

Supplementary Table 2. **Combined list of previously known and in this study identified CSPGs.** In a review of the literature, we found 39 previously known CSPGs. In the present study we identified 13 novel CS core proteins, where two of the core proteins were found with two CS-attachment sites (osteopontin and secretogranin-1). The novel identified CSPGs are indicated by asterisks (*). Bold serine residues depict probable attachment sites while bold and underlined are established CS-attachment sites. The identified CS-structure was a residual hexasaccharide structure (referred to as “6-mer”) with or without the following modifications; one sulfate group (SO₃⁻), two sulfate groups or one phosphate group (PO₃⁻). The modifications are referred to as “S” (SO₃⁻) and “P” (PO₃⁻), respectively.

Cell Surface CSPGs

<i>UniprotKB</i>	<i>Protein Name</i>	<i>Peptide Sequence</i>	<i>Glycan Modifications</i>	<i>References</i>	<i>Fluid</i>
P05067	Amyloid beta A4 protein	⁶⁵¹ TTRPG <u>S</u> GLTN ⁶⁶⁰	CS chain	Ref (1) (2)	
O60241	*Brain-specific angiogenesis inhibitor 2	²⁵⁰ ALVPGGPAPPAAEADLHSGSSNDLFTTEMR ²⁷⁸	6-mer	This study	CSF
P61073	C-X-C chemokine receptor type 4	¹⁰ DNYTEEMG <u>S</u> GDYDSMK ²⁵	CS chain	Ref (3)	
P16070	CD44 antigen	¹⁶³ TNPEDIYPSNPTDDDDVSSGSSSER ¹⁸⁶	6-mer, 6-mer + S	Ref (4); This study	Urine
H0Y4H3	*CD99L2 protein	¹³² YGSNDDPGSGGFLNPEAPVGHVK ¹⁵⁶	6-mer	This study	CSF
Q6UVK1	CSPG4	N.D.	CS chain	Ref (5)	
O95196	CSPG5	³⁵ EAG <u>S</u> AVEAEELVK ⁴⁶	6mer	Ref (6); This study	CSF
P04233	HLA class II histocompatibility antigen gamma chain	²⁷³ LELEDPS <u>S</u> GLGVTK ²⁸⁸	6-mer + S	Ref (7); This study	Urine

O14786	Neuropilin-1	⁶⁰⁴ DDDQANCHSGTGDDDFQL ⁶²⁰	CS chain / HS chain	Ref (8)	
Q8IUK5	*Plexin domain-containing protein 1	²⁴ GAGHDEGPGSGWAAK ⁴²	6-mer, 6-mer + P	This study	CSF, Urine
Q9NZ53	Podocalyxin-like protein 2	N.D.	CS chain	Ref (9)	
Q15256	Receptor-type tyrosine-protein phosphatase R	²² FSGNNDHFLAINQK ³⁵	6-mer	Ref (10); This study	CSF
P18827	Syndecan-1	N.D.	CS chain / HS chain	Ref (11)	
P07204	Thrombomodulin	⁴⁸⁹ DSGS _S GEPPPSPT ⁵⁰⁰	CS chain	Ref (12)	
Q9UIK5	Tomoregulin-2	N.D.	CS chain	Ref (13)	

Extracellular Matrix CSPGs

UniprotKB	Protein Name	Peptide Sequence	Glycan Modifications	References	
Q9UKP4	ADAMTS-7	N.D.	CS chain	Ref (14)	
P16112	Aggrecan core protein	N.D.	CS chain / KS chain	Ref (15)	
O00468	Agrin	N.D.	CS chain / HS chain	Ref (16)	
P98160	Basement membrane-specific HSPG core protein	⁴¹⁸⁷ DWHLEGS _S GGN ⁴¹⁹⁶	HS chain / 6-mer, 6-mer + S	Ref (17); This study	CSF, Urine
P21810	Biglycan	N.D.	CS or /DS chain	Ref (18)	

Q96GW7	Brevican core protein	⁴⁰³ GAIYSIPIMEDGGGGSSTPEDPAEAPR ⁴²⁹	6-mer + S/P	Ref (19); This study	CSF
Q99715	Collagen alpha-1(XII) chain	N.D.	CS Chain	Ref (20)	
P39059	Collagen alpha-1(XV) chain	²³¹ TPEELCDPEESSASGETSGLQE ²⁵¹	6-mer + S	Ref (21); This study	Urine
P07585	Decorin	³¹ DEASGIGPEVPDDR ⁴⁴	6-mer + S	Ref (22); This study	CSF, Urine
Q16363	*Laminin subunit alpha-4	²⁸ DDNAFPFDIEGSSAVGR ⁴⁴	6-mer, 6-mer + S	This study	Urine
Q9NR99	*Matrix-remodeling-associated protein 5	⁶⁹¹ VREDIVEDEGGSGMGDEENTSR ⁷¹²	6-mer	This study	CSF
O14594	Neurocan core protein	³⁵⁸ AHHPTSQHGDLETPSSGDEGEILSAEGPPVR ³⁸⁸	6-mer, 6-mer + S, 2x 6-mer, 2x 6-mer + S/P, 2x 6-mer + 2 x S/P	Ref (23); This study	CSF
O14594	Neurocan core protein	³⁸⁹ ELEPTLEEEVVTPDFQEPLVSSGEEETLILEEK ⁴²²	6-mer	Ref (23); This study	CSF
P10451	*Osteopontin	²²³ DSYETSQLDDQSAETHSHK ²⁴¹	6-mer, 6-mer + S	This study	CSF, Urine
P10451	*Osteopontin	³⁰² ISHELDSASSEVN ³¹⁴	6-mer + S	This study	CSF, Urine
Q63HQ2	Pikachurin	N.D.	CS chain; HS chain	Ref (24)	
Q32P28	Prolyl 3-hydroxylase 1	N.D.	CS chain	Ref (25)	
Q92752	Tenascin-R	N.D.	CS chain	Ref (26)	

Q08629	Testican-1	N.D	CS chain	Ref (27)	
Q92563	Testican-2	N.D.	CS chain; HS chain (inferred from Testican 1)	Ref (28)	
Q9BQ16	Testican-3	N.D.	CS chain; HS chain (inferred from Testican 1)	Ref (29)	
P13611	Versican core protein	¹⁵²⁶ DTEVGHQAHEHTEPVS LFPEESSGEIAIDQESQK ¹⁵⁵⁹	6-mer	Ref (30); This study	CSF

Intracellular Granules CSPGs

UniprotKB	Protein Name	Peptide Sequence	Glycan Modifications	References	
P13727	Bone marrow proteoglycan	⁵³ ELEEEEEWGSSEDASK ⁶⁹	6-mer, 6-mer + P, 6-mer + S/P	Ref (31); This study	Urine
P06307	*Cholecystokinin	²¹ QPVPPADPAGSGLQR ³⁵	6-mer	This study	CSF, Urine
P10645	Chromogranin-A	⁴¹⁹ KEEEGSANR ⁴²⁷	6-mer, 6-mer + P	Ref (32); This study	CSF, Urine
Q6UXH8	*Collagen and calcium-binding EGF domain-containing protein 1	³⁸² DLGSGDDHPR ³⁹¹	6-mer, 6-mer + S/P	This study	Urine
P81605	*Dermcidin	²⁰ YDPEAASAPGSGNPCHEASAAQK ⁴²	6-mer, 6-mer + S	Ref (33); This study	Urine
Q9NQ30	Endothelial cell-specific molecule 1	¹⁵⁰ TEHDMASGDGN ¹⁶⁰	CS chain	Ref (34)	
P09603	Macrophage colony-stimulating factor 1	N.D.	CS chain	Ref (35)	
Q8N729	*Neuropeptide W	¹²⁰ APEPALEPESLDFSGAGQR ¹³⁸	6-mer, 6-mer + P	This study	Urine, CSF

Q8TB73	Protein NDF	N.D.	CS chain / HS chain	Ref (24)	
Q92954	Proteoglycan 4	N.D.	CS chain / KS chain	Ref (36)	
P05060	*Secretogranin-1	⁸⁸ DPADASEAH ⁵⁵ SSR ¹⁰¹	6-mer + P	This study	CSF, Urine
P05060	*Secretogranin-1	²³⁵ SSQESGEETGSQENHPQESK ²⁵⁴	6-mer	This study	CSF
Q8WXD2	*Secretogranin-3	³⁵ ELSAERPLNEQIAEAEEDK ⁵³	6-mer	This study	CSF
P10124	Serglycin	⁹¹ EDYSGSGFGS ¹⁰⁰	CS chain / HS chain	Ref (37)	
P01266	Thyroglobulin	²⁷⁴¹ EEEELTAGSGLREDLLSLQE ²⁷⁶⁰	CS chain	Ref (38)	
Miscellaneous CSPGs					
Q9BUR5	Apolipoprotein O	N.D.	CS chain	Ref (39)	
P02760	Bikunin (Protein AMBP)	²⁰⁶ AVLPQEEEGSGGGQLVTEVTK ²²⁶	6-mer, 6-mer +P, 6-mer + P +S	Ref (40); This study	CSF, Urine
Q99574	*Neuroserpin	³⁹⁴ VMHPETMNTSGHDFEEL ⁴¹⁰	6-mer	This study	CSF

Supplementary references

1. Pangalos, M. N., Shioi, J., Efthimiopoulos, S., Wu, A., and Robakis, N. K. (1996) Characterization of appican, the chondroitin sulfate proteoglycan form of the Alzheimer amyloid precursor protein. *Neurodegeneration* **5**, 445-451
2. Tsuchida, K., Shioi, J., Yamada, S., Boghosian, G., Wu, A., Cai, H., Sugahara, K., and Robakis, N. K. (2001) Appican, the proteoglycan form of the amyloid precursor protein, contains chondroitin sulfate E in the repeating disaccharide region and 4-O-sulfated galactose in the linkage region. *J. Biol. Chem.* **276**, 37155-37160
3. Farzan, M., Babcock, G. J., Vasilieva, N., Wright, P. L., Kiprilov, E., Mirzabekov, T., and Choe, H. (2002) The role of post-translational modifications of the CXCR4 amino terminus in stromal-derived factor 1 alpha association and HIV-1 entry. *J. Biol. Chem.* **277**, 29484-29489
4. Jackson, D. G., Bell, J. I., Dickinson, R., Timans, J., Shields, J., and Whittle, N. (1995) Proteoglycan forms of the lymphocyte homing receptor CD44 are alternatively spliced variants containing the v3 exon. *J. Cell. Biol.* **128**, 673-685
5. Levine, J. M., and Nishiyama, A. (1996) The NG2 chondroitin sulfate proteoglycan: a multifunctional proteoglycan associated with immature cells. *Perspect. Dev. Neurobiol.* **3**, 245-259
6. Aono, S., Tokita, Y., Shuo, T., Yamauchi, S., Matsui, F., Nakanishi, K., Hirano, K., Sano, M., and Oohira, A. (2004) Glycosylation site for chondroitin sulfate on the neural part-time proteoglycan, neuroglycan C. *J. Biol. Chem.* **279**, 46536-46541
7. Miller, J., Hatch, J. A., Simonis, S., and Cullen, S. E. (1988) Identification of the glycosaminoglycan-attachment site of mouse invariant-chain proteoglycan core protein by site-directed mutagenesis. *Proc. Natl. Acad. Sci. U. S. A.* **85**, 1359-1363
8. Shintani, Y., Takashima, S., Asano, Y., Kato, H., Liao, Y., Yamazaki, S., Tsukamoto, O., Seguchi, O., Yamamoto, H., Fukushima, T., Sugahara, K., Kitakaze, M., and Hori, M. (2006) Glycosaminoglycan modification of neuropilin-1 modulates VEGFR2 signaling. *EMBO J.* **25**, 3045-3055
9. Sasseti, C., Van Zante, A., and Rosen, S. D. (2000) Identification of endoglycan, a member of the CD34/podocalyxin family of sialomucins. *J. Biol. Chem.* **275**, 9001-9010
10. Snyder, S. E., Li, J., Schauwecker, P. E., McNeill, T. H., and Salton, S. R. (1996) Comparison of RPTP zeta/beta, phosphacan, and trkB mRNA expression in the developing and adult rat nervous system and induction of RPTP zeta/beta and phosphacan mRNA following brain injury. *Brain Res. Mol. Brain Res.* **40**, 79-96
11. Kokenyesi, R., and Bernfield, M. (1994) Core protein structure and sequence determine the site and presence of heparan sulfate and chondroitin sulfate on syndecan-1. *J. Biol. Chem.* **269**, 12304-12309

12. Gerlitz, B., Hassell, T., Vlahos, C. J., Parkinson, J. F., Bang, N. U., and Grinnell, B. W. (1993) Identification of the predominant glycosaminoglycan-attachment site in soluble recombinant human thrombomodulin: potential regulation of functionality by glycosyltransferase competition for serine474. *Biochem. J.* **295** (Pt 1), 131-140
13. Glynne-Jones, E., Harper, M. E., Seery, L. T., James, R., Anglin, I., Morgan, H. E., Taylor, K. M., Gee, J. M., and Nicholson, R. I. (2001) TENB2, a proteoglycan identified in prostate cancer that is associated with disease progression and androgen independence. *Int. J. Cancer* **94**, 178-184
14. Somerville, R. P., Longpre, J. M., Apel, E. D., Lewis, R. M., Wang, L. W., Sanes, J. R., Leduc, R., and Apte, S. S. (2004) ADAMTS7B, the full-length product of the ADAMTS7 gene, is a chondroitin sulfate proteoglycan containing a mucin domain. *J. Biol. Chem.* **279**, 35159-35175
15. Sandy, J. D., Flannery, C. R., Neame, P. J., and Lohmander, L. S. (1992) The structure of aggrecan fragments in human synovial fluid. Evidence for the involvement in osteoarthritis of a novel proteinase which cleaves the Glu 373-Ala 374 bond of the interglobular domain. *J. Clin. Invest.* **89**, 1512-1516
16. Winzen, U., Cole, G. J., and Halfter, W. (2003) Agrin is a chimeric proteoglycan with the attachment sites for heparan sulfate/chondroitin sulfate located in two multiple serine-glycine clusters. *J. Biol. Chem.* **278**, 30106-30114
17. Knox, S., Fosang, A. J., Last, K., Melrose, J., and Whitelock, J. (2005) Perlecan from human epithelial cells is a hybrid heparan/chondroitin/keratan sulfate proteoglycan. *FEBS Lett.* **579**, 5019-5023
18. Roughley, P. J., and White, R. J. (1989) Dermatan sulphate proteoglycans of human articular cartilage. The properties of dermatan sulphate proteoglycans I and II. *Biochem. J.* **262**, 823-827
19. Seidenbecher, C. I., Richter, K., Rauch, U., Fassler, R., Garner, C. C., and Gundelfinger, E. D. (1995) Brevican, a chondroitin sulfate proteoglycan of rat brain, occurs as secreted and cell surface glycosylphosphatidylinositol-anchored isoforms. *J. Biol. Chem.* **270**, 27206-27212
20. Yamagata, M., Yamada, K. M., Yamada, S. S., Shinomura, T., Tanaka, H., Nishida, Y., Obara, M., and Kimata, K. (1991) The complete primary structure of type XII collagen shows a chimeric molecule with reiterated fibronectin type III motifs, von Willebrand factor A motifs, a domain homologous to a noncollagenous region of type IX collagen, and short collagenous domains with an Arg-Gly-Asp site. *J. Cell Biol.* **115**, 209-221
21. Li, D., Clark, C. C., and Myers, J. C. (2000) Basement membrane zone type XV collagen is a disulfide-bonded chondroitin sulfate proteoglycan in human tissues and cultured cells. *J. Biol. Chem.* **275**, 22339-22347

22. Seo, N. S., Hocking, A. M., Hook, M., and McQuillan, D. J. (2005) Decorin core protein secretion is regulated by N-linked oligosaccharide and glycosaminoglycan additions. *J. Biol. Chem.* **280**, 42774-42784
23. Rauch, U., Feng, K., and Zhou, X. H. (2001) Neurocan: a brain chondroitin sulfate proteoglycan. *Cell Mol. Life Sci.* **58**, 1842-1856
24. Manabe, R., Tsutsui, K., Yamada, T., Kimura, M., Nakano, I., Shimono, C., Sanzen, N., Furutani, Y., Fukuda, T., Oguri, Y., Shimamoto, K., Kiyozumi, D., Sato, Y., Sado, Y., Senoo, H., Yamashina, S., Fukuda, S., Kawai, J., Sugiura, N., Kimata, K., Hayashizaki, Y., and Sekiguchi, K. (2008) Transcriptome-based systematic identification of extracellular matrix proteins. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 12849-12854
25. Wassenhove-McCarthy, D. J., and McCarthy, K. J. (1999) Molecular characterization of a novel basement membrane-associated proteoglycan, leprecan. *J. Biol. Chem.* **274**, 25004-25017
26. Probstmeier, R., Stichel, C. C., Muller, H. W., Asou, H., and Pesheva, P. (2000) Chondroitin sulfates expressed on oligodendrocyte-derived tenascin-R are involved in neural cell recognition. Functional implications during CNS development and regeneration. *J. Neurosci. Res.* **60**, 21-36
27. Bonnet, F., Perin, J. P., Maillet, P., Jolles, P., and Alliel, P. M. (1992) Characterization of a human seminal plasma glycosaminoglycan-bearing polypeptide. *Biochem. J.* **288 (Pt 2)**, 565-569
28. Vannahme, C., Schubel, S., Herud, M., Gosling, S., Hulsmann, H., Paulsson, M., Hartmann, U., and Maurer, P. (1999) Molecular cloning of testican-2: defining a novel calcium-binding proteoglycan family expressed in brain. *J. Neurochem.* **73**, 12-20
29. Nakada, M., Yamada, A., Takino, T., Miyamori, H., Takahashi, T., Yamashita, J., and Sato, H. (2001) Suppression of membrane-type 1 matrix metalloproteinase (MMP)-mediated MMP-2 activation and tumor invasion by testican 3 and its splicing variant gene product, N-Tes. *Cancer Res.* **61**, 8896-8902
30. Dours-Zimmermann, M. T., and Zimmermann, D. R. (1994) A novel glycosaminoglycan attachment domain identified in two alternative splice variants of human versican. *J. Biol. Chem.* **269**, 32992-32998
31. Oxvig, C., Haaning, J., Hojrup, P., and Sottrup-Jensen, L. (1994) Location and nature of carbohydrate groups in proform of human major basic protein isolated from pregnancy serum. *Biochem. Mol. Biol. Int.* **33**, 329-336
32. Gowda, D. C., Hogue-Angeletti, R., Margolis, R. K., and Margolis, R. U. (1990) Chromaffin granule and PC12 cell chondroitin sulfate proteoglycans and their relation to chromogranin A. *Arch. Biochem. Biophys.* **281**, 219-224

33. Todorov, P. T., Deacon, M., and Tisdale, M. J. (1997) Structural analysis of a tumor-produced sulfated glycoprotein capable of initiating muscle protein degradation. *J. Biol. Chem.* **272**, 12279-12288
34. Sarrazin, S., Lyon, M., Deakin, J. A., Guerrini, M., Lassalle, P., Delehedde, M., and Lortat-Jacob, H. (2010) Characterization and binding activity of the chondroitin/dermatan sulfate chain from Endocan, a soluble endothelial proteoglycan. *Glycobiology* **20**, 1380-1388
35. Suzu, S., Ohtsuki, T., Yanai, N., Takatsu, Z., Kawashima, T., Takaku, F., Nagata, N., and Motoyoshi, K. (1992) Identification of a high molecular weight macrophage colony-stimulating factor as a glycosaminoglycan-containing species. *J. Biol. Chem.* **267**, 4345-4348
36. Schumacher, B. L., Block, J. A., Schmid, T. M., Aydelotte, M. B., and Kuettner, K. E. (1994) A novel proteoglycan synthesized and secreted by chondrocytes of the superficial zone of articular cartilage. *Arch. Biochem. Biophys.* **311**, 144-152
37. Kolset, S. O., and Tveit, H. (2008) Serglycin--structure and biology. *Cell Mol. Life Sci.* **65**, 1073-1085
38. Conte, M., Arcaro, A., D'Angelo, D., Gnata, A., Mamone, G., Ferranti, P., Formisano, S., and Gentile, F. (2006) A single chondroitin 6-sulfate oligosaccharide unit at Ser-2730 of human thyroglobulin enhances hormone formation and limits proteolytic accessibility at the carboxyl terminus. Potential insights into thyroid homeostasis and autoimmunity. *J. Biol. Chem.* **281**, 22200-22211
39. Lamant, M., Smih, F., Harmancey, R., Philip-Couderc, P., Pathak, A., Roncalli, J., Galinier, M., Collet, X., Massabuau, P., Senard, J. M., and Rouet, P. (2006) ApoO, a novel apolipoprotein, is an original glycoprotein up-regulated by diabetes in human heart. *J. Biol. Chem.* **281**, 36289-36302
40. Ly, M., Leach, F. E., 3rd, Laremore, T. N., Toida, T., Amster, I. J., and Linhardt, R. J. (2011) The proteoglycan bikunin has a defined sequence. *Nat. Chem. Biol.* **7**, 827-833