Proteomics, Glycomics, and Glycoproteomics of Matrisome Molecules

Authors
Rekha Raghunathan, Manveen K. Sethi, Joshua A. Klein, and Joseph Zaia

Correspondence
jzaia@bu.edu

Graphical Abstract

In Brief
Extracellular networks of matrisome proteins and their binding partners give rise to dynamic cell and tissue-specific microenvironments. The extreme complexity of matrisome molecule glycosylation and other post-translational modifications belies the need for specialized omics methods. It is necessary to map the modifications of matrisome molecules in detail in order to understand their roles in normal and pathological physiology. We review proteomics, glycomics and glycoproteomics methods for matrisome molecules toward the goal of achieving detailed matrisome maps.

Highlights
- Highly glycosylated matrisome proteins and their binding partners comprise extracellular networks that mediate tissue-specific cellular microenvironments.
- Elucidation of roles of matrisome molecules in disease mechanisms requires detailed mapping of matrisome glycosylation and other post-translational modifications.
- We review tissue workup methods for matrisome proteomics, glycomics and glycoproteomics.
- The combination of proteomics, glycomics and glycoproteomics profiles matrisome protein modifications distinct from those studied by immunohistochemistry.

Raghunathan et al., 2019, Molecular & Cellular Proteomics 18, 2138–2148
November 2019 © 2019 Raghunathan et al. Published under exclusive license by The American Society for Biochemistry and Molecular Biology, Inc.
https://doi.org/10.1074/mcp.R119.001543
Proteomics, Glycomics, and Glycoproteomics of Matrisome Molecules*

Rekha Raghunathan‡§, Manveen K. Sethi§, Joshua A. Klein¶, and Joseph Zaia‡§¶

The most straightforward applications of proteomics database searching involve intracellular proteins. Although intracellular gene products number in the thousands, their well-defined post-translational modifications (PTMs) makes database searching practical. By contrast, cell surface and extracellular matrisome proteins pass through the secretory pathway where many become glycosylated, modulating their physicochemical properties, adhesive interactions, and diversifying their functions. Although matrisome proteins number only a few hundred, their high degree of complex glycosylation multiplies the number of theoretical proteoforms by orders of magnitude. Given that extracellular networks that mediate cell-cell and cell-pathogen interactions in physiology depend on glycosylation, it is important to characterize the proteomes, glycomes, and glycoproteomes of matrisome molecules that exist in a given biological context. In this review, we summarize proteomics approaches for characterizing matrisome molecules, with an emphasis on applications to brain diseases. We demonstrate the availability of methods that should greatly increase the availability of information on matrisome molecular structure associated with health and disease.

Molecular & Cellular Proteomics 18: 2138–2148, 2019. DOI: 10.1074/mcp.R119.001543.

Viewed in evolutionary terms, the diversification of genes for intracellular versus extracellular proteins has separate drivers (1–3). Evolution of intracellular biology focuses on the regulation of signaling events, transcription and translation, through phosphorylation and other post-translational modifications (PTMs) that influence allosteric enzyme regulation and signaling cascades through activation/deactivation of recognition domains, for example, SH2, SH3, bromo-, chromo- and tudor domains (4, 5). Significantly, many of the PTMs that occur inside the cell produce well-defined molecular additions (phosphorylation, acetylation, acylation and methylation) that are compatible with the established database searching workflows (5–7). Ubiquitination leaves a recognizable peptide tag after tryptic digestion that is amenable to proteomics approaches (8–10). Evolution of such PTMs arose through the need for complex control of gene expression through regulated signaling networks. By contrast, complex glycosylation reflects the evolutionary response to pathogen pressure and the evolving need for multicellular complexity (2, 3). Organisms need a multicellular organization with the ability to distinguish self from non-self and orchestrated responses to infection and regulated tissue plasticity. Therefore, much of cellular biology responds to signals received from the extracellular environment. Complex glycosylation is heterogeneous as a rule at each protein site, multiplying the number of molecular forms and requiring specialized proteomics methods.

As summarized in recent reviews (11–13), the mechanisms of tissue homeostasis and most diseases include interactions with the extracellular microenvironment. The matrisome constitutes the non-cellular components that control biochemical and biomechanical cues, growth factor and morphogen gradients, and physical scaffolds that define tissue phenotypes including morphogenesis, differentiation, and homeostasis (12). Although each tissue has a unique extracellular environment, the number of gene products that code matrisome proteins in the entire body are limited (14–18). Cell surface receptors regulate adhesion and cytoskeletal connections to the matrisome. The structure and organization of the matrisome require maintenance as it adapts to tissue growth needs. Matrisome proteins become glycosylated in the secretory pathway, may be proteolytically processed and cross-linked. The resulting physical and biochemical characteristics reflect the organized networks that depend on numerous molecular interactions that arise from PTMs.

To date, the ability to apply established proteomics methods that depend on database searching to such highly modified and heterogeneous proteins remains far from adequate (19). Although we can detect many matrisome proteins using proteomics, the low sequence coverage leaves many structural elements of functional interest unidentified. In addition, the heterogeneity of glycosylation at each of the many glycosylation sites of matrisome proteins results in astronomically large numbers of possible proteoforms, if taken as multiples...
of the variants at each site. Although the existence of such large numbers of proteoforms seems unlikely, the number of functional proteoforms that exist in a given biological context remains largely undefined.

In this review, we summarize methods for proteomics, glycomics and glycoproteomics of matrisome molecules. The goal is to characterize matrisome molecular structure in the greatest detail possible using wide-angle omics experiments. The high extent of matrisome protein glycosylation and other post-translational modifications requires special consideration of sample workup and proteomics database searching. We summarize matrisome physiology with emphasis on brain diseases. We summarize experimental approaches for matrisome workup and mass spectrometric analysis.

**Extracellular Matrix Physiology and Pathophysiology**—Dysregulation of the cellular microenvironment occurs in cancers (20–22), neurodevelopmental and neuropsychiatric diseases (23, 24). Known as the matrisome, networks of extracellular matrix and cell surface molecules control the availability of growth factors to cellular receptors and the mechanical-physical properties of the cell microenvironment. Currently, the limited understanding of the regulation of matrisome glycosylation hinders understanding of the roles of glycosylation-dependent matrisome networks in the basic mechanisms necessary for the targeted intervention of many diseases.

The extracellular environment (14–16) consists of glycoproteins, proteoglycans, collagens, and their interacting partners. Matrisome protein functions are elaborated by biosynthetic enzymes of the secretory pathway that generate mature molecules with spatially and temporally regulated glycosylation. Thus, glycoproteins have context-specific structures and biological functions that remain largely undefined because of the lack of effective methods for quantifying changes to site-specific protein glycosylation. This means that it is necessary to achieve complete matrisome protein coverage in order to determine the changes to these molecules that occur during disease mechanisms.

Progress in developing treatments that target interactions among cells and their extracellular microenvironments, including dysregulated cell growth, morphogenesis, and host-pathogen interactions, is limited by the ability to quantify the extent to which changes in matrisome networks resulting from altered glycoprotein glycosylation determine disease mechanisms. Many matrisome proteins contain lectin domains that recognize glycan epitopes. Thus, the glycosylation of matrisome molecules determines their binding interactions with other matrisome molecules and with soluble growth factors. The result is organized assemblies of matrisome molecules that compose the spatially and temporally regulated microenvironments through which cells receive signals in physiology and pathophysiology.

The core matrisome consists of 195 glycoproteins, 44 collagens and 35 proteoglycans (16, 18, 25), many of which are glycosylated. Extensive proteomics studies have cataloged the abundances of matrisome molecules in different tissues (18); however, such studies have not defined the glycosylation states of matrisome molecules necessary to define networks of interactions with lectin-containing binding partners. Many matrisome molecules have several points of glycosylation, each with microheterogeneity, each representing a functional epitope that needs to be defined in order to characterize biological functions (26).

Glycan branching regulates glycoprotein dynamics and the residency of cell surface cytokine receptors (27). Interactions among cell surface and extracellular glycoproteins with lectins including galectins, C-type lectins, and siglecs, drive the clustering of cell surface molecules into networks that define tissue microenvironments. These organized assemblies of extracellular molecules (the matrisome) adapt cellular microenvironments to phenotypic needs. Although glycosylated proteins represent potential therapeutic targets, their macro- and micro-heterogeneities pose a significant challenge to exploitation. Because their functions depend on glycosylation and other PTMs, it is necessary to produce detailed proteolytic maps of matrisome proteins and follow the changes that occur during aging and disease development.

The **Dynamic Brain Matrisome**—The brain extracellular space has been referred to as the final frontier in neuroscience (28). Many of the matrisome molecules common to systemic organs are not found in brain (29). Matrisome functions depend on networks of interaction among glycosylated proteins and glycan-binding lectins. As illustrated in Fig. 1 for the brain, networks of cell surface and extracellular glycoproteins and proteoglycans bind many families of growth factors and growth factor receptors (30–32). They modulate receptor tyrosine kinase signaling pathways at the heart of mechanisms including tissue stiffness and growth factor transport.

In the brain, the extracellular space occupies ~20% of brain volume (28, 33). As shown in Fig. 1, matrisome in the central nervous system includes the interstitial matrix, basement membranes, and perineuronal nets (PNNs), the fine structures of which vary spatially and temporally (34, 35). The matrisome provides the environment necessary for cell homeostasis, repair, regeneration, and neural plasticity in a brain region-specific manner (36–38).

The basement membranes that line cerebral blood vessels consist of collagen IV, laminins and heparan sulfate proteoglycans (HSPGs) perlecan and agrin (39, 40). In the brain, neural interstitial matrix separates cells and consists of networks of chondroitin sulfate proteoglycans (CSPGs), tenascins, hyaluronan, and link proteins. The PNN consist of many of the same CSPGs, tenascins, hyaluronan, and link proteins condensed that surround some neuronal cell bodies and dendrites.
Dysregulation of matrisome molecular networks characterize pathophysiology of cancer (41, 42), autoimmune (43–45), cardiovascular (46–48), and neuropsychiatric diseases (23, 24, 34). In the brain, region-specific regulation of matrisome molecule glycosylation controls the neuronal microenvironment and becomes dysregulated in neuropsychiatric diseases (23, 24, 34, 49). For example, in neurodegeneration, proteoglycans bind to and play roles in the aggregation of proteins including Aβ, tau, prion protein, and α-synuclein (50, 51). In Alzheimer’s disease, proteoglycans participate in amyloid plaque formation, leading to disease pathology from altered proteolytic processing and accumulation of toxic aggregates (52, 53).

Perineuronal nets are lattices of matrisome molecules that surround the cell body and dendrites of neurons. They are thought to serve as a reservoir for cations and provide the connectional architecture that controls synaptic plasticity. Deficits in PNN structure appear to contribute to dysfunction in cortical circuitry in schizophrenia (54). The number of PNN in visual cortex increases during postnatal development, paralleling the critical period for synaptic plasticity and playing an important role in critical period closure. Significantly, the adult inability to repair spinal cord injury can be restored by the treatment of the injury site with chondroitinase enzymes (55, 56) with the implication that CSPGs maintain the extracellular environment in adult neural tissue that limits neural plasticity (37, 57–59). PNN structure differs spatially and temporally in the brain in association with injury, repair, development, aging, learning, memory, neuropsychiatric diseases, neurodegeneration, and in response to drug abuse (24, 60, 61).

Variation in Matrisome Molecular Structure Among Brain Regions, With Development, Aging, and Pathologies—Traditional antibody-based techniques including immunohistochemistry show spatial and temporal regulation of matrisome molecule expression in the brain (62–64). Although antibody binding indicates the levels of individual epitopes, the anti-
body specificity and underlying structure are assumed. Matrisome molecule glycosylation can also be stained using lectins, leaving the underlying matrisome site-specific glycosylation structure undefined (24, 54, 65).

Because antibodies bind to discrete structural epitopes on highly complex matrisome molecules, the changes in structure unrelated to such epitopes are not defined by antibody-based techniques. This is illustrated for the aggrecan proteoglycan in Fig. 2. Aggrecan contains three globular domains and an extended region modified with more than 100 CS chains. The C-terminal G3 domain has EGF-like repeats, a C-type lectin domain, and a complement regulatory protein domain (CRP) (66) and interacts with tenascins. The globular domains are N-glycosylated, and the extended domains also carry keratan sulfate, mucin-type O-GalNAc glycans, and O-Mannose glycans (67).

Although this provides a parts list for aggrecan, we have little information on how the variation of glycosylation modulates the functions of aggrecan in a weight-bearing tissue such as cartilage versus brain. Further, although it is clear that CS structure varies among brain regions, there is little information on the fine structures of the resulting matrisome molecular networks. The same reasoning applies to the other brain matrisome network glycoproteins, proteoglycans, and collagens.

It is clear, however, that matrisome glycosylation changes during development and with disease states. This can be seen from the alteration of CS sulfation during development in many tissues, including cartilage (68) and brain (69, 70). Further, staining of brain tissue with Wisteria floribunda agglutinin (WFA), a lectin that binds GalNAc residues, has been used to identify brain region pathologies associated with schizophrenia (71–74). The extreme complexities of matrisome proteins concerning glycosylation, cross-linking and other PTMs, drives the need for many more validated antibody reagents than are available to date (64). Such lectin and/or antibody staining studies do not define the glycosite changes that underlie dysregulated interactions with lectin-containing binding partners that give rise to pathologies.

**Matrisome Proteomics**

**Matrisome Sample Preparation Methods**—As summarized in Table I and reviewed in detail (75), methods for enrichment of matrisome proteins for proteomics studies include tissue decellularization and extraction of matrisome components from tissue homogenates. Proteomics researchers often use decellularization to remove cellular components prior to solubilization of matrisome molecules (76–84). The use of a chaotrope solubilizes some matrisome components, leaving an insoluble pellet rich in fibrillar collagens (85). The yield of such matrisome components improves with chemical digestion with cyanogen bromide (86). Alternatively, hydroxylamine cleavage at Asn-Gly sites has also been used to solubilize matrisome from insoluble pellets prior to tryptic digestion and proteomics (82, 86). For matrisome protein quantification among tissue sample cohorts, Hansen et al. have developed targeted proteomics (87–89).

**Proteomics Data Acquisition Methods for the Analysis of Matrisome Proteins**—Present discovery proteomics methods suffice to identify matrisome proteins based on the presence of minimally modified peptides using database searching. Such peptides have been used in targeted proteomics assays for quantification of matrisome molecules based on inferred core protein abundances (87–89). Data-independent analysis (DIA) has the advantage that all precursor ions are subjected to collisional dissociation. Using the sequential window acquisition of all theoretical fragment ion spectra (SWATH)-MS DIA method (90), fragment ion spectra for all precursors are acquired within the specified m/z range and retention time window. Interpretation of such datasets in which tandem mass spectra show product ions from co-eluting peptides requires the use of spectral libraries (91). Huang et al. devel-
oped a spectral library of 201 matrisome proteins and compared the performance of SWATH versus data-dependent acquisition (DDA) for analysis of unfractionated tissue extracts (92). They reported a 15–20% improvement in peptide reproducibility and a 54% increase in several matrisome proteins identified relative to DDA. Önnerfjord et al. used high pH reversed-phase fractionation of tryptic digests as a workflow for cartilage proteomics (93). Because of the additional fractionation step, they reported 653 proteins identified. They used DDA data to build spectral libraries for interpretation of a subset of identified proteins using DIA. They showed that DIA produced a more precise measurement of peptide abundances than DDA.

Naba et al. used a commercial Cytosol/Nucleus/Membrane/Cytoskeleton compartmental protein extraction kit to enrich intracellular and matrisome proteins in separate fractions from tissue (18, 94–96). The tryptic digests of the matrisome-enriched pellets were solubilized using urea prior to LC-tandem MS (97). This approach resulted in the identification of ~250 matrisome proteins from tissue (94, 95, 98, 99, 100).

Mayr et al. extracted matrisome proteins from vascular tissue using decellularization, solubilization using a three-stage extraction (salt, detergent, guanidine HCl) and MS-based quantification (101–105). In studies of human venous tissue, they reported the identification of ~150 matrisome proteins. They identified a proteomics 4-biomarker signature for atherosclerotic plaques from a comparison of vascular matrisome in human carotid artery specimens (106). In this work, they report 110 matrisome-associated proteins from guanidine HCl extraction and 87 from the salt fractions with an overlap of 51. They also performed matrisome proteomics studies of restenosis and thrombosis following coronary stent implantation in pigs, for which they report the identification of 151 matrisome proteins (107, 108).

Berretta et al. demonstrated, using a combination of immunohistochemistry and proteomics, that matrisome molecule expression is brain region-dependent (34). For this work, fresh rat brains were dissected, and regions were snap frozen. Tryptic peptides were fractionated using ERLIC, and the resulting fractions analyzed using LC-MS. The fold change abundances of a set of 17 matrisome molecules, including tenascins, hyalectans, link proteins, and others were reported for a set of five rat brain regions.

**MALDI Imaging of Matrisome Proteins**—As described in detail in recent reviews (109–111), MALDI imaging mass spectrometry (IMS) produces 2-dimensional maps of the distributions of ions desorbed from the surfaces of tissue slides. The advantage is that the maps can be produced at ~25 μm or better resolution and with impressive ion-specific spatial resolution patterns. The disadvantages are that in the absence of a separation step, the dynamic range of protein/peptide detection is limited, and identification of observed proteins or peptides can be cumbersome. Drake et al. have

| Enrichment method prior to LC-MS/MS analysis | Tissue type | Number of matrisome proteins identified | Reference | Limitations |
|---------------------------------------------|-------------|----------------------------------------|------------|-------------|
| High pH reversed phase fractionation          | Human knee articular cartilage | 196 (out of total 653 protein IDs) | Önnerfjord et al. (93) | Additional fractionation step, false positives and limited list of proteins |
| Commercial compartmental protein extraction kit and urea fractionation | Murine normal and melanoma tumor tissues | 200 | Naba et al. (18) | Semi-quantitative estimates, false positives and limited list of proteins |
| Decellularization, solubilization using a three-stage extraction (salt, detergent, guanidine HCl) | Atherosclerotic plaque tissue | ~200 | Mayr et al. (106) | False positives and limited list of proteins |
| Decellularization, solubilization using a three-stage extraction (salt, detergent, guanidine HCl) | Pig coronary arteries | 151 | Mayr et al. (105) | False positives and limited list of proteins |
| Decellularization, solubilization using a three-stage extraction (salt, detergent, guanidine HCl) | Rat brain (Striatum and Substantia nigra) | 17 | Berretta S et al. (34) | Additional fractionation step, low protein coverage |
| Decellularization, solubilization using a three-stage extraction (salt, detergent, guanidine HCl) | Rat brain | 15 | Zaia J et al. (130) | Additional fractionation step, low protein coverage |
demonstrated the use of MALDI-imaging to visualize proteins and peptides from matrisome-rich tissues, including heart (112). They demonstrated the use of matrix metalloproteinase enzymes to localize collagen and elastin peptides on the surfaces of the tumor and cardiac tissue slides (113). This group also pioneered MALDI-based imaging of glycans at ~25 μm spatial resolution on tissue slides (114, 115) for which they derivatized sialic acid residues to prevent dissociation resulting from the MALDI process.

Matrisome Glycomics and Proteomics from Histological Slides—We developed a workflow for profiling GAGs, N-glycans, and proteins from tissue slides (116, 117). We applied the method to comparative glycomics profiling from invertebrates (118–120), mammalian organ tissues (121–123), skeletal muscle (124), kidney tissues (125, 126) leukocytes (127), stem cell niche (128), and tumor tissues (117, 129). We have investigated brain aging (130) and neuropathological diseases including glioma (117). Our method provides a readout of GAG quantities, domain structures, and non-reducing end structures using simple enzyme digestions with minimal need for workup. The final proteomics of tryptic peptides identifies ~1200 proteins from the 10 nL tissue volume, providing deeper coverage than can be obtained from an MS imaging approach.

This approach requires small tissue volumes, minimal sample workup, and reduces the effort required per biospecimen for glycomics and proteomics studies. Fresh frozen slides are washed with a series of solvents, thereby denaturing tissue proteins. Formalin-fixed, paraffin embedded tissue slides require dewaxing, re-hydration, and high pH antigen retrieval prior to enzymatic digestion. Although proteins are denatured in both cases, the observed glycomics and proteomics profiles reflect tissue processing biases that remain to be studied in detail. Nonetheless, the analysis of matrisome molecules from tissue slides offers an attractive option to extraction from wet tissue in terms of lower sample quantities and effort required. For example, in a study of aging rat brain from tissue slides with no enrichment, we observed 9–11% of total proteins of extracellular origin, corresponding to 15 matrisome molecules (130).

Matrisome Glycoproteomics—Application of a conventional discovery proteomics workflow with database searching identified Matrisome molecules based on the presence of unmodified peptides. Although homogeneous PTMs including phosphorylation, acetylation, methylation, and ubiquitination are amenable to proteomics database searching (131, 132), glycosylation is heterogeneous as a rule. This multiplies the number of PTM forms of a given matrisome molecule glycopeptide, thereby dividing the precursor ion signals, and multiplying the size of the proteomics search space and the difficulty of assigning the glycopeptide with confidence (133–139). As shown in Fig. 3, the presence of complex glycosylation alters the collisional dissociation pattern of peptides significantly. Glycopeptide collisional dissociation tandem mass spectra show low m/z oxonium signature ions that indicate the presence of glycosylation. The spectra also show peptide + saccharide ions, the abundances of which depend on the extent of vibrational excitation of the precursor ions. If relatively low collision energies are used, then product ions resulting from losses of saccharide units are abundant. At higher collision energies, peptide plus from one to a few monosaccharide units are observed. Under such conditions, dissociation of the peptide backbone is often observed albeit at relatively low abundances. Thus, the most confident collisional tandem mass spectra for glycopeptide precursor ions contain all three ion types as shown for example for an aggrecan glycopeptide in Fig. 3 (140).

Lectin Enrichment of Glycopeptides—Investigators have used WFA and concanavalin A (ConA) lectin enrichment of guanidine HCl extracts to enrich glycoproteins from human cardiac tissue from which they reported identification of 65 glycosylation sites from 35 extracellular proteins (141, 142). Wheat germ agglutinin (WGA) was used to enrich glycoproteins from mouse brain to identify O-mannosylated peptides
from neurofascin 186 (143) and PNN associated hyalectan proteoglycans (67). O-Mannosylated peptides have been enriched from tissue extracts digested using trypsin and peptide-N-glycosidase F using ConA lectin chromatography (144, 145). A set of 16 O-mannosylated glycoproteins were identified, several belonging to the cadherin superfamily, using this approach.

Proteoglycan Glycoproteomics—Enzymatic digestion of GAG chains leaves a glycopeptide with linker saccharide attached to the core protein. Such linker glycopeptides can be identified by the presence of a diagnostic oxonium ion for CS and HS proteoglycans (146). The linker saccharide glycopeptides detected for CSPGs were modified with sulfate, phosphate, fucose and/or sialic acid (147–149). This approach has been used to analyze PGs from biological fluids including urine and cerebrospinal fluid (149). We used a similar approach for analysis of purified proteoglycans including aggrecan, decorin, brevican and neurocan (140). In order to interpret the glycopeptide tandem mass spectra automatically from the LC-tandem MS datasets, we optimized our GlycReSoft software (150, 151) for interpretation of linker-glycopeptides.

We have observed glycopeptide abundances too low for confident identification when analyzing proteolytic digests from tissue slides. It, therefore, appears that enrichment steps will be necessary to allow glycoproteomics from the tissue. Although such enrichment remains a challenge from small tissue volumes such as obtained from tissue slides, it seems feasible from wet tissue extracts.

CONCLUSIONS

For researchers interested in profiling abundances of matrisome core proteins, the use of decellularization or enrichment methods combined with targeted MS or DIA MS seems appropriate. As in other areas of proteomics, the use of multidimensional separations increases the number of proteins identified at the expense of analysis time and cost. One of these separation dimensions can be designed to enrich glycopeptides, thus increasing the ability to detect matrisome determinants of molecular networks. Such enrichment steps are most readily applied to tissue extracts. The analysis of tissue slides has potential benefits in terms of throughput, cost, and applicability to pathological workflows. The tissue volume, however, is rather low, making use of enrichment steps challenging. On the other hand, tissue slides can be micродissected (152), increasing the ability to select cell populations of interest for subsequent proteomics. Robotic approaches for manipulation of microdissected tissue have been described (153, 154). It may, therefore, be feasible to use glycopeptide enrichment in such robotic workflows to enable the application of glycoproteomics LC-MS methods to microdissected tissue. This will enable profiling of designated matrisome glycosylates as a means for assessing changes to extracellular networks during disease mechanisms.

* We acknowledge support from NIH grants P41GM104603 and U01CA221234.

Author contributions: R.R., M.K.S., J.A.K., and J.Z. wrote the paper.

|| To whom correspondence should be addressed: Boston University Medical Campus, 670 Albany St., Rm. 505, Boston, MA 02118. Tel.: 617-358-2429, E-mail: jzaia@bu.edu.

REFERENCES

1. Demetriou, M., Nabi, I. R., and Dennis, J. W. (2018) Galactinins as adaptors: linking glycosylation and metabolism with extracellular cues. Trends in Glycosci. Glycotechnol. 30, SE167–SE177
2. Dennis, J. W. (2017) Genetic code asymmetry supports diversity through experimentation with posttranslational modifications. Curr. Opin. Chem. Biol. 41, 1–11
3. Williams, R., Ma, X., Schott, R. K., Mohammad, N., Ho, C. Y., Li, C. F., Chang, B. S., Demetriou, M., and Dennis, J. W. (2014) Encoding asymmetry of the N-glycosylation motif facilitates glycoprotein evolution. PLoS ONE 9, e86088
4. Deribe, Y. L., Pawson, T., and Dikic, I. (2010) Post-translational modifications in signal integration. Nat. Struct. Mol. Biol. 17, 666–672
5. Reinhart, H. C., and Yaffe, M. B. (2013) Phospho-Ser/Thr-binding domains: navigating the cell cycle and DNA damage response. Nat. Rev. Mol. Cell Biol. 14, 563–580
6. Rogers, L. D., and Overall, C. M. (2013) Proteolytic post-translational modification of proteins: proteomic tools and methodology. Mol. Cell Proteomics 12, 3532–3542
7. Rahimi, N., and Costello, C. E. (2015) Emerging roles of post-translational modifications in signal transduction and angiogenesis. Proteomics 15, 300–309
8. Rose, C. M., Isasa, M., Orduz, A., Prado, M. A., Beausoleil, S. A., Jeffkowiski, M. P., Finley, D. J., Harper, J. W., and Gygi, S. P. (2016) Highly multiplexed quantitative mass spectrometry analysis of ubiquitylomes. Cell Syst. 3, 395–403.e394
9. Udeshi, N. D., Mertins, P., Svinkina, T., and Carr, S. A. (2013) Large-scale identification of ubiquitination sites by mass spectrometry. Nat. Protocols 8, 1950–1960
10. Udeshi, N. D., Mani, D. R., Eisenhaure, T., Mertins, P., Jaffe, J. D., Clauser, K. R., Hacohen, N., and Carr, S. A. (2012) Methods for quantification of in vivo changes in protein ubiquitination following proteasome and deubiquitinase inhibition. Molecular & Cellular Proteomics 11, 148–159
11. Sethi, M. K., and Zaia, J. (2017) Extracellular matrix proteomics in schizophrenia and Alzheimer’s disease. Anal. Bioanal. Chem. 409, 379–394
12. Frantz, C., Stewart, K. M., and Weaver, V. M. (2010) The extracellular matrix at a glance. J. Cell Sci. 123, 4195–4200
13. Iozzo, R. V., and Gubbiotti, M. A. (2018) Extracellular Matrix: The Driving force of mammalian diseases. Matrix Biol. 71–72, 1–9
14. Chautard, E., Fatoux-Ardore, M., Ballut, L., Thierry-Mieg, N., and Ricard-Blum, S. (2011) MatrixDB, the extracellular matrix interaction database. Nucleic Acids Res. 39, D235–D240
15. Chautard, E., Ballut, L., Thierry-Mieg, N., and Ricard-Blum, S. (2009) MatrixDB, a database focused on extracellular protein-protein and protein-carbohydrate interactions. Bioinformatics 25, 690–691
16. Hynes, R. O., and Naba, A. (2012) Overview of the matrixome—an inventory of extracellular matrix constituents and functions. Cold Spring Harb. Perspect. Biol. 4, a004903
17. Naba, A., Clauser, K. R., Ding, H., Whittaker, C. A., Carr, S. A., and Hynes, R. O. (2016) The extracellular matrix: Tools and Insights for the “omics” era. Matrix Biol. 49, 10–24
18. Naba, A., Clauser, K. R., Hoersch, S., Liu, H., Carr, S. A., and Hynes, R. O. (2012) The matrisome: in silico definition and in vivo characterization by proteomics of normal and tumor extracellular matrices. Mol. Cell Proteomics 11, M111.014647
19. Aebersold, R., Agar, J. N., Amster, I. J., Baker, M. S., Bortozzi, C. R., Boja, E. S., Costello, C. E., Cravatt, B. F., Fenselau, C., Garcia, B. A., Ge, Y., Gunawardena, J., Hendrickson, R. C., Hergenrother, P. J., Huber, C. G., Ivanov, A. R., Jensen, O. N., Ewett, M. C., Kelleher, N. L., Kiessling, L. L., Krogan, N. J., Larsen, M. R., Loo, J. A., Ogorezko Loo, R. R., Lundberg, E., MacCoss, M. J., Mallick, P., Mootsa, V. K., Mrksich, M., Muir, T. W., Patrie, S. M., Pesavento, J. J., Pitteri, S. J., Rodriguez, H., Saghatelian, A., Sandoval, W., Schluter, H., Sechi, S., Slavov, S. A.,
sulfate phage display antibodies identify distinct epitopes with complex binding characteristics: insights into protein binding specificities. J. Biol. Chem. 284, 35621–35631
64. Rickelt, S., and Hynes, R. O. (2018) Antibodies and methods for immunohistochemistry of extracellular matrix proteins. Matrix Biol. 71–72, 10–27
65. Rosu-Marianou, H., Boyer-Bolteau, A., Holbrook, E. H., Jang, W., Hahn, C. G., Arnold, S. E., and Berretta, S. (2013) Proteoglycan abnormalities in olfactory epithelium tissue from subjects diagnosed with schizophrenia. Schizophr. Res. 150, 366–372
66. Iozzo, R. V., and Schaefer, L. (2015) Proteoglycan form and function: A comprehensive nomenclature of proteoglycans. Matrix Biol. 42, 11–55
67. Parcharra, S., Hanisch, F. G., Muhlenhoff, M., Faissner, A., Rauch, U., and Breley, I. (2013) The lecithins of mammalian brain perineurial net are around parvalbumin-containing neurons. Brain Res. 1657–1668
68. Sorrell, J. M., and Caterson, B. (1989) Detection of age-related changes in the distributions of keratan sulfates and chondroitin sulfates in developing chick limbs: an immunocytochemical study. Development 106, 657–663
69. Miyata, S., and Kitagawa, H. (2016) Chondroitin sulfate and neuronal disorder. Front. Biosci. 21, 1330–1349
70. Lott, S. A. (2016) Formation and remodeling of the extracellular matrix in neural plasticity: Roles of chondroitin sulfate and hyaluronic Acid. J. Chem. Neuroanat. 33, 111–123
71. Pantazopoulos, H., Murray, E. A., and Berretta, S. (2008) Total number, distribution, and phenotype of cells expressing chondroitin sulfate proteoglycans in the normal human amygdala. Brain Res. 1207, 84–95
72. Ajmo, J. M., Eakin, A. K., Hamel, M. G., and Gottschall, P. E. (2008) Discordant localization of WFA reactivity and brevican/ADAMTS-derived fragment in rodent brain. BMC Neurosci. 9, 14
73. Costa, C., Tortosa, R., Domenech, A., Vidal, E., Pumarola, M., and Bas-sols, A. (2007) Mapping of aggrecan, hyaluronic acid, heparan sulfate proteoglycans and aquaporin 4 in the central nervous system of the mouse. J. Chem. Neuroanat. 33, 111–123
74. Hartig, W., Brauer, K., Bigl, V., and Bruckner, G. (1994) Chondroitin sulfate and hyaluronan. Biochem. Biophys. Acta Gen. Subj. 1691, 2420–2434
75. Pantazopoulos, H., Murray, E. A., and Berretta, S. (2008) Total number, distribution, and phenotype of cells expressing chondroitin sulfate proteoglycans in the normal human amygdala. Brain Res. 1207, 84–95
76. Ajmo, J. M., Eakin, A. K., Hamel, M. G., and Gottschall, P. E. (2008) Discordant localization of WFA reactivity and brevican/ADAMTS-de-}
islets during the angiogenic switch and insulinoma progression. Sci. Rep. 7, 40495

100. Hu, Z., Gu, H., Hu, J., Hu, S., Wang, X., Liu, X., Jiao, X., and Liu, X. (2018) Quantitative proteomics identify an association between extracellular matrix degradation and immunopathology of genotype VII Newcastle disease virus in the spleen in chickens. J. Proteomics 181, 201–212

101. Didangelos, A., Yin, X., Mandal, K., Baumert, M., Jahangiri, M., and Mayr, M. (2010) Proteomics characterization of extracellular space components in the human aorta. Mol. Cell Proteomics 9, 2048–2062

102. Didangelos, A., Yin, X., Mandal, K., Saje, A., Smith, A., Xu, Q., Jahangiri, M., and Mayr, M. (2011) Extracellular matrix composition and remodeling in human abdominal aortic aneurysms: a proteomics approach. Mol. Cell Proteomics 10, M111.008128

103. Schiller, H. B., Fernandez, I. E., Burgstaller, G., Schaab, C., Scheltema, R., Rau, S., Schirmer, R., Tron, T., Krall, T., and Zaia, J. (2015) Time- and compartment-resolved proteome profiling of the extracellular niche in lung injury and repair. Mol. Syst. Biol. 11, 819

104. Barallobre-Barreiro, J., Lynch, M., Yin, X., and Mayr, M. (2016) Systems biology-opportunities and challenges: the application of proteomics to study the cardiovascular extracellular matrix. Cardiovasc. Res. 112, 626–636

105. Lynch, M., Barallobre-Barreiro, J., Jahangiri, M., and Mayr, M. (2016) Vascular proteomics in metabolic and cardiovascular diseases. J. Intern. Med. 280, 325–338

106. Suna, G., Wojakowski, W., Lynch, M., Barallobre-Barreiro, J., Yin, X., Mayr, U., Baig, F., Lu, R., Barwari, T., Suna, G., Yin, X., Iglidseder, B., Paulweber, B., Willett, P., Shahlabou, J., Pasterkamp, G., Dickson, A. H., Hinneburg, H., Sven, N., Ng, K. H., Ho, J., Kreidberg, J. A., Taglenti, M., Royer-Pokora, B., and Ai, X. (2011) WT1-dependent sulfatase expression maintains the normal glomerular filtration barrier. J. Am. Soc. Nephrol. 22, 1286–1296

107. Reine, T. M., Kolseth, I. B., Meen, A. J., Lindahl, J. P., Jabs, W., Seeberger, P. H., Silva, D. V., Wuhrer, M., and Kolarich, D. (2016) Comparative glycomics of leukocyte glycosaminoglycans. The FEBS journal 280, 2447–2461

108. Langsdorf, A., Schumacher, V., Shi, X., Tran, T., Zaia, J., Jain, S., Taglenti, M., Kreidberg, J. A., Fine, A., and Ai, X. (2011) Expression regulation and function of heparan sulfate 6-O-endosulfatases in the spermatogonial niche. J. Biol. Chem. 286, 32651–32664

109. Armstead, J. S., Wilson, I. B., van Kuppevelt, T. H., and Dinglasan, R. R. (2011) A role for heparan sulfate proteoglycans in Plasmodium falciparum sporozoite invasion of anopheline mosquito salivary glands. Biochem. J. 438, 475–483

110. Dierker, T., Shao, C., Hatina, T., Zaia, J., Hinas, A., and Kjellén, L. (2016) Nemato ces join the family of chondroitin sulfite-synthesizing organisms: Identification of an active chondroitin sulfotransferase in Caenorhabditis elegans. Sci. Rep. 6, 34662

111. Shao, C., Shi, X., and Zaia, J. (2010) Extended NS domains reside at the non-reducing end of heparan sulfate chains. J. Biol. Chem. 285, 18336–18343

112. Didangelos, A., Yin, X., Mandal, K., Baumert, M., Jahangiri, M., and Mayr, M. (2010) Proteomics characterization of extracellular space components in the human aorta. Mol. Cell Proteomics 9, 2048–2062

113. Shao, C., Shi, X., Phillips, J. J., and Zaia, J. (2013) Mass spectral profiling of glycosaminoglycans from histological tissue surfaces. Anal. Chem. 85, 10984–10991
137. Hu, H., Khatri, K., Klein, J., Leymarie, N., and Zaia, J. (2016) A review of methods for interpretation of glycopeptide tandem mass spectral data. Glycoconj. J. 33, 285–296

138. Leymarie, N., and Zaia, J. (2012) Effective use of mass spectrometry for glycan and glycoepitope structural analysis. Anal. Chem. 84, 3040–3048

139. Thaysen-Andersen, M., Packer, N. H., and Schulz, B. L. (2016) Maturing glycoproteomics technologies provide unique structural insights into the N-glycoproteome and its regulation in health and disease. Mol. Cell Proteomics 15, 1773–1790

140. Klein, J. A., Meng, L., and Zaia, J. (2018) Deep sequencing of complex proteoglycans: a novel strategy for high coverage and site-specific identification of glycosaminoglycan-linked peptides. Mol. Cell Proteomics 17, 1578–1590

141. Barallobre-Barreiro, J., Gupta, S. K., Zoccarato, A., Kitazume-Taneike, R., Fava, M., Yin, X., and Mayr, M. (2017) O-mannosyl glycoproteomics of the extracellular matrix: a method for intact glycopeptide analysis using mass spectrometry. J. Vis. Exp. 122

142. Pacharra, S., Hanisch, F. G., and Breloy, I. (2012) Neurofascin 186 is O-mannosylated within and outside of the mucin domain. J. Proteome Res. 11, 3955–3964

143. Vester-Christensen, M. B., Hallim, A., Joshi, H. J., Steentoft, C., Bennett, E. P., Levery, S. B., Vakhruchev, S. Y., and Clausen, H. (2013) Mining the O-mannose glycoproteome reveals cadherins as major O-mannosylated glycoproteins. Proc. Natl. Acad. Sci. U.S.A. 110, 21018–21023

144. Bartels, M. F., Winterhalter, P. R., Yu, J., Liu, Y., Lommel, M., Mohrlen, F., Hu, H., Feizi, T., Westerlind, U., Ruppert, T., and Strahl, S. (2016) Protein O-mannosylation in the murine brain: occurrence of mono-O-mannosyl glycans and identification of new substrates. PLoS ONE 11, e0166119

145. Gomez Toledo, A., Nilsson, J., Noborn, F., Sihlbom, C., and Larson, G. (2015) Positive mode LC-MS/MS analysis of chondroitin sulfate modified glycopeptides derived from light and heavy chains of the human inter-alpha-trypsin inhibitor complex. Mol. Cell Proteomics 14, 3118–3131

146. Nilsson, J., Noborn, F., Gomez Toledo, A., Nasir, W., Sihlbom, C., and Larson, G. (2017) Characterization of glycan structures of chondroitin sulfate-glycopeptides facilitated by sodium ion-pairing and positive mode LC-MS/MS. J. Am. Soc. Mass. Spectrom. 28, 229–241

147. Noborn, F., Gomez Toledo, A., Green, A., Nasir, W., Sihlbom, C., Nilsson, J., and Larson, G. (2016) Site-specific identification of heparan and chondroitin sulfate glycosaminoglycans in hybrid proteoglycans. Sci. Rep. 6, 34537

148. Barallobre-Barreiro, J., Baig, F., Fava, M., Yin, X., and Mayr, M. (2017) Glycoproteomics of the extracellular matrix: a method for intact glycopeptide analysis using mass spectrometry. J. Vis. Exp. 122