Photoproximity Labeling of Sialylated Glycoproteins (GlycoMap) Reveals Sialylation-Dependent Regulation of Ion Transport

Claudio F. Meyer, Ciaran P. Seath, Steve D. Knutson, Wenyun Lu, Joshua D. Rabinowitz, and David W. C. MacMillan*

ABSTRACT: Sialylation, the addition of sialic acid to glycans, is a crucial post-translational modification of proteins, contributing to neurodevelopment, oncogenesis, and immune response. In cancer, sialylation is dramatically upregulated. Yet, the functional biochemical consequences of sialylation remain mysterious. Here, we establish a µMap proximity labeling platform that utilizes metabolically inserted azidosialic acid to introduce iridium-based photocatalysts on sialylated cell-surface glycoproteins as a means to profile local microenvironments across the sialylated proteome. In comparative experiments between primary cervical cells and a cancerous cell line (HeLa), we identify key differences in both the global sialome and proximal proteins, including solute carrier proteins that regulate metabolite and ion transport. In particular, we show that cell-surface interactions between receptors trafficking ethanolamine and zinc are sialylation-dependent and impact intracellular metabolite levels. These results establish a µMap method for interrogating proteoglycan function and support a role for sialylated glycoproteins in regulating cell-surface transporters.

INTRODUCTION

Glycosylation is one of the most common post-translational modifications (PTM) on proteins, occurring on at least 50% of all known mammalian proteins and dramatically increasing the functional proteome.1 Glycosylation can alter both protein localization and function, and misregulated deposition has been shown to contribute to varied disease phenotypes, such as cancer metastasis, viral immune escape, viral entry, and inflammation.2–4 Glycoproteins also play a critical role in overall cell-surface architecture, contributing to cell adhesion,3 cell signaling,4 viral docking,16 and cell–cell interactions.7 Among the array of cell-surface monosaccharides, sialic acid stands out as being particularly influential for cell function. This charged sugar is incorporated by sialyltransferases and commonly decorates the termini of polysaccharide chains (Figure 1). During oncogenesis, overexpression of sialyltransferases leads to hypersialylation,8 in turn promoting tumor progression through two different paradigms: (1) sialylation appears to inhibit apoptosis and allow the cell to evade the immune system9 and (2) the sialylglycoconjugate sialyl Lewis X facilitates metastasis via extravasation of cancer cells out of the bloodstream into nearby tissue.10,11

Despite these observations, the underlying biochemical mechanisms remain poorly understood, in part due to a lack of high-resolution tools to assess the functional roles of sialylation.12 As such, new methods to profile the biochemical impact of sialylation are needed. We hypothesized that chemoproteomic profiling of both the sialome (sialylated glycoproteins) and associated interactome could provide insights. Protein–protein interactions (PPIs) are typically identified by immunoprecipitation/mass spectrometry (IP/MS) workflows,13 whereby the target protein is enriched along with interactors. However, these strategies are challenging for the identification of transient or low-affinity PPIs, which are especially prevalent in glycoproteins and often suffer from poor signal-to-noise ratio.14,15 To surmount these challenges, pioneering work from the Kohler group, and others, has demonstrated that metabolically incorporated sialic acids bearing crosslinkers could be used for photoaffinity labeling (PAL) of sialic acid interactors.16–23 More recently, Huang and co-workers elegantly showed that galectin-3 fused to APEX2 could be used to interrogate the glycan interactome via peroxide-triggered proximity labeling, revealing novel galectin-3 glycoprotein receptors.24 To more comprehensively interrogate sialylation-dependent changes in cancer, we sought to augment our µMap photoaffinity labeling platform25 with the capacity to identify the interactome of sialylated cell–
surface glycoproteins. This strategy, which we term GlycoMap, employs precise introduction of iridium (Ir) photocatalysts onto sialylated cell—surface glycoproteins using the Bertozzi method\(^\text{26}\) for metabolic incorporation of azidosialic acid. Thereafter, strain-promoted alkyne-azide cycloaddition (SPAAC) with a DBCO-tethered iridium catalyst affords iridium-conjugated sialylated glycoproteins. Irradiation with blue light in the presence of a biotin-diazirine probe locally generates carbenes that crosslink with interacting proteins. This strategy offers several distinct advantages in that the introduced proximity labeling machinery is highly selective (iridium incorporation only localized to sialic acid), of relatively small size (compared to enzymatic proximity labeling strategies), catalytic in nature (allowing for signal amplification), and offers precise spatiotemporal control over labeling.\(^\text{28}\)

# RESULTS AND DISCUSSION

We began our study by optimizing the iridium catalyst incorporation into the sialome (Figure 2a). Incubation of HeLa cells with tetraacetyl-N-azidoacetylmannosamine (Ac\(_4\)ManNAz), followed by treatment with DBCO–iridium (S1) led to the incorporation of the iridium photocatalyst onto glycoproteins. Irradiation in the presence of biotin-diazirine (S2) resulted in cell–surface biotinylation as observed by western blot (Figure 2B). Control reactions displayed minimal biotinylation when omitting the azido-sugar, DBCO–iridium reagent, or blue light irradiation. Importantly, immunoprecipitation over streptavidin beads showed strong enrichment of known sialylated glycoproteins nicastrin (NCSTN) and complement decay-accelerating factor (CD55), providing confidence in our workflow (Figure 2C).

We also compared our method to direct biotinylation of Ac\(_4\)ManNAz using DBCO–biotin (S3) and observed labeling of the sialome via western blot. These results suggest that our
A catalytic labeling method can install ~0.7 tags per catalyst (see the Supporting Information), representing a significant improvement over existing PAL probes, which typically react with water (>95%) and display minimal protein labeling.28 Although the efficiency is slightly lower than direct DBCO-biotinylation, we reason that click-based loading of iridium photocatalysts onto metabolically incorporated Ac₄ManNAz is somewhat more sterically hindered but nonetheless affords photoactivatable biotinylation of glycoproteins and their interactors (Figure 2B). Furthermore, this labeling differential is optimal for distinguishing between glycoproteins and their interactors. Finally, this protocol was also applicable to both HEK293T cells and primary cervical cells (PCC) (Figure 2D), highlighting the versatility of our workflow. In all cases, confocal microscopy revealed strong biotinylation after treatment with either DBCO-iridium or DBCO-biotin and irradiation in the presence of biotin-diazirine. In the absence of azido-sugar, no labeling was observed in any case.

Encouraged by these results, we developed a tandem mass tag (TMT)-based quantitative chemoproteomics workflow to identify sialylated cell-surface glycoproteins and map their interactomes. We first examined this chemoproteomic workflow on HEK293T cells (Figure 3B). Using condition A (vs control) we found significant enrichment (>1.5 log₂(fold change); >1.5 −log₁₀(p-value)) of 363 proteins, of which 93% were known glycoproteins, including nicastrin (NCSTN), cadherin 2 (CDH2), small cell adhesion glycoprotein (SMAGP), cluster of differentiation 47 (CD47), basigin (BSG), cluster of differentiation 166 (CD166), cluster of differentiation 99 (CD99), and neuroplastin (NPTN). As predicted by Figure 3A, condition B (vs control) enriches both sialylated glycoproteins and proximal proteins and thus shares an ~65% overlap with the proteins enriched in condition A (vs control) (Figure S9). Analysis of the sialic acid interactome generated by GlycoMap showed 81% of the enriched proteins were membrane-associated, reflecting the accuracy of our labeling approach. In addition, several lysosomal proteins (21) were enriched, presumably arising via internalization of iridium-bound glycoproteins prior to proximity labeling.

In our initial analysis of the GlycoMap dataset, we examined a known membrane-bound protein complex, γ-secretase, to validate our methodology.
This heterotetramer of membrane proteins (NCSTN, APH1A, PSEN1, PEN-2) proteolytically cleaves many integral membrane proteins, but only NCSTN is directly sialylated. In our dataset, NCSTN is highly enriched (3.8 log2(FC)) in condition A (vs control), whereas the nonsialylated interactor, APH1A, is strongly enriched (3.3 log2(FC)) in the GlycoMap arm, demonstrating that the μMap workflow can delineate between sialylated glycoproteins and their interactors.

With an established workflow in hand, we next sought to investigate hypersialylation events in oncogenesis. To examine...
this, we performed comparative GlycoMap experiments on both primary cervical cancer cells (PCC) and the HeLa cervical adenocarcinoma cell line (Figure 4A). In agreement with previous observations of upregulated sialylation, chemo-proteomic analysis revealed significantly higher sialylation in HeLa cells (447 enriched proteins) than in PCC (223 enriched

Figure 4. (A) Comparative proteomics experiment of primary cervical cells (PCC) and HeLa cells. Top row: sialylated glycoproteins in PCC (left) and HeLa (right). Bottom row: interacting proteins in PCC (left) and HeLa cells (right). Middle: Venn diagram of the enriched proteins from each dataset. The same cutoffs (>1.5 \( \log_{2}(\text{fold change}) \); >1.5 \( -\log_{10}(\text{p-value}) \)) were used for the analysis of all data sets. (B) Gene ontology (GO) analysis of the identified sialylated glycoproteins (top) and their interactors (bottom). (C) Venn diagram of the enriched solute carrier proteins (SLCs) that interact with sialylated glycoproteins.
proteins) (Figure 4A). This sialome increase in HeLa cells consequently yielded a higher number of interacting proteins (166 enriched proteins in HeLa cells vs 63 enriched proteins in PCC).

Figure 5. (A) Workflow for the metabolomics analysis of HeLa cells. (B) Metabolite levels of selected small molecules. Experiments were performed in triplicate. (C) (Left) GO analysis suggests that cation homeostasis is affected by sialylation; (middle) enriched zinc transporters in the HeLa and PCC dataset; (right) colorimetric zinc assay, which shows a significant change in cellular zinc levels in response to desialylation. *p-Values determined by unpaired Students t-test. *p < 0.05, **p < 0.01.
We then performed global gene ontology (GO) analysis to categorize the enriched sialylated proteins and their interactors (Figure 4B). Functional enrichment from both cell types was in good agreement with the known roles of identified glycoproteins: cell adhesion, host–cell entry, and regulation of migration, death, and defense. Furthermore, examination of GO terms that differed significantly (≥10 log_{10}(p-value) change) between the primary and cancerous cervical cells validated that the sialylated glycoproteins in cancerous cells are related to typical oncological phenotypes, including cell morphogenesis, cell–cell adhesion, extracellular matrix organization, and tube morphogenesis.

Interestingly, when comparing the roles of identified sialic acid interacting proteins, terms related to small-molecule transport were clearly enriched in HeLa cells (organic ion transport, transport of small molecules, vitamin transport), and we specifically noted numerous enriched solute carrier proteins (SLCs) in this dataset (Figure 4C). In particular, interactions of sialylated proteins with SLCs associated with ethanolamine, carnitine, and zinc transport were all significantly enriched in HeLa cells over PCC.

To explore the potential consequences of these interactions, we looked for the metabolic consequences of enzymatically depleting them (Figure 5A). Treatment of HeLa cells with sialidase isolated from *Vibrio cholerae* (VC-Sia) efficiently cleaves sialic acids α2,3-, α2,6-, or α2,8-linked to cell–surface glycans, enabling us to modulate global sialylation status. We incubated HeLa cells either in the presence or absence of VC-Sia and then performed mass spectrometry-based metabolomic quantification on cellular metabolite extracts. While most metabolite levels were minimally affected by sialidase treatment, we found that levels of ethanolamine derivatives, including cytidine diphosphate ethanolamine (CDP-Etn), phosphate ethanolamine (P-Etn), and cytidine diphosphate choline (CDP-choline), were significantly increased in sialidase-treated cells (Figure 5B). The solute carrier protein responsible for the transport of ethanolamine, choline-like transporter 1 (SLC44A1), is not known to be glycosylated; however, our dataset suggests that its function could be regulated by neighboring sialylated glycoproteins. Based on these results, we hypothesize that cell–surface sialic acids may present a negatively charged surface around membrane-bound transporters, which in turn could affect the transport of ions, including metabolites.

Similarly, we also investigated the impact of sialylation on zinc uptake (Figure 5C). Zinc is imported through the cell membrane by a series of solute carrier proteins of the SLC39 family, four of which are shown in our HeLa dataset to be sialylated (SLC39A6, SLC39A8, SLC39A10, and SLC39A14) and one (SLC39A1) is suggested to interact with a sialylated glycoprotein. Zinc is a key micronutrient that plays a significant role in cell function, and its transport is dysregulated in many cancers. Using a colorimetric assay to determine the zinc level of untreated and sialidase-treated HeLa cells, we found that the zinc level was significantly higher in cells treated with VC-Sia. These data suggest that cell–surface sialylation and/or the interaction with sialylated glycoproteins play a role in the regulation of cellular zinc concentration.

**CONCLUSIONS**

In conclusion, hypersialylation in cancer has recently attracted significant interest from the academic and pharmaceutical sectors. However, tools to understand the biochemical consequences of hypersialylation remain limited. Here, we have described a novel proximity labeling platform to identify sialylated cell–surface glycoproteins and their interactors. This sensitive and precise method is robust and compatible with various cell lines, including primary cells. Our comparative proteomics studies between primary and cancerous cervical cell lines show a significant link between sialylation and solute carrier proteins, suggesting a new role for sialylation. Metabolomics data suggest that these interactions regulate the function of certain solute carriers. Together, our platform represents a powerful new approach to sialic acid interactome profiling, providing a systems-level tool for the elucidation of hypersialylation’s biochemical consequences.

**ASSOCIATED CONTENT**

**Supporting Information**

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacs.2c11094.

Proteomic data sets (XLSX)

Experimental procedures and supplementary figures (PDF)

**AUTHOR INFORMATION**

Corresponding Author

David W. C. MacMillan — Merck Center for Catalysis at Princeton University, Princeton, New Jersey 08544, United States; Department of Chemistry, Princeton University, Princeton, New Jersey 08544, United States; orcid.org/0000-0001-6447-0587; Email: dmacmill@princeton.edu

Authors

Claudio F. Meyer — Merck Center for Catalysis at Princeton University, Princeton, New Jersey 08544, United States; Department of Chemistry, Princeton University, Princeton, New Jersey 08544, United States

Ciaran P. Seath — Merck Center for Catalysis at Princeton University, Princeton, New Jersey 08544, United States; Department of Chemistry, Princeton University, Princeton, New Jersey 08544, United States

Steve D. Knutson — Merck Center for Catalysis at Princeton University, Princeton, New Jersey 08544, United States; Department of Chemistry, Princeton University, Princeton, New Jersey 08544, United States; Lewis-Sigler Institute for Integrative Genomics and Ludwig Institute for Cancer Research, Princeton University, Princeton, New Jersey 08544, United States; orcid.org/0000-0001-6822-464X

Wenyun Lu — Department of Chemistry, Princeton University, Princeton, New Jersey 08544, United States; Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, New Jersey 08544, United States; orcid.org/0000-0003-1787-2617

Joshua D. Rabinowitz — Department of Chemistry, Princeton University, Princeton, New Jersey 08544, United States; Lewis-Sigler Institute for Integrative Genomics and Ludwig Institute for Cancer Research, Princeton University, Princeton, New Jersey 08544, United States; orcid.org/0000-0002-1247-4727

Complete contact information is available at: https://pubs.acs.org/10.1021/jacs.2c11094

Notes

The authors declare the following competing financial interest(s): A provisional U.S. patent has been filed by DWCM and CPS based on materials used in this work, 62/
ACKNOWLEDGMENTS
This work was funded by the NIH National Institute of General Medical Sciences (R35-GM134897-02) and kind gifts from Merck, BMS, Pfizer, Janssen, Genentech, and Eli Lilly. The authors also acknowledge the Princeton Catalysis Initiative and the Ludwig Cancer Research for supporting this work. They also acknowledge the Swiss National Science Foundation (P2SKP2_199458, C.F.M.) and the NIH (1F32GM142206-01, S.D.K.; R50CA211437, W.L.) for fellowships. The authors thank Saw Kyin and Henry H. Shwe at the Princeton Proteomics Facility. The authors acknowledge the use of Princeton’s Imaging and Analysis Center, which is partially supported by the Princeton Center for Complex Materials, National Science Foundation/Materials Research Science and Engineering Centers program (DMR-1420541). Generalized schemes were created using Biorender.

REFERENCES