Electron-based dissociation is needed for O-glycopeptides derived from OpeRATOR proteolysis.

Nicholas M. Riley[1], Stacy A. Malaker[1], Carolyn R. Bertozzi[1,2]*

[1]Department of Chemistry and Stanford ChEM-H, Stanford University, Stanford, California, USA
[2]Howard Hughes Medical Institute, Stanford, California, USA

ABSTRACT: The recently described O-glycoprotease OpeRATOR presents exciting opportunities for O-glycoproteomics. This bacterial enzyme purified from *Akkermansia muciniphila* cleaves N-terminally to serine and threonine residues that are modified with (preferably asialylated) O-glycans. This provides orthogonal cleavage relative to canonical proteases (e.g., trypsin) for improved O-glycopeptide characterization with tandem mass spectrometry (MS/MS). O-glycopeptides with a modified N-terminal residue, such as those generated by OpeRATOR, present several potential benefits, perhaps the most notable being de facto O-glycosite localization without the need of glycan-retaining fragments in MS/MS spectra. Indeed, O-glycopeptides modified exclusively at the N-terminus would enable O-glycoproteomic methods to rely solely on collision-based fragmentation rather than electron-driven dissociation because glycan-retaining peptide fragments would not be required for localization. The caveat is that modified peptides would need to reliably contain only a single O-glycosite. Here we use methods that combine collision- and electron-based fragmentation to characterize the number of O-glycosites that are present in O-glycopeptides derived from OpeRATOR digestion of four known O-glycoproteins. Our data show that over 50% of O-glycopeptides generated from combined digestion using OpeRATOR and trypsin contain multiple O-glycosites, indicating that collision-based fragmentation alone is not sufficient. Electron-based dissociation methods are necessary to capture the O-glycopeptide diversity present in OpeRATOR digestions.

Mucin-type O-glycosylation is a prevalent post-translational modification on extracellular and secreted proteins that drives both biochemical and biophysical interactions at the cell surface.[1-5] This important modification, characterized by an initiating α-*N*-acetylgalactosamine (α-GalNAc) monosaccharide on serine and threonine residues, is challenging to study due to several inherent features. These include non-template driven elaboration of the α-GalNAc residue into four major core structures, the lack of a well-defined sequence motif, and occurrence in densely glycosylated regions rich in serine and threonine residues.[6] The combination of these attributes leads to O-glycosylated sequences that are heterogeneous both in the glycosites that are occupied and the glycans that modify them, requiring site-specific characterization. Tandem mass spectrometry (MS) is the premier method for O-glycosite mapping, but O-glycopeptides are largely intractable using standard glycoproteomic approaches, which have focused almost entirely on N-glycans. To this end, recent efforts have focused on improving O-glycopeptide analyses,[7-10] ranging from developments in sample preparation,[11-15] data acquisition,[16-21] and post-acquisition data analysis.[22-27]

One promising avenue to emerge from these investigations is the description of O-glycoproteases that can generate O-glycopeptides more amenable to MS characterization.[28-31] One such enyzme is OpeRATOR from *A. muciniphila*, which cleaves N-terminally to O-glycosylated

serine and threonines, generating O-glycopeptides that retain the modified residue at their N-terminus.[30-32] In the handful of studies to utilize OpeRATOR thus far, collision-based fragmentation, namely higher-energy collisional dissociation (HCD), has been used to identify O-glycopeptides. The operating presumption in these studies is that site-specific O-glycosite localization can be reported by attributing the entire glycan mass identified for the peptide to an O-glycan modifying the N-terminal serine or threonine. In order to preclude the need for electron-based fragmentation (such as electron transfer dissociation [ETD] and ETD with HCD supplemental activation [EThcD]), however, O-glycopeptides from OpeRATOR proteolysis must not have missed cleavages. Missed cleavages by this definition would mean that internal O-glycosites also exist within the peptide sequence rather than exclusively at the N-terminus, ultimately leading to misattribution of a portion of the identified glycan aggregate mass to the N-terminal O-glycosite.

Here we use methods employing both HCD and ETD/EThcD fragmentation to investigate O-glycopeptides derived from OpeRATOR proteolysis of four O-glycoprotein standards (fetuin, monocyte differentiation antigen [CD14], fibronectin, and P-selectin glycoprotein ligand 1 [PSGL-1]). Our data show that over half of O-glycopeptide identifications harbor multiple glycosites. As such, O-glycosite localization exclusively at N-terminal residues
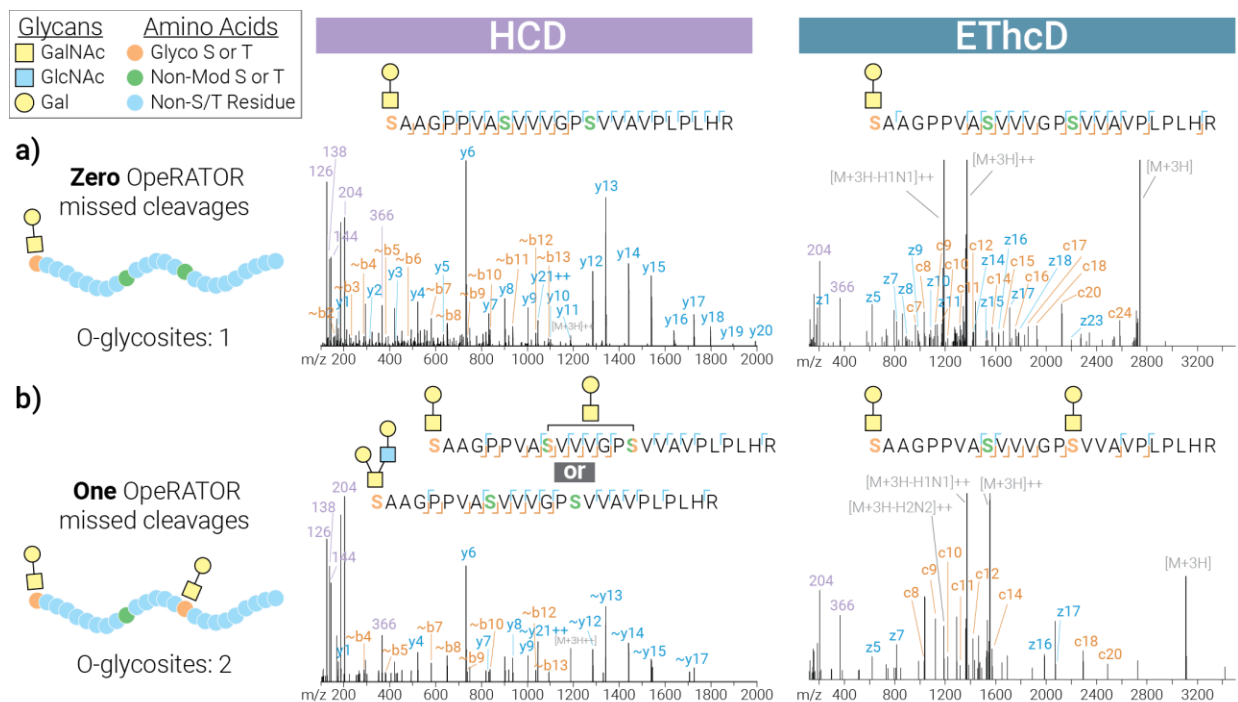
Figure 1. Bovine fetuin sequence ²⁸²SAAGPPVASVVVGPSVVAVPLPLHR³⁰⁶ is a tangible example of the benefits of electron-driven dissociation. Diagrams at the left present two glycopeptidoform possibilities of this sequence, which have both O-glycosylated (orange) and non-modified (green) serine and threonine residues. These two O-glycopeptidoforms were identified in our dataset using HCD-EThcD spectral pairs, including (**a**) a canonical OpeRATOR-derived O-glycopeptide with a single O-glycosite at the N-terminus and (**b**) a doubly glycosylated O-glycopeptide from a missed OpeRATOR cleavage that has both an N-terminal O-glycosite site and a second internal O-glycosite.

cannot be assumed. This renders collisional activation inadequate for site-specific analysis and underscores the need for electron-driven dissociation in O-glycopeptide characterization, even those generated using OpeRATOR.

EXPERIMENTAL

A mixture of O-glycopeptides generated by sequential OpeRATOR and trypsin digestion was generated using four glycoproteins: bovine fetuin (alpha-2-HS-glycoprotein, P12763), recombinant human CD14 (P08571), human fibronectin (P02751), and recombinant human P-selectin glycoprotein ligand 1 (PSGL1) (Q14242). OpeRATOR proteolysis was coupled with sialidase co-treatment, and PNGaseF was used to remove N-glycans. Following digestion and desalting, peptides were combined in equal parts by mass for the four proteins and analyzed by 90-min LC-MS/MS product-dependent methods constructed using ETD or EThcD triggered scans (HCD-pd-ETD and HCD-pd-EThcD methods).[33–36] All raw data were searched using O-Pair Search implemented in MetaMorpheus (0.0.308), which is available at https://github.com/smith-chem-wisc/MetaMorpheus.[27] A glycan database representing 12 common O-glycans was used, and the "Maximum OGlycan Allowed" parameter was set to 5. Only Level 1 and 1b identifications were retained for further analysis, meaning all identifications had spectral evidence in ETD or EThcD

spectra for localized glycosites (or had only one plausible glycoform). For glycan modifications shown as H#N#, H represents the number of hexose residues (galactose in O-glycans), and N represents the number of *N*-acetylhexosamines, which can be *N*-acetylgalactosamine or *N*-acetylglucosamine in O-glycans. Spectral annotation was aided by the Interactive Peptide Spectral Annotator (IPSA, http://www.interactivepeptide-spectralannotator.com).[37] Graphs were generated using OriginPro 2018 with the exception of the alluvial diagram, which was generated using the RAWGraphs web app, licensed under CC BY-NC-SA 4.0.[38] WebLogo was used for Logo plot generation (https://weblogo.berkeley.edu/logo.cgi).[39] More details are available in the **Supporting Information**.

RESULTS AND DISCUSSION

O-glycosite localization in OpeRATOR-derived O-glycopeptides has thus far relied on HCD fragmentation and the assumption that the total glycan mass observed can be attributed to a single glycosylated serine or threonine at the N-terminus. **Figure 1** provides an example of why collision-based fragmentation is insufficient for O-glycosite localization and instead why electron-driven dissociation is necessary. An O-glycopeptide from fetuin, ²⁸²SAAGPPVASVVVGPSVVAVPLPLHR³⁰⁶, represents two well-known O-glycosites, Ser-282 and Ser-296, in addition

to Ser-290, which was identified as O-glycosylated using OpeRATOR in the EXoO workflow.[30] Two glycopeptidoforms identified using an HCD-pd-EThcD25 method include the top example, with an O-glycosylated Ser-282 (orange) and unmodified Ser-290 and Ser-296 (green) (**Figure 1a**), and the bottom example, with O-glycosylated Ser-282 and Ser-296 (orange) and unmodified Ser-290 (green) (**Figure 1b**). In the top example, b-type peptide fragment ions in the HCD spectrum did not retain the O-glycan (as denoted by "~"), but the EThcD spectrum confirmed that the assumption of an N-terminal Ser-282 modified with H_1N_1 is correct. Here HCD was adequate, albeit less than ideal, for glycosite localization; the H_1N_1 glycan mass is presumably a single core-1 GalNAc-Gal, a common mucin-type O-glycan, but O-mannose (a hexose) can modify similar regions as O-GalNAc on the same glycan (albeit less likely to be the case).[40]

The doubly O-glycosylated peptide in **Figure 1b**, however, demonstrates a more problematic issue when relying on HCD. The total O-glycan mass observed is H_2N_2. Neither b- nor y-type fragments retain glycan modifications in the HCD spectrum (marked with "~"), leaving ambiguity as to which serine residues were modified. OpeRATOR digestion affords the assumption of the N-terminal serine as glycosylated, but H_2N_2 is equally as plausible of a glycan to observe at a single O-glycosite as H_1N_1. Even with the assumption that the N-terminal serine has the same H_1N_1 modification as the top example, ambiguity remained as to whether Ser-290 or Ser-296 was modified with the other H_1N_1 O-glycan. On the other hand, the EThcD spectrum in the bottom provided c- and z•-type fragment ions that unambiguously localized the H_1N_1 modification to both Ser-282 and Ser-296.

To understand the prevalence of multiply O-glycosylated peptides present in OpeRATOR digests, we generated a dataset comprising four different product-dependent methods with electron-driven dissociation as the triggered scan: HCD-pd-ETD, HCD-pd-EThcD15, HCD-pd-EThcD25, and HCD-pd-EThcD35 (collected in technical triplicate, 12 raw files total). **Figure 2a** provides the proportion of the total O-glycopeptide identifications that had one or multiple localized O-glycosites. Note that O-Pair Search returns a single identification representing two spectra, both an HCD and ETD/EThcD spectrum.[27] Approximately 46% of the glycopeptide spectral matches contained a single O-glycosite, indicating that over half of the identifications harbored two or more O-glycosites. We also note that our data were generated using a sialidase mixture to produce asialylated O-glycans, which is known to aid in OpeRATOR digestion. The number of missed cleavages and multiply O-glycosylated peptides would likely be higher if sialidase treatment was omitted. Our data also corroborates a recent reanalysis of the EXoO dataset using MSFragger-Glyco (a new search algorithm that functions similarly to the O-Pair Search strategy used here), where the authors observed both a substantial number of missed OpeRATOR cleavages and glycan mass modifications indicative of multiple O-glycosites.[41] In that study, Polasky *et*

al. also identified co-occurrence of phosphorylation sites on O-glycosylated sequences, furthering the need for electron-driven dissociation to properly localize multiple sites of modification.

Figure 2b examines what proportion of identifications represent O-glycopeptides that are either upstream/N-terminal or downstream/C-terminal from the OpeRATOR cleavage site. Note, "canonical" OpeRATOR glycopeptides with an O-glycosylated N-terminal residue are the downstream/C-terminal case. For O-glycopeptides with one O-glycosite, ~75% of identifications were canonical OpeRATOR peptides with a single N-terminal O-glycosylated serine or threonine residue, while nearly 20% of identified O-glycopeptides were from sequences that were upstream from an OpeRATOR cleavage site. These upstream O-glycosylated peptides would likely evade analysis in approaches that couple solid-supports for enrichment and OpeRATOR cleavage for O-glycopeptide release, because they remain tethered to the support matrix. Presence of internal glycosites both N- and C-terminal to OpeRATOR cleavage is supported by the presence of serine and threonine residues occurring 1 to >10 positions away from the cleavage site (both up- and downstream) (**Figure S1**). Additionally, search strategies that only consider peptides starting with serine or threonine in their searches, such as those used in previous OpeRATOR studies, would either be blind to these upstream sequences or could mis-assign their glycosites. A small percentage of O-glycopeptides passed filtering criteria, yet did not follow the expected OpeRATOR motif, instead showing only an internal O-glycosite (**Figure 2b**). These could be the presence of some non-glycosylation-dependent activity of OpeRATOR at serine and threonine residues or due to other non-specific cleavage (e.g., chemical degradation or from trace amounts of unknown proteases). Interestingly, approximately 48% of all identifications came from the mucin domain of PSGL-1, which could contribute to the proportion of multiply glycosylated sequences observed. That said, mucin O-glycoproteins were discussed as major components of the previous OpeRATOR datasets, making O-glycopeptides from densely O-glycosylated mucins a mainstay rather than an exception when digesting with OpeRATOR.

The alluvial diagram in **Figure 2c** maps the aggregate glycan masses that were identified for O-glycopeptides harboring a varying number of O-glycosites. The only two aggregate glycan masses that are exclusive to singly O-glycosylated peptides are N_1 and H_1N_1, meaning identification of any other glycan composition using HCD fragmentation can lead to ambiguity in O-glycosite assignment. As noted above, H_2N_2 is one such aggregate glycan mass with ambiguity, with possibilities including one glycan at a single glycosite or two glycans split between two sites. Here the majority of identifications with an aggregate composition of H_2N_2 were the result of two H_1N_1 O-glycosites rather than the total glycan mass modifying a single residue. Once multiple glycosites were considered on a peptide, we saw a substantial increase in the glycoforms (i.e., glycosites per peptide) that can explain an observed aggregate glycan
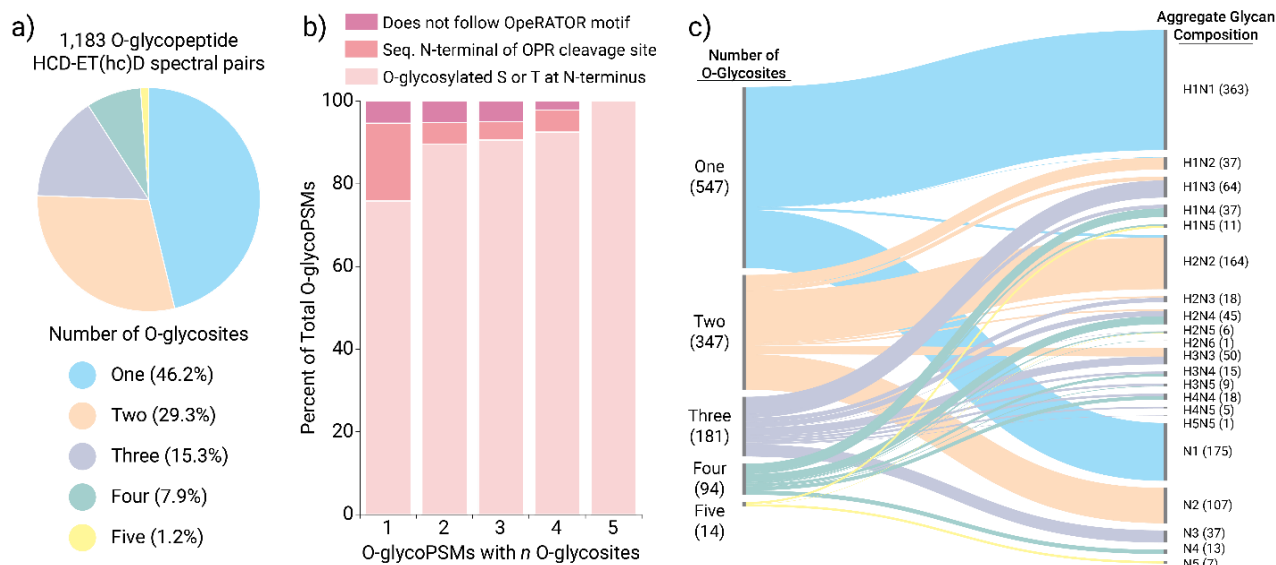
Figure 2. Characteristics of O-glycopeptides derived from OpeRATOR proteolysis. a) A total of 1,183 O-glycopeptide spectral pairs with confidently localized glycosites were detected using HCD-pd-ETD and HCD-pd-EThcD methods. The pie graph shows the proportion of identifications that had one or multiple O-glycosites. **b)** The bar graph delineates the proportion of O-glycopeptide spectral pairs that can be grouped into three main classes: 1) identifications that have an O-glycosylated serine or threonine at the N-terminus (light pink; bottom), 2) O-glycopeptides from sequences upstream (i.e., N-terminal) of the OpeRATOR (OPR) cleavage site that indicate a missed cleavage (pink; middle), and 3) O-glycopeptides that do not appear to follow the OpeRATOR cleavage motif (dark pink; top). **c)** The alluvial diagram relates the aggregate glycan compositions that were identified on O-glycopeptides with a varying number of localized O-glycosites. Numbers in parentheses provide the number of O-glycopeptide spectral pairs identified for each condition. For glycan modifications shown as H#N#, H represents the number of hexose residues (galactose in O-glycans), and N represents the number of *N*-acetylhexosamines, which can be *N*-acetylgalactosamine or *N*-acetylglucosamine in O-glycans.

mass, highlighting the need to localize O-glycans to specific residues with electron-driven methods. Presence of sialic acids, beyond the reduced OpeRATOR efficiency discussed above, would further complicate this issue because it drastically increases the number of possible glycan structures that can be observed, as noted by Polasky et al.[41] One potential option to eliminate the presence of core-2 structures like H_2N_2 would be to evaluate oxonium ion ratios, which can indicate the presence of GlcNAc residues rather than exclusively GalNAc.[21,42]

To test ETD and EThcD efficacy for OpeRATOR-based workflows, we compared identifications the four ETD-centric methods; EThcD with collision energies between 25 and 35 nce performed best (**Figure S2**). Glycans and serine/threonine-rich sequences of O-glycopeptides often generate low charge density precursor ions, which can be challenging for electron-driven fragmentation. More than a third of EThcD25 and EThcD35 identifications were from doubly charged precursor ions (**Figure S3**), likely contributing to better performance. Benefits of supplemental activation can be even more pronounced for multiply glycosylated peptides that have low charge density but require extensive fragmentation for site localization. As such, other hybrid electron-driven methods may also be well-suited for such analysis,[43–45] although they remain largely untested for O-glycopeptide characterizations.

We also collected a collisional dissociation-centric dataset to understand the prevalence of presumably singly O-glycosylated peptides (i.e., N_1 and H_1N_1 modified) compared to presumably multiply O-glycosylated peptides (i.e., those with other aggregate glycan masses) (**Figure S4**). Of those identifications purported to have one O-glycosite (approximately half of the total identifications), only 60% had glycan modification masses that exclusively indicate a single O-glycan (i.e., N_1 or H_1N_1). In combination with **Figure 2c** above, this means that 40% of the O-glycosylated identifications reported to be singly glycosylated (without localization evidence) are likely to have ambiguous or incorrect O-glycosite assignment if relying on the assumption of a N-terminal O-glycosite. These data further demonstrate the challenges with ambiguity that accompany reliance exclusively on HCD fragmentation with OpeRATOR digestion, and they show that HCD-only methods fail to properly identify a significant proportion of O-glycopeptides from the sample.

CONCLUSIONS

O-glycopeptides often contain multiple potential O-glycosites that can be modified by a heterogeneous pool of labile O-glycans. Thus, site-specific analysis of O-glycosylation typically requires alternative MS/MS fragmentation methods (namely electron-driven dissociation such as electron capture dissociation (ECD), electron transfer dissociation methods (ETD), and their derivate methods[46–48]) that can retain intact glycan moieties on peptide backbone fragments. Nevertheless, approaches that could make site-

specific O-glycopeptide analysis amenable to collisional-dissociation-centric methods would be valuable. One potential benefit of the recently described O-glycoprotease OpeRATOR is the reported ability to localize O-glycosites using collisional dissociation due to the N-terminal location of modified sites within the sequence. Here we investigated how often O-glycopeptides generated from OpeRATOR proteolysis contain only one N-terminal O-glycosite versus multiple glycosites. We show that over half of O-glycopeptides from OpeRATOR digestion contain multiple glycosites, where the total glycan mass addition cannot be solely attributed to the N-terminal residue. This alone indicates that collision-based fragmentation (e.g., HCD) is not sufficient for site-specific characterization of OpeRATOR O-glycopeptides. Additionally, we identified cases where glycan compositions that could ostensibly be attributed to a single O-glycosite were actually several smaller glycans distributed amongst multiple O-glycosites in OpeRATOR O-glycopeptides, as confirmed by electron-based fragmentation.

It is worth noting here that these data do not invalidate the results of previous studies using OpeRATOR and HCD-only methods. Their results appear to contain quality identifications and confidently localized O-glycosites, especially considering the high proportion of reported H_1N_1 modifications. Creative alterations to workflows can improve O-glycosite localization for HCD-centric methods, e.g., the recently reported EXoO-Tn study that can indicate the number of O-glycans present in the sequence using heavy labeled monosaccharides.[32] Yet, ambiguities can still remain. Rather than discard previous work, this study points out 1) ambiguity is possible when using OpeRATOR for O-glycosite localization in a considerable number of identifications when solely utilizing HCD, and 2) that many multiply O-glycosylated peptides generated by OpeRATOR proteolysis go undetected when HCD-only methods are used to presume a single O-glycosite at the peptide N-terminus. As such, it is clear that use of the O-glycoprotease OpeRATOR is a valuable tool for O-glycosite mapping, but methods relying on this approach cannot forgo electron-based dissociation, especially EThcD methods, for localizing O-glycosites.

ASSOCIATED CONTENT
Supporting Information
The following Supporting Information is available for this manuscript in an associated document. This includes Supplemental Experimental Methods, Supplemental Discussion, Supplemental References, and Supplemental Figures: **Figure S1.** Logo plot generated using O-glycopeptides generated from OpeRATOR proteolysis; **Figure S2.** Performance summary for ETD and EThcD methods; **Figure S3.** Charge state distributions of O-glycopeptide identifications from various methods; **Figure S4.** HCD and sceHCD methods for analyzing OpeRATOR O-glycopeptides.

AUTHOR INFORMATION
Corresponding Author

*Correspondence should be to C.R.B., email: bertozzi@stanford.edu

Notes

ACKNOWLEDGMENT

REFERENCES
(1)   Möckl, L. The Emerging Role of the Mammalian Glycocalyx in Functional Membrane Organization and Immune System Regulation. *Front. Cell Dev. Biol.* **2020**, *8*, 253. https://doi.org/10.3389/fcell.2020.00253.

(2)   Woods, E. C.; Kai, F.; Barnes, J. M.; Pedram, K.; Pickup, M. W.; Hollander, M. J.; Weaver, V. M.; Bertozzi, C. R. A Bulky Glycocalyx Fosters Metastasis Formation by Promoting G1 Cell Cycle Progression. *Elife* **2017**, *6*. https://doi.org/10.7554/eLife.25752.

(3)   Reily, C.; Stewart, T. J.; Renfrow, M. B.; Novak, J. Glycosylation in Health and Disease. *Nature Reviews Nephrology*. Nature Publishing Group June 1, 2019, pp 346–366. https://doi.org/10.1038/s41581-019-0129-4.

(4)   Kuo, J. C. H.; Gandhi, J. G.; Zia, R. N.; Paszek, M. J. Physical Biology of the Cancer Cell Glycocalyx. *Nat. Phys.* **2018**, *14* (7), 658–669. https://doi.org/10.1038/s41567-018-0186-9.

(5)   Paszek, M. J.; Dufort, C. C.; Rossier, O.; Bainer, R.; Mouw, J. K.; Godula, K.; Hudak, J. E.; Lakins, J. N.; Wijekoon, A. C.; Cassereau, L.; et al. The Cancer Glycocalyx Mechanically Primes Integrin-Mediated Growth and Survival. *Nature* **2014**, *511* (7509), 319–325. https://doi.org/10.1038/nature13535.

(6)   Brockhausen, I.; Stanley, P. Chapter 10 O-GalNAc Glycans. *Essentials Glycobiol.* **2017**, *1*, 1–9. https://doi.org/10.1101/glycobiology.3e.010.

(7)   Khoo, K. H. Advances toward Mapping the Full Extent of Protein Site-Specific O-GalNAc Glycosylation That Better Reflects Underlying Glycomic Complexity. *Current Opinion in Structural Biology*. Elsevier Ltd June 1, 2019, pp 146–154. https://doi.org/10.1016/j.sbi.2019.02.007.

(8)   You, X.; Qin, H.; Ye, M. Recent Advances in Methods for the Analysis of Protein O-Glycosylation at Proteome Level. *Journal of Separation Science*. January 2018, pp 248–261. https://doi.org/10.1002/jssc.201700834.

(9)   Darula, Z.; Medzihradszky, K. F. Analysis of Mammalian O-Glycopeptides - We Have Made a Good Start, but There Is a Long Way to Go. *Molecular and Cellular Proteomics*. American Society for Biochemistry and Molecular Biology Inc. January 1, 2018, pp 2–17. https://doi.org/10.1074/mcp.MR117.000126.

(10) Levery, S. B.; Steentoft, C.; Halim, A.; Narimatsu, Y.; Clausen, H.; Vakhrushev, S. Y. Advances in Mass Spectrometry Driven O-Glycoproteomics. *Biochimica et Biophysica Acta - General Subjects*. Elsevier B.V. 2015, pp 33–42. https://doi.org/10.1016/j.bbagen.2014.09.026.

(11) Stavenhagen, K.; Hinneburg, H.; Kolarich, D.; Wuhrer, M. Site-Specific N- and O-Glycopeptide Analysis Using an Integrated C18-PGC-LC-ESI-QTOF-MS/MS Approach. In *Methods in Molecular Biology*; Humana Press Inc., 2017; Vol. 1503, pp 109–119. https://doi.org/10.1007/978-1-4939-6493-2_9.

(12) Woo, C. M.; Lund, P. J.; Huang, A. C.; Davis, M. M.; Bertozzi, C. R.; Pitteri, S. J. Mapping and Quantification of over 2000 O-Linked Glycopeptides in Activated Human T Cells with Isotope-Targeted Glycoproteomics (Isotag). *Mol. Cell. Proteomics* **2018**, *17* (4), 764–775. https://doi.org/10.1074/mcp.RA117.000261.

(13) Schumann, B.; Malaker, S. A.; Wisnovsky, S. P.; Debets, M. F.; Agbay, A. J.; Fernandez, D.; Wagner, L. J. S.; Lin, L.; Li, Z.; Choi, J.; et al. Bump-and-Hole Engineering Identifies Specific Substrates of Glycosyltransferases in

Living Cells. *Mol. Cell* **2020**, *78* (5), 824-834.e15. https://doi.org/10.1016/j.molcel.2020.03.030.

(14) Narimatsu, Y.; Joshi, H. J.; Schjoldager, K. T.; Hintze, J.; Halim, A.; Steentoft, C.; Nason, R.; Mandel, U.; Bennett, E. P.; Clausen, H.; et al. Exploring Regulation of Protein O-Glycosylation in Isogenic Human HEK293 Cells by Differential O-Glycoproteomics. *Mol. Cell. Proteomics* **2019**, *18* (7), 1396–1409. https://doi.org/10.1074/mcp.RA118.001121.

(15) Steentoft, C.; Vakhrushev, S. Y.; Joshi, H. J.; Kong, Y.; Vester-Christensen, M. B.; Schjoldager, K. T.-B. G.; Lavrsen, K.; Dabelsteen, S.; Pedersen, N. B.; Marcos-Silva, L.; et al. Precision Mapping of the Human O-GalNAc Glycoproteome through SimpleCell Technology. *EMBO J.* **2013**, *32* (10), 1478–1488. https://doi.org/10.1038/emboj.2013.79.

(16) Pap, A.; Klement, E.; Hunyadi-Gulyas, E.; Darula, Z.; Medzihradszky, K. F. Status Report on the High-Throughput Characterization of Complex Intact O-Glycopeptide Mixtures. *J. Am. Soc. Mass Spectrom.* **2018**, *29* (6), 1210–1220. https://doi.org/10.1007/s13361-018-1945-7.

(17) Yu, Q.; Canales, A.; Glover, M. S.; Das, R.; Shi, X.; Liu, Y.; Keller, M. P.; Attie, A. D.; Li, L. Targeted Mass Spectrometry Approach Enabled Discovery of O- Glycosylated Insulin and Related Signaling Peptides in Mouse and Human Pancreatic Islets. *Anal. Chem.* **2017**, *89* (17), 9184–9191. https://doi.org/10.1021/acs.analchem.7b01926.

(18) Kelly, M. I.; Dodds, E. D. Parallel Determination of Polypeptide and Oligosaccharide Connectivities by Energy-Resolved Collison-Induced Dissociation of Protonated O-Glycopeptides Derived from Nonspecific Proteolysis. *J. Am. Soc. Mass Spectrom.* **2020**, *31* (3), 624–632. https://doi.org/10.1021/jasms.9b00065.

(19) Darula, Z.; Pap, Á.; Medzihradszky, K. F. Extended Sialylated O-Glycan Repertoire of Human Urinary Glycoproteins Discovered and Characterized Using Electron-Transfer/Higher-Energy Collision Dissociation. *J. Proteome Res.* **2019**, *18* (1), 280–291. https://doi.org/10.1021/acs.jproteome.8b00587.

(20) Zhang, Y.; Xie, X.; Zhao, X.; Tian, F.; Lv, J.; Ying, W.; Qian, X. Systems Analysis of Singly and Multiply O-Glycosylated Peptides in the Human Serum Glycoproteome via EThcD and HCD Mass Spectrometry. *J. Proteomics* **2018**, *170*, 14–27. https://doi.org/10.1016/j.jprot.2017.09.014.

(21) Riley, N. M.; Malaker, S. A.; Driessen, M.; Bertozzi, C. R. Optimal Dissociation Methods Differ for N- and O-Glycopeptides. *J. Proteome Res.* **2020**. https://doi.org/10.1021/acs.jproteome.0c00218.

(22) Pap, A.; Medzihradszky, K. F.; Darula, Z. Using "Spectral Families" to Assess the Reproducibility of Glycopeptide Enrichment: Human Serum O-Glycosylation Revisited. *Anal. Bioanal. Chem.* **2017**, *409* (2), 539–550. https://doi.org/10.1007/s00216-016-9960-7.

(23) Huang, J.; Jiang, B.; Zhao, H.; Wu, M.; Kong, S.; Liu, M.; Yang, P.; Cao, W. Development of a Computational Tool for Automated Interpretation of Intact O -Glycopeptide Tandem Mass Spectra from Single Proteins. *Anal. Chem.* **2020**. https://doi.org/10.1021/acs.analchem.0c01091.

(24) Chalkley, R. J.; Medzihradszky, K. F.; Darula, Z.; Pap, A.; Baker, P. R. The Effectiveness of Filtering Glycopeptide Peak List Files for Y Ions. *Mol. Omi.* **2020**, *16* (2), 147–155. https://doi.org/10.1039/c9mo00178f.

(25) Mao, J.; You, X.; Qin, H.; Wang, C.; Wang, L.; Ye, M. A New Searching Strategy for the Identification of O-Linked Glycopeptides. *Anal. Chem.* **2019**, *91* (6), 3852–3859. https://doi.org/10.1021/acs.analchem.8b04184.

(26) Zhao, X.; Zheng, S.; Li, Y.; Huang, J.; Zhang, W.; Xie, Y.; Qin, W.; Qian, X. An Integrated Mass Spectroscopy Data Processing Strategy for Fast Identification, In-Depth, and Reproducible Quantification of Protein O-Glycosylation in a Large Cohort of Human Urine Samples. *Anal. Chem.* **2019**, *92* (1), 690–698. https://doi.org/10.1021/acs.analchem.9b02228.

(27) Lu, L.; Riley, N. M.; Shortreed, M. R.; Bertozzi, C. R.; Smith, L. M. O-Pair Search with MetaMorpheus for O-Glycopeptide Characterization. *bioRxiv* **2020**, 2020.05.18.102327. https://doi.org/10.1101/2020.05.18.102327.

(28) Malaker, S. A.; Pedram, K.; Ferracane, M. J.; Bensing, B. A.; Krishnan, V.; Pett, C.; Yu, J.; Woods, E. C.; Kramer, J. R.; Westerlind, U.; et al. The Mucin-Selective Protease StcE Enables Molecular and Functional Analysis of Human Cancer-Associated Mucins. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116* (15), 7278–7287. https://doi.org/10.1073/pnas.1813020116.

(29) Shon, D. J.; Malaker, S.; Pedram, K.; Yang, E.; Krishnan, V.; Dorigo, O.; Bertozzi, C. An Enzymatic Toolkit for Selective Proteolysis, Detection, and Visualization of Mucin-Domain Glycoproteins. **2020**. https://doi.org/10.26434/CHEMRXIV.12476369.V1.

(30) Yang, W.; Ao, M.; Hu, Y.; Li, Q. K.; Zhang, H. Mapping the O-glycoproteome Using Site-specific Extraction of O-linked Glycopeptides (EXoO). *Mol. Syst. Biol.* **2018**, *14* (11). https://doi.org/10.15252/msb.20188486.

(31) Yang, S.; Onigman, P.; Wu, W. W.; Sjogren, J.; Nyhlen, H.; Shen, R. F.; Cipollo, J. Deciphering Protein O-Glycosylation: Solid-Phase Chemoenzymatic Cleavage and Enrichment. *Anal. Chem.* **2018**, *90* (13), 8261–8269. https://doi.org/10.1021/acs.analchem.8b01834.

(32) Yang, W.; Ao, M.; Song, A.; Xu, Y.; Sokoll, L.; Zhang, H. Mass Spectrometric Mapping of Glycoproteins Modified by Tn-Antigen Using Solid-Phase Capture and Enzymatic Release. *Anal. Chem.* **2020**, acs.analchem.0c01564. https://doi.org/10.1021/acs.analchem.0c01564.

(33) Singh, C.; Zampronio, C. G.; Creese, A. J.; Cooper, H. J. Higher Energy Collision Dissociation (HCD) Product Ion-Triggered Electron Transfer Dissociation (ETD) Mass Spectrometry for the Analysis of N-Linked Glycoproteins. *J. Proteome Res.* **2012**, *11* (9), 4517–4525. https://doi.org/10.1021/pr300257c.

(34) Saba, J.; Dutta, S.; Hemenway, E.; Viner, R. Increasing the Productivity of Glycopeptides Analysis by Using Higher-Energy Collision Dissociation-Accurate Mass-Product-Dependent Electron Transfer Dissociation. *Int. J. Proteomics* **2012**, *2012*, 1–7. https://doi.org/10.1155/2012/560391.

(35) Wu, S. W.; Pu, T. H.; Viner, R.; Khoo, K. H. Novel LC-MS2 Product Dependent Parallel Data Acquisition Function and Data Analysis Workflow for Sequencing and Identification of Intact Glycopeptides. *Anal. Chem.* **2014**, *86* (11), 5478–5486. https://doi.org/10.1021/ac500945m.

(36) Rose, C. M.; Rush, M. J. P.; Riley, N. M.; Merrill, A. E.; Kwiecien, N. W.; Holden, D. D.; Mullen, C.; Westphall, M. S.; Coon, J. J. A Calibration Routine for Efficient ETD in Large-Scale Proteomics. *J. Am. Soc. Mass Spectrom.* **2015**, *26* (11), 1848–1857. https://doi.org/10.1007/s13361-015-1183-1.

(37) Brademan, D. R.; Riley, N. M.; Kwiecien, N. W.; Coon, J. J. Interactive Peptide Spectral Annotator: A Versatile Web-Based Tool for Proteomic Applications. *Mol. Cell. Proteomics* **2019**, *18* (8), S193–S201. https://doi.org/10.1074/mcp.TIR118.001209.

(38) Mauri, M.; Elli, T.; Caviglia, G.; Uboldi, G.; Azzi, M. RAWGraphs: A Visualisation Platform to Create Open Outputs. In *Proceedings of the 12th Biannual Conference on Italian SIGCHI Chapter - CHItaly '17*; ACM Press: New York, New York, USA, 2017; Vol. Part F131371, pp 28:1-28:5. https://doi.org/10.1145/3125571.3125585.

(39) Crooks, G. E.; Hon, G.; Chandonia, J. M.; Brenner, S. E. WebLogo: A Sequence Logo Generator. *Genome Res.* **2004**, *14* (6), 1188–1190. https://doi.org/10.1101/gr.849004.

(40) Vester-Christensen, M. B.; Halim, A.; Joshi, H. J.; Steentoft, C.; Bennett, E. P.; Levery, S. B.; Vakhrushev, S. Y.; Clausen, H. Mining the O-Mannose Glycoproteome Reveals Cadherins as Major O-Mannosylated Glycoproteins. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (52), 21018–21023. https://doi.org/10.1073/pnas.1313446110.

(41) Polasky, D. A.; Yu, F.; Teo, G. C.; Nesvizhskii, A. I. Fast and Comprehensive N- and O-Glycoproteomics Analysis with MSFragger-Glyco. *bioRxiv* **2020**, 2020.05.18.102665. https://doi.org/10.1101/2020.05.18.102665.

(42) Halim, A.; Westerlind, U.; Pett, C.; Schorlemer, M.; Rüetschi, U.; Brinkmalm, G.; Sihlbom, C.; Lengqvist, J.; Larson, G.; Nilsson, J. Assignment of Saccharide Identities through Analysis of Oxonium Ion Fragmentation Profiles in LC–MS/MS of Glycopeptides. *J. Proteome Res.* **2014**, *13* (12), 6024–6032. https://doi.org/10.1021/pr500898r.

(43) Riley, N. M.; Hebert, A. S.; Westphall, M. S.; Coon, J. J. Capturing Site-Specific Heterogeneity with Large-Scale N-Glycoproteome Analysis. *Nat. Commun.* **2019**, *10* (1). https://doi.org/10.1038/s41467-019-09222-w.

(44) Shaw, J. B.; Malhan, N.; Vasil'Ev, Y. V.; Lopez, N. I.; Makarov, A.; Beckman, J. S.; Voinov, V. G. Sequencing Grade Tandem Mass Spectrometry for Top-Down Proteomics Using Hybrid Electron Capture Dissociation Methods in a Benchtop Orbitrap Mass Spectrometer. *Anal. Chem.* **2018**, *90* (18), 10819–10827. https://doi.org/10.1021/acs.analchem.8b01901.

(45) Khatri, K.; Pu, Y.; Klein, J. A.; Wei, J.; Costello, C. E.; Lin, C.; Zaia, J. Comparison of Collisional and Electron-Based Dissociation Modes for Middle-Down Analysis of Multiply Glycosylated Peptides. *J. Am. Soc. Mass Spectrom.* **2018**, *29* (6), 1075–1085. https://doi.org/10.1007/s13361-018-1909-y.

(46) Riley, N. M.; Coon, J. J. The Role of Electron Transfer Dissociation in Modern Proteomics. *Analytical Chemistry*. American Chemical Society January 2, 2018, pp 40–64. https://doi.org/10.1021/acs.analchem.7b04810.

(47) Zubarev, R.; Kelleher, N. L.; McLafferty, F. W. Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process. *J. Am. Chem. Soc.* **1998**, *120*, 3265–3266.

(48) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. Peptide and Protein Sequence Analysis by Electron Transfer Dissociation Mass Spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (26), 9528–9533. https://doi.org/10.1073/pnas.0402700101.

(49) Perez-Riverol, Y.; Csordas, A.; Bai, J.; Bernal-Llinares, M.; Hewapathirana, S.; Kundu, D. J.; Inuganti, A.; Griss, J.; Mayer, G.; Eisenacher, M.; et al. The PRIDE Database and Related Tools and Resources in 2019: Improving Support for Quantification Data. *Nucleic Acids Res.* **2019**, *47* (D1), D442–D450. https://doi.org/10.1093/nar/gky1106.