 695970, 2021. doi: 10.3389/fcell.2021.695970.
61. **Bu C, Jin L.** NMR Characterization of the Interactions Between Glycosaminoglycans and Proteins. *Front Mol Biosci* 8: 646808, 2021. doi: 10.3389/fmolb.2021.646808.
62. **Szekeres GP, Krekic S, Miller RL, Mero M, Pagel K, Heiner Z.** The interaction of chondroitin sulfate with a lipid monolayer observed by using nonlinear vibrational spectroscopy. *Phys Chem Chem Phys* 23: 13389–13395, 2021. doi: 10.1039/d1cp01975a.
63. **Szekeres GP, Pagel K, Heiner Z.** Analytical challenges of glycosaminoglycans at biological interfaces. *Anal Bioanal Chem* 414: 85–93, 2022. doi: 10.1007/s00216-021-03705-w.
64. **Gray CJ, Migas LG, Barran PE, Pagel K, Seeberger PH, Evers CE, Boons G-J, Pohl NLB, Compagnon I, Widmalm G, Flitsch SL.** Advancing Solutions to the Carbohydrate Sequencing Challenge. *J Am Chem Soc* 141: 14463–14479, 2019. doi: 10.1021/jacs.9b06406.

FIGURES

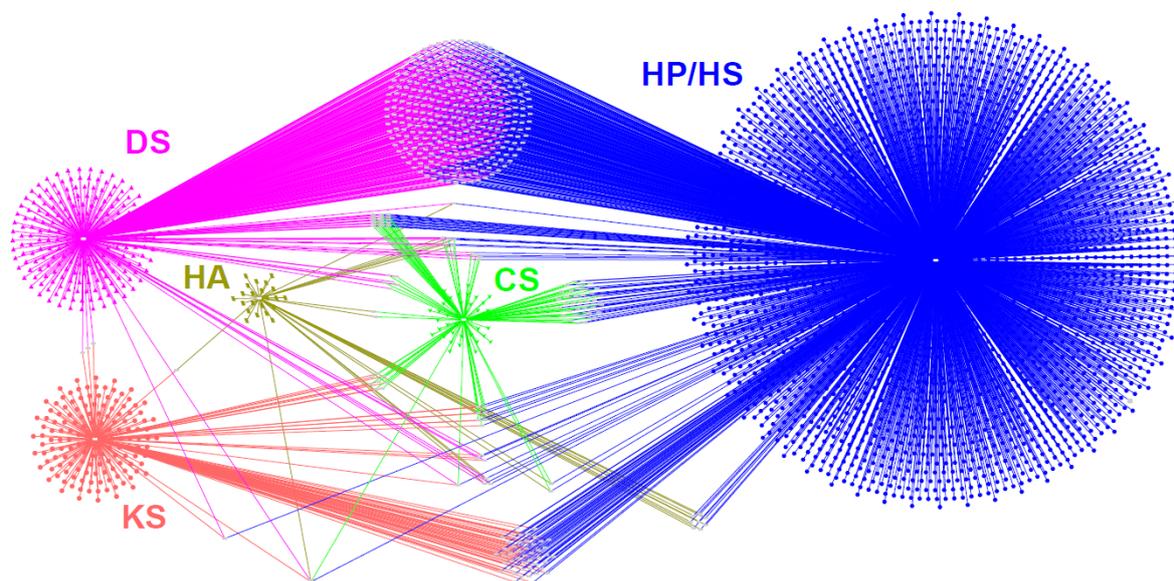


Figure 1: Interaction network of the six natural glycosaminoglycans. Glycosaminoglycan partners are represented as symbols, and interactions as edges. Abbreviations: CS, chondroitin sulfate (green); DS, dermatan sulfate (pink); HA, hyaluronan (dark yellow); HP/HS, heparin/heparan sulfate (blue); KS, keratan sulfate (light red). Partners shared by two or more GAGs are in light grey.

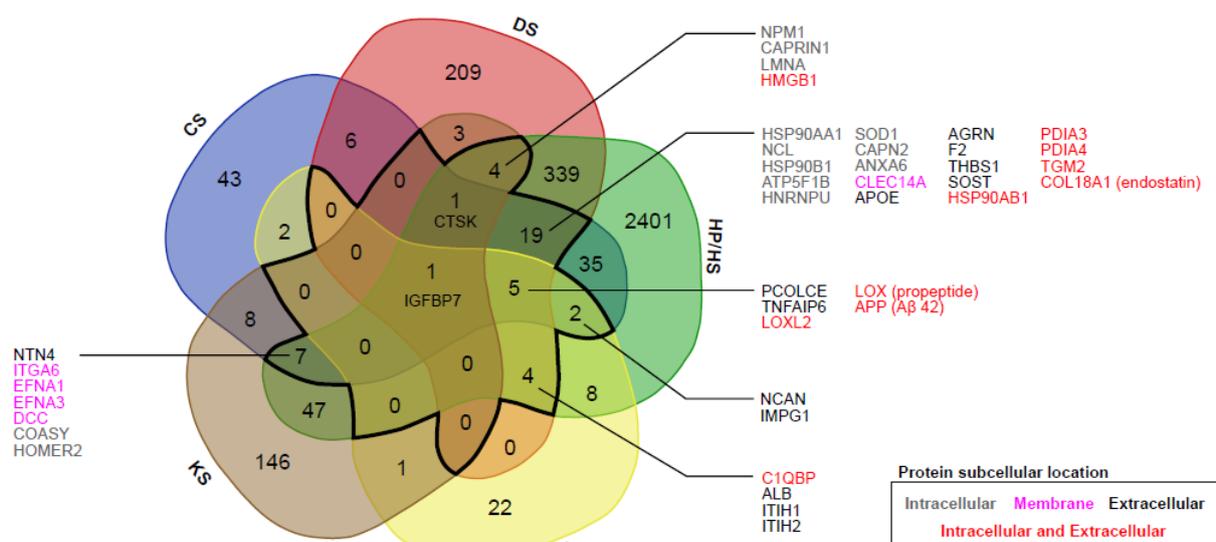


Figure 2: The Venn diagram of proteins binding to natural glycosaminoglycans. The number of specific partners of individual GAGs, and the number of the partners they share are indicated. Common and specific partners of GAGs are listed in **Supplemental Table S2** (<https://figshare.com/s/523df8467338eecd4a4>). The black bold contour includes partners shared by at least three GAGs. The gene names of the common partners are listed and color-coded according to the protein location (grey: intracellular, pink: membrane, black: extracellular, red: intracellular and extracellular). The diagram was built with a web-service (<http://bioinformatics.psb.ugent.be/webtools/Venn/>). Abbreviations: CS, chondroitin sulfate; DS, dermatan sulfate; HA, hyaluronan; HP, heparin; HS, heparan sulfate; KS, keratan sulfate.

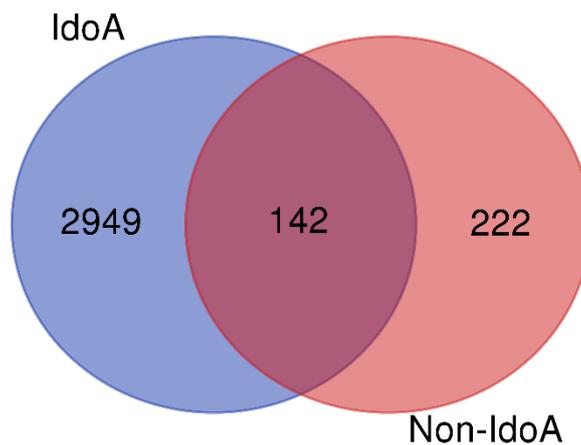


Figure 3: The Venn diagram of proteins binding to IdoA-GAGs and to non-IdoA-GAGs. The number of their specific and common protein partners are indicated. Specific and common partners of each GAG category are listed in **Supplemental Table S2** (<https://figshare.com/s/523df8467338eecd4a4>). The diagram was built with a web-service (<http://bioinformatics.psb.ugent.be/webtools/Venn/>). Abbreviations: IdoA, iduronic acid. Only unique proteins are represented in the diagram.

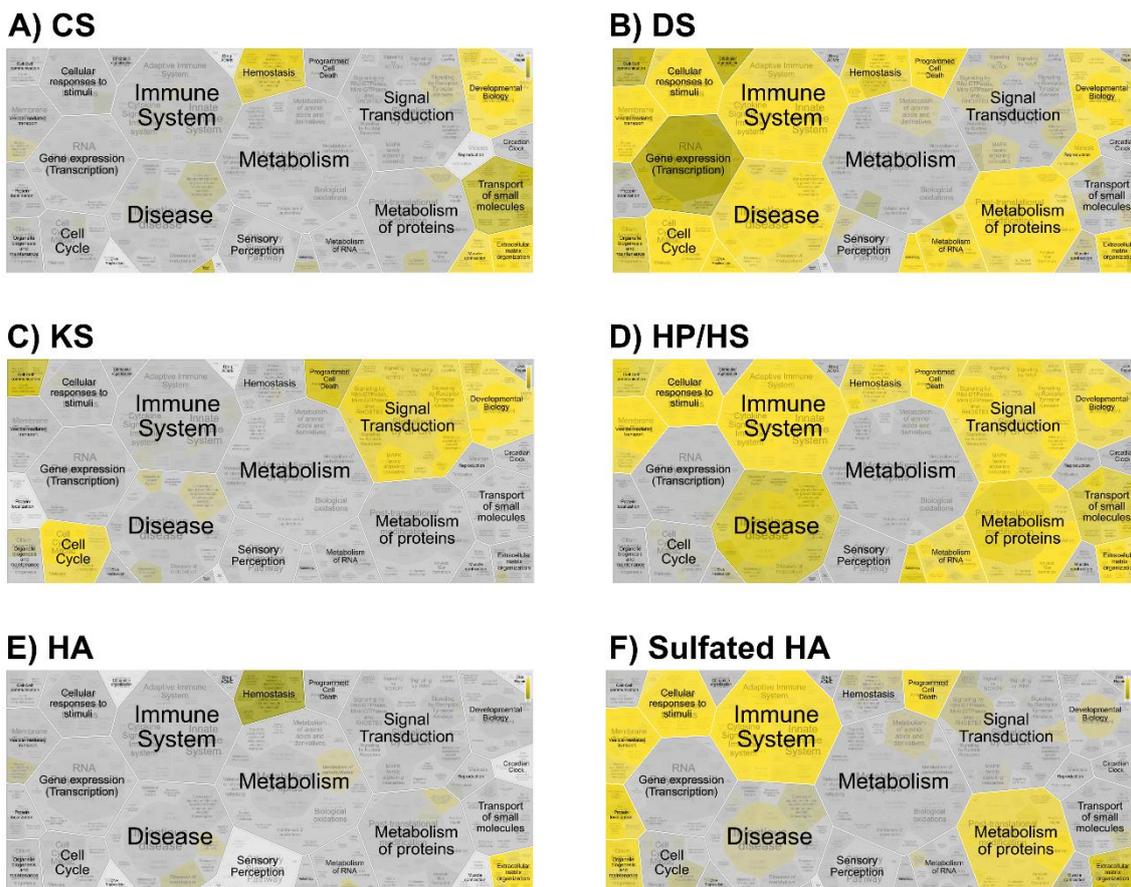
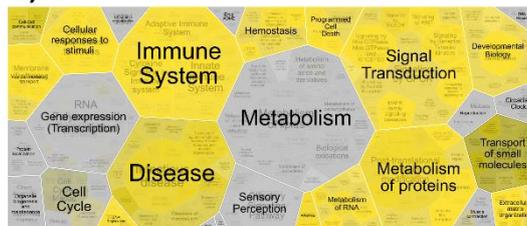
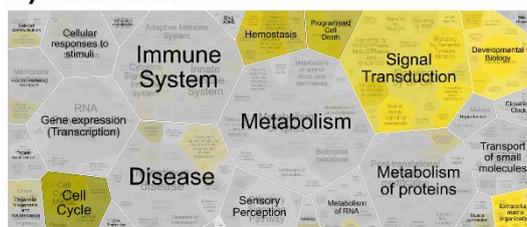


Figure 4: Pathway analysis of the glycosaminoglycan-binding proteins. A) CSBPs, B) DSBPs, C) KSBPs, D) HP/HSBPs, E) HABPs, F) Sulfated HABPs. *Homo sapiens* pathways represented as proportionately packed cells, each cell representing a Reactome pathway. The cell area is proportional to the number of proteins annotated in the pathway and its sub-pathways in humans (<https://www.ebi.ac.uk/training/online/courses/reactome-exploring-biological-pathways/understanding-the-pathway-browser/diagram-panel/pathway-overview/>). The color scale ranges from light yellow (the smallest p-value) to dark. Accessed on February, 6th 2022, and March 16th, 2022.

A) IdoA



B) Non-IdoA



C) IdoA and non-IdoA

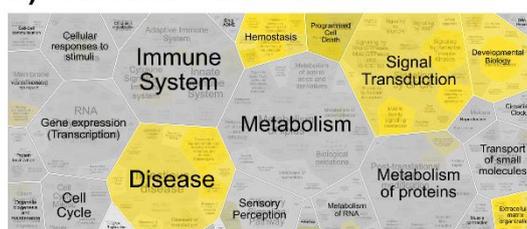


Figure 5: Pathway analysis of iduronic acid-GAGBs and non-iduronic-GAGBs. A) Iduronic acid-GAGBs, B) Non-iduronic acid-GAGBs, C) Pathway analysis of proteins binding to both IdoA-GAGs and non-IdoA-GAGs. *Homo sapiens* pathways represented as proportionately packed cells, each cell representing a Reactome pathway. The cell area is proportional to the number of proteins annotated in the pathway and its sub-pathways in humans (<https://www.ebi.ac.uk/training/online/courses/reactome-exploring-biological-pathways/understanding-the-pathway-browser/diagram-panel/pathway-overview/>). The color scale ranges from light yellow (the smallest p-value) to dark. Accessed on February, 6th 2022, and March 16th, 2022. IdoA, iduronic acid.

SUPPLEMENTAL MATERIAL (Supplemental Tables S1-S3 and Supplemental Figures S1-S6) are available as a Figshare collection at <https://figshare.com/s/523df8467338eecd4a4>