Anionic Glycan Diversity in Waterfowl Egg Whites through Glycoblotting-based Sulphoglycomics Approach

(グライコブロッティング連動型スルフォグライコミクス法による
水鳥卵白中のアニオン性糖鎖多様性に関する研究)

Doctoral Dissertation
September 2023

Bryan M. Montalban

Laboratory of Advance Chemical Biology
Graduate School of Life Science
Hokkaido University

ABSTRACT

Sulfated *N*- and *O*-glycans exist in trace levels which are challenging to detect, particularly in the presence of abundant neutral and sialylated glycans. Current MALDI-TOF MS-based sulfoglycomics approaches employ permethylation to discriminate sulfated glycans from sialylated glycans and charge-based separation to isolate the sulfated glycans from the rest of the permethylated neutral and sialyl-glycans. However, these methods suffer from concomitant sample losses during cleanup steps. In this study, we describe Glycoblotting as a straightforward complementary method that offers a seamless platform for glycan purification, enrichment, methylation, and labeling to address sulfated glycan enrichment, sialic acid methylation, and sample loss. Glycoblotting's on-bead chemoselective ligation of reducing sugars with hydrazide demonstrates excellent recovery of sulfated glycans and allows the detection of more sulfated glycan species. The on-bead methyl esterification of sialic acid using 3-methyl-1-*p*-tolyltriazene (MTT) effectively discriminates sulfated glycans from sialylated glycans. Furthermore, MTT facilitates simultaneous detection and differentiation of sulfate and phosphate groups in isobaric *N*-glycan species.

Additionally, we investigate the expression of acidic *N*-glycans, specifically sulfated and phosphorylated *N*-glycans, in the egg whites of 72 avian species belonging to the Order Anseriformes (waterfowls). Employing the Glycoblotting-based sulphoglycomics approach, we elucidated the diversity of acidic *N*-glycans and their implication in protecting embryos from infections. Our findings revealed family-specific variations in waterfowl egg whites sulfated and phosphorylated *N*-glycan profiles. Different waterfowl species exhibit distinct expressions of sulfated trans-Gal(+) and trans-Gal(-) *N*-glycan structures. Moreover, species-specific expression of phosphorylated *N*-glycans was also observed. Notably, waterfowl species with a high virus prevalence expressed a higher abundance of phosphorylated hybrid and high-mannose *N*-glycans on their egg whites.

The Glycoblotting-based sulphoglycomics approach presents a significant breakthrough in sulfated glycan analysis, simplifying the existing MALDI-TOF MS-based sulphoglycomics workflow and enabling comprehensive exploration of the complex glycome of biological samples. Furthermore, the findings of this study shed light on the importance of phosphorylated and sulfated *N*-glycans in understanding the role of acidic glycans in the Influenza A virus (IAV) propagation in waterfowl. These results hold immense potential for advancing our understanding of IAV propagation in avian species and guiding the development of targeted interventions to combat influenza.

ACKNOWLEDGEMENT

These past three years have been a period of intense learning not only as a researcher but also as an individual. I want to thank the people and institutions who contributed not only to the completion of this paper but to this Ph.D. journey.

This work would not have been possible without the scholarship grant and financial support of the **Japanese Government Ministry of Education, Culture, Sports, Science and Technology (MEXT)**, enabling me to pursue graduate studies at Hokkaido University.

To my adviser, **Professor Hiroshi Hinou**, I would like to express my gratitude for unselfishly sharing your time and expertise. Thank you very much for your never-ending support and guidance.

To **Prof. Uehara** and **Prof. Kitamura** for your valuable comments and suggestions for improving this study.

To **Professor Shin-Ichiro Nishimura**, for your valued inputs on the Glycoblotting approach and your constant encouragement every time my experiments failed.

Thank you very much to **Morita-san**, **Takahashi-san**, and **Hirane-san** for the technical assistance and material support during the research.

To **Shogo-san**, **Kohki-san**, **Koki-san**, **Yuki-san**, **Taichi-san**, and **Sawada-san** thank you!

To **Larry**, **Ren**, **June**, and **Feleke**, thank you aren't enough for all the time we spent together, for sharing your smiles, laughter, support, encouragement, and even criticisms.

To **HAFS**, all of you made the last three years of my life very memorable.

Nobody has been more important to me in the pursuit of this research than the members of my family. I want to thank my parents and siblings, whose love and guidance are with me in whatever endeavor I pursue.

Maraming salamat po! Domo arigatou gozaimasu!

TABLE OF CONTENTS

Chapter 1 General Introduction

Influenza A virus (IAV) continues to pose a significant threat to global public health, causing seasonal epidemics and occasional pandemics with severe morbidity and mortality. The constant evolution and adaptability of IAV strains and the ability to infect a wide range of avian and mammalian species present ongoing challenges for disease control and prevention. IAV belongs to the family of Orthomyxoviridae, and its virion is covered with three viral proteins – hemagglutinin (HA), neuraminidase (NA), and M2 ion channel. Wherein HA and NA modulate the host range, pathogenicity, and immunogenicity of IAVs[1]. Waterfowl (ducks, geese, and swans) are generally considered the natural reservoirs of avian influenza viruses (AIV) and play a crucial role in the transmission and evolution of the virus. Understanding the molecular interactions between AIVs and their avian hosts is essential for comprehending AIVs' pathogenesis and transmission dynamics[2].

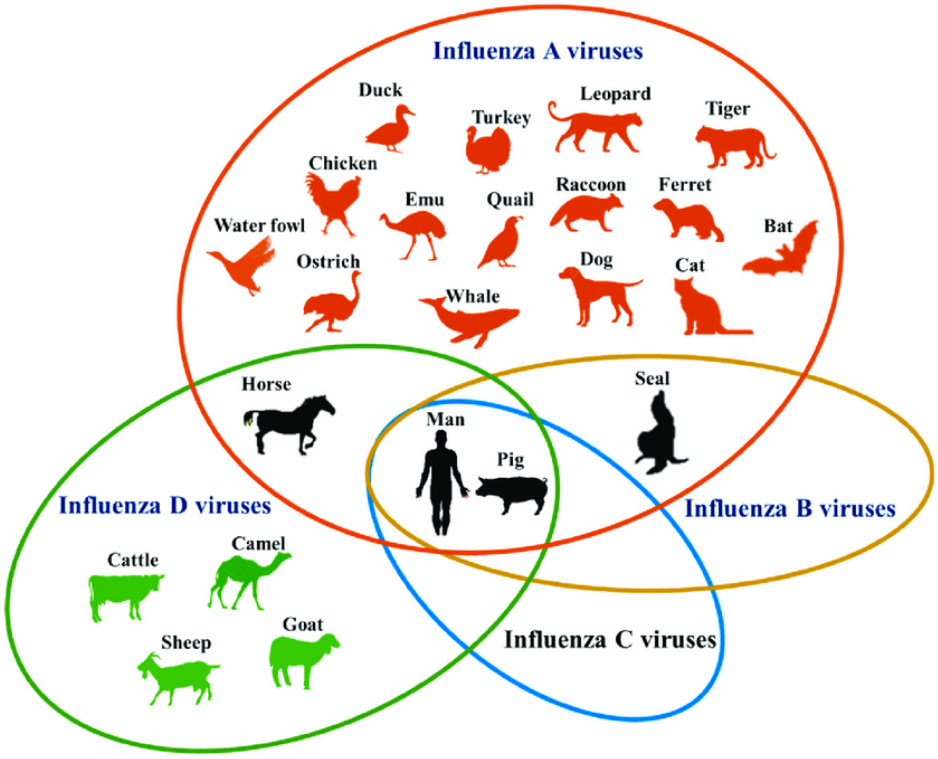


Figure 1.1 Host-range of influenza viruses[2].

In recent years, glycans have emerged as critical factors in the interactions between IAVs and their hosts. Glycans are complex biomolecules that play diverse roles in biological systems. They are present on the surface of cells and proteins, forming a glycan shield that influences various biological processes, including cell adhesion, immune responses, and viral infections. In the context of IAV, glycans serve as receptors or attachment factors for IAVs, mediating their entry into host cells and facilitating viral replication. The glycan-binding specificity of IAVs determines their host range and tissue tropism. In particular, human-adapted IAVs exhibit binding specificity towards sialyl-LacNAc, characterized by an α2,6 linkage between Sia and Gal, while avian IAVs bind to sialyl-LacNAc with α2,3 linkages[3, 4]. However, recently it was found that receptor specificity of every hemagglutinin subtype varies and greatly depends not only on α2,3/6-linked sialoside but also on the underlying glycan structures[5, 6].

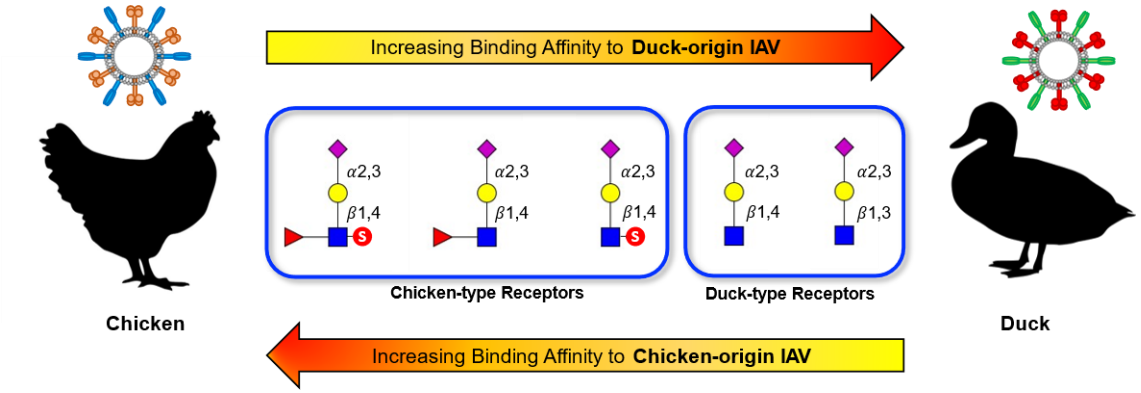


Figure 1.2 Binding specificity of avian Has based on synthetic glycan library[5, 7].

Virus receptor binding specificity also correlated with the expression level of relevant sialic acid determinants on the target cells of different host species. A poultry H9N2 virus was observed to recognize Neu5Acα2-6Gal-terminated sialyloligosaccharides, indicating that AIV

may display human-virus-like receptor specificity. It was also found that chicken and quail intestinal cells express both Neu5Acα2-3Gal and Neu5Acα2-6Gal sialyloligosaccharides, in contrast to ducks that express only Neu5Acα2-3Gal. Although most avian viruses share the Neu5Acα2-3Gal receptor specificity, viruses adapted to different avian species can differ in their ability to recognize the third saccharide and more distant moieties of Neu5Acα2-3Gal-terminated receptors. For example, duck viruses preferentially bind to Neu5Acα2-3Galβ1-6GalNAc, while chicken viruses prefer receptors with inner β-N-acetylglucosamine moiety, Neu5Acα2-3Galβ1-4GlcNAc[5, 8, 9].

Furthermore, Ichimiya *et. al.*[10] reported the role of sulfation on LacNAc moieties in IAV replication. They demonstrated that IAV inoculation of MDCK cells with overexpressed sulfotransferase increases viral replication by 90-fold. Additionally, they have shown that 6-sulfo sialyl Lewis X and 6-sulfo sialyl LacNAc moieties are highly expressed in chicken embryos, suggesting their involvement in the efficient propagation of human H1N1. Human influenza was also found to bind to phosphorylated glycans in human lungs. Interestingly, Byrd-Leotis, *et al.*[11, 12] suggested an alternative phosphorylated glycans binding site may exist. The discovery of sulfated and phosphorylated glycans as crucial determinants for IAV HA showed the importance of acidic glycans on IAV infection.

Despite the numerous biological significances of sulfated and phosphorylated glycans, their analysis remains a challenging task. An inherent technical problem why sulfated *N*- and *O*-glycans are often not detected by current approaches in mass spectrometry-based glycomics, is due to their naturally lower abundance, compounded further by their negatively charged nature, which adversely disfavors their ionization and detection amid a sea of often much more abundant non-sulfated and sialylated glycans[13–15]. Recognizing these limitations, the general approach for MS-based sulfoglycomics is permethylation followed by anion exchange separation. This approach was devised, since sulfate would remain the only substituent carrying the negative

charge, fully methylated sulfated glycans can then be selectively detected by MS in negative mode and readily separated from the more abundant non-sulfated glycans which is difficult to achieve with native glycans[16, 17].

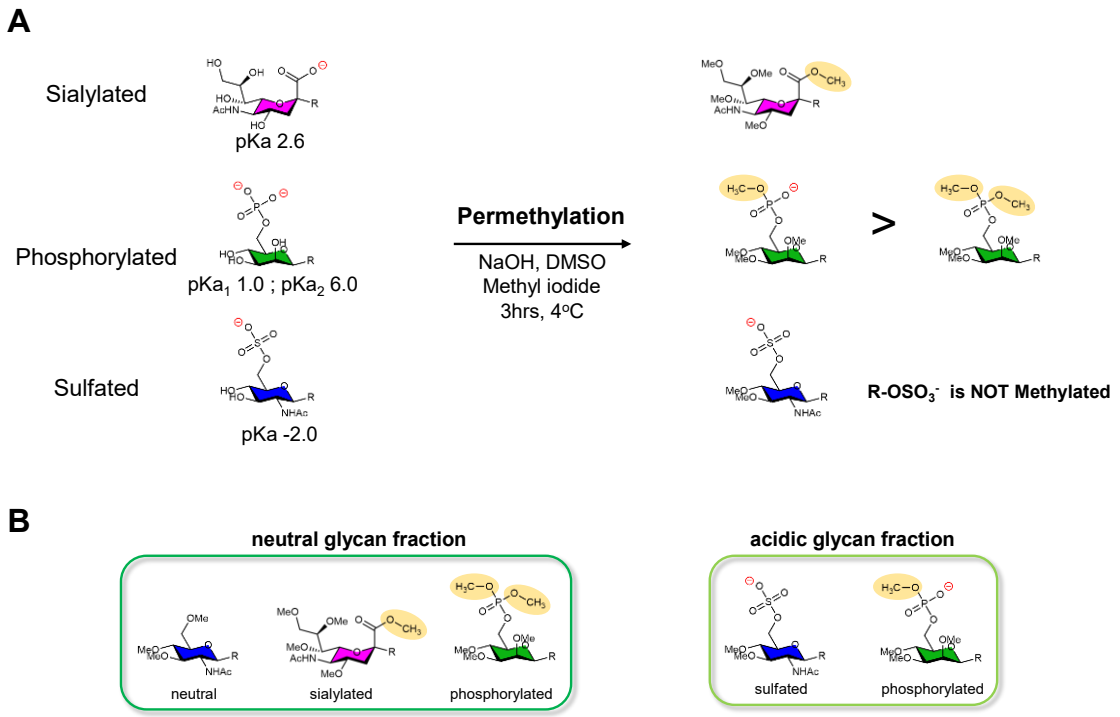


Figure 1.3 (A) Permethylation of sialylated, phosphorylated, and sulfated glycans. (B) Neutral glycan and acidic glycan species obtained after weak anion exchange (WAX) separation of permethylated glycans.

The general objective of this study is to develop a straightforward workflow for the analysis of sulfated *N*- and *O*-glycans in biological samples. Specifically, this study aimed:

i. To develop a glycoblotting-based sulfoglycomics approach to analyze sulfated *N*- and *O*-glycans.

ii. To elucidate the acidic *N*-glycan expression profiles of avian egg whites from Order Anseriformes (waterfowls).

The objectives mentioned above were pursued due to the heavy reliance on current MALDI-TOF MS-based sulfoglycomics approaches on permethylation to effectively distinguish sulfated glycans from abundant neutral and sialyl glycans. In contrast, this study introduces the Glycoblotting protocol as a complementary method for the MALDI-TOF MS-based sulfoglycomics workflow. The Glycoblotting protocol offers a streamlined process for on-bead glycan purification, enrichment, methylation, and labeling on a single platform, addressing the three main challenges in sulfoglycomics: trace abundance, sample loss, and the presence of sialic acid. Moreover, the Glycoblotting-based sulfoglycomics workflow enables the differentiation of sulfate from phosphate groups in isobaric glycan species, expanding the scope of sulfoglycomics to include phosphorylated glycans. Additionally, the elucidation of acidic *N*-glycan expression in waterfowl egg whites, undertaken for the first time, will not only shed light on glycan diversity but also provide valuable insights into the potential roles of acidic *N*-glycans in understanding influenza infections in waterfowl species.

References

1. Hiono, T., Matsuda, A., Wagatsuma, T., Okamatsu, M., Sakoda, Y., & Kuno, A. (2019). Lectin microarray analyses reveal host cell-specific glycan profiles of the hemagglutinins of influenza A viruses. *Virology* 527, 132–140

2. Kuchipudi, S. V., & Nissly, R. H. (2018). Novel flu viruses in bats and cattle: "Pushing the envelope" of influenza infection. *Vet Sci* 5,

3. Suzuki, N., Abe, T., & Natsuka, S. (2022). Structural analysis of N-glycans in chicken trachea and lung reveals potential receptors of chicken influenza viruses. *Sci Rep* 12,

4. Zhao, C., & Pu, J. (2022). Influence of Host Sialic Acid Receptors Structure on the Host Specificity of Influenza Viruses. *Viruses* 14,

5. Gambaryan, A., Yamnikova, S., Lvov, D., Tuzikov, A., Chinarev, A., Pazynina, G., Webster, R., Matrosovich, M., & Bovin, N. (2005). Receptor specificity of influenza viruses from birds and mammals: New data on involvement of the inner fragments of the carbohydrate chain. *Virology* 334, 276–283

6. Thompson, A. J., & Paulson, J. C. (2021). Adaptation of influenza viruses to human airway receptors. *Journal of Biological Chemistry* 296, 100017

7. Hiono, T., Okamatsu, M., Nishihara, S., Takase-Yoden, S., Sakoda, Y., & Kida, H. (2014). A chicken influenza virus recognizes fucosylated α2,3 sialoglycan receptors on the epithelial cells lining upper respiratory tracts of chickens. *Virology* 456–457, 131–138

8. Gambaryan, A. S., Tuzikov, A. B., Pazynina, G. V., Desheva, J. A., Bovin, N. V., Matrosovich, M. N., & Klimov, A. I. (2008). 6-sulfo sialyl Lewis X is the common receptor determinant recognized by H5, H6, H7 and H9 influenza viruses of terrestrial poultry. *Virol J* 5,

9. Gambaryan, A. S., Matrosovich, T. Y., Philipp, J., Munster, V. J., Fouchier, R. A. M., Cattoli, G., Capua, I., Krauss, S. L., Webster, R. G., Banks, J., Bovin, N. V., Klenk, H.-D., & Matrosovich, M. N. (2012). Receptor-Binding Profiles of H7 Subtype Influenza Viruses in Different Host Species. *J Virol* 86, 4370–4379

10. Ichimiya, T., Okamatsu, M., Kinoshita, T., Kobayashi, D., Ichii, O., Yamamoto, N., Sakoda, Y., Kida, H., Kawashima, H., Yamamoto, K., Takase-Yoden, S., & Nishihara, S. (2021). Sulfated glycans containing NeuAcα2-3Gal facilitate the propagation of human H1N1 influenza A viruses in eggs. *Virology* 562, 29–39

11. Byrd-Leotis, L., Jia, N., Dutta, S., Trost, J. F., Gao, C., Cummings, S. F., Braulke, T., Müller-Loennies, S., Heimburg-Molinaro, J., Steinhauer, D. A., & Cummings, R. D. (2019). Influenza binds phosphorylated glycans from human lung. *Sci Adv* 5, 1–10

12. Jia, N., Byrd-Leotis, L., Matsumoto, Y., Gao, C., Wein, A. N., Lobby, J. L., Kohlmeier, J. E., Steinhauer, D. A., & Cummings, R. D. (2020). The Human Lung Glycome Reveals Novel Glycan Ligands for Influenza A Virus. *Sci Rep* 10, 1–14

13. Khoo, K. H., & Yu, S. Y. (2010). Mass spectrometric analysis of sulfated N- and O-glycans. *Methods Enzymol* 478, 3–26

14. Yu, S.-Y., Snovida, S., & Khoo, K.-H. (2020). Permethylation and Microfractionation of Sulfated Glycans for MS Analysis. *Bio Protoc* 10, 1–13

15. Cheng, P. F., Snovida, S., Ho, M. Y., Cheng, C. W., Wu, A. M., & Khoo, K. H. (2013). Increasing the depth of mass spectrometry-based glycomic coverage by additional dimensions of sulfoglycomics and target analysis of permethylated glycans. *Anal Bioanal Chem* 405, 6683–6695

16. Cheng, C. W., Chou, C. C., Hsieh, H. W., Tu, Z., Lin, C. H., Nycholat, C., Fukuda, M., & Khoo, K. H. (2015). Efficient Mapping of Sulfated Glycotopes by Negative Ion Mode nanoLC-MS/MS-Based Sulfoglycomic Analysis of Permethylated Glycans. *Anal Chem* 87, 6380–6388

17. Yu, S. Y., Wu, S. W., Hsiao, H. H., & Khoo, K. H. (2009). Enabling techniques and strategic workflow for sulfoglycomics based on mass spectrometry mapping and sequencing of permethylated sulfated glycans. *Glycobiology* 19, 1136–1149

Chapter 2 Development of a Glycoblotting-based Sulfoglycomics Workflow for the Analysis of Sulfated *N*- and *O*-glycans

2.1 Introduction

Studies have shown the significance of sulfation of the non-reducing terminal epitopes of *N*- and *O*-glycans, which modifies their physicochemical characteristics and ability to function as cognate ligands for pathogens and endogenous glycan-binding proteins. However, the presence of sulfated glycotopes in physiologically relevant situations is generally less corroborated[1].

Sulfated glycans, such as those found in tissues[1, 2], blood[3, 4], and urine[2, 5], are often present in low abundance in biological samples, making their purification and enrichment challenging. In addition, sulfated glycans are difficult to ionize and detect because of their hydrophilicity and labile sulfate groups. Furthermore, sulfate groups may be lost during ionization, resulting in the loss of structural information[6, 7]. Given these limitations, the general approach for mass spectrometry (MS)-based sulfoglycomics is permethylation[8]. Fully permethylated sulfated glycans maintain their negative charge, allowing the selective detection of sulfated glycans using MS in negative ion mode, and are readily separated from the more abundant non-sulfated glycans, which is challenging to achieve with native glycans[8–10]. In contrast, the analysis of native sulfated glycans usually employs specialized columns, such as serotonin-immobilized[4, 11] columns and TiO_2-PGC[3] columns for liquid chromatography (LC) separation, followed by electrospray ionization mass spectrometry (ESI-MS)[1, 9].

Permethylated glycans have greater hydrophobicity and reduced charge repulsion, thus improving ionization efficiency and increasing MS sensitivity[12]. The structural elucidation of sulfated glycans can also be facilitated by permethylation, mainly because methylation can produce distinctive fragmentation patterns in MS/MS analysis[8, 10]. However, due to the possibility of partial methyl-esterification[13] and sulfate loss[6], permethylation reaction conditions must be carefully optimized[7, 14]. Although this approach has succeeded in several MS-based structural analyses of complex oligosaccharides, its multi-step process requires skilled analysts and time-consuming sample handling and clean-up stages, which may result in

successive sample losses[12]. To address these challenges while maintaining the merits of methylation, thus this work proposed a Glycoblotting-based approach.

Glycoblotting is a high-throughput method for quantitative large-scale glycomics that allows efficient glycan enrichment and quantification that our group developed. In addition, we established that the glycan recovery efficiency of the Glycoblotting method is superior to that of hydrophilic interaction liquid chromatography (HILIC) and solid-phase extraction (SPE) for glycan separation and enrichment[15, 16]. In this study, we describe the utilization of the Glycoblotting protocol as a complementary method with a seamless sequential process of on-bead glycan purification, enrichment, methylation, and labeling techniques on a single platform. Followed by weak anion exchange (WAX) separation of sulfated glycans, then MALDI-TOF MS analysis (Fig. S2.1).

2.2 Methodology

2.2.1 ***Materials.*** Porcine stomach mucin (PSM) type III and ovomucoid (OVM) type III-O were purchased from Sigma-Aldrich Corp. (St. Louis, MO, USA). Bovine thyroglobulin (BTG) was purchased from EMD Millipore Corp. (USA). Peptide *N*-glycosidase F (PNGase F) was acquired from New England BioLabs (Ipswich, MA, USA), proteinase K was from Roche (Germany), trypsin was from Sigma-Aldrich Corp. (St. Louis, MO, USA), and the bacterial alkaline phosphatase (BAP) was from Nippon Gene, Ltd. (Tokyo, Japan). Ammonium carbamate, benzyloxyamine hydrochloride (BOA), 3-methyl-1-*p*-tolyltriazene (MTT), disialyloctasaccharide (SGP-10), hexa-*N*-acetylchitohexaose, 2,5-dihydroxybenzoic acid (DHB), sodium bicarbonate ($NaHCO_3$), 3,4-diaminobenzophenone (DABP), and trifluoroacetic acid (TFA) were purchased from Tokyo Chemical Industry Co. (Tokyo, Japan). BlotGlycoH beads were acquired from Sumitomo Bakelite Co., Ltd. (Tokyo, Japan).

2.2.2 *O*-glycan Release. *O*-glycans were chemically liberated using the ammonium carbamate method, as previously described[17]. Briefly, to a 30 mg dry ammonium carbamate powder, 10 μL of PSM suspension (10 μg/μL) was added in an Eppendorf tube. The mixture was incubated at 60°C for 20 hr. After incubation, 500 μL water was added and evaporated at 60°C *in vacuo*. The residue was then reconstituted with 500 μL of 150 mM acetic acid (AcOH) solution. The sample was dried at 60°C in a SpeedVac and stored at -20°C until use.

2.2.3 *N*-glycan Release[18, 19]**.** Eight microliters of 50 μg/μL bovine thyroglobulins (BTG) (approximately 400 μg) was dissolved in 50 μL of 200 mM NH_4HCO_3, followed by the addition of 4 μL denaturation buffer (5% SDS, 0.4 M DTT). The mixture was denatured for 10 mins at 100°C. After denaturation, 10 μL of 123 mM iodoacetamide was added to the mixture, which was then incubated in the dark at room temperature for 1 hr. Tryptic digestion was achieved by adding 10 μL of 40 U/μL sequence-grade Trypsin (Sigma-Aldrich) in 1 mM HCl; the mixture was then incubated overnight at 37°C, followed by heat inactivation of the enzyme at 90°C for 10 min. The tryptic digest was allowed to cool at room temperature, then 8 μL reaction buffer (0.5 M Na_3PO_4, pH 7.5), 8 μL 10% NP-40, and 2 μL of 5 U/μL PNGase F (New England BioLabs) were added and incubated overnight at 37°C. The mixture was further digested with 10 μL of 0.5 U/μL Proteinase K (Roche, Germany) at 37°C for 3 hr, followed by heat inactivation of the enzyme at 90°C for 10 min. The sample was dried in a SpeedVac and stored at -20°C until use.

2.2.4 Glycan Enrichment Using Glycoblotting[18–20]**.** A 250 μL aliquot of 10 mg/mL BlotGlycoH bead (Sumitomo Bakelite, Co.) suspension was dispensed into each well of the 96-well multiScreen Solvinert filter plate (Millipore, Billerica, MA, USA). The filter plate was then attached to a vacuum manifold to remove water. The dried samples containing released glycans (*N*- or *O*-glycans) from the glycoproteins were reconstituted with 40 μL MilliQ water. A 20 μL

aliquot of the reconstituted sample was added to the wells with 10 μL of 100 μM disialyloctasaccharide, SGP-10 (Tokyo Chemical Industry Co., Ltd.) internal standard, and 180 μL of 2% AcOH in acetonitrile (MeCN). The 96-well filter plate was incubated at 80°C for 45 min until dry. Each sample well was sequentially washed with 200 μL of 2 M guanidine-HCl in 16 mM NH_4HCO_3, water, and 1% triethylamine in methanol (MeOH). Each solvent washing was performed twice and vacuumed after every washing step. The unreacted hydrazide functional groups on the beads were capped with acetyl groups by incubating each sample well with 100 μL of 10% acetic anhydride in MeOH for 30 min at room temperature. The capping solution was then removed by vacuum and sequentially washed twice with 200 μL of 10 mM HCl, MeOH, and dioxane. On-bead methyl esterification of the carboxyl groups of acidic glycans (e.g., sialic acid) was performed by adding 100 μL of 100 mM 3-methyl-1-*p*-tolyltriazene (MTT) in dioxane into the sample and incubated at 60°C for 90 min until dry[21]. The 96-well plate was sequentially washed twice with 200 μL of dioxane, water, MeOH, and water. The captured glycans on the BlotGlycoH beads were labeled with benzyloxyamine (BOA) via trans-iminization reaction. The labeling was performed by adding 20 μL of 50 mM BOA-HCl and 180 μL of 2% AcOH in MeCN at 80°C for 45 min. BOA-labeled glycans were eluted twice with 150 μL water. The sample was dried in a SpeedVac and stored at -20°C until use.

2.2.5 Anionic-glycan Separation using WAX[10, 22]**.** Fifty microliters of 100 mg/mL 3-aminopropyl silica gel suspension (1 mmol/mg, Tokyo Chemical Industry Co. Ltd.) were packed into a 200 μL micropipette tip with a cotton plug. The packed weak anion exchange (WAX) microcolumn was conditioned and washed sequentially with 100 μL water, MeCN, and 1% AcOH in 95% MeCN twice. After each conditioning and washing step, the column was centrifugated at 500 rpm for 2 min. BOA-labeled *N*-glycans (or *O*-glycans) were reconstituted with 20 μL water. A 5 μL sample aliquot was dissolved in 150 μL 1% AcOH in 95% MeCN and then loaded into

the column. The sample was allowed to elute by gravity, and the collected eluate was reloaded back into the column; this step was done three times. The column was washed with 1% AcOH in 95% MeCN to remove unbound and hydrophobic contaminants. BOA-labeled neutral and mono-methylated sialyl *N*-glycans (or *O*-glycans) were eluted with 1% AcOH in 50% MeCN, while sulfated *N*-glycans (or *O*-glycans) were eluted with 1% NH_4OH in 5% MeCN (pH 10.5). The eluates were dried in a SpeedVac and stored at -20ºC until use.

2.2.6 Mass Spectrometric Analysis. MALDI-TOF MS analysis of BOA-labeled *N*- and *O*-glycans was performed using Ultraflex III (Bruker, Bremen, Germany) operated in reflectron mode in positive and negative ion acquisition modes. Neutral and monomethylated sialyl-glycans were analyzed in positive ion mode using 10 mg/mL DHB/$NaHCO_3$ (10:1) in 50% MeCN matrix[23, 24]. While sulfated glycans were analyzed in negative ion mode using the DABP matrix (3,4-diaminobenzophenone, 10 mg/mL in 75% MeCN with 0.1% TFA)[1, 10, 22, 25]. MALDI-TOF and MALDI-TOF/TOF MS data were annotated using the Bruker FlexAnalysis 3.0 software package. Experimental *m/z* was used to predict the possible glycan composition using the Expasy GlycoMod Tool and Glyconnect Database of the Swiss Institute of Bioinformatics (https://web.expasy.org/glycomod/) and GlycoWorkbench[26, 27].

2.2.7 Alkaline Phosphatase Digestion. The *N*-glycans released from chicken egg white ovomucoid (OVM) were reconstituted with 50 µL BAP buffer (100 mM Tris-HCl, 1 mM $MgSO_4$, pH 8.0). A 1 µL aliquot of 0.8 U/µL alkaline phosphatase (Nippon Gene) was added. The mixture was incubated at 37ºC for 24 hr.

2.3 Results and Discussion

2.3.1 Direct Analysis of Native Sulfated *O*-glycans. Our group has previously described a method for the direct analysis of neutral and sialylated native *O*-glycans**.** The technique utilizes a

novel matrix, wherein a mixture of benzyloxyamine (BOA), 2,5-dihydroxybenzoic acid (DHB), and sodium bicarbonate ($NaHCO_3$) was used[24]. This method demonstrated good sensitivity and stable modification of the glycan-reducing end by BOA, owing to oxime bond formation. Furthermore, sialic acid loss during MALDI-TOF MS analysis was also suppressed. Thus, the method showed that native glycans could be analyzed by MALDI-TOF MS devoid of sialic acid loss, even without additional glycan modifications, such as permethylation. Using the described novel matrix, we attempted to simultaneously analyze unmodified sulfated and sialylated *O*-glycans from porcine stomach mucin (PSM). Anionic *O*-glycans (*i.e.,* sulfated and sialylated) were analyzed from PSM by chemically liberating *O*-glycans using the ammonium carbamate method[17], followed by glycan enrichment and purification by cotton HILIC[28]; anionic glycans were then fractionated using WAX[10, 22] and finally, MALDI-TOF MS (Method S1). Cotton HILIC afforded satisfactory recovery of unmodified native *O*-glycans from PSM after elution with water. Furthermore, after cotton HILIC enrichment, WAX separation successfully fractionated sulfated *O*-glycans from neutral *O*-glycans (Fig 2.1A). Anionic WAX fractionation of sulfated *O*-glycans was achieved by first eluting neutral *O*-glycans with 1% AcOH in 50% MeCN. Sulfated *O*-glycans were then eluted using 1% NH_4OH in 5% MeCN (pH 10.5) solution. Despite the repeatability of the method, poor sulfated *O*-glycan detection was observed compared to that of neutral *O*-glycans. Thus, this limitation warrants a more efficient glycan enrichment protocol to effectively concentrate abundant neutral glycans, especially trace sulfated glycans, from glycoproteins. Therefore, instead of using cotton HILIC, we explored the practicality of using the Glycoblotting method for sulfoglycomics workflow.

2.3.2 Glycoblotting-based Analysis of Sulfated *O*-glycans. We previously established that the efficiency of *N*-glycan recovery from whole human serum using the Glycoblotting method was 72.5±5.0%[15]. Thus, we evaluated the feasibility of the Glycoblotting protocol as a complementary glycan purification, enrichment, methylation, and labeling method for sulfated

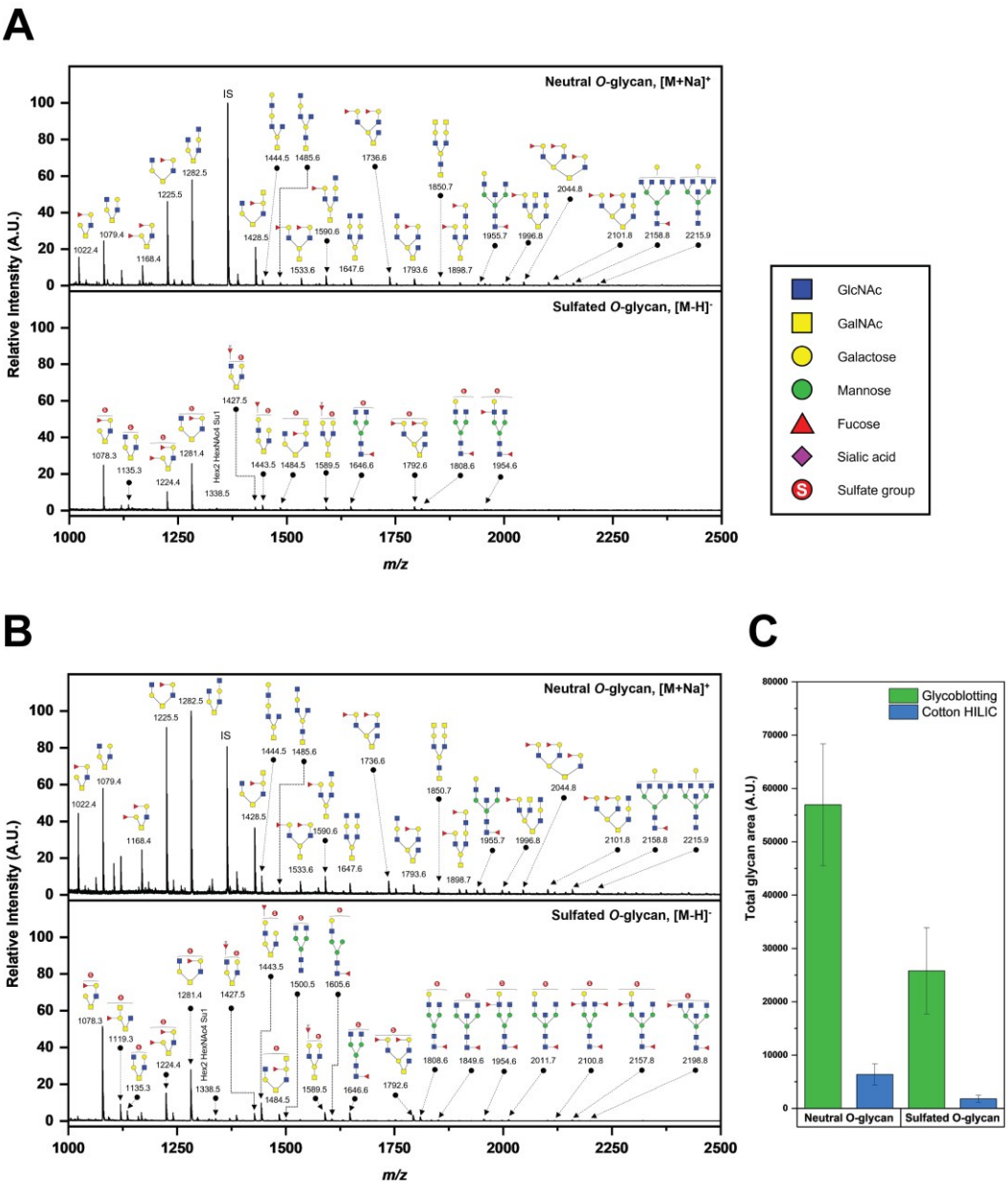


Figure 2.1. MALDI-TOF MS profile of *O*-glycans from porcine stomach mucin (PSM) enriched using (**A**) cotton HILIC and (**B**) Glycoblotting method. (**C**) Comparison of qualitative estimation of glycan enrichment between cotton HILIC and Glycoblotting method for both neutral and sulfated *O*-glycans from PSM. The total glycan area was calculated from the summation of all glycan peak areas identified from each spectrum shown in (A) and (B). *O*-glycan structures were inferred using the experimental *m/z* from the Expasy GlycoMod Tool and Glyconnect Database of the Swiss Institute of Bioinformatics and GlycoWorkbench. Fucose residues explicitly linked to the *O*-glycan backbone (e.g., Le^y) were based on the MS/MS analysis of porcine stomach mucin *O*-glycans deduced by Cheng, P. *et al.*[29]. Absolute quantification and linkage-specific MS/MS structural analysis of each glycan species were not attempted and are beyond the scope of this work. I.S. indicates internal standard, Hexa-*N*-acetylchitohexaose (200 pmol)

glycan analysis.

Figure 2.1B shows the MS profile of *O*-glycans from PSM enriched using the Glycoblotting method. The neutral *O*-glycan profile observed was of similar molecular ion peaks detected on the cotton HILIC enrichment. However, using Glycoblotting enrichment, the relative glycan peak intensities were higher compared to the cotton HILIC-enriched glycans with respect to the internal standard, Hexa-*N*-acetylchitohexaose (200 pmol). Also, Glycoblotting enrichment allowed the detection of more sulfated glycan species. To qualitatively estimate the glycan enrichment of cotton HILIC and Glycoblotting, the total glycan peak area was calculated from the summation of *O*-glycan peak areas identified in each spectrum (Fig 2.1C). From the total glycan peak areas, cotton HILIC enrichment was approximately 10% of that of the Glycoblotting method for neutral and sulfated glycans. This indicates that the Glycoblotting method is more suitable for enriching trace sulfated glycans in glycoproteins.

It is worth noting that during MALDI-TOF MS analysis, not only sulfated *O*-glycan species were detected but also sulfated *N*-glycans, especially in the high-mass region (*m/z* 1400–2500). This is because of the ammonium carbamate method used to chemically liberate *O*-glycans through nonreductive β-elimination, which minimizes peeling reactions and significant deletion of acid-labile sialic acid moieties as previously reported by Miura, Y. *et al.*[17]. Furthermore, in previous studies using WAX separation, sulfated glycans were eluted at low pH using either 0.1% FA (pH 2.5), 0.25% FA (pH 1.5), or 0.1% TFA in MeCN solution[10]. However, in this study, WAX fractionation of sulfated glycans was achieved using 1% NH_4OH in 5% MeCN (pH 10.5) solution. The 1% NH_4OH in 5% MeCN eluent solution was considered over the previously reported acidic eluent solutions[10], primarily because at low pH, partial hydrolysis of BOA from the BOA-labeled glycans was observed (data not shown). Partially hydrolyzed glycan species during MALDI-TOF analysis make the MS profiles complicated to analyze.

2.3.3 Glycoblotting-based Analysis of Sulfated-Sialyl *N*-glycans. Bovine thyroglobulin (BTG) contains 14 possible *N*-glycosylation sites, 13 of which have been confirmed to be *N*-glycosylated[25]. Besides having sialylated *N*-glycans, BTG has been reported to contain sulfated *N*-glycans[11, 30] and sulfated-sialyl *N*-glycans[25, 31]. Since BTG contains a mixture of sialylated, sulfated, and sulfated-sialyl *N*-glycans, we employed BTG to evaluate the feasibility of the Glycoblotting method to discriminate between sulfated *N*-glycans and sialylated *N*-glycans. To analyze sulfated *N*-glycans in the presence of sialylated *N*-glycans, MS-based sulfoglycomics workflow generally employs permethylation to methyl-esterify the carboxyl functional group of sialic acid, apart from the *O*-methylation of the hydroxyl moieties within the glycan backbone. Previous works have shown that sulfated *N*-glycans can be permethylated using the modified Hakomori method without losing sulfate groups[14]. Also, sulfated-sialyl glycans were permethylated efficiently using the NaOH/DMSO slurry method at low temperatures. Furthermore, the sulfate groups remained unmethylated under these reaction conditions[7, 10]. On the other hand, the Glycoblotting method utilizes the 3-methyl-1-*p*-tolyltriazene (MTT) to methyl esterify the carboxyl group of sialic acid while the hydroxyl groups remain unmodified. The methyl esterification protocol using MTT also allows the mono-methylation of phosphates, while sulfates remain unmethylated[16, 32].

High mannose (Man5, Man6, Man7, Man8, and Man9) and mono-, di-, and tri-sialylated multi-antennary complex *N*-glycans were detected in the neutral *N*-glycan fraction after the WAX separation of BOA-labeled *N*-glycans from BTG (Fig. 2.2). The acidic *N*-glycan fraction also showed a mixture of mono-sulfated Man5 and complex *N*-glycans and mono- and di-sialylated multi-antennary complex *N*-glycans with single sulfation. Interestingly, several neutral *N*-glycan peaks (*1362.6, 1914.8, 2219.9, 2381.9, 2525.0, 2585.0, 2747.1, 2890.2, 3052.2, 3274.2, and 3417.4 m/z*) have corresponding sulfated *N*-glycan peaks in the MS profile of the acidic glycan fraction, having a mass difference of *m/z 56* [$SO_3 - Na^+ - H^+$], indicating the presence of a sulfate moiety. Furthermore, the results showed the possible presence of terminal sialyl-LacNAc, sulfo-

LacNAc, and sulfo sialyl-LacNAc epitopes on the *N*-glycans expressed on BTG. Lastly, we demonstrated that the Glycoblotting method using MTT effectively discriminates sulfated *N*-glycans from sialylated *N*-glycans through the utilization of the on-bead methyl esterification method. The on-bead methyl esterification protocol also addresses the concomitant sample losses due to several clean-up steps accompanying the permethylation method, resulting in higher glycan recovery[16].

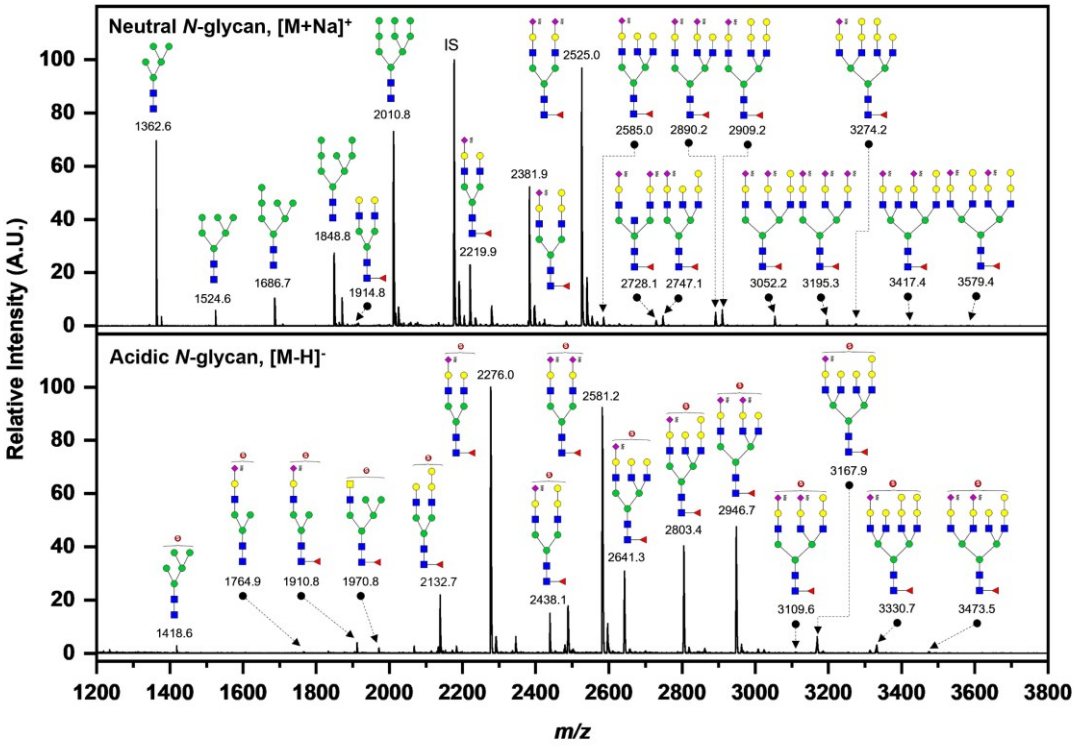


Figure 2.2 MALDI-TOF MS profile of BOA-labeled *N*-glycans from bovine thyroglobulin (BTG) enriched using Glycoblotting method then fractionated using WAX. *N*-glycan structures were inferred using the experimental *m/z* from the Expasy GlycoMod Tool and Glyconnect Database of the Swiss Institute of Bioinformatics, GlycoWorkbench, and previous reports[25, 30, 31]. Linkage-specific MS/MS structural analysis of each glycan species was not attempted. I.S. indicates internal standard, disialyloctasaccharide, SGP-10 (1000 pmol).

2.3.4 Chicken Egg White Ovomucoid Contains Phosphorylated High-mannose *N*-glycan. Ovomucoid (OVM) is one of the major glycoproteins found in chicken egg white and is heavily glycosylated. Around 25-30% of the entire molecule is composed of oligosaccharides. Bisecting multi-antennary *N*-glycans with and without terminal galactose has been reported to be the major glycan structure in ovomucoid. Acidic *N*-glycans such as mono- and di-sialylated bi-antennary *N*-glycans were also found together with sulfated *N*-glycans[33–35]. In addition to the previously reported *N*-glycans expressed in ovomucoid, in this study, we report for the first time the detection of a mono-phosphorylated high mannose (Man6) *N*-glycan from ovomucoid. Whether glycans are phosphorylated or sulfated, it is difficult to distinguish based on MALDI-TOF MS alone primarily because phosphate ($–HPO_3^-$, 79.9799 *m/z*) and sulfate ($–SO_3^-$, 80.0632 *m/z*) moieties have almost similar masses[4, 11]. Previously, we showed that the on-bead methyl esterification incorporated with the Glycoblotting method using MTT as a methylating agent successfully *O*-methylated phosphate groups on phospholipids[32]. In this study, we demonstrated that phosphate groups on *N*-glycans could also be methylated using MTT, which allowed us to discriminate phosphate moieties from sulfate groups on *N*-glycans.

Accompanying the sulfated *N*-glycans detected was a phosphorylated high-mannose *N*-glycan (1594.7 *m/z*, $[M – H]^-$) (Fig 2.3A, sulfated *N*-glycan spectra). A neutral *N*-glycan peak at 1524.7 *m/z*, $[M+Na]^+$ which corresponds to Man6 without phosphorylation, was also present (Fig 2.3A, neutral glycan spectra), having a mass difference of *m/z* 70 [$HPO_3Me – Na^+ – H^+$], indicating the presence of a phosphate group. MALDI-TOF/TOF MS analysis of the phosphorylated Man6 in positive ion reflectron mode was performed on its corresponding di-sodiated molecular ion adduct $[M+2Na – H]^+$ having 1640.6 *m/z* (Fig 2.3B). Fragment ions bearing phosphate groups were detected in the TOF/TOF MS spectra. The fragment ion peak at *m/z* 140.67 corresponds to the loss of the mono-methylated phosphate group [$PO_3Me+2Na^+$]. The fragment ions at 462.93, 949.42, and 1111.27 *m/z* correspond to [$PO_3Me+Hex_2+2Na^+$], [$PO_3Me+Hex_5+2Na^+$], and [$PO_3Me+Hex_6+2Na^+$], respectively. In addition, neutral losses corresponding to Hex, HexHex,

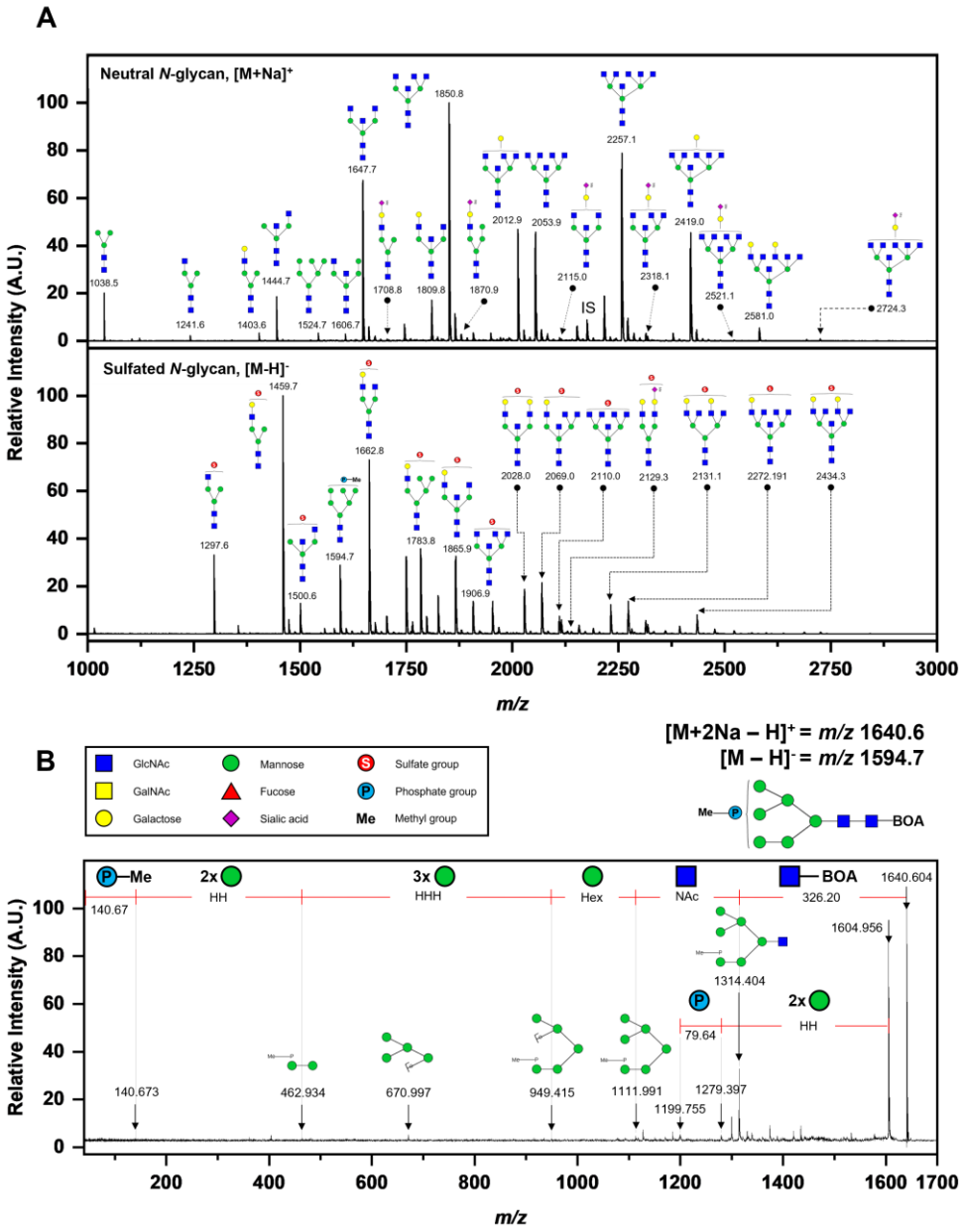


Figure 2.3. MALDI-TOF MS profile of BOA-labeled *N*-glycans from chicken egg white ovomucoid (OVM) enriched using Glycoblotting method then fractionated using WAX (**A**). MALDI LIFT TOF/TOF MS spectra of BOA-labeled mono-methylated phosphorylated Man6 *N*-glycan (1640.6 *m/z*) from OVM in positive ion mode having an $[M+2Na - H]^+$ ion adduct (**B**). *N*-glycan structures were inferred using the experimental *m/z* from the Expasy GlycoMod Tool and Glyconnect Database of the Swiss Institute of Bioinformatics. Fragment ion assignment was manually annotated using Bruker FlexAnalysis 3.0 and GlycoWorkbench. I.S. indicates internal standard, disialyloctasaccharide, SGP-10 (1000 pmol).

HexHexHex, and HexNAc were observed. The reducing-end GlcNAc terminus labeled with BOA was also observed, corresponding to a neutral loss of 326.20 *m/z*. Furthermore, *N*-glycans from

ovomucoid were digested with 1 µL of 0.8 U/µL alkaline phosphatase (Nippon Gene) in BAP buffer (100 mM Tris-HCl, 1 mM $MgSO_4$, pH 8.0) at 37°C for 24 hr. Alkaline phosphatase digestion was performed to confirm that the molecular ion peak at 1594.7 *m/z*, $[M - H]^-$ is a mono-phosphorylated high-mannose *N*-glycan. It can be seen in Figure 2.4 that the peak intensity of 1594.7 *m/z* diminished after phosphatase treatment, which suggests that it is indeed a phosphorylated *N*-glycan. Furthermore, on-bead methyl-esterification using MTT allowed us to differentiate between *N*-glycan isomers that carry either sulfate or phosphate groups since MTT

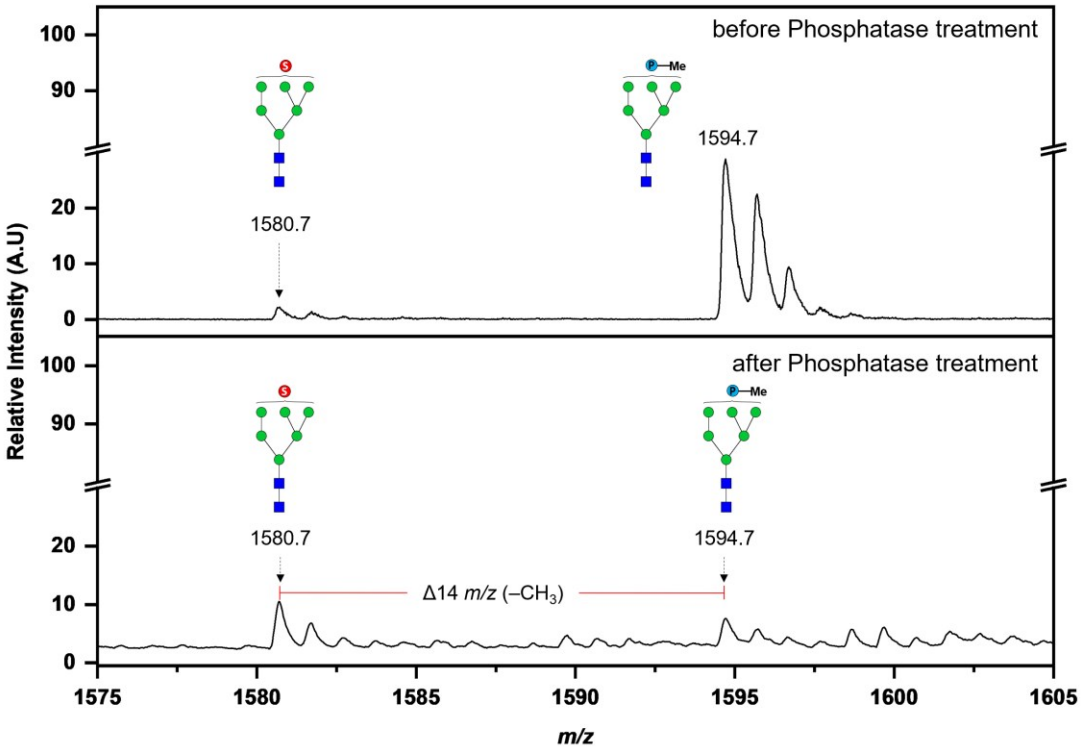


Figure 2.4. MALDI-TOF MS profile of BOA-labeled mono-phosphorylated high-mannose N-glycan from chicken egg white ovomucoid (OVM) before and after alkaline phosphatase treatment. N-glycan structures were inferred using the experimental m/z from the Expasy GlycoMod Tool and Glyconnect Database of the Swiss Institute of Bioinformatics and GlycoWorkbench.

can methylate phosphate, which gives a mass difference of 14 *m/z* corresponding to a methyl group. The molecular ion peak at 1580.7 *m/z* was identified as mono-sulfated Man6, which was

14 *m/z* less than 1594.7 *m/z.* Thus, this is the first report that uses MTT as a methylating agent to simultaneously detect and differentiate sulfated *N*-glycan from its corresponding isobaric phosphorylated *N*-glycan isomer using MALDI-TOF MS. Lastly, the workflow described for the simultaneous analysis of sulfated (Sul) and phosphorylated (Pho) glycans using MALDI-TOF MS is now referred to as Sulphoglycomics.

2.3.5 Glycoblotting-based Sulphoglycomics of Human Saliva and Chicken Egg White. Human saliva is a highly complex secretion containing mucus, glycoproteins, and electrolytes from the parotid, submaxillary, sublingual, and submandibular glands. Glycoproteins from saliva are highly glycosylated, like mucins (MUC5B and MUC7), salivary agglutinin (gp340), and secretory immunoglobulins (IgA). The primary functions of saliva are lubrication, digestion, and the regulation of oral microflora. However, increasing evidence has shown that peptides and glycoproteins in the saliva play an essential role in the first line of oral defense against pathogens. Furthermore, salivary glycosylation is highly sensitive to changes in biological conditions and has been implicated in various diseases[36–40]. With this, we explored the feasibility of the Glycoblotting-based sulphoglycomics workflow to elucidate the glycan profile of a highly complex sample, such as saliva.

Unstimulated saliva samples were collected from a healthy volunteer by passive drooling into a 50 mL Falcon tube with 2 mL PBS and a protease inhibitor cocktail. The saliva was clarified and concentrated using an Amicon Ultra-15 Centrifugal Filter device with Ultracel 10k membrane (Merck Millipore, Cork, IRL). Protein quantification of stock saliva was performed using Micro BCA Protein Assay Kit (Thermo Scientific, Rockford, IL, USA). *N*- and *O*-glycans were subsequently released and Glycoblotted. Sulfated glycans were separated using WAX and analyzed using MALDI-TOF MS.

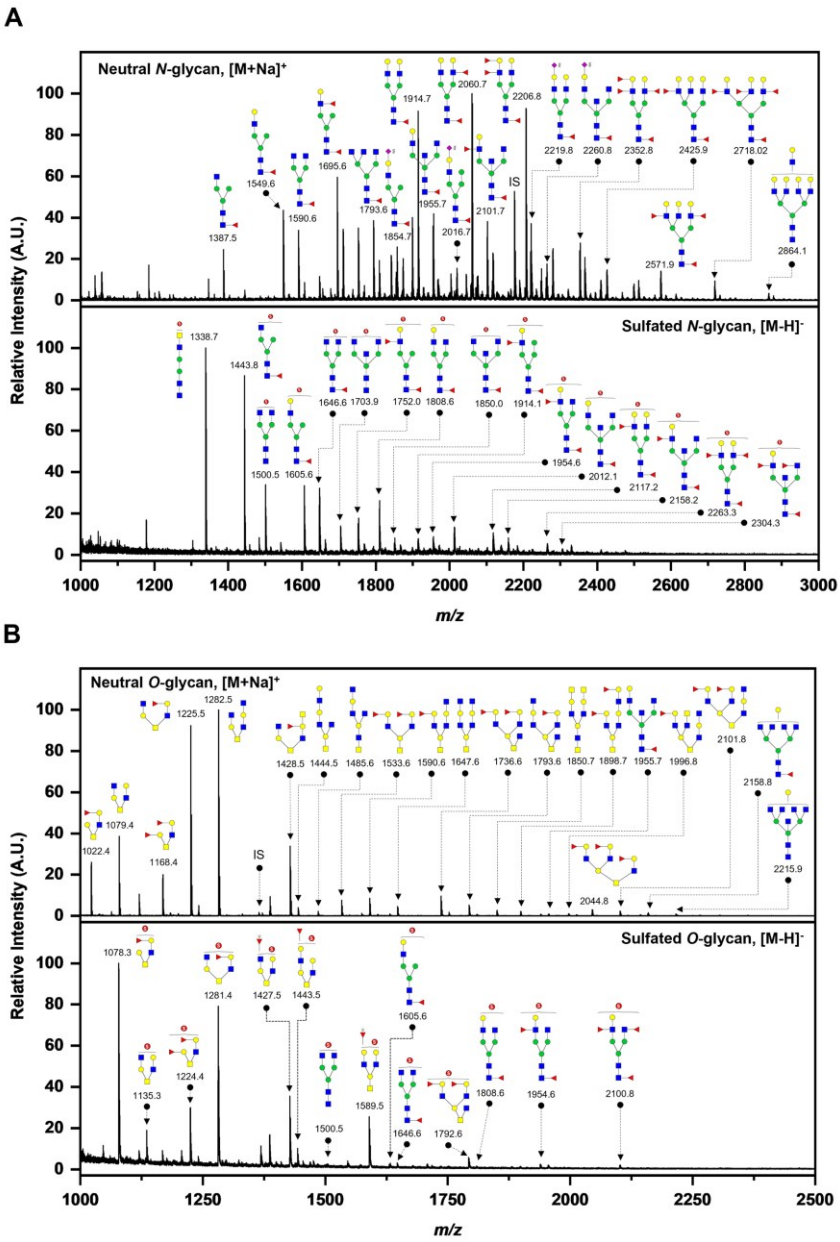


Figure 2.5. MALDI-TOF MS profile of BOA-labeled neutral and sulfated glycans from human saliva. (**A**) *N*-glycans and (**B**) *O*-glycans. *N*- and *O*-glycan structures were inferred using the experimental *m/z* from the Expasy GlycoMod Tool and Glyconnect Database of the Swiss Institute of Bioinformatics and GlycoWorkbench. Linkage-specific MS/MS structural analysis of each glycan species was not attempted to verify the structure. I.S. indicates internal standard, disialyloctasaccharide, SGP-10 (1000 pmol).

As previously described, most *N*-glycans observed in human saliva of healthy individuals have a bi-antennary structure with single or multiple fucosylations. Core and branch-fucosylation have also been reported, and terminal epitopes such as Le^x, Le^y, and H-antigens exist. Core fucosylated sialyl *N*-glycans were also observed [39, 41]. Conversely, mono-sulfated complex *N*-glycans with mono-, di-, or tri-fucosylation were detected in the acidic *N*-glycan fraction of the WAX separation. *N*-glycans with possible sulfated GlcNAc, sulfated LacNAc, and sulfated Le^x terminal epitopes were inferred from the *m/z* values of the observed molecular ion peaks (Fig 2.5A).

Similarly, Everest-Dass *et al*. reported that human salivary *O*-glycans are fucosylated and that approximately 54% of these *O*-glycans carry the blood group H-antigen. In addition, sulfation was exclusively observed in *O*-glycans, which may be attributed to salivary mucins, which are typically modified with sulfate moieties[36]. Interestingly, sulfated *O*-glycans have been reported to be almost undetectable in the saliva of healthy individuals. However, the expression of sulfated *O*-glycans in the saliva of patients with ocular rosacea is indicative of disease progression. Sulfated *O*-glycans detected have typical *core 2* structures, in which the sulfate group is generally attached to GlcNAc[37]. In addition to sulfated *O*-glycans, sulfated *N*-glycans were also observed in the high-mass region of the MS profile, owing to the ammonium carbamate method used to liberate the *O*-glycans chemically (Fig 2.5B).

N-glycans from chicken egg white were also profiled using the Glycoblotting-based sulphoglycomics approach. Avian egg white is interesting because it plays a vital role in embryonic development by acting as a mechanism for thermal and mechanical stress and providing the necessary water source and nutrients during development[42, 43]. In addition, egg white glycoproteins such as lysozyme, ovotransferrin, ovomucoid, and ovostatin elicit antimicrobial activity, thus protecting the embryo from infection. These proteins are heavily glycosylated; therefore, glycosylation may be vital during infection[44–46].

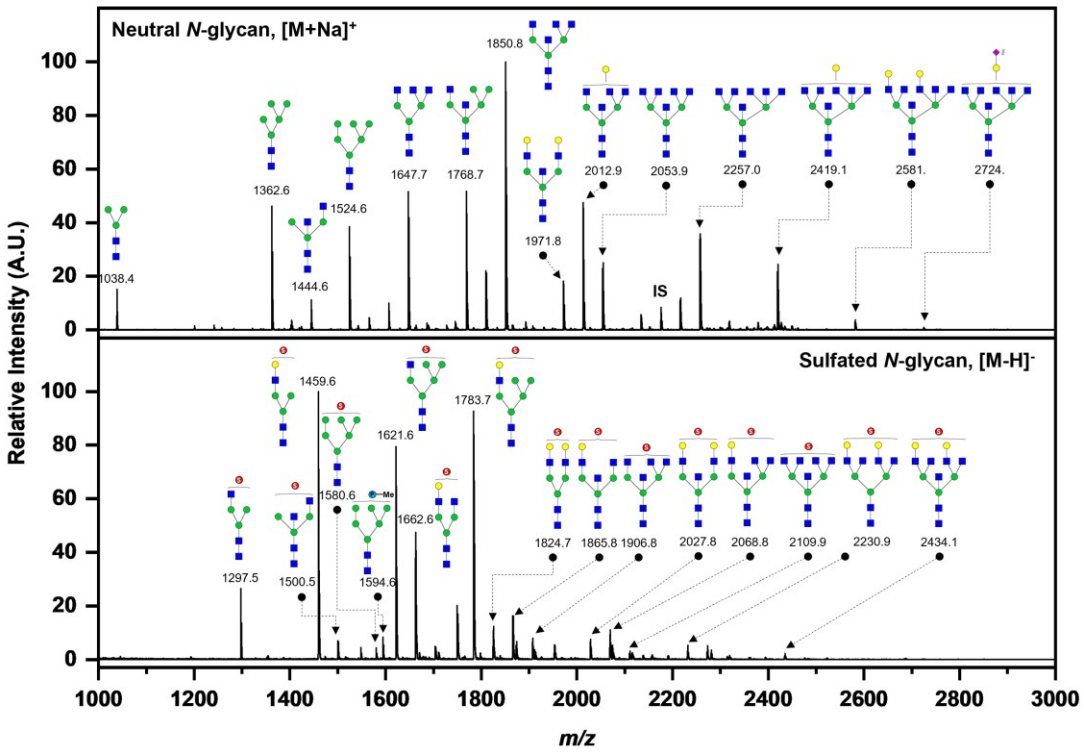


Figure 2.6. MALDI-TOF MS profile of BOA-labeled *N*-glycans from chicken egg white (CEW) enriched using Glycoblotting method then fractionated using WAX. *N*-glycan structures were inferred using the experimental *m/z* from the Expasy GlycoMod Tool and Glyconnect Database of the Swiss Institute of Bioinformatics and GlycoWorkbench. No attempt was made to verify the *N*-glycan structures using a linkage-specific MS/MS analysis. I.S. indicates internal standard, disialyloctasaccharide, SGP-10 (1000 pmol).

The neutral *N*-glycan profile showed (Fig. 2.6) that the major glycan structures inferred from the masses of the monoisotopic ions $[M+Na]^+$ observed are high-mannose and bisecting multi-antennary complex *N*-glycans. Unlike human salivary *N*-glycans, no fucosylation was observed in egg white *N*-glycans. On the other hand, sulfated *N*-glycan structures were mono-sulfated high-mannose, hybrid, and complex *N*-glycans with GlcNAc and LacNAc terminal epitopes (Fig. 2.6 and S2). Two isomeric high-mannose structures with sulfate (1580.6 m/z) and phosphate (1594.6 m/z) groups were also detected, similar to the ovomucoid shown in Figure 2.4.

Furthermore, chicken egg white's ammonium carbamate-released glycan profiles (neutral and acidic) were remarkably similar to its' PNGase F-released *N*-glycan profiles (data not shown). Previous work noted that significant amounts of *N*-glycans were also concurrently released from glycoproteins upon treatment with ammonium carbamate[17]. Direct comparison to the *N*-glycan profile (PNGase F-released) may suggest that the monoisotopic peaks observed in the ammonium carbamate-released glycan profiles of chicken egg white are probably *N*-glycans.

We have shown that trace sulfated and phosphorylated *N*- and *O*-glycans present in complex biological samples can be detected using MALDI-TOF MS with an efficient complementary enrichment technique like the Glycoblotting method with a seamless methylation protocol and a charge-based separation technique, such as WAX. Finally, the anionic glycan expression profiles of human saliva and chicken egg whites provide insights into the importance of sulfated[47] and phosphorylated[48] glycans as critical receptors in the propagation of the influenza virus. Especially both species (humans and chickens) are essential hosts in understanding influenza infections.

2.4 Conclusion

To address the current challenges in the analysis of sulfated glycans in complex biological samples, here we described a straightforward seamless workflow utilizing the advantages of the Glycoblotting method as a complementary enrichment and methylation technique for MALDI-TOF MS-based sulphoglycomics. We have shown that Glycoblotting provides an efficient glycan enrichment platform, and together with its on-bead methyl esterification using MTT, we have addressed the three main hurdles in sulphoglycomics – trace abundance, sample loss, and the presence of sialic acid. Glycoblottings' on-bead methyl esterification step proved useful in discriminating sulfated glycans from sialylated glycans and differentiating isomeric glycans bearing either sulfate or phosphate groups. Thus, this workflow offers efficient enrichment and detection of trace sulfated and phosphorylated glycans. We believe this approach would

significantly contribute to simplifying the current MALDI-TOF MS-based sulphoglycomics workflow.

2.5 References

1. Chen, J. Y., Huang, H. H., Yu, S. Y., Wu, S. J., Kannagi, R., & Khoo, K. H. (2018). Concerted mass spectrometry-based glycomic approach for precision mapping of sulfo sialylated N-glycans on human peripheral blood mononuclear cells and lymphocytes. *Glycobiology* 28, 9–20

2. Song, W., Zhou, X., Benktander, J. D., Gaunitz, S., Zou, G., Wang, Z., Novotny, M. V., & Jacobson, S. C. (2019). In-Depth Compositional and Structural Characterization of N-Glycans Derived from Human Urinary Exosomes. *Anal Chem* 91, 13528–13537

3. Wang, J. R., Gao, W. N., Grimm, R., Jiang, S., Liang, Y., Ye, H., Li, Z. G., Yau, L. F., Huang, H., Liu, J., Jiang, M., Meng, Q., Tong, T. T., Huang, H. H., Lee, S., Zeng, X., Liu, L., & Jiang, Z. H. (2017). A method to identify trace sulfated IgG N-glycans as biomarkers for rheumatoid arthritis. *Nat Commun* 8, 1–13

4. Yamada, K., Suzuki, K., Hirohata, Y., & Kinoshita, M. (2020). Analysis of Minor Acidic N-Glycans in Human Serum. *J Proteome Res* 19, 3033–3043

5. Kawashima, H., & Fukuda, M. (2012). Sulfated glycans control lymphocyte homing. *Ann N Y Acad Sci* 1253, 112–121

6. Scudder, P., Tang, P. W., Hounsell, E. F., Mehmet, H., Feizi, T., & Lawson, A. M. (1986). Isolation and characterization of sulphated oligosaccharides released from bovine corneal keratan sulphate by the action of endo-β-galactosidase. *Eur J Biochem* 157, 365–373

7. Ijuin, T., Kitajima, K., Song, Y., Kitazume, S., Inoue, S., Haslam, S., Morris, H., Dell, A., & Inoue, Y. (1996). Isolation and identification of novel sulfated and nonsulfated oligosialyl glycosphingolipids from sea urchin sperm. *Glycoconj J* 9, 401–413

8. Khoo, K. H., & Yu, S. Y. (2010). Mass spectrometric analysis of sulfated N- and O-glycans. *Methods Enzymol* 478, 3–26

9. Cheng, C. W., Chou, C. C., Hsieh, H. W., Tu, Z., Lin, C. H., Nycholat, C., Fukuda, M., & Khoo, K. H. (2015). Efficient Mapping of Sulfated Glycotopes by Negative Ion Mode nanoLC-MS/MS-Based Sulfoglycomic Analysis of Permethylated Glycans. *Anal Chem* 87, 6380–6388

10. Yu, S. Y., Wu, S. W., Hsiao, H. H., & Khoo, K. H. (2009). Enabling techniques and strategic workflow for sulfoglycomics based on mass spectrometry mapping and sequencing of permethylated sulfated glycans. *Glycobiology* 19, 1136–1149

11. Yamada, K., Kayahara, H., Kinoshita, M., & Suzuki, S. (2018). Simultaneous Analysis of Sulfated and Phosphorylated Glycans by Serotonin-Immobilized Column Enrichment and Hydrophilic Interaction Chromatography. *Anal Chem* 90, 8387–8395

12. Shajahan, A., Supekar, N. T., Chapla, D., Heiss, C., Moremen, K. W., & Azadi, P. (2020). Simplifying Glycan Profiling through a High-Throughput Micropermethylation Strategy. *SLAS Technol* 25, 367–379

13. Costello, C. E., Contado-Miller, J. M., & Cipollo, J. F. (2007). A Glycomics Platform for the Analysis of Permethylated Oligosaccharide Alditols. *J Am Soc Mass Spectrom* 18, 1799–1812

14. Dell, A., Rogers, M. E., Thomas-Oates, J. E., Huckerby, T. N., Sanderson, P. N., & Nieduszynski, I. A. (1988). Fast-atom-bombardment mass-spectrometric strategies for sequencing sulfated oligosaccharides. *Carbohydr Res* 179, 7

15. Furukawa, J. I., Shinohara, Y., Kuramoto, H., Miura, Y., Shimaoka, H., Kurogochi, M., Nakano, M., & Nishimura, S. I. (2008). Comprehensive approach to structural and functional glycomics based on chemoselective glycoblotting and sequential tag conversion. *Anal Chem* 80, 1094–1101

16. Amano, M., & Nishimura, S. I. (2010). Large-Scale Glycomics for Discovering Cancer-Associated N-Glycans by Integrating Glycoblotting and Mass Spectrometry. *Methods Enzymol* 478, 109–125

17. Miura, Y., Kato, K., Takegawa, Y., Kurogochi, M., Furukawa, J. I., Shinohara, Y., Nagahori, N., Amano, M., Hinou, H., & Nishimura, S. I. (2010). Glycoblotting-assisted O-glycomics: Ammonium carbamate allows for highly efficient o-glycan release from glycoproteins. *Anal Chem* 82, 10021–10029

18. Sanes, J. T., Hinou, H., Lee, Y. C., & Nishimura, S. I. (2019). Glycoblotting of Egg White Reveals Diverse N-Glycan Expression in Quail Species. *J Agric Food Chem* 67, 531–540

19. Hirose, K., Amano, M., Hashimoto, R., Lee, Y. C., & Nishimura, S. I. (2011). Insight into glycan diversity and evolutionary lineage based on comparative avio-N-glycomics and sialic acid analysis of 88 egg whites of galloanserae. *Biochemistry* 50, 4757–4774

20. Gizaw, S. T., Ohashi, T., Tanaka, M., Hinou, H., & Nishimura, S. I. (2016). Glycoblotting method allows for rapid and efficient glycome profiling of human Alzheimer's disease brain, serum and cerebrospinal fluid towards potential biomarker discovery. *Biochim Biophys Acta Gen Subj* 1860, 1716–1727

21. Miura, Y., Shinohara, Y., Furukawa, J. I., Nagahori, N., & Nishimura, S. I. (2007). Rapid and simple solid-phase esterification of sialic acid residues for quantitative glycomics by mass spectrometry. *Chemistry - A European Journal* 13, 4797–4804

22. Yu, S.-Y., Snovida, S., & Khoo, K.-H. (2020). Permethylation and Microfractionation of Sulfated Glycans for MS Analysis. *Bio Protoc* 10, 1–13

23. Hinou, H. (2019). Aniline derivative/DHB/alkali metal matrices for reflectron mode MALDI-TOF and TOF/TOF MS analysis of unmodified sialylated oligosaccharides and glycopeptides. *Int J Mass Spectrom* 443, 109–115

24. Barada, E., & Hinou, H. (2022). BOA/DHB/Na: An Efficient UV-MALDI Matrix for High-Sensitivity and Auto-Tagging Glycomics. *Int J Mol Sci* 23, 12510

25. Kuo, C. W., Guu, S. Y., & Khoo, K. H. (2018). Distinctive and Complementary MS 2 Fragmentation Characteristics for Identification of Sulfated Sialylated N-Glycopeptides by nanoLC-MS/MS Workflow. *J Am Soc Mass Spectrom* 29, 1166–1178

26. Ceroni, A., Maass, K., Geyer, H., Geyer, R., Dell, A., & Haslam, S. M. (2008). GlycoWorkbench: A tool for the computer-assisted annotation of mass spectra of glycans. *J Proteome Res* 7, 1650–1659

27. Cooper, C. A., Gasteiger, E., & Packer, N. H. (2001). GlycoMod-A software tool for determining glycosylation compositions from mass spectro-metric data. *Proteomics* 1, 340–349

28. Selman, M. H. J., Hemayatkar, M., Deelder, A. M., & Wuhrer, M. (2011). Cotton HILIC SPE microtips for microscale purification and enrichment of glycans and glycopeptides. *Anal Chem* 83, 2492–2499

29. Cheng, P. F., Snovida, S., Ho, M. Y., Cheng, C. W., Wu, A. M., & Khoo, K. H. (2013). Increasing the depth of mass spectrometry-based glycomic coverage by additional dimensions of sulfoglycomics and target analysis of permethylated glycans. *Anal Bioanal Chem* 405, 6683–6695

30. Lei, M., Novotny, M. V., & Mechref, Y. (2010). Sequential Enrichment of Sulfated Glycans by Strong Anion-Exchange Chromatography Prior to Mass Spectrometric Measurements. *J Am Soc Mass Spectrom* 21, 348–357

31. Wheeler, S. F., & Harvey, D. J. (2001). Extension of the In-Gel Release Method for Structural Analysis of Neutral and Sialylated N-Linked Glycans to the Analysis of Sulfated Glycans: Application to the Glycans from Bovine Thyroid-Stimulating Hormone. *Anal Biochem* 296, 92–100

32. Furukawa, T., Hinou, H., Takeda, S., Chiba, H., Nishimura, S. I., & Hui, S. P. (2017). An Efficient Glycoblotting-Based Analysis of Oxidized Lipids in Liposomes and a Lipoprotein. *ChemBioChem* 18, 1903–1909

33. Yamashita, K., Tachibana, Y., Hitoi, A., & Kobata, A. (1984). Sialic acid-containing sugar chains of hen ovalbumin and ovomucoid. *Carbohydr Res* 130, 271–288

34. Yamashita, K., Kamerling, J. P., & Kobata, A. (1983). Structural studies of the sugar chains of hen ovomucoid. Evidence indicating that they are formed mainly by the alternate biosynthetic pathway of asparagine-linked sugar chains. *Journal of Biological Chemistry* 258, 3099–3106

35. Yamashita, K., Kamerling, J. P., & Kobata, A. (1982). Structural study of the carbohydrate moiety of hen ovomucoid. Occurrence of a series of pentaantennary complex-type asparagine-linked sugar chains. *Journal of Biological Chemistry* 257, 12809–12814

36. Everest-Dass, A. v., Jin, D., Thaysen-Andersen, M., Nevalainen, H., Kolarich, D., & Packer, N. H. (2012). Comparative structural analysis of the glycosylation of salivary and buccal cell proteins: Innate protection against infection by Candida albicans. *Glycobiology* 22, 1465–1479

37. Vieira, A. C., An, H. J., Ozcan, S., Kim, J. H., Lebrilla, C. B., & Mannis, M. J. (2012). Glycomic analysis of tear and saliva in ocular rosacea patients: The search for a biomarker. *Ocular Surface* 10, 184–192

38. Kozak, R. P., Urbanowicz, P. A., Punyadeera, C., Reiding, K. R., Jansen, B. C., Royle, L., Spencer, D. I., Fernandes, D. L., & Wuhrer, M. (2016). Variation of human salivary O-glycome. *PLoS One* 11, 1–15

39. Sinevici, N., Mittermayr, S., Davey, G. P., Bones, J., & O'Sullivan, J. (2019). Salivary N-glycosylation as a biomarker of oral cancer: A pilot study. *Glycobiology* 29, 726–734

40. Cross, B. W., & Ruhl, S. (2018). Glycan recognition at the saliva – oral microbiome interface. *Cell Immunol* 333, 19–33

41. Guile, G. R., Harvey, D. J., O'Donnell, N., Powell, A. K., Hunter, A. P., Zamze, S., Fernandes, D. L., Dwek, R. A., & Wing, D. R. (1998). Identification of highly fucosylated N-linked oligosaccharides from the human parotid gland. *Eur J Biochem* 258, 623–656

42. Krkavcová, E., Kreisinger, J., Hyánková, L., Hyršl, P., & Javůrková, V. (2018). The hidden function of egg white antimicrobials: Egg weightdependent effects of avidin on avian embryo survival and hatchling phenotype. *Biol Open* 7,

43. Qiu, N., Ma, M., Cai, Z., Jin, Y., Huang, X., Huang, Q., & Sun, S. (2012). Proteomic analysis of egg white proteins during the early phase of embryonic development. *J Proteomics* 75, 1895–1905

44. Asperger, A., Marx, K., Albers, C., Molin, L., Pinato, O., & Silliker, C. (2015). Low Abundant N-linked Glycosylation in Hen Egg White Lysozyme Is Localized at Nonconsensus Sites. *Journal of Proteome* 14, 2633–2641

45. Offengenden, M., Fentabil, M. A., & Wu, J. (2011). N-glycosylation of ovomucin from hen egg white. *Glycoconj J* 28, 113–123

46. Haltiwanger, R. S., & Lowe, J. B. (2004). Role of Glycosylation in Development. *Annu. Rev. Biochem.* 73, 491–537

47. Ichimiya, T., Okamatsu, M., Kinoshita, T., Kobayashi, D., Ichii, O., Yamamoto, N., Sakoda, Y., Kida, H., Kawashima, H., Yamamoto, K., Takase-Yoden, S., & Nishihara, S. (2021). Sulfated glycans containing NeuAcα2-3Gal facilitate the propagation of human H1N1 influenza A viruses in eggs. *Virology* 562, 29–39

48. Byrd-Leotis, L., Jia, N., Dutta, S., Trost, J. F., Gao, C., Cummings, S. F., Braulke, T., Müller-Loennies, S., Heimburg-Molinaro, J., Steinhauer, D. A., & Cummings, R. D. (2019). Influenza binds phosphorylated glycans from human lung. *Sci Adv* 5, 1–10

49. Karlsson, N. G., Nordman, H., Karlsson, H., Carlstedt, I., & Hansson, G. C. (1997). Glycosylation differences between pig gastric mucin populations : a comparative study of the neutral oligosaccharides using mass spectrometry. *Biochem. J* 326, 911–917

50. Hwang, H. S., Kim, B. S., Park, H., Park, H. Y., Choi, H. D., & Kim, H. H. (2014). Type and branched pattern of N-glycans and their structural effect on the chicken egg allergen ovotransferrin: A comparison with ovomucoid. *Glycoconj J* 31, 41–50

51. Harvey, D. J., Wing, D. R., Kuster, B., & Wilson, I. B. H. (2000). Composition of N-linked carbohydrates from ovalbumin and co-purified glycoproteins. *J Am Soc Mass Spectrom* 11, 564–571

52. Lattova, E., Perreault, H., & Krokhin, O. (2004). Matrix-assisted laser desorption/ionization tandem mass spectrometry and post-source decay fragmentation study of phenylhydrazones of N-linked oligosaccharides from ovalbumin. *J Am Soc Mass Spectrom* 15, 725–735

53. da Silva, M. L. C., Stubbs, H., Tamura, T., & Rice, K. (1995). 1H NMR Characterization of a Hen Ovalbumin Tyrosinamide N-Linked Oligosaccharide Library. *Arch Biochem Biophys* 318, 465–475

54. Yamashita, K., Ueda, I., & Kobatas, A. (1983). *Sulfated Asparagine-linked Sugar Chains of Hen Egg Albumin**

55. Zhu, J., Lin, Y.-H., Dingess, K. A., Mank, M., Stahl, B., & Heck, A. J. R. (2020). Quantitative Longitudinal Inventory of the N-Glycoproteome of Human Milk from a Single Donor Reveals the Highly Variable Repertoire and Dynamic Site-Specific Changes. *Cite This: J. Proteome Res* 19,

56. Sumiyoshi, W., Nakakita, S. I., Miyanishi, N., & Hirabayashi, J. (2009). Strategic glycan elution map for the production of human-type N-linked oligosaccharides: The case of hen egg yolk and white. *Biosci Biotechnol Biochem* 73, 543–551

2.6 Supplementary Information

Method S1: Direct analysis of Native Sulfated *O*-glycans from Porcine Stomach Mucin.

Cotton HILIC Micro-SPE of Native *O*-glycans from Porcine Stomach Mucin (PSM)[24, 28]**.** Approximately 4 mg of cotton wool (100% absorbent cotton) was inserted into a 200 μL micropipette tip. The micro-SPE column was washed with 100 μL water five times and conditioned with 100 μL of 1% AcOH in 95% MeCN three times. The native *O*-glycans released from PSM were reconstituted with 40 μL MilliQ water. A 20 μL aliquot of the sample and 6 μL of internal standard (Hexa-*N*-acetylchitohexaose, 35 μM) were diluted with 150 μL of 1% AcOH in 95% MeCN and then loaded into the column. The sample was allowed to elute by gravity, and the collected eluate was reloaded back into the column; this step was done three times. The column was washed with 95% MeCN to remove unbound and hydrophobic contaminants. Native *O*-glycans were eluted with 100 μL water twice. The eluates were then dried in a SpeedVac and stored at -20ºC until use.

Anionic-glycan Separation of Native *O*-glycans from PSM using WAX. Fifty microliters of 100 mg/mL 3-aminopropyl silica gel suspension (1 mmol/mg, Tokyo Chemical Industry Co. Ltd.) were packed into a 200 μL micropipette tip with a cotton plug. The packed weak anion exchange (WAX) microcolumn was conditioned and washed sequentially with 100 μL water, MeCN, and 1% AcOH in 95% MeCN twice. After every conditioning and washing step, the column was centrifugated at 500 rpm for 2 mins. Native *O*-glycans were reconstituted with 20 μL water. A 5 μL sample aliquot was dissolved in 150 μL 1% AcOH in 95% MeCN and then loaded into the column. The sample was allowed to elute by gravity, and the collected eluate was reloaded back into the column; this step was done three times. The column was washed with 1% AcOH in 95% MeCN to remove unbound and hydrophobic contaminants. Native neutral *O*-glycans were eluted

with 1% AcOH in 50% MeCN, while sulfated and sialylated *O*-glycans were eluted with 1% NH_4OH in 5% MeCN (pH 10.5). The eluates were then dried in a SpeedVac and stored at -20ºC until use.

MALDI-TOF MS Analysis of Native *O*-glycans from PSM[24]**.** MALDI-TOF MS analysis of native *O*-glycans was achieved using the BOA/DHB/Na matrix in reflectron mode on both positive and negative ion acquisition mode. The matrix was prepared by mixing 60 μL of 1 M benzyloxyamine (BOA), 100 μL of 0.5 M DHB, and 10 μL of 0.5 M $NaHCO_3$ diluted to 1 mL with 50% MeCN. Matrix solutions (0.5 μL) were spotted on the MALDI target plate, and a 0.5 μL analyte solution was deposited on top of the matrix spots. The target plate was then incubated at 60ºC for 1 hr.

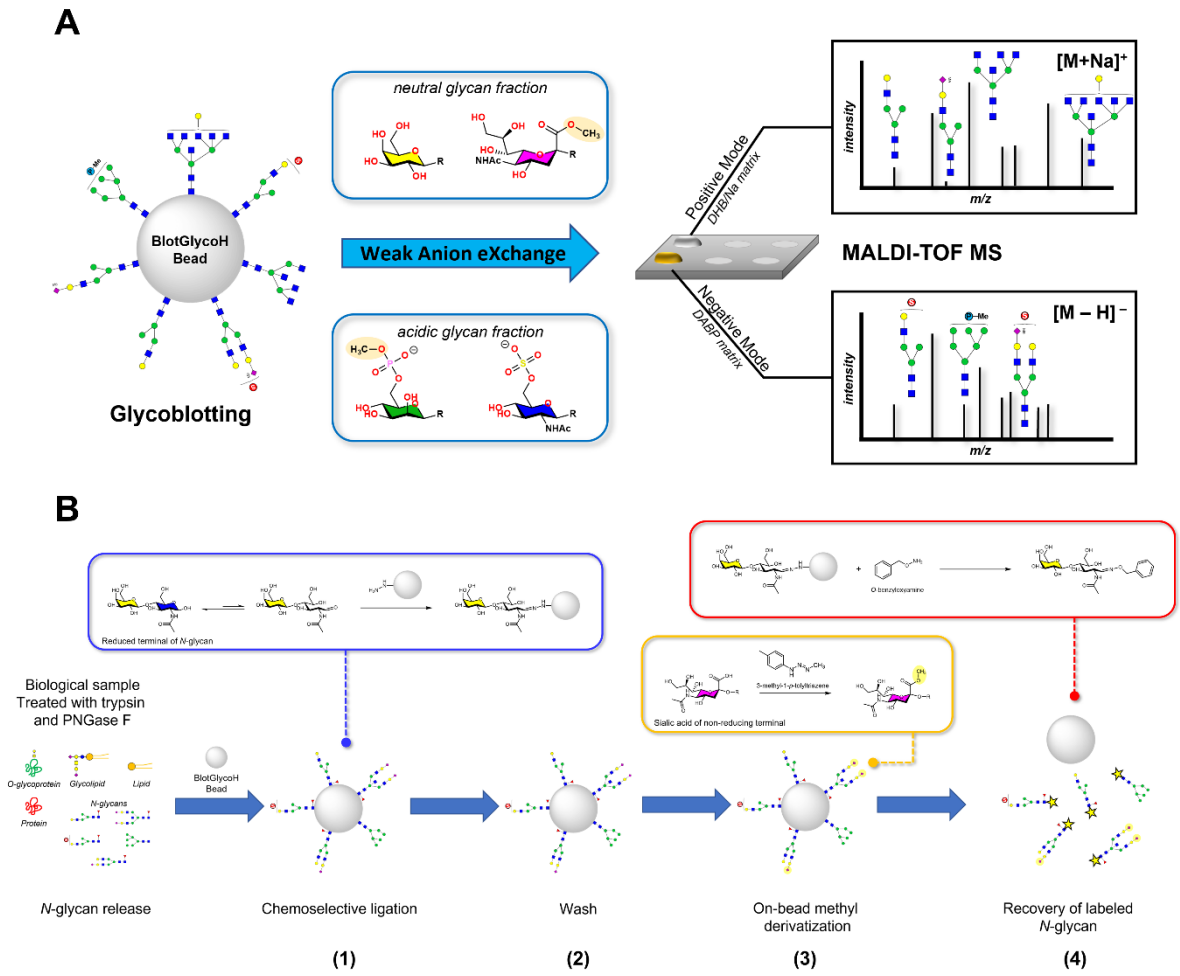


Figure S2.1. (**A**) Simplified schematic diagram of the Glycoblotting-based Sulphoglycomics workflow. (**B**) General Glycoblotting protocol for *N*-glycan analysis showing its seamless process starting from (1) chemoselective purification and enrichment of glycans by capturing reducing sugars onto the hydrazide-functionalized BlotGlycoH beads, (2) washing step to remove impurities, (3) on-bead methyl esterification of -COOH of sialic acid residues followed by (4) trans-iminization reaction labeling using benzyloxyamine (BOA) and subsequent recovery of BOA-labeled glycans.

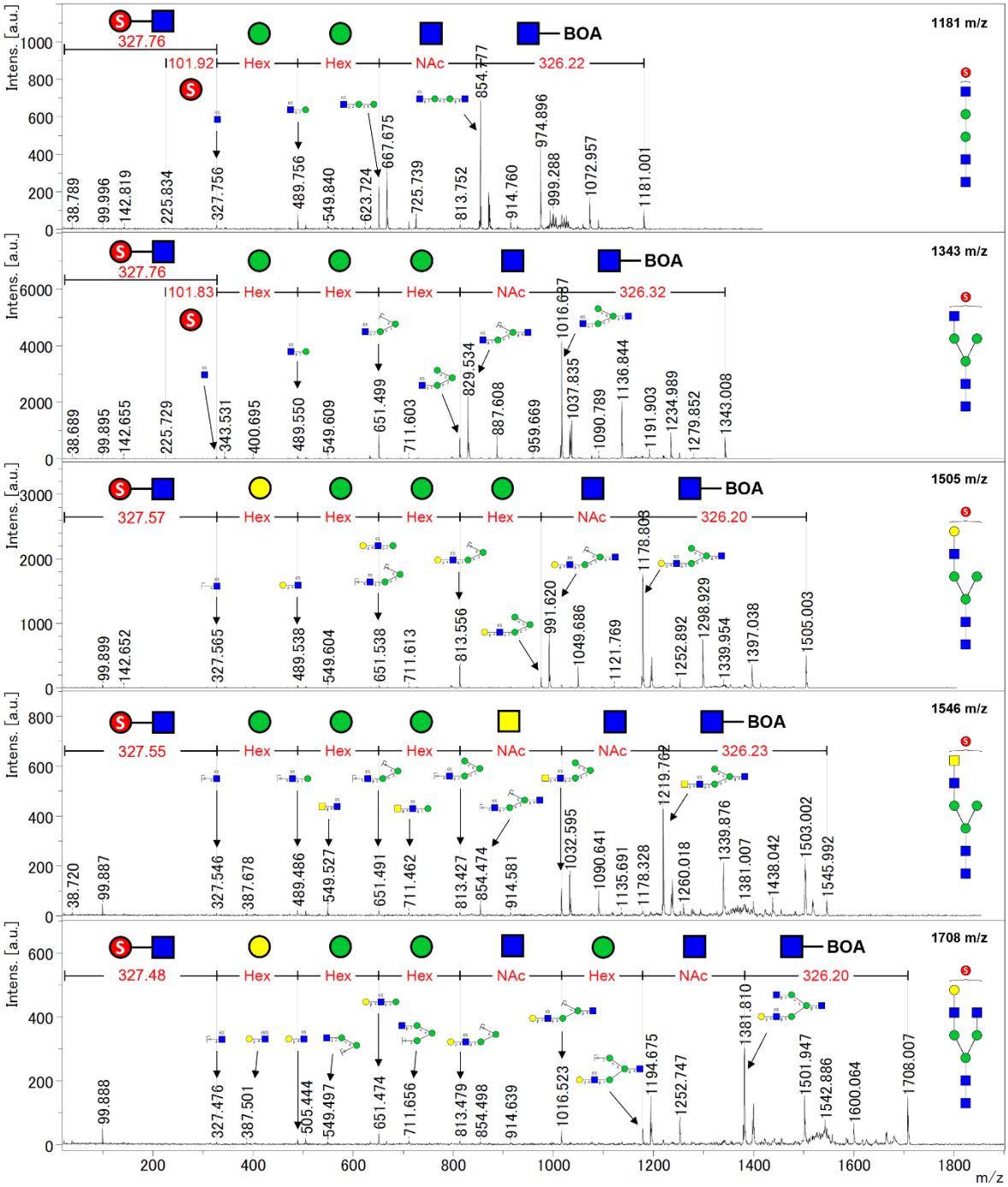
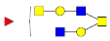
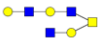
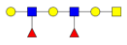
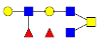
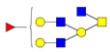
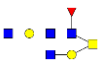
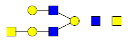
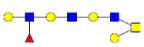
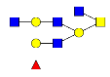
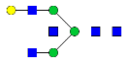


Figure S2.2. MALDI TOF/TOF MS profile of selected BOA-labeled sulfated *N*-glycans in positive ion mode having an $[M+2Na - H]^+$ ion adduct. *N*-glycan structures were inferred using the experimental *m/z* from the Expasy GlycoMod Tool and Glyconnect Database of the Swiss Institute of Bioinformatics. Fragment ion assignment was manually annotated using Bruker FlexAnalysis 3.0 and GlycoWorkbench.

Table S2.1. Neutral *O*-glycans from porcine stomach mucin (PSM) labeled with BOA.

Observed mass, m/z $[M+Na]^+$	**Calculated mass, m/z $[M+Na]^+$**	**Error ppm**	**Monosaccharide composition**	**O-glycan Probable Structures** [26, 27, 49]	**N-glycan Probable Structures** [26, 27]	**Glyconnect Database Links**
860.34	860.34	-0.2	Hex1 HexNAc2 dHex1			Glyconnect
876.35	876.33	-23.3	Hex2 HexNAc2			Glyconnect
917.38	917.36	-23.4	Hex1 HexNAc3			Glyconnect
1022.39	1022.39	-3.5	Hex2 HexNAc2 dHex1			Glyconnect
1063.45	1063.42	-28.0	Hex1 HexNAc3 dHex1			Glyconnect
1079.43	1079.41	-13.7	Hex2 HexNAc3			Glyconnect
1120.45	1120.44	-7.9	Hex1 HexNAc4			Glyconnect
1168.44	1168.45	-3.1	Hex2 HexNAc2 dHex2			Glyconnect
1225.47	1225.47	-1.5	Hex2 HexNAc3 dHex1			Glyconnect
1241.48	1241.46	-11.4	Hex3 HexNAc3			Glyconnect
1282.49	1282.49	-2.5	Hex2 HexNAc4			Glyconnect
1387.52	1387.52	0.1	Hex3 HexNAc3 dHex1			Glyconnect
1428.54	1428.55	7.2	Hex2 HexNAc4 dHex1			Glyconnect
1444.54	1444.54	4.4	Hex3 HexNAc4			Glyconnect
1485.55	1485.57	11.9	Hex2 HexNAc5			Glyconnect
1533.55	1533.58	17.4	Hex3 HexNAc3 dHex2			Glyconnect
1574.63	1574.61	-14.0	Hex2 HexNAc4 dHex2			Glyconnect
1590.58	1590.60	12.2	Hex3 HexNAc4 dHex1			Glyconnect
1606.60	1606.60	-3.5	Hex4 HexNAc4			Glyconnect
1631.62	1631.63	5.5	Hex2 HexNAc5 dHex1			Glyconnect
1647.60	1647.62	14.7	Hex3 HexNAc5			Glyconnect
1736.64	1736.66	12.7	Hex3 HexNAc4 dHex2			Glyconnect
1752.64	1752.65	9.8	Hex4 HexNAc4 dHex1			Glyconnect
1793.65	1793.68	18.3	Hex3 HexNAc5 dHex1			Glyconnect
1809.67	1809.68	1.1	Hex4 HexNAc5			Glyconnect

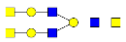
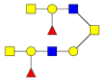
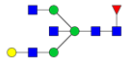
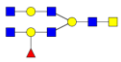
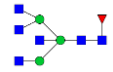
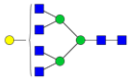
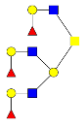
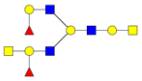
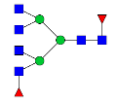
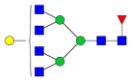
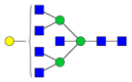
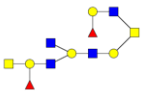
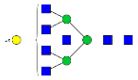
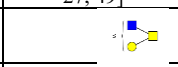
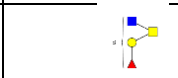
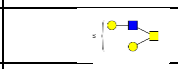
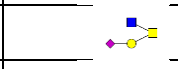
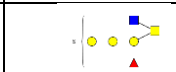
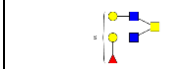
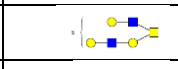
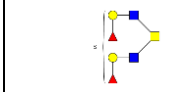
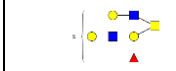
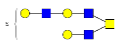
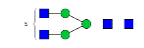
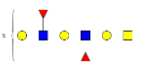
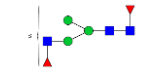
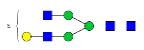
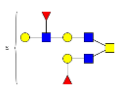
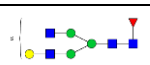
1850.67	1850.70	19.3	Hex3 HexNAc6			Glyconnect
1898.69	1898.71	12.7	Hex4 HexNAc4 dHex2			Glyconnect
1939.70	1939.74	1.1	Hex3 HexNAc5 dHex2			Glyconnect
1955.72	1955.73	9.7	Hex4 HexNAc5 dHex1			Glyconnect
1996.76	1996.76	2.9	Hex3 HexNAc6 dHex1			Glyconnect
2012.73	2012.76	11.7	Hex4 HexNAc6			Glyconnect
2044.75	2044.77	-27.6	Hex4 HexNAc4 dHex3			Glyconnect
2101.78	2101.76	7.0	Hex4 HexNAc5 dHex2			Glyconnect
2142.81	2142.82	1.8	Hex3 HexNAc6 dHex2			Glyconnect
2158.79	2158.81	10.6	Hex4 HexNAc6 dHex1			Glyconnect
2215.85	2215.84	-6.8	Hex4 HexNAc7			Glyconnect
2304.90	2304.87	-11.5	Hex4 HexNAc6 dHex2			Glyconnect
2362.03	2361.89	-59.1	Hex4 HexNAc7 dHex1			Glyconnect

Table S2.2. Sulfated and sialylated *O*-glycans from porcine stomach mucin (PSM) labeled with BOA enriched using cotton HILIC.

Observed mass, m/z $[M+2Na-H]^+$	Calculated mass, m/z $[M+2Na-H]^+$	Error ppm	Monosaccharide composition	O-glycan Probable Structures [26, 27, 49]	N-glycan Probable Structures [26, 27]	Glyconnect Database Links
816.30	816.22	-105.0	Hex1 HexNAc2 Su1			Glyconnect
962.37	962.28	-101.9	Hex1 HexNAc2 Su1 dHex1			Glyconnect
978.35	978.27	-81.3	Hex2 HexNAc2 Su1			Glyconnect
1027.44	1027.36	-85.1	Hex1 HexNAc2 NeuAc1			Glyconnect
1124.44	1124.33	-94.5	Hex2 HexNAc2 Su1 dHex1			Glyconnect
1165.45	1165.36	-79.5	Hex1 HexNAc3 Su1 dHex1			Glyconnect
1181.43	1181.35	-66.9	Hex2 HexNAc3 Su1			Glyconnect
1189.50	1189.41	-77.9	Hex2 HexNAc2 NeuAc1			Glyconnect
1270.50	1270.39	-88.4	Hex2 HexNAc2 Su1 dHex2			Glyconnect
1286.47	1286.38	-66.9	Hex3 HexNAc2 Su1 dHex1			Glyconnect
1327.52	1327.41	-84.5	Hex2 HexNAc3 Su1 dHex1			Glyconnect
1343.49	1343.40	-67.0	Hex3 HexNAc3 Su1			Glyconnect
1392.59	1392.49	-71.1	Hex2 HexNAc3 NeuAc1			Glyconnect
1473.57	1473.47	-71.9	Hex2 HexNAc3 Su1 dHex2			Glyconnect
1489.56	1489.46	-67.6	Hex3 HexNAc3 Su1 dHex1			Glyconnect
1505.52	1505.46	-40.5	Hex4 HexNAc3 Su1			
1530.59	1530.49	-69.4	Hex2 HexNAc4 Su1 dHex1			
1546.60	1546.48	-73.1	Hex3 HexNAc4 Su1			Glyconnect
1635.64	1635.52	-71.1	Hex3 HexNAc3 Su1 dHex2			Glyconnect
1692.65	1692.54	-65.8	Hex3 HexNAc4 Su1 dHex1			Glyconnect
1708.64	1708.54	-59.1	Hex4 HexNAc4 Su1			Glyconnect
1838.72	1838.60	-64.7	Hex3 HexNAc4 Su1 dHex2			Glyconnect
1854.71	1854.59	-61.4	Hex4 HexNAc4 Su1 dHex1			Glyconnect

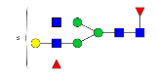
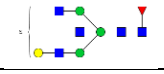
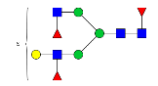
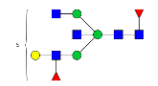
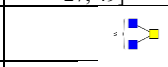
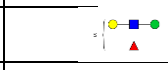
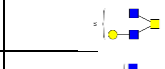
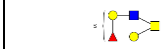
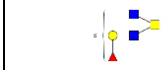
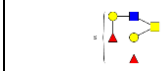
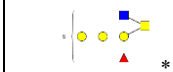
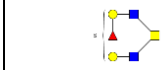
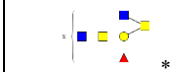
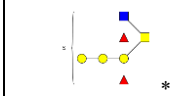
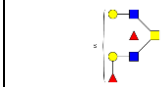
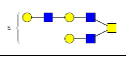
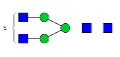
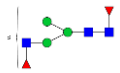
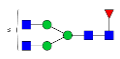
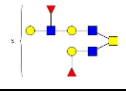
2000.79	2000.65	-71.7	Hex4 HexNAc4 Su1 dHex2			Glyconnect
2057.85	2057.63	-109.0	Hex4 HexNAc5 Su1 dHex1			Glyconnect
2146.87	2146.71	-74.8	Hex4 HexNAc4 Su1 dHex3			Glyconnect
2203.91	2203.69	-101.6	Hex1 HexNAc7 Su2 dHex2			Glyconnect

Table S2.3. Sulfated *O*-glycans from porcine stomach mucin (PSM) labeled with BOA enriched using Glycoblotting method.

Observed mass, m/z $[M-H]^-$	Calculated mass, m/z $[M-H]^-$	Error ppm	Monosaccharide composition	O-glycan Probable Structures [26, 27, 49]	N-glycan Probable Structures [26, 27]	Glyconnect Database Links
811.247	811.267	-30.5	HexNAc3 Su1			Glyconnect
875.278	875.271	1.4	Hex2 HexNAc1 dHex1 Su1			Glyconnect
916.314	916.298	12.3	Hex1 HexNAc2 dHex1 Su1			Glyconnect
932.314	932.293	17.4	Hex2 HexNAc2 Su1			Glyconnect
973.346	973.319	21.8	Hex1 HexNAc3 Su1			Glyconnect
1062.363	1062.356	2.1	Hex1 HexNAc2 dHex2 Su1			Glyconnect
1078.338	1078.351	-16.5	Hex2 HexNAc2 dHex1 Su1			Glyconnect
1119.371	1119.377	-10.5	Hex1 HexNAc3 dHex1 Su1			Glyconnect
1135.376	1135.372	-1.6	Hex2 HexNAc3 Su1			Glyconnect
1224.422	1224.409	6.7	Hex2 HexNAc2 dHex2 Su1			Glyconnect
1240.410	1240.404	1.0	Hex3 HexNAc2 dHex1 Su1	 *		Glyconnect
1281.446	1281.446	8.0	Hex2 HexNAc3 dHex1 Su1			Glyconnect
1297.447	1297.425	13.3	Hex3 HexNAc3 Su1			Glyconnect
1322.472	1322.457	7.7	Hex1 HexNAc4 dHex1 Su1	 *		Glyconnect
1386.479	1386.462	8.8	Hex3 HexNAc2 dHex2 Su1	 *		Glyconnect
1427.516	1427.488	16.3	Hex2 HexNAc3 dHex2 Su1			Glyconnect
1443.515	1443.483	18.9	Hex3 HexNAc3 dHex1 Su1			Glyconnect
1500.527	1500.505	11.5	Hex3 HexNAc4 Su1			Glyconnect
1589.579	1589.541	20.9	Hex3 HexNAc3 dHex2 Su1			Glyconnect
1605.566	1605.536	15.7	Hex4 HexNAc3 dHex1 Su1			Glyconnect
1646.603	1646.541	21.4	Hex3 HexNAc4 dHex1 Su1			Glyconnect
1792.651	1792.620	14.1	Hex3 HexNAc4 dHex2 Su1			Glyconnect

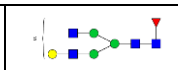
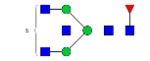
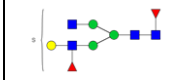
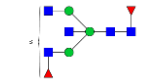
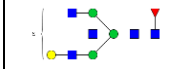
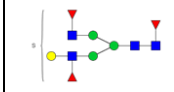
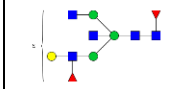
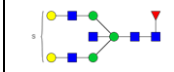
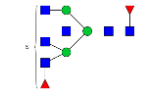
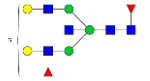
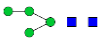
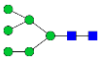
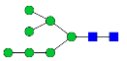
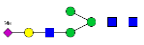
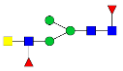
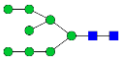
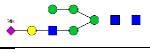
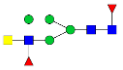
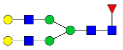
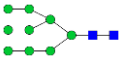
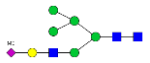
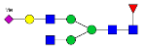
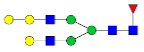
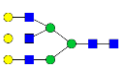
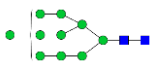
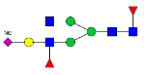
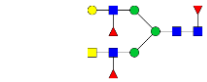
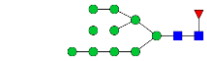
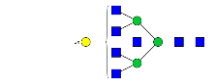
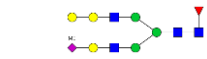
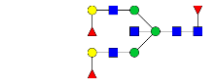
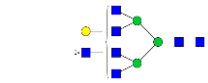
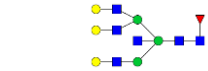
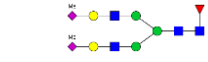
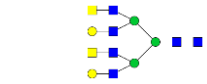
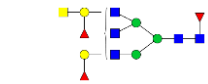
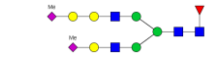
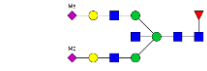
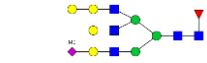
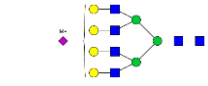
1808.657	1808.615	20.0	Hex4 HexNAc4 dHex1 Su1			Glyconnect
1849.711	1849.642	34.7	Hex3 HexNAc5 dHex1 Su1			Glyconnect
1954.696	1954.673	9.3	Hex4 HexNAc4 dHex2 Su1			Glyconnect
1995.739	1995.699	17.2	Hex3 HexNAc5 dHex2 Su1			
2011.744	2011.695	22.0	Hex4 HexNAc5 dHex1 Su1			Glyconnect
2100.782	2100.731	22.0	Hex4 HexNAc4 dHex3 Su1			Glyconnect
2157.799	2157.753	19.1	Hex4 HexNAc5 dHex2 Su1			Glyconnect
2173.782	2173.747	13.4	Hex5 HexNAc5 dHex1 Su1			
2198.829	2198.779	20.6	Hex3 HexNAc6 dHex2 Su1			Glyconnect
2319.883	2139.805	31.1	Hex5 HexNAc5 dHex2 Su1			

Table S2.4. Neutral *N*-glycans from bovine thyroglobulin (BTG) labeled with BOA.

Observed mass, m/z $[M+Na]^+$	Calculated mass, m/z $[M+Na]^+$	Error ppm	Monosaccharide composition	Probable Structure [25–27]	Glyconnect Database Links
1022.52	1022.39	121.3	Hex2 HexNAc2 dHex1		Glyconnect
1200.58	1200.44	110.0	Hex4 HexNAc2		Glyconnect
1362.62	1362.49	88.1	Hex2 + Man3 GlcNAc2		Glyconnect
1524.67	1524.54	80.0	Hex3 + Man3 GlcNAc2		Glyconnect
1686.71	1686.60	65.8	Hex4 + Man3 GlcNAc2		Glyconnect
1708.75	1708.63	77.8	Hex1 HexNAc1 NeuAc1 + Man3 GlcNAc2		Glyconnect
1736.77	1736.66	61.0	HexNAc2 dHex2 + Man3 GlcNAc2		Glyconnect
1848.76	1848.65	54.6	Hex5 + Man3 GlcNAc2		Glyconnect
1854.79	1854.69	62.0	Hex1 HexNAc1 dHex1 NeuAc1 + Man3 GlcNAc2		Glyconnect
1870.79	1870.68	62.5	Hex2 HexNAc1 NeuAc1 + Man3 GlcNAc2		Glyconnect
1898.81	1898.71	47.4	Hex1 HexNAc2 dHex2 + Man3 GlcNAc2		Glyconnect
1914.81	1914.71	50.7	Hex2 HexNAc2 dHex1 + Man3 GlcNAc2		Glyconnect
2010.80	2010.70	45.3	Hex6 + Man3 GlcNAc2		Glyconnect
2016.83	2016.74	47.1	Hex2 HexNAc1 dHex1 NeuAc1 + Man3 GlcNAc2		Glyconnect
2032.83	2032.73	53.1	Hex3 HexNAc1 NeuAc1 + Man3 GlcNAc2		Glyconnect
2057.86	2057.77	50.5	Hex1 HexNAc2 dHex1 NeuAc1 + Man3 GlcNAc2		Glyconnect
2073.87	2073.76	56.4	Hex2 HexNAc2 NeuAc1 + Man3 GlcNAc2		Glyconnect
2076.86	2076.76	46.7	Hex3 HexNAc2 dHex1 + Man3 GlcNAc2		Glyconnect
2114.85	2114.79	34.0	Hex1 HexNAc3 NeuAc1 + Man3 GlcNAc2		Glyconnect
2133.88	2133.78	41.2	Hex3 HexNAc3 + Man3 GlcNAc2		Glyconnect
2172.92	2172.76	72.3	Hex7 + Man3 GlcNAc2		Glyconnect
2203.93	2203.82	50.8	Hex1 HexNAc2 dHex2 NeuAc1 + Man3 GlcNAc2		Glyconnect

2216.00	2215.84	70.9	Hex1 HexNAc5 + Man3 GlcNAc2		Glyconnect
2219.92	2219.82	51.4	Hex2 HexNAc2 dHex1 NeuAc1 + Man3 GlcNAc2		Glyconnect
2247.93	2247.85	33.4	Hex1 HexNAc3 dHex3 + Man3 GlcNAc2		Glyconnect
2260.95	2260.85	50.4	Hex1 HexNAc3 dHex1 NeuAc1 + Man3 GlcNAc2		Glyconnect
2318.93	2318.81	45.7	Hex7 dHex1 + Man3 GlcNAc3		Glyconnect
2377.95	2377.89	24.8	Hex2 HexNAc5 + Man3 GlcNAc2		Glyconnect
2381.97	2381.87	46.2	Hex3 HexNAc2 dHex1 NeuAc1 + Man3 GlcNAc2		Glyconnect
2410.00	2409.90	36.5	Hex2 HexNAc3 dHex3 + Man3 GlcNAc2		Glyconnect
2418.97	2418.91	18.6	Hex1 HexNAc6 + Man3 GlcNAc2		Glyconnect
2423.01	2422.90	48.3	Hex2 HexNAc3 dHex1 NeuAc1 + Man3 GlcNAc2		Glyconnect
2483.02	2482.92	38.7	Hex3 HexNAc4 dHex1 + Man3 GlcNAc2		Glyconnect
2525.05	2524.93	56.2	Hex2 HexNAc2 dHex1 NeuAc2 + Man3 GlcNAc2		Glyconnect
2581.09	2580.97	44.9	Hex2 HexNAc6 + Man3 GlcNAc2		Glyconnect
2585.08	2584.95	51.8	Hex3 HexNAc3 dHex1 NeuAc1 + Man3 GlcNAc2		Glyconnect
2613.07	2612.98	32.5	Hex2 HexNAc4 dHex3 + Man3 GlcNAc2		Glyconnect
2661.10	2660.97	47.0	Hex5 HexNAc4 + Man3 GlcNAc2		Glyconnect
2687.07	2686.98	41.3	Hex3 HexNAc2 dHex1 NeuAc2 + Man3 GlcNAc2		Glyconnect
2728.13	2728.01	53.9	Hex2 HexNAc3 dHex1 NeuAc2 + Man3 GlcNAc2		Glyconnect
2747.13	2747.00	49.9	Hex4 HexNAc3 dHex1 NeuAc1 + Man3 GlcNAc2		Glyconnect
2804.14	2804.03	44.2	Hex4 HexNAc4 NeuAc1 + Man3 GlcNAc2		Glyconnect

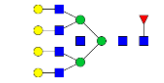
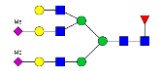
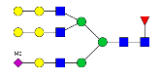
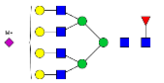
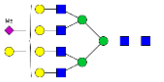
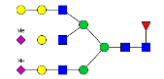
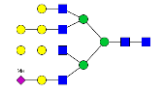
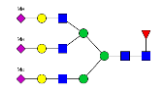
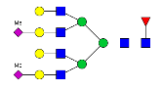
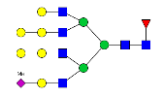
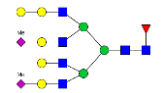
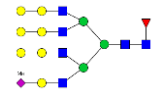
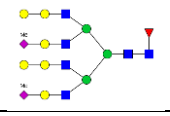
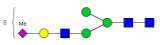
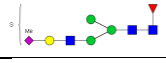
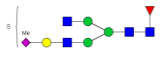
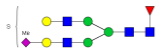
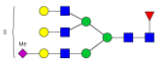
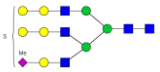
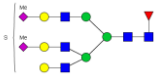
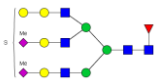
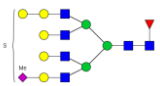
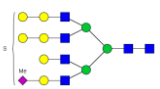
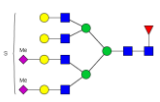
2848.05	2848.05	-2.8	Hex4 HexNAc5 dHex1 + Man3 GlcNAc2		Glyconnect
2890.19	2890.06	52.2	Hex3 HexNAc3 dHex1 NeuAc2 + Man3 GlcNAc2		Glyconnect
2909.18	2909.06	44.7	Hex5 HexNAc3 dHex1 NeuAc1 + Man3 GlcNAc2		Glyconnect
2950.20	2950.08	41.7	Hex4 HexNAc4 dHex1 NeuAc1 + Man3 GlcNAc2		Glyconnect
2966.23	2966.08	53.6	Hex5 HexNAc4 NeuAc1 + Man3 GlcNAc2		Glyconnect
3052.24	3052.11	49.5	Hex4 HexNAc3 dHex1 NeuAc2 + Man3 GlcNAc2		Glyconnect
3128.30	3128.13	57.9	Hex6 HexNAc4 NeuAc1 + Man3 GlcNAc2		Glyconnect
3195.31	3195.17	56.0	Hex3 HexNAc3 dHex1 NeuAc3 + Man3 GlcNAc2		Glyconnect
3255.33	3255.19	49.5	Hex4 HexNAc4 dHex1 NeuAc2 + Man3 GlcNAc2		Glyconnect
3274.28	3274.19	26.9	Hex6 HexNAc4 dHex1 NeuAc1 + Man3 GlcNAc2		Glyconnect
3417.41	3417.25	54.1	Hex5 HexNAc4 dHex1 NeuAc2 + Man3 GlcNAc2		
3435.38	3436.24	56.2	Hex7 HexNAc4 dHex1 NeuAc1 + Man3 GlcNAc2		Glyconnect
3578.40	3579.30	21.8	Hex6 HexNAc4 dHex1 NeuAc2 + Man3 GlcNAc2		Glyconnect

Table S2.5. Sulfated *N*-glycans from bovine thyroglobulin (BTG) labeled with BOA.

Observed mass, m/z $[M-H]^-$	Calculated mass, m/z $[M-H]^-$	Error ppm	Monosaccharide composition	Probable Structure [25–27, 31]	Glyconnect Database Links
1418.63	1418.45	118.6	Hex2 Su1 + Man3 GlcNAc2		
1764.90	1764.59	182.0	Hex1 HexNAc1 NeuAc1 Su1 + Man3 GlcNAc2		
1910.87	1910.65	123.1	Hex1 HexNAc1 dHex1 NeuAc1 Su1 + Man3 GlcNAc2		Glyconnect
1970.84	1970.67	85.4	Hex2 HexNAc2 dHex1 Su1 + Man3 GlcNAc2		Glyconnect
2066.87	2066.66	99.3	Hex6 Su1 + Man3 GlcNAc2		
2113.96	2113.73	113.2	Hex1 HexNAc2 dHex1 NeuAc1 Su1 + Man3 GlcNAc2		Glyconnect
2132.98	2132.72	117.3	Hex3 HexNAc2 dHex1 Su1 + Man3 GlcNAc2		Glyconnect
2170.94	2170.75	91.3	Hex1 HexNAc3 NeuAc1 Su1 + Man3 GlcNAc2		Glyconnect
2276.06	2275.78	127.1	Hex2 HexNAc2 dHex1 NeuAc1 Su1 + Man3 GlcNAc2		Glyconnect
2438.11	2437.83	117.4	Hex3 HexNAc2 dHex1 NeuAc1 Su1 + Man3 GlcNAc2		Glyconnect
2581.25	2580.89	150.0	Hex2 HexNAc2 dHex1 NeuAc2 Su1 + Man3 GlcNAc2		Glyconnect
2596.15	2595.90	92.5	Hex3 HexNAc5 Su1 + Man3 GlcNAc2		
2641.29	2640.91	146.2	Hex3 HexNAc3 dHex1 NeuAc1 Su1 + Man3 GlcNAc2		
2803.47	2802.96	182.7	Hex4 HexNAc3 dHex1 NeuAc1 Su1 + Man3 GlcNAc2		
2818.11	2818.96	-299.1	Hex5 HexNAc3 NeuAc1 Su1 + Man3 GlcNAc2		
2861.13	2860.93	67.9	Hex10 dHex1 Su1 + Man3 GlcNAc2		
2946.73	2946.02	248.2	Hex3 HexNAc3 dHex1 NeuAc2 Su1 + Man3 GlcNAc2		Glyconnect
2961.18	2961.03	47.0	Hex4 HexNAc6 Su1 + Man3 GlcNAc2		
3108.70	3108.08	209.8	Hex4 HexNAc3 dHex1 NeuAc2 Su1 + Man3 GlcNAc2		Glyconnect
3167.98	3168.10	-34.0	Hex5 HexNAc4 dHex1 NeuAc1 Su1 + Man3 GlcNAc2		
3184.17	3184.09	27.4	Hex6 HexNAc4 NeuAc1 Su1 + Man3 GlcNAc2		
3311.33	3311.15	59.9	Hex4 HexNAc4 dHex1 NeuAc2 Su1 + Man3 GlcNAc2		

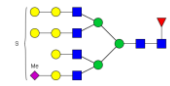
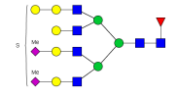
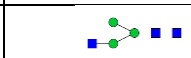
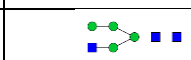
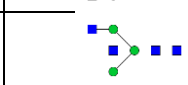
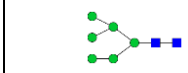
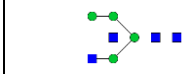
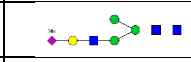
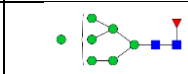
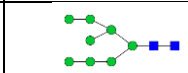
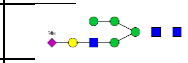
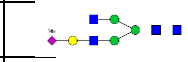
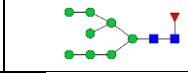
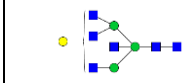
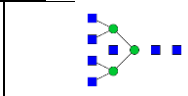
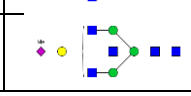
3330.70	3330.15	48.1	Hex6 HexNAc4 dHex1 Su1 + Man3 GlcNAc2		
3474.65	3473.21	-168.6	HexNAc8 dHex1 NeuAc2 Su1 + Man3 GlcNAc2		

Table S2.6. Neutral *N*-glycans from chicken ovomucoid (OVM) labeled with BOA.

Observed mass, m/z $[M+Na]^+$	Calculated mass, m/z $[M+Na]^+$	Error ppm	Monosaccharide composition	Probable Structure [26, 27, 50–53]	Glyconnect Database Links
876.42	876.33	95.6	Hex2 HexNAc2		Glyconnect
1038.51	1038.39	119.5	Hex3 HexNAc2		Glyconnect
1241.61	1241.46	118.1	HexNAc1 + Man3 GlcNAc2		Glyconnect
1403.68	1403.52	113.3	Hex1 HexNAc1 + Man3 GlcNAc2		Glyconnect
1444.71	1444.54	112.3	HexNAc2 + Man3 GlcNAc2		Glyconnect
1524.73	1524.54	123.1	Hex3 + Man3 GlcNAc2		Glyconnect
1606.77	1606.60	106.0	Hex1 HexNAc2 + Man3 GlcNAc2		Glyconnect
1647.79	1647.62	103.4	HexNAc3 + Man3 GlcNAc2		Glyconnect
1708.80	1708.63	100.9	Hex1 HexNAc1 NeuAc1 + Man3 GlcNAc2		Glyconnect
1727.83	1727.62	118.0	Hex3 HexNAc1 + Man3 GlcNAc2		Glyconnect
1768.83	1768.65	99.7	Hex2 HexNAc2 + Man3 GlcNAc2		Glyconnect
1809.86	1809.68	100.0	Hex1 HexNAc3 + Man3 GlcNAc2		Glyconnect
1832.86	1832.65	112.1	Hex4 dHex1 + Man3 GlcNAc2		Glyconnect
1848.88	1848.65	125.8	Hex5 + Man3 GlcNAc2		Glyconnect
1850.89	1850.70	101.7	HexNAc4 + Man3 GlcNAc2		Glyconnect
1870.91	1870.68	120.7	Hex2 HexNAc1 NeuAc1 + Man3 GlcNAc2		Glyconnect
1911.90	1911.71	99.8	Hex1 HexNAc2 NeuAc1 + Man3 GlcNAc2		Glyconnect
1930.91	1930.70	104.9	Hex3 HexNAc2 + Man3 GlcNAc2		Glyconnect
1994.93	1994.71	110.0	Hex5 dHex1 + Man3 GlcNAc2		Glyconnect
2012.96	2012.76	101.3	Hex1 HexNAc4 + Man3 GlcNAc2		Glyconnect
2054.00	2053.78	104.3	HexNAc5 + Man3 GlcNAc2		Glyconnect
2115.06	2114.79	130.8	Hex1 HexNAc3 NeuAc1 + Man3 GlcNAc2		Glyconnect

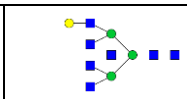
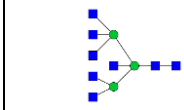
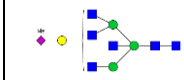
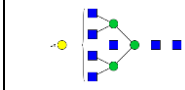
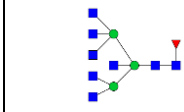
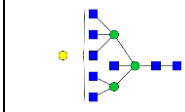
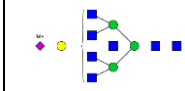
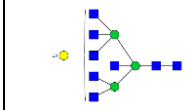
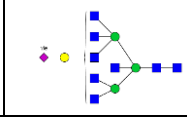
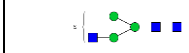
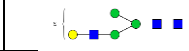
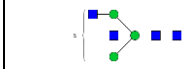
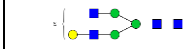
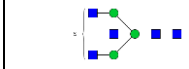
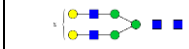
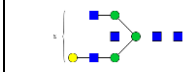
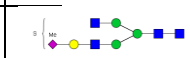
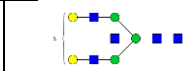
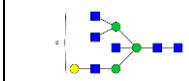
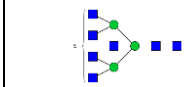
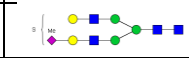
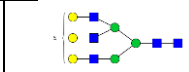
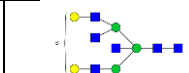
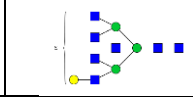
2216.06	2215.84	102.2	Hex1 HexNAc5 + Man3 GlcNAc2		Glyconnect
2257.11	2256.86	110.2	HexNAc6 + Man3 GlcNAc2		Glyconnect
2318.12	2317.87	108.5	Hex1 HexNAc4 NeuAc1 + Man3 GlcNAc2		Glyconnect
2378.16	2377.89	114.7	Hex2 HexNAc5 + Man3 GlcNAc2		Glyconnect
2403.17	2402.92	105.5	HexNAc6 dHex1 + Man3 GlcNAc2		Glyconnect
2419.19	2418.91	112.4	Hex1 HexNAc6 + Man3 GlcNAc2		Glyconnect
2521.19	2520.95	97.3	Hex1 HexNAc5 NeuAc1 + Man3 GlcNAc2		Glyconnect
2581.25	2580.97	111.2	Hex2 HexNAc6 + Man3 GlcNAc2		Glyconnect
2724.33	2724.03	112.7	Hex1 HexNAc6 NeuAc1 + Man3 GlcNAc2		

Table S2.7. Sulfated *N*-glycans from chicken ovomucoid (OVM) labeled with BOA.

Observed mass, m/z $[M - H]^-$	Calculated mass, m/z $[M - H]^-$	Error ppm	Monosaccharide composition	Probable Structure [26, 27, 33, 54]	Glyconnect Database Links
1297.60	1297.43	129.6	HexNAc1 Su1 + Man3 GlcNAc2		Glyconnect
1459.67	1459.48	129.6	Hex1 HexNAc1 Su1 + Man3 GlcNAc2		
1500.69	1500.50	118.1	HexNAc2 Su1 + Man3 GlcNAc2		Glyconnect
1580.69	1580.50	115.9	Hex3 Su1 + Man3 GlcNAc2		Glyconnect
1594.72	1594.53	124.9	Hex3 PMe1 + Man3 GlcNAc2		Glyconnect
1621.75	1621.53	133.9	Hex2 HexNAc1 Su1 + Man3 GlcNAc2		
1662.77	1662.56	125.2	Hex1 HexNAc2 Su1 + Man3 GlcNAc2		Glyconnect
1703.78	1703.58	111.1	HexNAc3 Su1 + Man3 GlcNAc2		Glyconnect
1783.83	1783.58	136.4	Hex3 HexNAc1 Su1 + Man3 GlcNAc2		Glyconnect
1797.83	1797.61	132.0	Hex3 HexNAc1 PMe1 + Man3 GlcNAc2		Glyconnect
1824.85	1824.61	128.4	Hex2 HexNAc2 Su1 + Man3 GlcNAc2		Glyconnect
1865.89	1865.64	132.0	Hex1 HexNAc3 Su1 + Man3 GlcNAc2		Glyconnect
1906.91	1906.66	126.5	HexNAc4 Su1 + Man3 GlcNAc2		Glyconnect
1967.97	1967.67	156.6	Hex1 HexNAc2 NeuAc1 Su1 + Man3 GlcNAc2		
2027.98	2027.69	142.1	Hex2 HexNAc3 Su1 + Man3 GlcNAc2		
2069.01	2068.72	139.3	Hex1 HexNAc4 Su1 + Man3 GlcNAc2		
2110.01	2109.74	125.2	HexNAc5 Su1 + Man3 GlcNAc2		Glyconnect
2129.31	2129.72	-189.6	Hex2 HexNAc2 NeuAc1 Su1 + Man3 GlcNAc2		Glyconnect
2190.11	2189.74	166.8	Hex3 HexNAc3 Su1 + Man3 GlcNAc2		
2231.10	2230.77	148.5	Hex2 HexNAc4 Su1 + Man3 GlcNAc2		
2272.14	2271.80	150.2	Hex1 HexNAc5 Su1 + Man3 GlcNAc2		

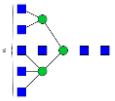
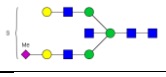
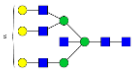
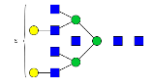
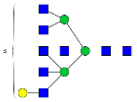
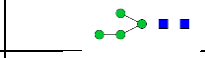
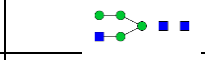
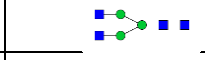
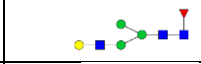
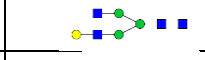
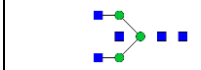
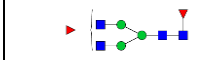
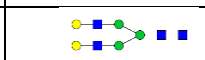
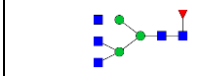
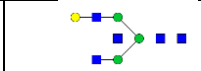
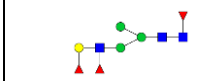
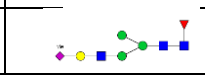
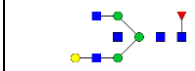
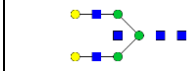
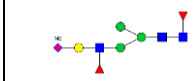
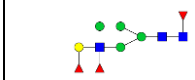
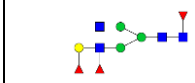
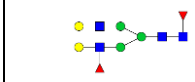
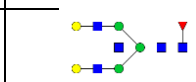
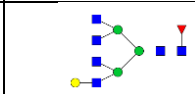
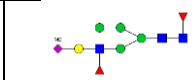
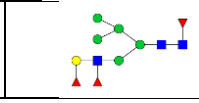
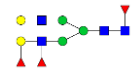
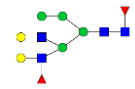
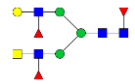
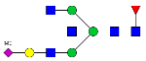
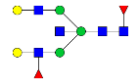
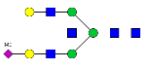
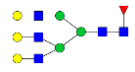
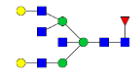
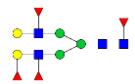
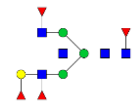
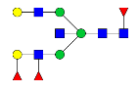
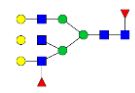
2313.21	2312.82	164.4	HexNAc6 Su1 + Man3 GlcNAc2		
2333.07	2332.80	122.3	Hex2 HexNAc3 NeuAc1 Su1 + Man3 GlcNAc2		
2393.22	2392.82	164.3	Hex3 HexNAc4 Su1 + Man3 GlcNAc2		
2434.26	2433.85	167.7	Hex2 HexNAc5 Su1 + Man3 GlcNAc2		
2475.25	2475.87	149.2	Hex1 HexNAc6 Su1 + Man3 GlcNAc2		

Table S2.8. Neutral *N*-glycans from human saliva (HS) labeled with BOA.

Observed mass, m/z $[M+Na]^+$	Calculated mass, m/z $[M+Na]^+$	Error ppm	Monosaccharide composition	Probable Structure [26, 27, 36, 39, 41, 55]	Glyconnect Database Links
1038.44	1038.39	-53.5	Hex3 HexNAc2		Glyconnect
1200.50	1200.44	-54.8	Hex4 HexNAc2		Glyconnect
1282.54	1282.49	-39.5	Hex2 HexNAc4		Glyconnect
1346.53	1346.50	23.7	Hex4 HexNAc2 dHex1		Glyconnect
1362.53	1362.49	-27.1	Hex2 + Man3 GlcNAc2		Glyconnect
1387.56	1387.52	-28.3	HexNAc1 dHex1 + Man3 GlcNAc2		Glyconnect
1403.54	1403.52	-18.8	Hex1 HexNAc1 + Man3 GlcNAc2		Glyconnect
1444.57	1444.54	-15.8	HexNAc2 + Man3 GlcNAc2		Glyconnect
1508.58	1508.55	-21.2	Hex2 dHex1 + Man3 GlcNAc2		Glyconnect
1524.58	1524.54	-25.0	Hex3 + Man3 GlcNAc2		Glyconnect
1549.60	1549.58	-17.1	Hex1 HexNAc1 dHex1 + Man3 GlcNAc2		Glyconnect
1565.66	1565.57	-58.5	Hex2 HexNAc1 + Man3 GlcNAc2		Glyconnect
1590.61	1590.60	-6.9	HexNAc2 dHex1 + Man3 GlcNAc2		Glyconnect
1606.60	1606.60	0.6	Hex1 HexNAc2 + Man3 GlcNAc2		Glyconnect
1647.63	1647.62	-2.7	HexNAc3 + Man3 GlcNAc2		Glyconnect
1695.64	1695.63	-4.5	Hex1 HexNAc1 dHex2 + Man3 GlcNAc2		Glyconnect
1711.64	1711.63	-5.1	Hex2 HexNAc1 dHex1 + Man3 GlcNAc2		Glyconnect
1736.63	1736.66	15.5	HexNAc2 dHex2 + Man3 GlcNAc2		Glyconnect
1752.66	1752.65	-3.5	Hex1 HexNAc2 dHex1 + Man3 GlcNAc2		Glyconnect
1768.65	1768.65	-1.2	Hex2 HexNAc2 + Man3 GlcNAc2		Glyconnect
1793.68	1793.68	3.0	HexNAc3 dHex1 + Man3 GlcNAc2		Glyconnect
1809.67	1809.68	3.5	Hex1 HexNAc3 + Man3 GlcNAc2		Glyconnect
1841.69	1841.69	0.2	Hex1 HexNAc1 dHex3 + Man3 GlcNAc2		
1854.68	1854.67	-7.6	Hex1 HexNAc1 dHex1 NeuAc1 + Man3 GlcNAc2		Glyconnect

1857.68	1857.69	1.2	Hex2 HexNAc1 dHex2 + Man3 GlcNAc2		Glyconnect
1873.68	1873.68	1.1	Hex3 HexNAc1 dHex1 + Man3 GlcNAc2		Glyconnect
1898.71	1898.71	2.0	Hex1 HexNAc2 dHex2 + Man3 GlcNAc2		Glyconnect
1914.71	1914.71	1.4	Hex2 HexNAc2 dHex1 + Man3 GlcNAc2		Glyconnect
1930.68	1930.70	13.3	Hex3 HexNAc2 + Man3 GlcNAc2		Glyconnect
1955.72	1955.73	8.8	Hex1 HexNAc3 dHex1 + Man3 GlcNAc2		Glyconnect
1971.70	1971.73	16.8	Hex2 HexNAc3 + Man3 GlcNAc2		Glyconnect
1996.78	1996.76	-10.1	HexNAc4 dHex1 + Man3 GlcNAc2		Glyconnect
2000.77	2000.74	-18.1	Hex1 HexNAc1 dHex2 NeuAc1 + Man3 GlcNAc2		
2003.73	2003.74	6.1	Hex2 HexNAc1 dHex3 + Man3 GlcNAc2		
2012.74	2012.76	6.3	Hex1 HexNAc4 + Man3 GlcNAc2		Glyconnect
2016.74	2016.72	-7.6	Hex2 HexNAc1 dHex1 NeuAc1 + Man3 GlcNAc2		Glyconnect
2019.73	2019.74	3.0	Hex3 HexNAc1 dHex2 + Man3 GlcNAc2		
2044.76	2044.77	7.2	Hex1 HexNAc2 dHex3 + Man3 GlcNAc2		Glyconnect
2060.76	2060.77	2.2	Hex2 HexNAc2 dHex2 + Man3 GlcNAc2		Glyconnect
2076.74	2076.76	10.8	Hex3 HexNAc2 dHex1 + Man3 GlcNAc2		Glyconnect
2101.77	2101.79	8.6	Hex1 HexNAc3 dHex2 + Man3 GlcNAc2		Glyconnect
2117.78	2117.79	4.3	Hex2 HexNAc3 dHex1 + Man3 GlcNAc2		Glyconnect
2158.80	2158.81	5.8	Hex1 HexNAc4 dHex1 + Man3 GlcNAc2		Glyconnect
2162.78	2162.78	0.8	Hex2 HexNAc1 dHex2 NeuAc1 + Man3 GlcNAc2		
2165.81	2165.80	-7.9	Hex3 HexNAc1 dHex3 + Man3 GlcNAc2		

2187.82	2187.83	4.1	HexNAc2 dHex3 NeuAc1 + Man3 GlcNAc2		
2206.82	2206.82	-0.2	Hex2 HexNAc2 dHex3 + Man3 GlcNAc2		Glyconnect
2219.81	2219.80	-4.5	Hex2 HexNAc2 dHex1 NeuAc1 + Man3 GlcNAc2		Glyconnect
2222.82	2222.82	0.2	Hex3 HexNAc2 dHex2 + Man3 GlcNAc2		Glyconnect
2247.82	2247.85	11.6	Hex1 HexNAc3 dHex3 + Man3 GlcNAc2		Glyconnect
2260.85	2260.83	-9.0	Hex1 HexNAc3 dHex1 NeuAc1 + Man3 GlcNAc2		Glyconnect
2263.86	2263.85	-5.3	Hex2 HexNAc3 dHex2 + Man3 GlcNAc2		Glyconnect
2276.81	2276.81	-2.4	Hex2 HexNAc3 NeuAc1 + Man3 GlcNAc2		
2279.84	2279.84	-0.5	Hex3 HexNAc3 dHex1 + Man3 GlcNAc2		Glyconnect
2292.83	2292.82	-3.3	Hex1 HexNAc1 dHex4 NeuAc1 + Man3 GlcNAc2		
2295.94	2295.83	-47.1	Hex4 HexNAc3 + Man3 GlcNAc2		Glyconnect
2304.83	2304.87	17.2	Hex1 HexNAc4 dHex2 + Man3 GlcNAc2		Glyconnect
2320.84	2320.87	11.0	Hex2 HexNAc4 dHex1 + Man3 GlcNAc2		Glyconnect
2334.88	2334.81	-29.6	Hex8 + Man3 GlcNAc2		Glyconnect
2352.90	2352.88	-6.6	Hex2 HexNAc2 dHex4 + Man3 GlcNAc2		
2365.89	2365.86	-10.1	Hex2 HexNAc2 dHex2 NeuAc1 + Man3 GlcNAc2		Glyconnect
2381.93	2381.86	-32.8	Hex3 HexNAc2 dHex1 NeuAc1 + Man3 GlcNAc2		Glyconnect
2393.87	2393.91	14.2	Hex1 HexNAc3 dHex4 + Man3 GlcNAc2		
2409.92	2407.90	-7.5	Hex2 HexNAc3 dHex3 + Man3 GlcNAc2		Glyconnect
2425.92	2425.90	-11.2	Hex3 HexNAc3 dHex2 + Man3 GlcNAc2		Glyconnect

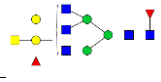
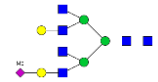
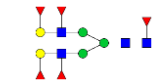
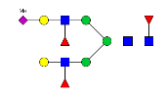
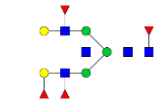
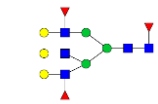
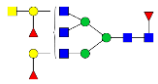
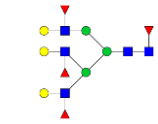
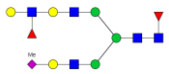
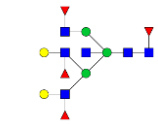
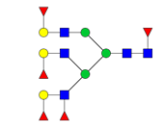
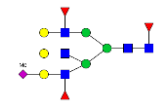
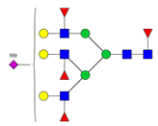
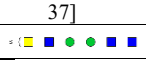
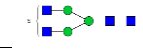
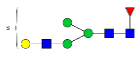
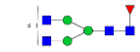
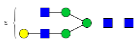
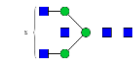
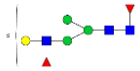
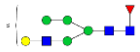
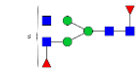
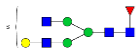
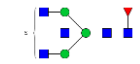
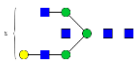
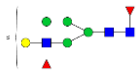
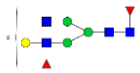
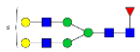
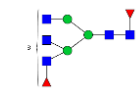
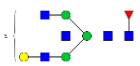
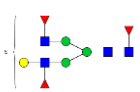
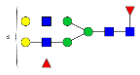
2466.91	2466.92	4.6	Hex2 HexNAc4 dHex2 + Man3 GlcNAc2		Glyconnect
2479.90	2479.90	1.2	Hex2 HexNAc4 NeuAc1 + Man3 GlcNAc2		Glyconnect
2498.95	2498.94	-14.7	Hex2 HexNAc2 dHex5 + Man3 GlcNAc2		Glyconnect
2511.94	2511.93	-10.0	Hex2 HexNAc2 dHex3 NeuAc1 + Man3 GlcNAc2		Glyconnect
2555.96	2555.96	1.5	Hex2 HexNAc3 dHex4 + Man3 GlcNAc2		
2571.98	2571.96	-10.6	Hex3 HexNAc3 dHex3 + Man3 GlcNAc2		Glyconnect
2584.96	2584.94	-10.3	Hex3 HexNAc3 dHex1 NeuAc1 + Man3 GlcNAc2		Glyconnect
2612.96	2612.98	8.5	Hex2 HexNAc4 dHex3 + Man3 GlcNAc2		Glyconnect
2718.05	2718.05	-14.5	Hex3 HexNAc3 dHex4 + Man3 GlcNAc2		Glyconnect
2731.05	2730.99	-20.4	Hex3 HexNAc3 dHex2 NeuAc1 + Man3 GlcNAc2		Glyconnect
2759.04	2759.04	-0.3	Hex2 HexNAc4 dHex4 + Man3 GlcNAc2		
2864.10	2864.05	-17.3	Hex3 HexNAc3 dHex5 + Man3 GlcNAc2		Glyconnect
2877.11	2877.05	-20.5	Hex3 HexNAc3 dHex3 NeuAc1 + Man3 GlcNAc2		Glyconnect
3011.21	3011.09	-40.4	Hex5 HexNAc2 dHex1 NeuAc2 + Man3 GlcNAc2		
3023.22	3023.11	-35.0	Hex3 HexNAc3 dHex4 NeuAc1 + Man3 GlcNAc2		

Table S2.9. Sulfated *N*-glycans from human saliva (HS) labeled with BOA.

Observed mass, m/z $[M - H]^-$	Calculated mass, m/z $[M - H]^-$	Error ppm	Monosaccharide composition	Probable Structure [26, 27, 37]	Glyconnect Database Links
1338.78	1338.45	-245.0	Hex2 HexNAc4 Su1		
1443.86	1443.48	-263.7	HexNAc1 dHex1 Su1 + Man3 GlcNAc2		
1500.83	1500.50	-215.8	HexNAc2 Su1 + Man3 GlcNAc2		Glyconnect
1605.94	1605.54	-254.0	Hex1 HexNAc1 dHex1 Su1 + Man3 GlcNAc2		Glyconnect
1646.95	1646.56	-238.2	HexNAc2 dHex1 Su1 + Man3 GlcNAc2		Glyconnect
1662.94	1662.56	-230.2	Hex1 HexNAc2 Su1 + Man3 GlcNAc2		Glyconnect
1703.95	1703.58	-214.3	HexNAc3 Su1 + Man3 GlcNAc2		Glyconnect
1752.03	1751.59	-248.3	Hex1 HexNAc1 dHex2 Su1 + Man3 GlcNAc2		
1768.02	1767.59	-243.7	Hex2 HexNAc1 dHex1 Su1 + Man3 GlcNAc2		Glyconnect
1793.00	1792.62	-210.7	HexNAc2 dHex2 Su1 + Man3 GlcNAc2		
1809.03	1808.62	-227.5	Hex1 HexNAc2 dHex1 Su1 + Man3 GlcNAc2		Glyconnect
1850.07	1849.64	-229.8	HexNAc3 dHex1 Su1 + Man3 GlcNAc2		Glyconnect
1865.99	1865.64	-188.5	Hex1 HexNAc3 Su1 + Man3 GlcNAc2		Glyconnect
1914.12	1913.65	-249.3	Hex2 HexNAc1 dHex2 Su1 + Man3 GlcNAc2		
1955.15	1954.67	-243.8	Hex1 HexNAc2 dHex2 Su1 + Man3 GlcNAc2		Glyconnect
1971.14	1970.67	-239.2	Hex2 HexNAc2 dHex1 Su1 + Man3 GlcNAc2		Glyconnect
1996.07	1995.70	-183.6	HexNAc3 dHex2 Su1 + Man3 GlcNAc2		
2012.11	2011.69	-208.8	Hex1 HexNAc3 dHex1 Su1 + Man3 GlcNAc2		Glyconnect
2076.16	2075.70	-221.9	Hex3 HexNAc1 dHex2 Su1 + Man3 GlcNAc2		
2101.22	2100.73	-231.9	Hex1 HexNAc2 dHex3 Su1 + Man3 GlcNAc2		Glyconnect
2117.21	2116.73	-230.4	Hex2 HexNAc2 dHex2 Su1 + Man3 GlcNAc2		

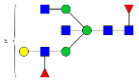
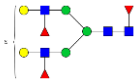
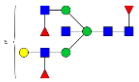
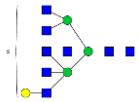
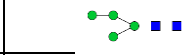
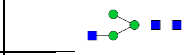
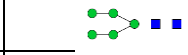
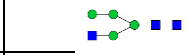
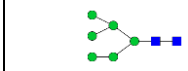
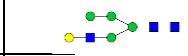
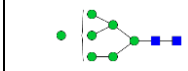
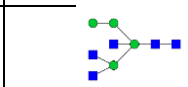
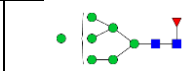
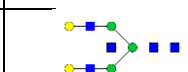
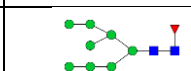
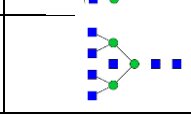
2158.22	2157.75	-218.4	Hex1 HexNAc3 dHex2 Su1 + Man3 GlcNAc2		Glyconnect
2263.31	2262.78	219.7	Hex2 HexNAc2 dHex3 Su1 + Man3 GlcNAc2		
2304.31	2303.81	-218.0	Hex1 HexNAc3 dHex3 Su1 + Man3 GlcNAc2		
2474.46	2474.87	166.9	Hex1 HexNAc6 Su1 + Man3 GlcNAc2		

Table S2.10. Neutral *N*-glycans from chicken egg white (CEW) labeled with BOA.

Observed mass, m/z $[M+Na]^+$	Calculated mass, m/z $[M+Na]^+$	Error ppm	Monosaccharide composition	Probable Structure [19, 26, 27, 50–53, 56]	Glyconnect Database Links
1038.45	1038.39	119.5	Hex3 HexNAc2		Glyconnect
1200.50	1200.44	45.1	Hex4 HexNAc2		Glyconnect
1241.53	1241.46	118.1	HexNAc1 + Man3 GlcNAc2		Glyconnect
1282.56	1282.44	50.1	Hex2 HexNAc4		Glyconnect
1362.56	1362.49	44.9	Hex2 + Man3 HexNAc2		Glyconnect
1403.58	1403.52	113.3	Hex1 HexNAc1 + Man3 GlcNAc2		Glyconnect
1444.61	1444.54	112.3	HexNAc2 + Man3 GlcNAc2		Glyconnect
1524.61	1524.54	123.1	Hex3 + Man3 GlcNAc2		Glyconnect
1565.64	1565.57	41.6	Hex2 HexNAc1 + Man3 GlcNAc2		Glyconnect
1606.66	1606.60	106.0	Hex1 HexNAc2 + Man3 GlcNAc2		Glyconnect
1647.69	1647.62	103.4	HexNAc3 + Man3 GlcNAc2		Glyconnect
1686.68	1686.60	46.4	Hex4 + Man3 GlcNAc2		Glyconnect
1727.70	1727.62	118.0	Hex3 HexNAc1 + Man3 GlcNAc2		Glyconnect
1768.72	1768.65	99.7	Hex2 HexNAc2 + Man3 GlcNAc2		Glyconnect
1809.75	1809.68	100.0	Hex1 HexNAc3 + Man3 GlcNAc2		Glyconnect
1832.76	1832.65	112.1	Hex4 dHex1 + Man3 GlcNAc2		Glyconnect
1850.77	1850.70	101.7	HexNAc4 + Man3 GlcNAc2		Glyconnect
1930.79	1930.70	104.9	Hex3 HexNAc2 + Man3 GlcNAc2		Glyconnect
1971.81	1971.73	36.6	Hex2 HexNAc3 + Man3 GlcNAc2		Glyconnect
1994.83	1994.71	110.0	Hex5 dHex1 + Man3 GlcNAc2		Glyconnect
2012.83	2012.76	101.3	Hex1 HexNAc4 + Man3 GlcNAc2		Glyconnect
2053.87	2053.78	104.3	HexNAc5 + Man3 GlcNAc2		Glyconnect

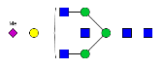
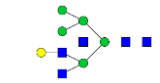
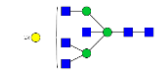
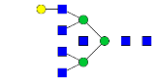
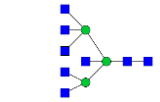
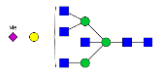
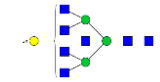
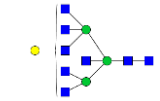
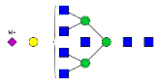
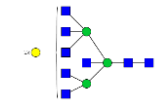
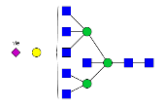
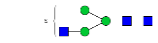
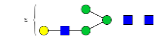
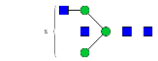
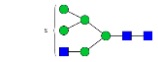
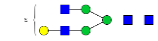
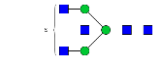
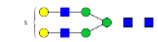
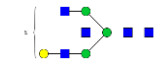
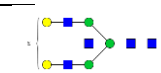
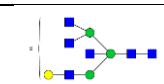
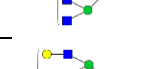
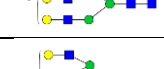
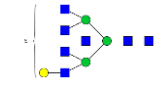
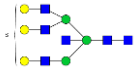
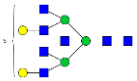
2114.87	2114.79	130.8	Hex1 HexNAc3 NeuAc1 + Man3 GlcNAc2		Glyconnect
2133.87	2133.78	39.0	Hex3 HexNAc3 + Man3 GlcNAc2		Glyconnect
2174.89	2174.81	33.7	Hex2 HexNAc4 + Man3 GlcNAc2		Glyconnect
2215.92	2215.84	102.2	Hex1 HexNAc5 + Man3 GlcNAc2		Glyconnect
2256.95	2256.86	110.2	HexNAc6 + Man3 GlcNAc2		Glyconnect
2317.97	2317.87	108.5	Hex1 HexNAc4 NeuAc1 + Man3 GlcNAc2		Glyconnect
2377.99	2377.89	114.7	Hex2 HexNAc5 + Man3 GlcNAc2		Glyconnect
2419.02	2418.91	112.4	Hex1 HexNAc6 + Man3 GlcNAc2		Glyconnect
2521.05	2520.95	97.3	Hex1 HexNAc5 NeuAc1 + Man3 GlcNAc2		Glyconnect
2581.10	2580.97	111.2	Hex2 HexNAc6 + Man3 GlcNAc2		Glyconnect
2724.17	2724.03	112.7	Hex1 HexNAc6 NeuAc1 + Man3 GlcNAc2		

Table S2.11. Sulfated *N*-glycans from chicken egg white (CEW) labeled with BOA.

Observed mass, m/z $[M - H]^-$	Calculated mass, m/z $[M - H]^-$	Error ppm	Monosaccharide composition	Probable Structure [26, 27, 33, 54]	Glyconnect Database Links
1297.47	1297.43	36.3	HexNAc1 Su1 + Man3 GlcNAc2		Glyconnect
1459.56	1459.48	55.4	Hex1 HexNAc1 Su1 + Man3 GlcNAc2		
1500.54	1500.50	24.9	HexNAc2 Su1 + Man3 GlcNAc2		Glyconnect
1580.55	1580.50	30.8	Hex3 Su1 + Man3 GlcNAc2		Glyconnect
1594.59	1594.53	38.0	Hex3 PMe1 + Man3 GlcNAc2		Glyconnect
1621.63	1621.53	60.7	Hex2 HexNAc1 Su1 + Man3 GlcNAc2		
1662.66	1662.56	61.8	Hex1 HexNAc2 Su1 + Man3 GlcNAc2		Glyconnect
1703.62	1703.58	18.7	HexNAc3 Su1 + Man3 GlcNAc2		Glyconnect
1783.74	1783.58	85.3	Hex3 HexNAc1 Su1 + Man3 GlcNAc2		Glyconnect
1797.69	1797.61	47.3	Hex3 HexNAc1 PMe1 + Man3 GlcNAc2		Glyconnect
1824.72	1824.61	62.8	Hex2 HexNAc2 Su1 + Man3 GlcNAc2		Glyconnect
1865.78	1865.64	74.4	Hex1 HexNAc3 Su1 + Man3 GlcNAc2		Glyconnect
1906.78	1906.66	60.9	HexNAc4 Su1 + Man3 GlcNAc2		Glyconnect
2027.84	2027.69	76.7	Hex2 HexNAc3 Su1 + Man3 GlcNAc2		
2068.88	2068.72	80.0	Hex1 HexNAc4 Su1 + Man3 GlcNAc2		
2109.88	2109.74	67.0	HexNAc5 Su1 + Man3 GlcNAc2		Glyconnect
2189.95	2189.74	96.5	Hex3 HexNAc3 Su1 + Man3 GlcNAc2		
2230.95	2230.77	83.4	Hex2 HexNAc4 Su1 + Man3 GlcNAc2		
2272.00	2271.80	90.3	Hex1 HexNAc5 Su1 + Man3 GlcNAc2		
2313.05	2312.82	98.6	HexNAc6 Su1 + Man3 GlcNAc2		

2393.09	2392.82	113.4	Hex3 HexNAc4 Su1 + Man3 GlcNAc2		
2434.12	2433.85	112.5	Hex2 HexNAc5 Su1 + Man3 GlcNAc2		

Chapter 3 Glycoblotting-based SulPhoglycomics Analysis of Sulfated and Phosphorylated *N*-glycans from the Egg Whites of Anseriformes Species (Waterfowls)

3.1 Introduction

Influenza A virus (IAV) membrane proteins, hemagglutinin (HA) and neuraminidase (NA), play crucial roles in virus infectivity, transmissibility, pathogenicity, and host specificity[1]. Combinations of different HA and NA subtypes give rise to some of the deadliest IAVs that have caused major pandemics[2]. IAVs possess the remarkable ability to evolve and evade neutralization by antibodies and vaccinations through antigenic evolution[3]. In particular, human-adapted IAVs exhibit binding specificity towards sialyl-LacNAc, characterized by an α2,6 linkage between Sia and Gal, while avian IAVs bind to sialyl-LacNAc with α2,3 linkages[4, 5].

IAV demonstrates a wide range of host species, with avian species belonging to the Order Anseriformes (waterfowls) serving as its natural reservoir[2]. Waterfowl are resistant to avian influenza virus (AIV) and show no clinical symptoms despite harboring almost all subtypes of IAV[6, 7]. Avian species primarily experience IAV infections in the respiratory and intestinal tracts. In chickens, the respiratory tract serves as the initial site of infection, facilitating viral transmission through aerosol droplets. On the other hand, ducks transmit the virus through the oral-fecal route, as IAV replication occurs in their intestine, colon, and cloaca, typically observed in birds infected with low pathogenic avian influenza virus (LPAIV). In contrast, highly pathogenic avian influenza virus (HPAIV) causes systemic infections, affecting various tissues such as the heart, brain, spleen, liver, and oviduct[6–8].

HPAIV has been observed and isolated in the oviducts of certain avian species. Additionally, LPIAV infection of avian oviduct explants has been demonstrated in chickens, turkeys, and ducks, with susceptibility observed in all sections of the oviduct, particularly the magnum cells[9]. This susceptibility explains the detection of HPIAV in the yolk and albumen of eggs, suggesting a potential role of the reproductive tract in the pathobiology of IAV in avian species. Furthermore, egg-borne influenza viruses not only impact wild and domesticated birds but also pose serious implications for viral dissemination to humans[10, 11].

Avian IAVs exhibits species-specific differences in the receptor binding specificities of HA. While Sia-α2,3Gal is considered the minimum essential glycan structure for binding, the fine details of HA specificity vary depending on the original host species. These binding specificities of avian IAVs are determined through synthetic glycan library evaluations using microarray or histochemical analyses[6, 12, 13]. Therefore, it is crucial to determine the glycan structures expressed in host cells and tissues, as the receptor binding specificities of HA play a crucial role in determining tissue and species tropism of IAV. Information about the glycan structures present in these tissues is valuable for identifying critical structures that IAVs would bind to, highlighting the significance of host cell glycans as natural barriers for transmission between different species[8].

Recent studies have shed light on the importance of sulfated-glycan structures on IAV hosts. Notably, the inoculation of IAV into MDCK cells overexpressing sulfotransferase resulted in a 90-fold increase in viral replication. Fuc- and sulfated Sialyl-LacNAc moieties were found to be expressed in chicken embryos, suggesting their involvement in the efficient propagation of human H1N1 in chicken embryos[14, 15]. These findings underscore the importance of thoroughly evaluating different sulfated-glycan structures on IAV hosts, considering that the receptor binding affinity of each hemagglutinin subtype varies and greatly depends not only on α2,3/6-linked sialosides but also on the underlying glycan structures[13, 14, 16].

Egg white samples offer valuable insights into glycan diversity and their roles in protecting embryos from infections. The structural diversity of egg white glycans is species-specific, resulting from environmental pressures such as pathogenic invasions. Additionally, since embryos are unable to produce antibodies, immunoglobulins from the hen are transferred into the egg, providing a snapshot of the maternal immune system's specific defense against pathogens[17]. Hence, the diversity of glycans in egg whites can be inferred as a product of the evolutionary history of antipathogenic offense and defense[18]. In this study, we conducted a large-scale analysis of avian egg whites from Order Anseriformes (Waterfowls) using the

Glycoblotting-based sulphoglycomics approach we previously described. Our findings unveiled a wide range of sulfated and phosphorylated *N*-glycan species that exhibit distinct expression patterns in waterfowl egg whites. Importantly, we inferred the potential significance of these glycan variations concerning influenza infection in waterfowl.

3.2 Methodology

3.2.1 Egg Whites. Egg whites from various species of Order Anseriformes (4 families, 27 genera, 66 species) were collected by M. Laskowski, Jr.[19–21], and were maintained at -20°C. The scientific names of the birds were adapted from Sibley and Monroe. The DNA-DNA hybridization of Sibley *et al*. was used as our primary reference because their classification of birds worldwide is complementary to the phylogenetic analysis[22, 23].

3.2.2 Materials. Peptide *N*-glycosidase F (PNGase F) was acquired from New England BioLabs (Ipswich, MA, USA), proteinase K was from Roche (Germany), trypsin was from Sigma-Aldrich Corp. (St. Loius, MO, USA), and the bacterial alkaline phosphatase (BAP) was from Nippon Gene, Ltd. (Tokyo, Japan). Ammonium carbamate, benzyloxyamine hydrochloride (BOA), 3-methyl-1-*p*-tolyltriazene (MTT), disialyloctasaccharide (SGP-10), hexa-*N*-acetylchitohexaose, 2,5-dihydroxybenzoic acid (DHB), sodium bicarbonate ($NaHCO_3$), 3,4-diaminobenzophenone (DABP), and trifluoroacetic acid (TFA) were obtained from Tokyo Chemical Industry Co. (Tokyo, Japan). BlotGlycoH bead was acquired from Sumitomo Bakelite, Co., Ltd. (Tokyo, Japan).

3.2.3 *N*-glycan Release[18, 24] **.** Lyophilized egg whites (approx. 1 mg) were dissolved in 50 μL of 200 mM NH_4HCO_3, followed by the addition of 4 μL denaturation buffer (5% SDS, 0.4 M DTT). The mixture was denatured for 10 mins at 100°C. After denaturation, 10 μL of 123 mM iodoacetamide was added to the mix and incubated in the dark at room temperature for 1 hr. Tryptic digestion was achieved by adding 10 μL of 40 U/μL sequence-grade Trypsin (Sigma-

Aldrich) in 1 mM HCl; the mixture was then incubated overnight at 37°C, followed by heat inactivation of the enzyme at 90°C for 10 mins. The tryptic digest was allowed to cool at room temperature, then 8 μL reaction buffer (0.5 M Na_3PO_4, pH 7.5), 8 μL 10% NP-40, and 2 μL of 5 U/μL PNGase F (New England BioLabs) were added, and incubated overnight at 37°C. The mixture was further digested with 10 μL of 0.5 U/μL Proteinase K (Roche, Germany) at 37°C for 3 hrs, followed by heat inactivation of the enzyme at 90°C for 10 mins. The sample was dried in a SpeedVac and stored at -20°C until use.

3.2.4 Glycan Enrichment Using Glycoblotting [18, 24, 25]**.** A 250 μL aliquot of 10 mg/mL BlotGlycoH bead (Sumitomo Bakelite, Co.) suspension was dispensed into each well of the 96-well multiScreen Solvinert filter plate (Millipore, Billerica, MA). The filter plate was then attached to a vacuum manifold to remove water. The dried sample containing released *N*-glycans from egg whites was reconstituted with 40 μL MilliQ water. A 20 μL aliquot of the reconstituted sample was added into the wells with 10 μL of 100 μM disialyloctasaccharide, SGP-10 (Tokyo Chemical Industry Co., Ltd.) internal standard, and 180 μL of 2% AcOH in acetonitrile (MeCN). The 96-well filter plate was incubated at 80°C for 45 mins until dry. Each sample well was washed with 200 μL of 2 M guanidine-HCl in 16 mM NH_4HCO_3, water, and 1% triethylamine in methanol (MeOH) sequentially. Each solvent washing was performed twice and vacuumed after every washing step. Unreacted hydrazide functional groups on the beads were capped with an acetyl group by incubating each sample well with 100 μL of 10% acetic anhydride in MeOH for 30 mins at room temperature. The capping solution was then removed by vacuum and sequentially washed twice with 200 μL of 10 mM HCl, MeOH, and dioxane. On-bead methyl esterification of the carboxyl groups of acidic glycans (e.g., sialic acid) was performed by adding 100 μL of 100 mM 3-methyl-1-*p*-tolyltriazene (MTT) in dioxane into the sample and incubated at 60°C for 90 mins until dry [26]. The 96-well plate was washed twice with 200 μL dioxane, water, MeOH, and water sequentially. The captured glycans on the BlotGlycoH beads were labeled with benzyloxyamine

(BOA) via trans-iminization reaction. The labeling was performed by adding 20 μL of 50 mM BOA-HCl and 180 μL of 2% AcOH in MeCN at 80°C for 45 mins. BOA-labeled glycans were eluted with 150 μL water twice. The sample was dried in a SpeedVac and stored at -20°C until use.

3.2.5 Anionic-glycan Separation using WAX [27, 28]**.** Fifty microliters of 100 mg/mL 3-aminopropyl silica gel suspension (1 mmol/mg, Tokyo Chemical Industry Co. Ltd.) were packed into a 200 μL micropipette tip with a cotton plug. The packed weak anion exchange (WAX) microcolumn was conditioned and washed sequentially with 100 μL water, MeCN, and 1% AcOH in 95% MeCN twice. After every conditioning and washing step, the column was centrifugated at 500 rpm for 2 mins. BOA-labeled *N*-glycans were reconstituted with 20 μL water. A 5 μL sample aliquot was dissolved in 150 μL 1% AcOH in 95% MeCN and then loaded into the column. The sample was allowed to elute by gravity, and the collected eluate was reloaded back into the column; this step was done three times. The column was washed with 1% AcOH in 95% MeCN to remove unbound and hydrophobic contaminants. BOA-labeled neutral and mono-methylated sialyl *N*-glycans were eluted with 1% AcOH in 50% MeCN (Neutral *N*-glycan Fraction), while BOA-labeled sulfated and phosphorylated *N*-glycans were eluted with 1% NH_4OH in 5% MeCN (pH 10.5) (Acidic *N*-glycan Fraction). The eluates were then dried in a SpeedVac and stored at -20°C until use.

3.2.6 Mass Spectrometric Analysis. MALDI-TOF MS analysis of BOA-labeled *N*-glycans was performed using Ultraflex III (Bruker, Bremen, Germany) operated on reflectron mode on both positive and negative ion acquisition mode. Neutral and mono-methylated sialyl-glycans were analyzed in positive ion mode using 10 mg/mL DHB/$NaHCO_3$ (10:1) in 50% MeCN matrix [29, 30]. While sulfated glycans were analyzed in negative ion mode using the DABP matrix (3,4-diaminobenzophenone, 10 mg/mL in 75% MeCN with 0.1% TFA) [27, 28, 31, 32]. The MALDI-

TOF and MALDI-TOF/TOF MS data were annotated using Bruker FlexAnalysis 3.0 software package. Experimental *m/z* were used to predict possible glycan composition using the Expasy GlycoMod Tool and Glyconnect Database of the Swiss Institute of Bioinformatics (https://web.expasy.org/glycomod/) and GlycoWorkbench [33, 34].

3.2.7 Statistical Analysis. Multivariate analysis was performed using OriginPro statistical software (OriginLab Corp.). The *N*-glycan peaks detected on MALDI-TOF MS spectra were picked and annotated using FlexAnalysis software (Bruker Daltonics). The peak area of each *N*-glycan monoisotopic peak was normalized to the area of the most abundant *N*-glycan in the MS spectra and the peak areas were then expressed as relative abundance. MS data sets used for the subsequent multivariate analyses contain the relative abundance of monoisotopic peaks (total 89 *m/z*) identified as *N*-glycan peaks from each egg white of 72 waterfowl species. Hierarchical clustering analysis (HCA) was performed to investigate the similarity of the glycan expression profiles based on statistical distance. Principal component analysis (PCA) was also used to explore the glycan profiles based on their variation-covariance (information on each glycan), score plots are provided for visual inspection of the relationships of principal components. While MetaboAnalyst ver. 5.0 was used to perform partial least square-discriminant analysis (PLS-DA) and determined variable importance to projection (VIP) scores to identify *N*-glycan species that strongly influence the group separation in the PLS-DA score plot.

3.3 Results and Discussion

3.3.1 Sulphoglycomics Revealed Diverse Acidic *N*-glycans in Waterfowl Egg Whites. We analyzed egg whites from 72 avian species that belong to the Order Anseriformes (waterfowls). The species were classified into different Families based on Sibley's DNA-DNA hybridization, wherein sixty-four (64) of the species belong to Anatidae, six (6) from Dendrocygnidae, one (1) from Anhimidae, and one (1) from Anseranatidae (Table S3.1). Acidic *N*-glycans (*i.e.,* sulfated

and phosphorylated) from egg whites were analyzed using the Glycoblotting-based sulphoglycomics workflow we described previously [35]. Briefly, egg white *N*-glycans were released using PNGase F, followed by glycan enrichment, methyl esterification, and BOA-labeling by Glycoblotting. BOA-labeled *N*-glycans were then fractionated using an amine-functionalized weak anion exchange (WAX) microcolumn. Neutral and Sialylated *N*-glycans were eluted first with 1% AcOH in 50% MeCN (neutral fraction), while sulfated and phosphorylated *N*-glycans were eluted next with 1% NH_4OH in 5% MeCN (pH 10.5) (acidic fraction). MALDI-TOF MS analysis of the acidic fractions from the waterfowl egg whites (72 species) detected a total of 89 monoisotopic peaks in negative ion mode $[M-H]^-$ identified as acidic *N*-glycan peaks. However, 89 does not correspond to the number of *N*-glycan structures due to the presence of structural isomers for every *m/z* value observed. Figure 3.1A shows the representative acidic *N*-glycan profiles of each waterfowl family. Major acidic *N*-glycans expressed by Anhimidae (*Chauna torquata*) and Anseranatidae (*Anseranas semipalmata*) have monoisotopic masses above 1800 *m/z*. While acidic *N*-glycans with low *m/z* values (below 1800 *m/z*) are observed for Anatidae (*Anas platyrhynchos*) and Dendrocygnidae (*Dendrocygna viduata*). The relative abundance of each monoisotopic peak (*m/z* values) identified as *N*-glycans expressed in the egg whites of the 72 waterfowl species described in this work are shown as Heat Map in Figure 3.1B. It can be noticed that two acidic *N*-glycans (A8 and A16) are highly expressed across all 72 waterfowl species. A8 and A16 have monoisotopic masses of 1297 *m/z* and 1459 *m/z* respectively. A8 (1297 *m/z*) was inferred to have a glycan composition of HexNAc1Su1 + Hex3GlcNAc2 (complex *N*-glycan) while A16 (1459 *m/z*) have Hex1HexNAc1Su1 + Hex3GlcNAc2 (hybrid/complex *N*-glycan). On the other hand, A37 (1783 *m/z*) was also noted to be abundant in some species. A37 glycan composition was inferred to be Hex3HexNAc1Su1 + Man3GlcNAc2, multiple hexoses with a single HexNAc residue on its antennae indicate that probably it has a mono-sulfated hybrid *N*-glycan structure. The complete list of the inferred structures of the 89 acidic *N*-glycans based on their glycan composition are shown in Table S3.2.

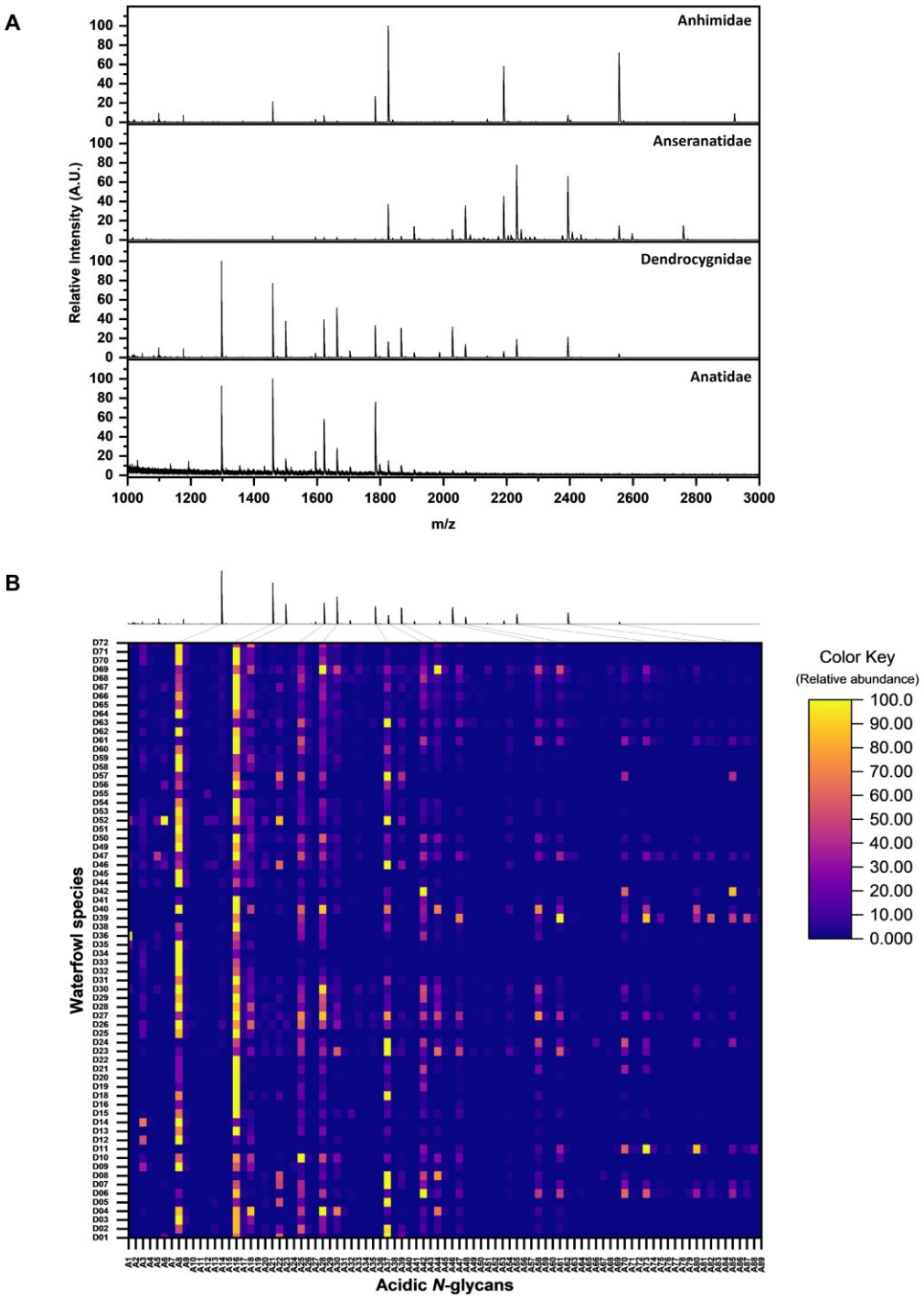
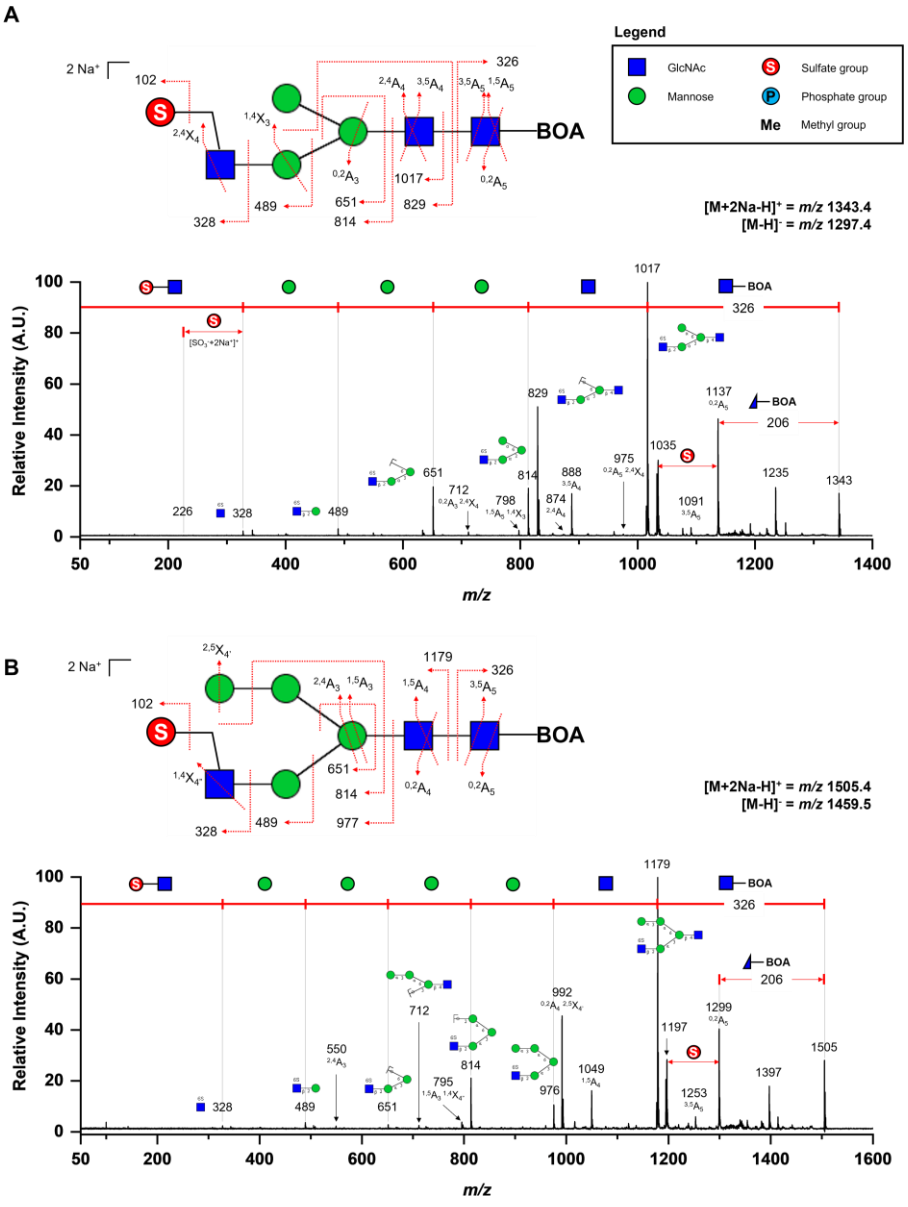


Figure 3.1. (A) Representative MALDI-TOF MS acidic *N*-glycan profiles (in negative ion mode) of the four waterfowl families (anhimidae, anseranatidae, dendrocygnidae and anatidae) belonged to Order Anseriformes. **(B)** Heat Map representation of the area normalized acidic *N*-glycan profiles of the 72 waterfowl species depicting the relative abundance of each *m/z* value (total 89) identified as *N*-glycan peaks. MALDI-TOF MS profile of *Dendrocygna viduata* is shown on top of the heat map.

3.3.2 MALDI-TOF/TOF MS Analysis of Acidic *N*-glycans. MALDI-TOF/TOF MS analysis of the acidic *N*-glycan species was performed in positive ion reflectron mode on their corresponding di-sodiated molecular ion adducts $[M+2Na-H]^+$. The common glycosidic bond cleavages (B and Y ions) and cross-ring cleavages (A and X ions) observed on the TOF/TOF spectra of acidic *N*-glycans enables an informative sequence assignment of their glycan structures. Diagnostic fragment ions and neutral losses implicating a BOA-labeled sulfated *N*-glycan structures were observed. A neutral loss of 326 *m/z* was found across all TOF/TOF MS spectra indicating the loss of the reducing end GlcNAc terminus labeled with BOA. Followed by a subsequent neutral loss of another GlcNAc residue (*m/z* 203) at the *N*-glycan core structure. Also, a cross-ring fragment (^{0,2}A) of the GlcNAc residue at the reducing end terminus was observed with a concurrent 206 *m/z* neutral loss. Fragment ions bearing the sulfate moiety were also detected on the TOF/TOF spectra. The fragment ion peak at *m/z* 328 corresponds to the loss of a mono-sulfated GlcNAc residue $[SO_3+GlcNAc+2Na]^+$ at the non-reducing end of the *N*-glycan structure. Fragment ions at *m/z* 489, 651, and 814 correspond to $[SO_3+GlcNAc+Hex+2Na]^+$, $[SO_3+GlcNAc+Hex_2+2Na]^+$, $[SO_3+GlcNAc+Hex_3+2Na]^+$ respectively, indicating the sequential cleavages of the mannose residues of the *N*-glycan core. Furthermore, a neutral loss of sodium sulfite (*m/z* 102) occurs readily[28] and observed cross-ring fragments (^{1,4}X and ^{2,4}X) of the GlcNAc antennae may suggest that the sulfate group is attached at C6 position (Figure 3.2A and 3.2B).

Phosphorylated *N*-glycan structures were also observed on egg whites as previously reported by Montalban, B. *et al* [35]. Figure 3.2C shows the MALDI-TOF/TOF MS profile of a mono-methylated phosphorylated high mannose (Man6). Similar to the sulfated *N*-glycans TOF/TOF profiles, a neutral loss of *m/z* 326 and subsequent neutral loss of the second GlcNAc residue on the *N*-glycan core was also observed. The cross-ring fragment ion ^{0,2}A was also present. Which suggests that these neutral losses (*m/z* 326, 203, and 206) are characteristics of a BOA-labeled *N*-glycan TOF/TOF MS fragmentation [18, 24]. On the other hand, diagnostic fragmentation ions bearing the phosphate group were detected. The molecular ion peak of the

mono-methylated phosphate group [$PO_3Me+2Na^+$] was observed at *m/z* 140.7. While [$PO_3Me+Hex_2+2Na^+$], [$PO_3Me+Hex_5+2Na^+$], and [$PO_3Me+Hex_6+2Na^+$] were also observed at *m/z* 462.9, 949.4 and 1112.0 respectively. Cross-ring fragment ions $^{0,4}X_4$, and $^{2,5}X_4$, of mannose residue located at the α1-6 antennae, may suggest that the phosphate moiety is attached to the C6 position.



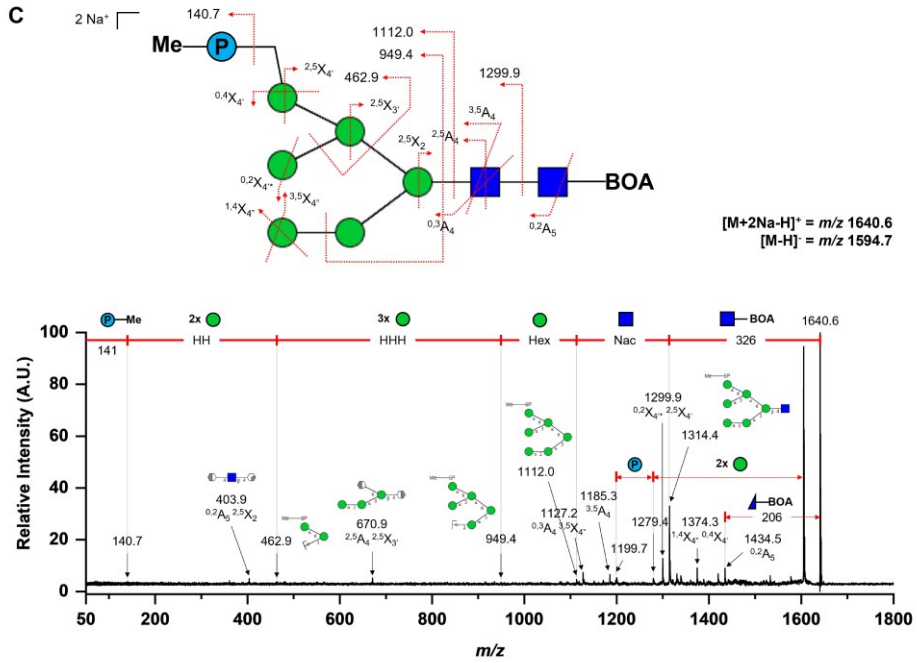


Figure 3.2. MALDI-TOF/TOF MS profiles of selected BOA-labeled acidic *N*-glycans from waterfowl's egg whites. TOF/TOF analysis was performed in positive ion mode on the $[M+2Na - H]^+$ ion adducts of each *N*-glycans. (A) TOF/TOF MS profile of a complex *N*-glycan type – A8 (1297.4 *m/z*, $[M-H]^-$; 1343.4 m/z, $[M+2Na - H]^+$). (B) TOF/TOF MS profile of a hybrid *N*-glycan type – A16 (1459.5 *m/z*, $[M-H]^-$; 1505.4 m/z, $[M+2Na - H]^+$). Wherein both A8 and A16 were identified to have mono-sulfation on its terminal GlcNAc. (C) TOF/TOF MS profile of a phosphorylated high mannose type *N*-glycan – A22 (1594.7 *m/z*, $[M-H]^-$; 1640.6 m/z, $[M+2Na - H]^+$). Linkage specific MS/MS structural analysis of each glycan species was not attempted.

It should be noted that there are few unassigned fragment ions on each MALDI-TOF/TOF MS spectra. Thus, care must be taken when interpreting these TOF/TOF data due to the possibility that the parent ion is a composite of several isomeric forms of each glycan structure which are commonly found in biological samples. Lastly, an MS/MS linkage specific analysis of the glycan structures was not attempted due to the low signal intensities of cross-ring fragments.

3.3.3 Differential Expression Profiles of Acidic *N*-glycans in Waterfowls Egg Whites. Multivariate analysis was employed to explore and understand the relationships between the different acidic *N*-glycan expressions of egg whites among the Anseriformes species. Agglomerative hierarchical clustering analysis (HCA) employing the Ward linkage method and

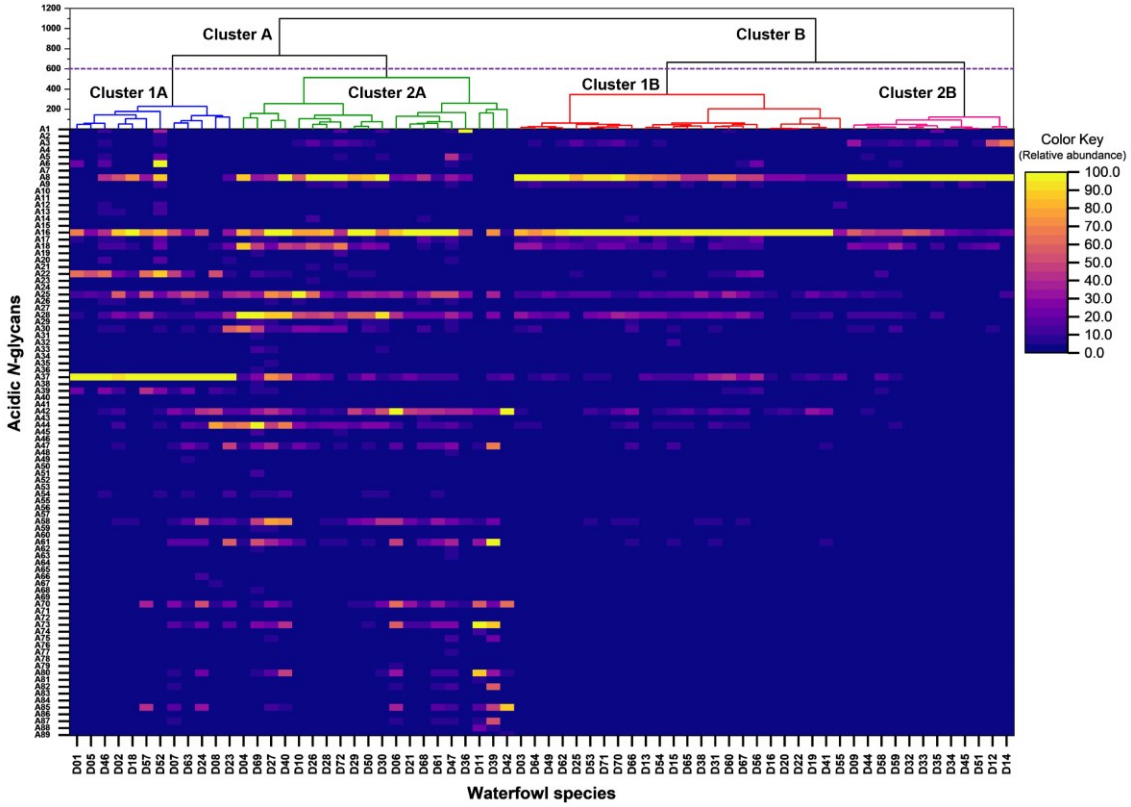
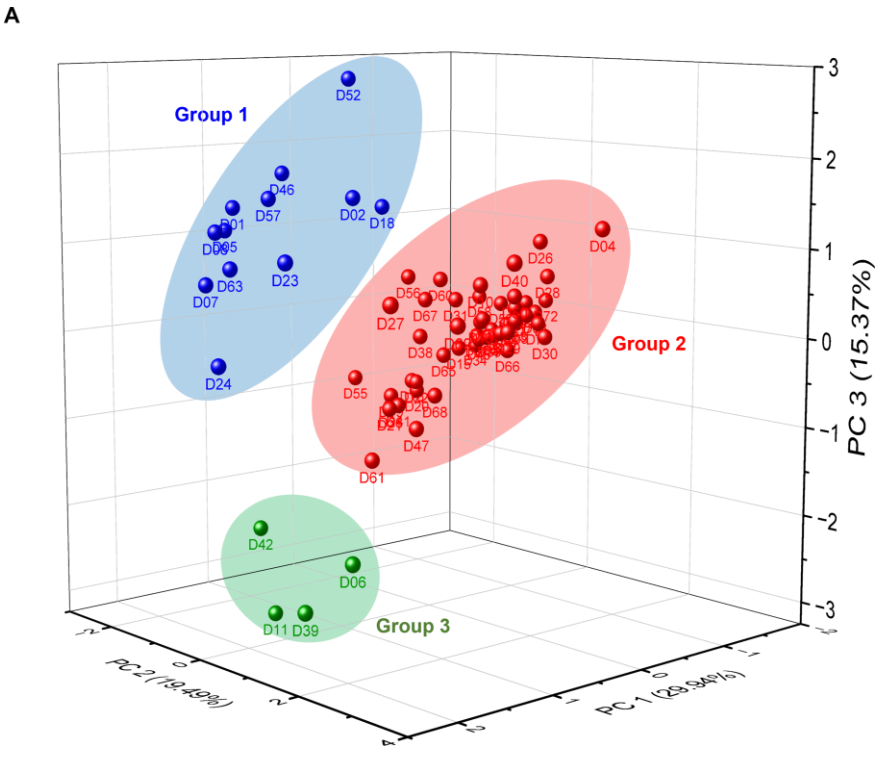


Figure 3.3. Multivariate hierarchical clustering analysis (HCA) of the 72 egg white samples from Anseriformes species. Waterfowl species were clustered according to the differential expression of acidic *N*-glycan (*i.e.,* sulfated and phosphorylated) expressed on egg whites. Relative abundance of the 89 monoisotopic masses identified as acidic *N*-glycans on egg white samples are shown as heat map.

Euclidean distance metric was used to group the egg whites from 72 waterfowl species based on the relative abundance of the 89 monoisotopic peaks (acidic *N*-glycan peaks) (see Table S3.2). The dendrogram (Fig. 3.3) shows that the 72 waterfowl species are grouped into two major clusters (A and B) based on their acidic *N*-glycan expressions. Waterfowl species belonging to Cluster A expressed acidic *N*-glycans having monoisotopic masses above 1800 *m/z*, whereas Cluster B expresses little to none. Waterfowl species in Cluster B were grouped based on the expressed *N*-glycans having *m/z* values below 1800 *m/z*. These two major clusters were further sub-grouped into 4 different clusters (Clusters 1A, 2A, 1B and 2B). Clusters 1A and 2A split from

the branch of Cluster A, while Clusters 1B and 2B split from Cluster B. The branch differentiation between Clusters 1A and 2A is maybe due to the expression of the mono-sulfated hybrid *N*-glycan (A37) having a monoisotopic mass of 1783 *m/z*. Clusters 1B and 2B were grouped according to the relative abundances of 1297 *m/z* (A8) and 1459 *m/z* (A16) expressions. Interestingly, Anseriformes species belonged to Cluster A is a mixture of the four (4) waterfowl families, while Cluster B was exclusively Anatidae family species.

Principal component analysis (PCA) was performed to visualize the variance of the acidic *N*-glycans expressed on waterfowl egg whites. The PCA is an unsupervised model that finds natural variation in the data without over-fitting on new vectors (principal components, PCs). The observations (waterfowl species) are displayed in the score plots while the variables (acidic *N*-glycans) are shown as loading plots. The analysis resulted in a PCA model explaining 90.52% of the variance within the dataset using 8 principal components. Figure 3.4A displays the score plot using PC1 (29.9%), PC2 (19.49%) and PC3 (15.37%) which accounts for 64.80% of the total



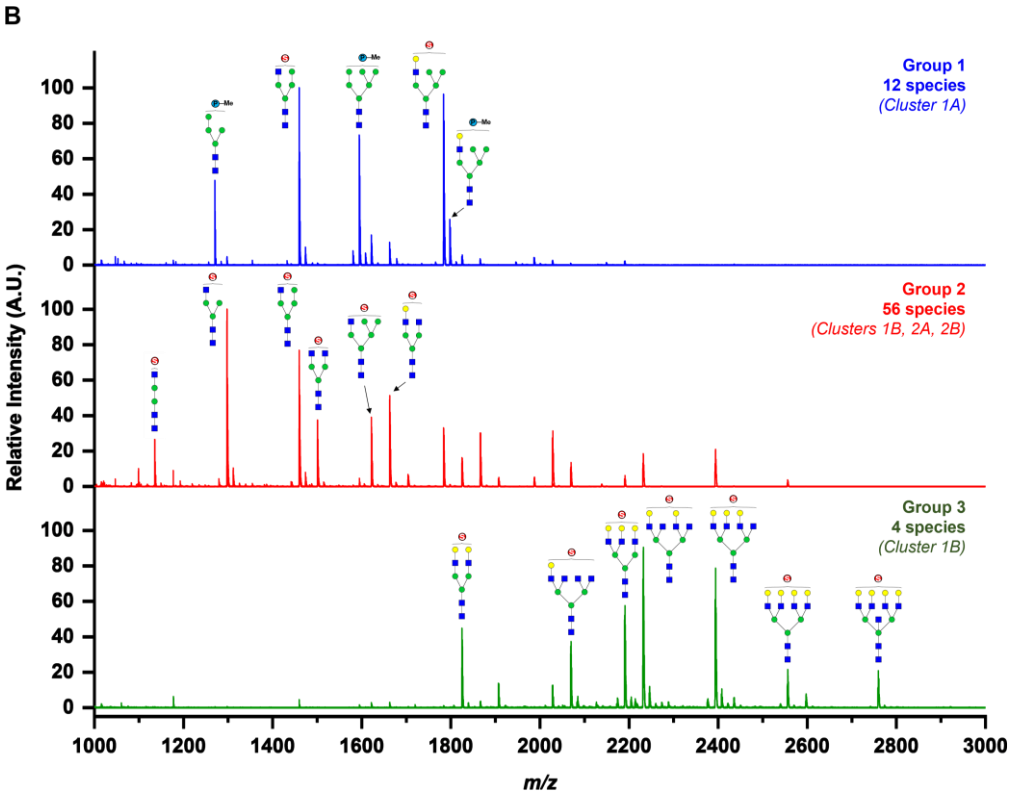
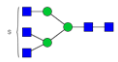
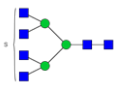
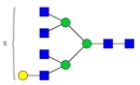
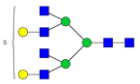
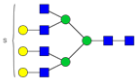
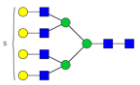


Figure 3.4. (A) Principal component analysis of the acidic *N*-glycans expressed on egg whites from 72 waterfowl species. The model resulted in 8 principal components (PCs) explaining 90.52% of the variance within the dataset. The 72 waterfowl species were also classified into 3 groups based on their acidic *N*-glycan expressions. (B) Representative MALDI-TOF MS spectra of each groups revealed specific acidic *N*-glycan species that cause the variation between groups.

variance between the differential expressions of acidic *N*-glycans in egg whites from the 79 waterfowl species. The 79 waterfowl species were classified into three major groups, wherein 12 species clustered in Group 1 (Blue), while Group 2 (Red) had 56 species and Group 3 (Green) only had 4 species. Interestingly, waterfowl species belonging to Group 1 are the same species that belonged to Cluster 1A from hierarchical clustering analysis. Group 2 is a mix of species that belong to Clusters 2A, 1B and 2B. While Group 3 species belonged to a specific branch of Cluster

Table 3.1. List of acidic *N*-glycans that gave the variation between groups in PCA.

PCA Group No.	Glycan ID	Observed mass, *m/z* $[M-H]^-$	[1]Glycan composition	Probable structures	[2]trans-Gal structure type
1	A6	1270.431	Hex1 PMe1 + Core		trans-Gal(-)
	A22	1594.528	Hex3 PMe1 + Core		trans-Gal(-)
	A37	1783.585	Hex3 HexNAc1 Su1 + Core		trans-Gal(+)
	A39	1797.610	Hex3 HexNAc1 PMe1 + Core		trans-Gal(+)
2	A3	1135.373	Hex2 HexNAc3 Su1		trans-Gal(-)
	A8	1297.420	HexNAc1 Su1 + Core		trans-Gal(-)
	A16	1459.475	Hex1 HexNAc1 Su1 + Core		trans-Gal(-)
	A18	1500.497	HexNAc2 Su1 + Core		trans-Gal(-)
	A25	1621.533	Hex2 HexNAc1 Su1 + Core		trans-Gal(-)
	A28	1662.556	Hex1 HexNAc2 Su1 + Core		trans-Gal(+)
	A30	1703.584	HexNac3 Su1 + Core		trans-Gal(-)
	A44	1865.620	Hex1 HexNAc3 Su1 + Core		trans-Gal(+)
3	A42	1824.608	Hex2 HexNAc2 Su1 + Core		trans-Gal(+)
	A47	1906.655	HexNAc4 Su1 + Core		trans-Gal(-)
	A58	2027.682	Hex2 HexNAc3 Su1 + Core		trans-Gal(+)
	A61	2068.719	Hex1 HexNAc4 Su1 + Core		trans-Gal(+)
	A70	2189.759	Hex3 HexNAc3 Su1 + Core		trans-Gal(+)
	A73	2230.794	Hex2 HexNAc4 Su1 + Core		trans-Gal(+)
	A80	2392.895	Hex3 HexNAc4 Su1 + Core		trans-Gal(+)
	A85	2555.030	Hex4 HexNAc4 Su1 + Core		trans-Gal(+)

[1]Monosaccharide nomenclatures are based on the SNFG: Hexose (Hex), *N*-acetyl hexosamine (HexNAc), Sulfate (Su), methylated Phosphate (PMe) and *N*-glycan core Man3GlcNAc2 (Core). The number of units corresponding to each monosaccharide are indicated after each abbreviation.

[2]The definition of *trans*-Gal(+/-) classifications were based on Hirose, *et al.* [18].

2A. The group classification of every waterfowl species was due to some specific acidic *N*-glycans expressed on their respective egg whites. Figure 3.4B displays a representative MALDI-TOF MS spectra of Groups 1, 2 and 3 which clearly shows the acidic *N*-glycans that cause the variation of each group. Phosphorylated high-mannose and sulfated hybrid *N*-glycans were observed in Group 1. Complex and hybrid type sulfated *N*-glycans without terminal galactose were the abundant glycan structures present in Group 2. Group 3 expressed multi-antennary complex *N*-glycans having terminal galactose on their antennae. It was also noted that some waterfowl species belonged to Group 2 expressed multi-antennary complex *N*-glycans similar to Group 3 but of lower abundance. The list of acidic *N*-glycan species that gave the variance between each group on PCA analysis is listed in Table 3.1. Furthermore, each *N*-glycan structure in Table 3.1 was classified as *trans*-Gal(+/−) based on Hirose *et. al.*[18] classification of *N*-glycans expressed in avian egg whites. Accordingly, *trans*-Gal(+) are *N*-glycan structures with terminal galactose, which are abundantly expressed in large waterfowls as hyperbranched structures. While *trans*-Gal(−) are *N*-glycans with little to no Gal-modified structures on their reducing terminus, commonly expressed in small waterfowl egg whites.

3.3.4 Phosphorylated *N*-glycans may suggest IAV prevalence in waterfowl. Previous studies revealed that host phylogeny is a crucial driver in the influenza A virus (IAV) host range. Indicating that phylogeny may be an essential factor in host-virus co-evolution which may explain the variability in the host response to infection[36]. Furthermore, species variation in IAV prevalence is often associated with host susceptibility and ecology[37, 38]. Since glycans are traits that varies between organisms as a direct response to physiological and ecological conditions, their structures are physical records due to genetic and environmental influences[39]. Here, we scrutinized the acidic *N*-glycan expressions of egg whites from waterfowl species with respect to the species virus prevalence and lineage. The virus prevalence (VP) data used in this study was taken from the published work of Wille, M. *et al.*[36] and Olsen, B. *et al.*[40]. Virus prevalence

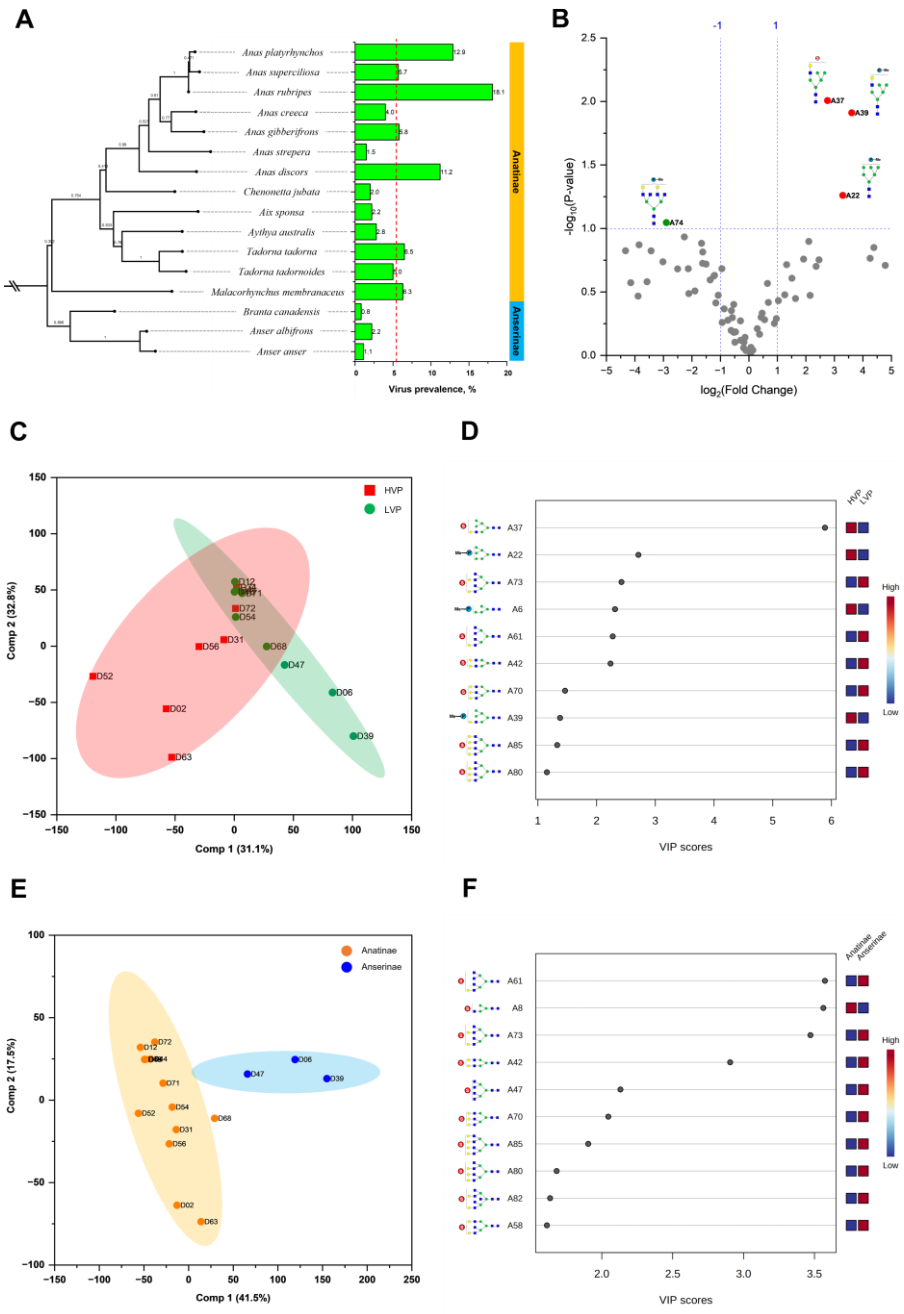


Figure 3.5. (A) Phylogenetic tree of 16 waterfowl species was reconstructed based on the mitochondrial Cytochrome *b* (*cty b*) gene using Maximum Likelihood method and Tamura-Nei model (TN93 G) in MEGA11. Influenza virus prevalence (VP) of each species is shown as bar-graph[36, 40]. (B) Differential expression of acidic *N*-glycans between waterfowl species with low viral prevalence (LVP) and high viral prevalence (HVP) is shown as volcano plot. Partial least square-discriminant analysis (PLS-DA) of the 16 waterfowl species based on viral prevalence (C and D); and based on their lineage classification (E and F). Variable important in projection (VIP) scores show important glycoforms that strongly influenced the PLS-DA score plot.

data of Anseriformes species with greater than 50 samples were used. A total of 16 Anatidae species belonging to Anatinae (13) and Anserinae (3) sub-families were included. The phylogenetic tree of these 16 waterfowl species was then reconstructed based on the mitochondrial Cytochrome *b* (cty *b*) gene[41–43] (Figure S3.2). The gene sequences were obtained from NCBI GenBank then analyzed using Maximum Likelihood method and Tamura-Nei model (TN93 G) in MEGA11[44, 45]. The tree was rooted using *Struthio camelus* (ostrich) as an out-group. Waterfowl species were then classified based on their VPs as low viral prevalence (LVP) and high viral prevalence (HVP). Classification was made by taking the average value of the 16 VPs. Based on the average VP (5.50%), species with VPs above 5.50% was classified as HVP, while below it was classified as LVP. It can be clearly inferred from Figure 3.5A that Anatinae species generally have higher viral prevalence compared to Anserinae species. Furthermore, *Anas rubripes* (D34, 18.1%), *Anas platyrhynchos* (D2, 12.9%) and *Anas discors* (D52, 11.2%) are the top 3 species with HVP and belonged to genus Anas.

Acidic *N*-glycan expressions of egg whites from the 16 waterfowl species were further analyzed according to their VP classifications (LVP and HVP). Fold change (FC) analysis of the glycan expression between HVP and LVP groups shown as volcano plot (Fig 3.5B) revealed three acidic *N*-glycans that were highly expressed in egg whites of waterfowl species with high virus prevalence. These differentially expressed *N*-glycans are 1594 *m/z* (A22, phosphorylated high-mannose), 1783 *m/z* (A37, sulfated hybrid *N*-glycan) and 1797 *m/z* (A39, phosphorylated hybrid *N*-glycan). Furthermore, partial least square-discriminant analysis (PLS-DA) was used to assess the significant difference of *N*-glycans expressed between species belonging to HVP and LVP groups. Figure 3.5C shows the PLS-DA score plot using PC1 (31.1%) and PC2 (32.8%) which explains the 63.9% of the total variance between the groups. Variable importance to projection (VIP) scores was also determined to identify the acidic *N*-glycans that strongly influence the separation of the two groups in the PLS-DA score plot. Acidic *N*-glycans with VIP scores > 1 are shown in Figure 3.5D. The expression of these acidic *N*-glycans in waterfowl egg whites provided

the discriminating information between HVP and LVP groups. Similar to FC analysis, VIP scores identified *m/z* 1594, 1783 and 1797 as important discriminating variables for species belonging to HVP. In addition, another phosphorylated high-mannose *N*-glycan (A6, 1270 *m/z*) was also identified as an important variable for HVP. While VIP scores identified bi-, tri- and tetra-antennary sulfated *trans*-Gal(+) *N*-glycans as important variables for waterfowl species with low virus prevalence. Interestingly, these sulfated multi-antennary *trans*-Gal(+) *N*-glycans are identified as characteristic *N*-glycan structures expressed by waterfowl species belonging to Anserinae sub-family (Fig 3.5E and 3.5F) with relatively low percentage of viral prevalence (Fig 3.5A).

The identification of phosphorylated high-mannose and hybrid *N*-glycan structures as possible determinants of influenza A virus prevalence in waterfowl species offers a profound understanding of the crucial role acidic *N*-glycans (*i.e.*, sulfated and phosphorylated) play as receptors for IAV tropism and infection. In Figure 3.6, when combined with phylogenetic and virus prevalence information, glycomics data sheds light on the intricate dynamics of IAV in waterfowl populations.

Notably, the expression of acidic *N*-glycans in waterfowl egg whites implies a fascinating process of host and virus evolutionary adaptation. Waterfowl species with a high virus prevalence have evolved to express sulfated *trans*-Gal(−) *N*-glycan structures, an adaptation that aids the species in evading IAV infection. The absence of terminal galactose expression in small waterfowls prevents efficient binding of IAV to sulfated *trans*-Gal(−) *N*-glycans, further reinforcing the evolutionary need for alternative interactions.

Consequently, IAV has adapted and selectively binds to phosphorylated *N*-glycans, as these glycan structures are predominantly expressed by Anatidae species exhibiting high virus prevalence. This fascinating co-evolutionary relationship between the virus and its hosts highlights the complex interplay of molecular interactions governing viral infection.

In contrast, large waterfowls continuously express sulfated *trans*-Gal(+) *N*-glycans due to low virus prevalence in these species. Intriguingly, these glycan structures are expressed since large waterfowls may have a robust defense mechanism and thus can mount a more vigorous immune defense than smaller waterfowls, as IAV effectively binds to these structures, rendering further viral adaptation unnecessary.

However, care must be taken when interpreting these results because the virus prevalence data used in the study was limited to 16 Anatidae species. Furthermore, the glycomics data was limited to waterfowl species with available VP data. We believe that these findings are significant in understanding the dynamics of IAV infection in waterfowl, but it may not completely reflect the 174 waterfowl species (53 Genera) belonging to Order Anseriformes.

Furthermore, implications of this work extend beyond avian populations, offering valuable insights into influenza transmission and infection dynamics, potentially affecting humans as well. This study may improve surveillance and control strategies by understanding the intricate relationships between viral infection and glycan structures, mitigating the risk of zoonotic transmission from birds to humans.

Finally, the study of acidic *N*-glycan structures in waterfowl species unravels the important role of acidic *N*-glycans as determinants of IAV prevalence. Phylogenetic and prevalence data complemented by glycomics data, this research strengthens our grasp of the co-evolutionary adaptations between the virus and its avian hosts, illuminating the molecular intricacies governing IAV dynamics in waterfowl populations.

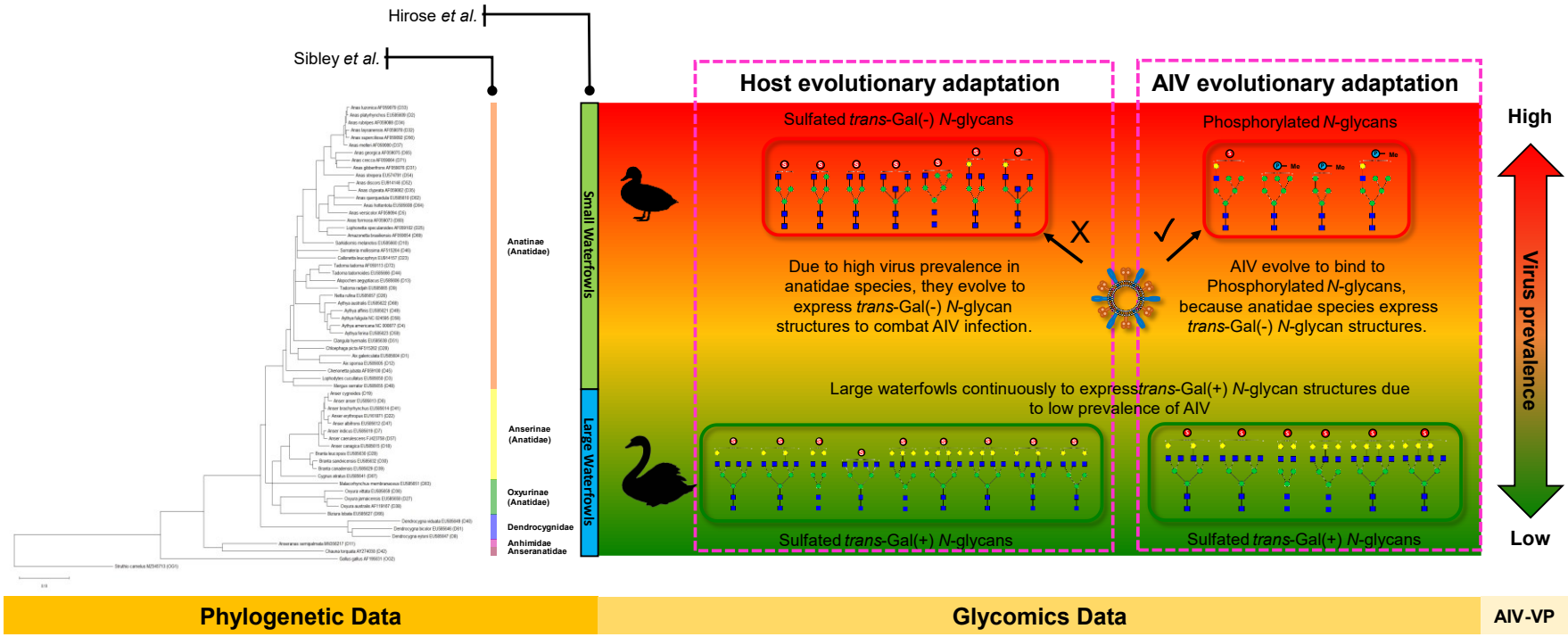


Figure 3.6. Glycomics data complements phylogenetic and virus prevalence data and together provides important insights on the dynamics of IAV in waterfowl species. The phylogenetic tree was reconstructed based on the mitochondrial Cytochrome *b* (*Cty b*) gene of 60 Anseriformes species obtained from NCBI GenBank.

3.4 Conclusion

The Glycoblotting-based sulphoglycomics approach revealed a diverse array of sulfated and phosphorylated *N*-glycans in waterfowl egg whites, providing meaningful insights into influenza A virus (IAV) dynamics. Distinct variations in acidic *N*-glycan expressions were observed among the four families (Anhimidae, Anseranatidae, Dendrocygnidae, and Anatinae) within the order Anseriformes. Waterfowl species were differentiated based on their expressions of sulfated trans-Gal(+) and trans-Gal(-) *N*-glycan structures, as well as phosphorylated *N*-glycans. Moreover, phosphorylated hybrid and high-mannose *N*-glycans were found to be highly expressed in the egg whites of waterfowl species with a high prevalence of the virus. These findings underscore the importance of phosphorylated *N*-glycans, in addition to sulfated *N*-glycans, in comprehending the dynamics of IAV in waterfowl species. Understanding these glycan structures provides valuable insights into the factors influencing the transmission and evolution of IAV within avian populations.

3.5 References

1. Sakai, T., Nishimura, S. I., Naito, T., & Saito, M. (2017). Influenza A virus hemagglutinin and neuraminidase act as novel motile machinery. *Sci Rep* 7,

2. Horimoto, T., & Kawaoka, Y. (2005). Influenza: Lessons from past pandemics, warnings from current incidents. *Nat Rev Microbiol* 3, 591–600

3. Broszeit, F., van Beek, R. J., Unione, L., Bestebroer, T. M., Chapla, D., Yang, J. Y., Moremen, K. W., Herfst, S., Fouchier, R. A. M., de Vries, R. P., & Boons, G. J. (2021). Glycan remodeled erythrocytes facilitate antigenic characterization of recent A/H3N2 influenza viruses. *Nat Commun* 12,

4. Suzuki, N., Abe, T., & Natsuka, S. (2022). Structural analysis of N-glycans in chicken trachea and lung reveals potential receptors of chicken influenza viruses. *Sci Rep* 12,

5. Zhao, C., & Pu, J. (2022). Influence of Host Sialic Acid Receptors Structure on the Host Specificity of Influenza Viruses. *Viruses* 14,

6. Kobayashi, D., Hiono, T., Ichii, O., Nishihara, S., Takase-Yoden, S., Yamamoto, K., Kawashima, H., Isoda, N., & Sakoda, Y. (2022). Turkeys possess diverse Siaα2-3Gal glycans that facilitate their dual susceptibility to avian influenza viruses isolated from ducks and chickens. *Virus Res* 315,

7. Yang, J., Cui, H., Teng, Q., Ma, W., Li, X., Wang, B., Yan, D., Chen, H., Liu, Q., & Li, Z. (2019). Ducks induce rapid and robust antibody responses than chickens at early time after intravenous infection with H9N2 avian influenza virus. *Virol J* 16,

8. Vreman, S., Bergervoet, S. A., Zwart, R., Stockhofe-Zurwieden, N., & Beerens, N. (2022). Tissue tropism and pathology of highly pathogenic avian influenza H5N6 virus in chickens and Pekin ducks. *Res Vet Sci* 146, 1–4

9. Sid, H., Hartmann, S., Winter, C., & Rautenschlein, S. (2017). Interaction of influenza A viruses with oviduct explants of different avian species. *Front Microbiol* 8, 1–11

10. Pillai, S. P. S., Saif, Y. M., & Lee, C. W. (2010). Detection of Influenza A Viruses in Eggs Laid by Infected Turkeys. *Avian Disease* 54, 830–833

11. Uchida, Y., Takemae, N., Tanikawa, T., Kanehira, K., & Saito, T. (2016). Transmission of an H5N8-Subtype Highly Pathogenic Avian Influenza Virus from Infected Hens to Laid Eggs. *Avian Dis* 60, 450–453

12. Hiono, T., Okamatsu, M., Nishihara, S., Takase-Yoden, S., Sakoda, Y., & Kida, H. (2014). A chicken influenza virus recognizes fucosylated α2,3 sialoglycan receptors on the epithelial cells lining upper respiratory tracts of chickens. *Virology* 456–457, 131–138

13. Gambaryan, A., Yamnikova, S., Lvov, D., Tuzikov, A., Chinarev, A., Pazynina, G., Webster, R., Matrosovich, M., & Bovin, N. (2005). Receptor specificity of influenza viruses from birds and mammals: New data on involvement of the inner fragments of the carbohydrate chain. *Virology* 334, 276–283

14. Ichimiya, T., Okamatsu, M., Kinoshita, T., Kobayashi, D., Ichii, O., Yamamoto, N., Sakoda, Y., Kida, H., Kawashima, H., Yamamoto, K., Takase-Yoden, S., & Nishihara, S. (2021). Sulfated glycans containing NeuAcα2-3Gal facilitate the propagation of human H1N1 influenza A viruses in eggs. *Virology* 562, 29–39

15. Ichimiya, T., Nishihara, S., Takase-Yoden, S., Kida, H., & Aoki-Kinoshita, K. (2014). Frequent glycan structure mining of influenza virus data revealed a sulfated glycan motif that increased viral infection. *Bioinformatics* 30, 706–711

16. Broszeit, F., Tzarum, N., Zhu, X., Nemanichvili, N., Eggink, D., Leenders, T., Li, Z., Liu, L., Wolfert, M. A., Papanikolaou, A., Martínez-Romero, C., Gagarinov, I. A., Yu, W., García-Sastre, A., Wennekes, T., Okamatsu, M., Verheije, M. H., Wilson, I. A., Boons, G. J., & de Vries, R. P. (2019). N-Glycolylneuraminic Acid as a Receptor for Influenza A Viruses. *Cell Rep* 27, 3284-3294.e6

17. Hincke, M. T., Da Silva, M., Guyot, N., Gautron, J., McKee, M. D., Guabiraba-Brito, R., & Réhault-Godbert, S. (2019). Dynamics of Structural Barriers and Innate Immune Components during Incubation of the Avian Egg: Critical Interplay between Autonomous Embryonic Development and Maternal Anticipation. *J Innate Immun* 11, 111–124

18. Hirose, K., Amano, M., Hashimoto, R., Lee, Y. C., & Nishimura, S. I. (2011). Insight into glycan diversity and evolutionary lineage based on comparative avio-N-glycomics and sialic acid analysis of 88 egg whites of galloanserae. *Biochemistry* 50, 4757–4774

19. Laskowski, M., Kato, I., Ardelt, W., Cook, J., Denton, A., Empie, M. W., Kohr, W. J., Soon, #, Park, J., Parks, K., Schatzley, B. L., Schoenberger, O. L., Tashiro, M., Vichot, G., Whatley, H. E., Wieczorek, A., & Wieczorek, M. (1987). Ovomucoid Third Domains from 100 Avian Species: Isolation, Sequences, and Hypervariability of Enzyme-Inhibitor Contact Residues. *Biochemistry* 26, 202–221

20. Laskowski, M., Apostol, I., Ardelt, W., Cook, J., Giletto, A., Kelly, C. A., Lu, W., Park, S. J., Qasim, M. A., Whatley, H. E., Wieczorek, A., & Wynn, R. (1990). Amino Acid Sequences of Ovomucoid Third Domain from 25 Additional Species of Birds. *J Protein Chem* 9,

21. Apostol, I., Giletto, A., Komiyama, T., Zhang, W., & Laskowski, M. (1993). Amino Acid Sequences of Ovomucoid Third Domains from 27 Additional Species of Birds. *J Protein Chem* 12,

22. Sibley, C. G., & Monroe, B. L. (1990). *Distribution and Taxonomy of Birds of the World* (Yale University Press, New Haven, CT.)

23. Sibley, C. G., & Monroe, B. L. (1993). *A Supplement to Distribution and Taxonomy of Birds of the Worlds* (Yale University Press, New Haven, CT)

24. Sanes, J. T., Hinou, H., Lee, Y. C., & Nishimura, S. I. (2019). Glycoblotting of Egg White Reveals Diverse N-Glycan Expression in Quail Species. *J Agric Food Chem* 67, 531–540

25. Gizaw, S. T., Ohashi, T., Tanaka, M., Hinou, H., & Nishimura, S. I. (2016). Glycoblotting method allows for rapid and efficient glycome profiling of human Alzheimer's disease brain, serum and cerebrospinal fluid towards potential biomarker discovery. *Biochim Biophys Acta Gen Subj* 1860, 1716–1727

26. Miura, Y., Shinohara, Y., Furukawa, J. I., Nagahori, N., & Nishimura, S. I. (2007). Rapid and simple solid-phase esterification of sialic acid residues for quantitative glycomics by mass spectrometry. *Chemistry - A European Journal* 13, 4797–4804

27. Yu, S.-Y., Snovida, S., & Khoo, K.-H. (2020). Permethylation and Microfractionation of Sulfated Glycans for MS Analysis. *Bio Protoc* 10, 1–13

28. Yu, S. Y., Wu, S. W., Hsiao, H. H., & Khoo, K. H. (2009). Enabling techniques and strategic workflow for sulfoglycomics based on mass spectrometry mapping and sequencing of permethylated sulfated glycans. *Glycobiology* 19, 1136–1149

29. Hinou, H. (2019). Aniline derivative/DHB/alkali metal matrices for reflectron mode MALDI-TOF and TOF/TOF MS analysis of unmodified sialylated oligosaccharides and glycopeptides. *Int J Mass Spectrom* 443, 109–115

30. Barada, E., & Hinou, H. (2022). BOA/DHB/Na: An Efficient UV-MALDI Matrix for High-Sensitivity and Auto-Tagging Glycomics. *Int J Mol Sci* 23, 12510

31. Chen, J. Y., Huang, H. H., Yu, S. Y., Wu, S. J., Kannagi, R., & Khoo, K. H. (2018). Concerted mass spectrometry-based glycomic approach for precision mapping of sulfo sialylated N-glycans on human peripheral blood mononuclear cells and lymphocytes. *Glycobiology* 28, 9–20

32. Kuo, C. W., Guu, S. Y., & Khoo, K. H. (2018). Distinctive and Complementary MS 2 Fragmentation Characteristics for Identification of Sulfated Sialylated N-Glycopeptides by nanoLC-MS/MS Workflow. *J Am Soc Mass Spectrom* 29, 1166–1178

33. Ceroni, A., Maass, K., Geyer, H., Geyer, R., Dell, A., & Haslam, S. M. (2008). GlycoWorkbench: A tool for the computer-assisted annotation of mass spectra of glycans. *J Proteome Res* 7, 1650–1659

34. Cooper, C. A., Gasteiger, E., & Packer, N. H. (2001). GlycoMod-A software tool for determining glycosylation compositions from mass spectro-metric data. *Proteomics* 1, 340–349

35. Montalban, B. M., & Hinou, H. (2023). Glycoblotting enables seamless and straightforward workflow for MALDI-TOF/MS-based sulphoglycomics of N- and O-glycans. *Proteomics*, 1–10

36. Wille, M., Lisovski, S., Roshier, D., Ferenczi, M., Hoye, B. J., Leen, T., Warner, S., Fouchier, R. A. M., Hurt, A. C., Holmes, E. C., & Klaassen, M. (2023). Strong host phylogenetic and ecological effects on host competency for avian influenza in Australian wild birds. *Proceedings of the Royal Society B: Biological Sciences* 290, 1–9

37. van Dijk, J. G., Verhagen, J. H., Wille, M., & Waldenström, J. (2018). Host and virus ecology as determinants of influenza A virus transmission in wild birds. *Curr Opin Virol* 28, 26–36

38. Garamszegi, L. Z., & Møller, A. P. (2007). Prevalence of avian influenza and host ecology. *Proceedings of the Royal Society B: Biological Sciences* 274, 2003–2012

39. Springer, S. A., & Gagneux, P. (2016). Glycomics: revealing the dynamic ecology and evolution of sugar molecules. *J Proteomics* 135, 90–100

40. Olsen, B., Munster, V. J., Wallensten, A., Waldenström, J., Osterhaus, A. D. M. E., & Fouchier, R. A. M. (2006). Global Patterns of Influenza A Virus in Wild Birds. *Science (1979)* 312, 384–388

41. Gonzalez, J., Düttmann, H., & Wink, M. (2009). Phylogenetic relationships based on two mitochondrial genes and hybridization patterns in Anatidae. *J Zool* 279, 310–318

42. Donne-Gousse, C., Laudet, V., & Hanni, C. (2002). A molecular phylogeny of anseriformes based on mitochondrial DNA analysis. *Molecular Phylogenetics and Evolution* 23, 339–356

43. Sun, Z., Pan, T., Hu, C., Sun, L., Ding, H., Wang, H., Zhang, C., Jin, H., Chang, Q., Kan, X., & Zhang, B. (2017). Rapid and recent diversification patterns in Anseriformes birds: Inferred from molecular phylogeny and diversification analyses. *PLoS One* 12, 1–21

44. Tamura, K., & Nei, M. (1993). Estimation of the Number of Nucleotide Substitutions in the Control Region of Mitochondrial DNA in Humans and Chimpanzees ’. *Mol Biol Evol* 10, 512–526

45. Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol Biol Evol* 38, 3022–3027

3.6 Supplementary Information

Table S3.1. The list of egg whites from various species of Order Anseriformes (4 families, 27 genera, 66 species) used in this study. Classification was based on Sibley's DNA-DNA hybridization[22, 23].

Sample ID	Scientific name	Family	Sub-Family	Common Name
D1	*Aix galericulata*	Anatidae	Anatinae	Mandarin Duck
D2	*Anas platyrhynchos*	Anatidae	Anatinae	Mallard Duck
D3	*Lophodytes cucullatus*	Anatidae	Anatinae	Hooded Mergenser
D4	*Aythya americana*	Anatidae	Anatinae	Red head
D5	*Anas versicolor*	Anatidae	Anatinae	Silver Teal
D6	*Anser anser*	Anatidae	Anserinae	Graylag Goose
D7	*Anser indicus*	Anatidae	Anserinae	Bar Headed Goose
D8	*Dendrocygna eytoni*	Dendrocygnidae	Dendrocygninae	Plumed Whistling Duck
D9	*Tadorna radjah*	Anatidae	Anserinae	White Headed Shelduck
D10	*Sarkidiornis melanotos*	Anatidae	Anserinae	Knob-billed Goose
D11	*Anseranas semipalmata*	Anseranatidae	Anseranatidae	Magpie Goose
D12	*Aix sponsa*	Anatidae	Anatinae	Wood Duck
D13	*Alopochen aegyptiaca*	Anatidae	Anserinae	Egyptian Goose
D14	*Anas platyrhynchos domesticus*	Anatidae	Anatinae	White Call Duck
D15	*Anser anser domesticus (America)*	Anatidae	Anserinae	Buff Goose
D16	*Anser anser domesticus (France)*	Anatidae	Anserinae	Dewlap Toulouse Goose
D17	*Anser anser domesticus (Germany)*	Anatidae	Anserinae	Embdens Goose
D18	*Anser canagicus*	Anatidae	Anserinae	Emperor Goose
D19	*Anser cygnoides domesticus*	Anatidae	Anserinae	African Goose
D20	*Anser cygnoides domesticus*	Anatidae	Anserinae	White China Goose
D21	*Anser cygnoides domesticus*	Anatidae	Anserinae	Brown China Goose
D22	*Anser erythropus*	Anatidae	Anserinae	Lesser white-fronted Goose
D23	*Callonetta leucophrys*	Anatidae	Anatinae	Ringed Teal
D24	*Dendrocygna arborea*	Dendrocygnidae	Dendrocygninae	West Indian Whistling Duck
D25	*Lophonetta speculariodes*	Anatidae	Anatinae	Crested Duck
D26	*Netta rufina*	Anatidae	Anatinae	Red Crested Pochard
D27	*Oxyura jamaicensis*	Anatidae	Oxyurinae	Ruddy duck
D28	*Chloephaga picta picta*	Anatidae	Anserinae	Magellan Goose
D29	*Branta leucopsis*	Anatidae	Anserinae	Barnacle Goose
D30	*Branta sandvicensis*	Anatidae	Anserinae	Hawaiian Goose
D31	*Anas gibberifrons*	Anatidae	Anatinae	Indonesian Teal
D32	*Anas laysanensis*	Anatidae	Anatinae	Laysan Duck
D33	*Anas luzonica*	Anatidae	Anatinae	Philippine Duck
D34	*Anas rubripes*	Anatidae	Anatinae	American Black Duck
D35	*Anas clypeata*	Anatidae	Anatinae	Northern Shoveler
D36	*Oxyura vittata*	Anatidae	Oxyurinae	Lake Duck
D37	*Anas melleri*	Anatidae	Anatinae	Meller's Duck
D38	*Oxyura australis*	Anatidae	Oxyurinae	Blue-billed Duck
D39	*Branta canadensis maxima*	Anatidae	Anserinae	Canada Goose
D40	*Dendrocygna viduata*	Dendrocygnidae	Dendrocygninae	White-faced Whistling Duck
D41	*Anser brachyrhynchus*	Anatidae	Anserinae	Pink-footed Goose
D42	*Chauna torquata*	Anhimidae	Anhimidae	Southern Screamer
D43	*Thalassornis leuconotos*	Dendrocygnidae	Dendrocygninae	White-backed Duck
D44	*Tadorna tadornoides*	Anatidae	Anserinae	Australian Shelduck
D45	*Chenonetta jubata*	Anatidae	Anatinae	Australian Wood Duck
D46	*Somateria mollissima*	Anatidae	Anatinae	Common Eider
D47	*Anser albifrons*	Anatidae	Anserinae	Greater whitefronted Goose
D48	*Mergus serrator*	Anatidae	Anatinae	Red-breasted Merganser
D49	*Aythya affinis*	Anatidae	Anatinae	Lesser Scaup
D50	*Dendrocygna autumnalis*	Dendrocygnidae	Dendrocygninae	Black-bellied whistling Duck
D51	*Clangula hyemalis*	Anatidae	Anatinae	Longtailed Duck
D52	*Anas discors*	Anatidae	Anatinae	Blue-winged Teal
D53	*Oxyura punctata*	Anatidae	Oxyurinae	
D54	*Anas strepera*	Anatidae	Anatinae	Gadwall
D55	*Heteronetta atricapilla*	Anatidae	Anatinae	Black-headed Duck
D56	*Anas superciliosa*	Anatidae	Anatinae	Pacific black Duck

D57	*Anser caerulescens*	Anatidae	Anserinae	Snow Goose
D58	*Aythya fuligula*	Anatidae	Anatinae	Tufted Duck
D59	*Aythya ferina*	Anatidae	Anatinae	Common Pochard
D60	*Anas formosa*	Anatidae	Anatinae	Baikal Teal
D61	*Dendrocygna bicolor*	Dendrocygnidae	Dendrocygninae	Fulvous Whistling Duck
D62	*Anas querquedula*	Anatidae	Anatinae	Garganey
D63	*Malacorhynchus membranaceus*	Anatidae	Anatinae	Pink-eared Duck
D64	*Anas hottentota*	Anatidae	Anatinae	Hottentot Teal
D65	*Anas georgica*	Anatidae	Anatinae	Yellow-billed Pintail
D66	*Biziura lobata*	Anatidae	Oxyurinae	Musk Duck
D67	*Cygnus atratus*	Anatidae	Cygninae	Black swan
D68	*Aythya australis*	Anatidae	Anatinae	Hardhead
D69	*Amazonetta brasiliensis*	Anatidae	Anatinae	Brazilian Teal
D70	*Lophonetta cristata*	Anatidae	Anatinae	
D71	*Anas crecca*	Anatidae	Anatinae	Common Teal
D72	*Tadorna tadorna*	Anatidae	Anserinae	Common Shelduck

Table S3.2. List of 89 acidic *N*-glycans identified from the egg whites of 79 waterfowl species. Glycan structures were inferred from glycan composition based on the observed monoisotopic masses.

Glycan ID	Observed mass, m/z $[M-H]^-$	Calculated Glycoform Mass [M – BOA]	Mass Error, ppm	Monosaccharide composition	Glyconnect Database Links
A1	1094.339	972.274	-10.8	Hex3 HexNAc2 Su1	
A2	1108.327	972.284	-30.5	Hex3 HexNAc2 Pho1	Glyconnect
A3	1135.373	1013.301	-4.2	Hex2 HexNAc3 Su1	
A4	1149.392	1013.310	4.8	Hex2 HexNAc3 Pho1	
A5	1256.320	1134.327	-66.7	Hex4 HexNAc2 Su1	
A6	1270.431	1134.336	14.1	Hex4 HexNAc2 Pho1	
A7	1295.303	1159.368	-109.5	Hex2 HexNAc3 dHex1 Pho1	
A8	1297.420	1175.353	-7.5	HexNAc1 Su1 + Man3 GlcNAc2	Glyconnect
A9	1311.445	1175.363	3.7	HexNAc1 Pho1 + Man3 GlcNAc2	
A10	1338.419	1216.380	-28.2	Hex2 HexNAc4 Su1	
A11	1352.412	1216.390	-40.5	Hex2 HexNAc4 Pho1	
A12	1418.430	1296.380	-18.9	Hex2 Su1 + Man3 GlcNAc2	
A13	1432.469	1293.389	2.4	Hex2 Pho1 + Man3 GlcNAc2	Glyconnect
A14	1441.498	1305.426	-3.3	Unknown structure	
A15	1443.485	1321.411	-1.9	HexNAc1 dHex1 Su1 + Man3 GlcNAc2	
A16	1459.475	1337.406	-5.3	Hex1 HexNAc1 Su1 + Man3 GlcNAc2	
A17	1473.505	1337.416	8.1	Hex1 HexNAc1 Pho1 + Man3 GlcNAc2	
A18	1500.497	1378.433	-8.5	HexNAc2 Su1 + Man3 GlcNAc2	Glyconnect
A19	1514.530	1378.442	7.5	HexNAc2 Pho1 + Man3 GlcNAc2	
A20	1580.496	1458.432	-8.1	Hex3 Su1 + Man3 GlcNAc2	Glyconnect
A21	1589.525	1467.469	-13.1	HexNAc1 dHex2 Su1 + Man3 GlcNAc2	
A22	1594.528	1458.442	5.5	Hex3 Pho1 + Man3 GlcNAc2	Glyconnect
A23	1603.574	1467.479	11.4	HexNAc1 dHex2 Pho1 + Man3 GlcNAc2	
A24	1605.533	1483.464	-4.8	Hex1 HexNAc1 dHex1 Su1 + Man3 GlcNAc2	Glyconnect
A25	1621.533	1499.459	-1.7	Hex2 HexNAc1 Su1 + Man3 GlcNAc2	
A26	1635.542	1499.469	-2.2	Hex2 HexNAc1 Pho1 + Man3 GlcNAc2	
A27	1646.558	1524.491	-5.9	HexNAc2 dHex1 Su1 + Man3 GlcNAc2	Glyconnect
A28	1662.556	1540.486	-4.1	Hex1 HexNAc2 Su1 + Man3 GlcNAc2	Glyconnect
A29	1676.605	1540.495	19.8	Hex1 HexNAc2 Pho1 + Man3 GlcNAc2	
A30	1703.584	1581.512	-2.8	HexNac3 Su1 + Man3 GlcNAc2	Glyconnect
A31	1717.608	1581.522	5.6	HexNac3 Pho1 + Man3 GlcNAc2	
A32	1749.602	1613.537	-6.8	Unknown structure	
A33	1751.599	1629.522	0.1	Hex1 HexNAc1 dHex2 Su1 + Man3 GlcNAc2	
A34	1764.667	1628.502	50.0	Hex1 HexNAc1 NeuAc1 Su1 + Man3 GlcNAc2	
A35	1765.618	1629.532	5.3	Hex1 HexNAc1 dHex2 Pho1 + Man3 GlcNAc2	
A36	1767.620	1645.517	14.8	Hex2 HexNAc1 dHex1 Su1 + Man3 GlcNAc2	Glyconnect
A37	1783.585	1661.512	-2.1	Hex3 HexNAc1 Su1 + Man3 GlcNAc2	Glyconnect
A38	1792.814	1670.549	105.0	HexNAc2 dHex2 Su1 + Man3 GlcNAc2	
A39	1797.610	1661.521	6.6	Hex3 HexNAc1 Pho1 + Man3 GlcNAc2	Glyconnect
A40	1806.657	1670.558	12.5	HexNAc2 dHex2 Pho1 + Man3 GlcNAc2	
A41	1808.607	1686.544	-7.6	Hex1 HexNAc2 dHex1 Su1 + Man3 GlcNAc2	Glyconnect
A42	1824.608	1702.538	-3.7	Hex2 HexNAc2 Su1 + Man3 GlcNAc2	Glyconnect
A43	1838.661	1702.548	19.5	Hex2 HexNAc2 Pho1 + Man3 GlcNAc2	
A44	1865.620	1743.565	-11.7	Hex1 HexNAc3 Su1 + Man3 GlcNAc2	Glyconnect
A45	1879.661	1743.575	4.8	Hex1 HexNAc3 Pho1 + Man3 GlcNAc2	
A46	1888.924	1766.543	161.1	Hex4 dHex1 Su1 + Man3 GlcNAc2	
A47	1906.655	1784.592	-7.2	HexNAc4 Su1 + Man3 GlcNAc2	Glyconnect
A48	1920.682	1784.601	2.3	HexNAc4 Pho1 + Man3 GlcNAc2	
A49	1945.612	1823.565	-15.3	Hex4 HexNAc1 Su1 + Man3 GlcNAc2	
A50	1952.677	1816.616	-7.9	HexNAc2 dHex3 Pho1 + Man3 GlcNAc2	
A51	1954.866	1832.601	96.3	Hex1 HexNAc2 dHex2 Su1 + Man3 GlcNAc2	Glyconnect
A52	1959.599	1823.574	-26.5	Hex4 HexNAc1 Pho1 + Man3 GlcNAc2	Glyconnect
A53	1970.726	1848.596	26.9	Hex2 HexNAc2 dHex1 Su1 + Man3 GlcNAc2	Glyconnect
A54	1986.667	1864.591	-0.5	Hex3 HexNAc2 Su1 + Man3 GlcNAc2	
A55	2000.577	1864.601	-50.2	Hex3 HexNAc2 Pho1 + Man3 GlcNAc2	
A56	2009.653	1873.638	-30.7	HexNAc3 dHex2 Pho1 + Man3 GlcNAc2	
A57	2011.663	1889.623	-18.1	Hex1 HexNAc3 dHex1 Su1 + Man3 GlcNAc2	Glyconnect
A58	2027.682	1905.618	-6.1	Hex2 HexNAc3 Su1 + Man3 GlcNAc2	

A59	2041.713	1905.627	4.3	Hex2 HexNAc3 Pho1 + Man3 GlcNAc2	
A60	2050.704	1928.596	15.2	Hex5 dHex1 Su1 + Man3 GlcNAc2	
A61	2068.719	1946.644	-1.0	Hex1 HexNAc4 Su1 + Man3 GlcNAc2	
A62	2082.738	1946.654	3.3	Hex1 HexNAc4 Pho1 + Man3 GlcNAc2	
A63	2109.758	1987.671	5.0	HexNAc5 Su1 + Man3 GlcNAc2	Glyconnect
A64	2114.753	1978.669	3.4	Hex1 HexNAc2 dHex3 Pho1 + Man3 GlcNAc2	
A65	2129.999	1993.634	135.3	Hex2 HexNAc2 NeuAc1 Su1 + Man3 GlcNAc2	Glyconnect
A66	2146.364	2010.659	-173.2	Hex3 HexNAc2 dHex1 Pho1 + Man3 GlcNAc2	
A67	2148.660	2026.644	-28.3	Hex4 HexNAc2 Su1 + Man3 GlcNAc2	
A68	2158.155	2035.681	184.1	Hex1 HexNAc3 dHex2 Su1 + Man3 GlcNAc2	Glyconnect
A69	2171.782	2035.690	6.9	Hex1 HexNAc3 dHex2 Pho1 + Man3 GlcNAc2	
A70	2189.759	2067.671	5.1	Hex3 HexNAc3 Su1 + Man3 GlcNAc2	
A71	2203.771	2067.680	6.5	Hex3 HexNAc3 Pho1 + Man3 GlcNAc2	
A72	2212.764	2090.649	17.2	Hex6 dHex1 Su1 + Man3 GlcNAc2	
A73	2230.794	2108.697	8.8	Hex2 HexNAc4 Su1 + Man3 GlcNAc2	
A74	2244.834	2108.707	22.5	Hex2 HexNAc4 Pho1 + Man3 GlcNAc2	
A75	2271.838	2149.724	16.2	Hex1 HexNAc5 Su1 + Man3 GlcNAc2	
A76	2278.900	2156.707	51.0	Hex3 HexNAc2 dHex2 Su1 + Man3 GlcNAc2	
A77	2287.809	2165.744	-5.2	HexNAc3 dHex4 Su1 + Man3 GlcNAc2	
A78	2320.103	2197.734	126.0	Hex2 HexNAc3 dHex2 Su1 + Man3 GlcNAc2	
A79	2374.818	2252.702	16.6	Hex7 dHex1 Su1 + Man3 GlcNAc2	
A80	2392.895	2270.750	28.3	Hex3 HexNAc4 Su1 + Man3 GlcNAc2	
A81	2406.937	2270.760	41.8	Hex3 HexNAc4 Pho1 + Man3 GlcNAc2	
A82	2433.956	2311.777	42.1	Hex2 HexNAc5 Su1 + Man3 GlcNAc2	
A83	2449.933	2327.797	24.4	Hex1 HexNAc3 dHex4 Su1 + Man3 GlcNAc2	
A84	2479.957	2343.801	31.9	Hex2 HexNAc3 dHex3 Pho1 + Man3 GlcNAc2	
A85	2555.030	2432.803	58.9	Hex4 HexNAc4 Su1 + Man3 GlcNAc2	
A86	2569.031	2432.812	55.4	Hex4 HexNAc4 Pho1 + Man3 GlcNAc2	
A87	2596.066	2473.829	61.6	Hex3 HexNAc5 Su1 + Man3 GlcNAc2	
A88	2758.164	2635.882	74.5	Hex4 HexNAc5 Su1 + Man3 GlcNAc2	
A89	2920.151	2797.935	47.7	Hex5 HexNAc5 Su1 + Man3 GlcNAc2	

*Monosaccharide nomenclatures are based on the SNFG: Hexose (Hex), *N*-acetyl hexosamine (HexNAc), Mannose (Man), *N*-acetyl glucosamine (GlcNAc), Fucose (dHex), Sulfate (Su), and Phosphate (Pho). The number of units corresponding to each monosaccharide are indicated after each abbreviation.

*The links to the Glyconnect database of the Swiss Institute of Bioinformatics are provided for selected monoisotopic peaks found in the database.

*From the 89 monoisotopic masses, 55 sulfated and 34 phosphorylated *N*-glycans were identified based on their glycan composition and MS/MS analysis. Fucosylated acidic *N*-glycan structures were also found in trace abundance relative to un-fucosylated acidic *N*-glycans.

*Glycoform mass is the mass of unlabeled *N*-glycan structure denoted as [M-BOA], BOA is benzyloxyamine with a molecular mass of 123.0684 Da.

Table S3.3. Waterfowl classification based on their virus prevalence.

Duck ID	Species	PCA Group	VP Values	VP Classification
D2	*Anas platyrhynchos*	1	12.9	HVP
D56	*Anas superciliosa*	2	5.7	HVP
D34	*Anas rubripes*	2	18.1	HVP
D71	*Anas creeca*	2	4.0	LVP
D31	*Anas gibberifrons*	2	5.8	HVP
D54	*Anas strepera*	2	1.5	LVP
D52	*Anas discors*	1	11.2	HVP
D45	*Chenonetta jubata*	2	2.0	LVP
D12	*Aix sponsa*	2	2.2	LVP
D68	*Aythya australis*	2	2.8	LVP
D72	*Tadorna tadorna*	2	6.5	HVP
D44	*Tadorna tadornoides*	2	5.0	LVP
D63	*M. Membranaceus*	1	6.3	HVP
D39	*Branta canadensis*	3	0.8	LVP
D47	*Anser albifrons*	2	2.2	LVP
D06	*Anser anser*	3	1.1	LVP

*Each waterfowl species was classified either as a high virus prevalence (HVP) or low virus prevalence (LVP). Classification was based on the average virus prevalence (5.5%). LVP < 5.50% < HVP.
*Virus prevalence data of the 16 species shown on the table was taken from the work of Wille, M. *et al.* [36] and Olsen, B. *et al.* [40].

Table S3.4. GenBank accession numbers for various genes of the 72 Anseriformes species in this study.

Sample ID	Scientific name	CO1	Cty *b*	ND2	Complete mtDNA
D01	*Aix galericulata*	JN703260	EU585604	EU585667	KF437906
D02	*Anas platyrhynchos*	Mk262361	EU585609	EU585672	MN720361
D03	*Lophodytes cucullatus*		EU585650	EU585713	
D04	*Aythya americana*	DQ434316	NC_000877	NC_000877	NC_000877
D05	*Anas versicolor*	FJ027121	AF059094	AF059154	
D06	*Anser anser*	GU571243	EU585613	EU585676	NC_011196
D07	*Anser indicus*	GU571246	EU585619	EU585682	NC_025654
D08	*Dendrocygna eytoni*	MZ153330	EU585647	EU585710	
D09	*Tadorna radjah*		EU585665	EU585728	
D10	*Sarkidiornis melanotos*	FJ028237	EU585660	EU585723	
D11	*Anseranas semipalmata*	MN356217	NC_005933		MN356217
D12	*Aix sponsa*	AY666569	EU585605	EU585668	
D13	*Alopochen aegyptiaca*	Mf580159	EU585606	EU585669	
D14	*Anas platyrhynchos domesticus*				
D15	*Anser anser domesticus (America)*				
D16	*Anser anser domesticus (France)*				
D17	*Anser anser domesticus (Germany)*				
D18	*Anser canagica*	DQ432849	EU585615	EU585678	
D19	*Anser cygnoides domesticus*	LC145060	EU585616	EU585679	NC_023832
D20	*Anser cygnoides domesticus*				
D21	*Anser cygnoides domesticus*				
D22	*Anser erythropus*	GU571729	EU161871	EU585680	
D23	*Callonetta leucophrys*	FJ027277	EU914157	AF059157	
D24	*Dendrocygna arborea*				
D25	*Lophonetta speculariodes*	JN801488	AF059102	AF059162	
D26	*Netta rufina*	GQ482234	EU585657	EU585720	NC_024922
D27	*Oxyura jamaicensis*	AY666448	EU585658	EU585721	MW574354
D28	*Chloephaga picta*	FJ027353	AF515262	AF515266	
D29	*Branta leucopsis*	GU571283	EU585630	EU585693	
D30	*Branta sandvicensis*	JF498832	EU585632	EU585695	
D31	*Anas gibberifrons*	JQ174015	AF059076	AF059136	
D32	*Anas laysanensis*	JF498830	AF059078	AF059138	
D33	*Anas luzonica*	KT151721	AF059079	AF059139	
D34	*Anas rubripes*	AY666211	AF059088	AF059148	
D35	*Anas clypeata*	GU571236	AF059062	AF059122	NC_028346
D36	*Oxyura vittata*	JQ175648	EU585659	EU585722	
D37	*Anas melleri*		AF059080	AF059140	
D38	*Oxyura australis*		AF119167	AY747867	
D39	*Branta canadensis*	GU571280	EU585629	EU585692	NC_007011
D40	*Dendrocygna viduata*	FJ027502	EU585649	EU585712	
D41	*Anser brachyrhynchus*	GU571244	EU585614	EU585677	
D42	*Chauna torquata*	AY140730	AY274030	AY274053	NC_052807
D43	*Thalassornis leuconotos*	U97738			
D44	*Tadorna tadornoides*		EU585666	EU585729	
D45	*Chenonetta jubata*	JN801436	AF059100	AF059160	
D46	*Somateria mollissima*	GU571620	EU585661	EU585724	MW849292
D47	*Anser albifrons*	DQ433314	EU585612	EU585675	NC_004539
D48	*Mergus serrator*	GU571482	EU585655	EU585718	MZ365040
D49	*Aythya affinis*	DQ434308	EU585621	EU585684	
D50	*Dendrocygna autumnalis*	FJ027495			
D51	*Clangula hyemalis*	GU571339	EU585638	EU585701	MW849278
D52	*Anas discors*	AY666325	EU914146	AF059128	
D53	*Oxyura punctata*				

D54	*Anas strepera (Mareca strepera)*	GQ481327	EU574791	AF059169	NC_045373
D55	*Heteronetta atricapilla*	FJ027649			
D56	*Anas superciliosa*	JN801396	AF059092	AF059152	
D57	*Anser caerulescens*	DQ434537	FJ423758		
D58	*Aythya fuligula*	JF499099	KU697802	EU585687	NC_024595
D59	*Aythya ferina*	JF499098	EU585623	EU585686	NC_024602
D60	*Anas formosa*	JN703250	AF059073	AF059133	NC_015482
D61	*Dendrocygna bicolor*		EU585646	EU585709	
D62	*Anas querquedula*	GQ481326	EU585610	EU585673	
D63	*Malacorhynchus membranaceus*		EU585651	EU585714	
D64	*Anas hottentota (Anas punctata)*		EU585608	EU585671	
D65	*Anas georgica*	FJ027096	AF059075	AF059135	
D66	*Biziura lobata*		EU585627	EU585690	
D68	*Cygnus atratus*	NC_012843	EU585641	EU585704	NC_012843
D67	*Aythya australis*	MW151626	EU585622	EU585685	
D69	*Amazonetta brasiliensis*	FJ027059	AF059054	AF059115	
D70	*Lophonetta cristata*				
D71	*Anas crecca*	KC771255	AF059064	EU585670	NC_022452
D72	*Tadorna tadorna*	KU140668	AF059113	AF059173	NC_024750
OG1	*Gallus gallus*		AF195631		NC_040902
OG2	*Struthio camelus*	LC145063	MZ545713		NC_002785

*Accession numbers of the mitochondrial gene sequences of the cytochrome *b* (Cty *b*), cytochrome oxidase subunit 1 (CO1), NADH dehydrogenase subunit 2 (ND2) and the complete mitochondrial DNA (mtDNA) was taken from GenBank.

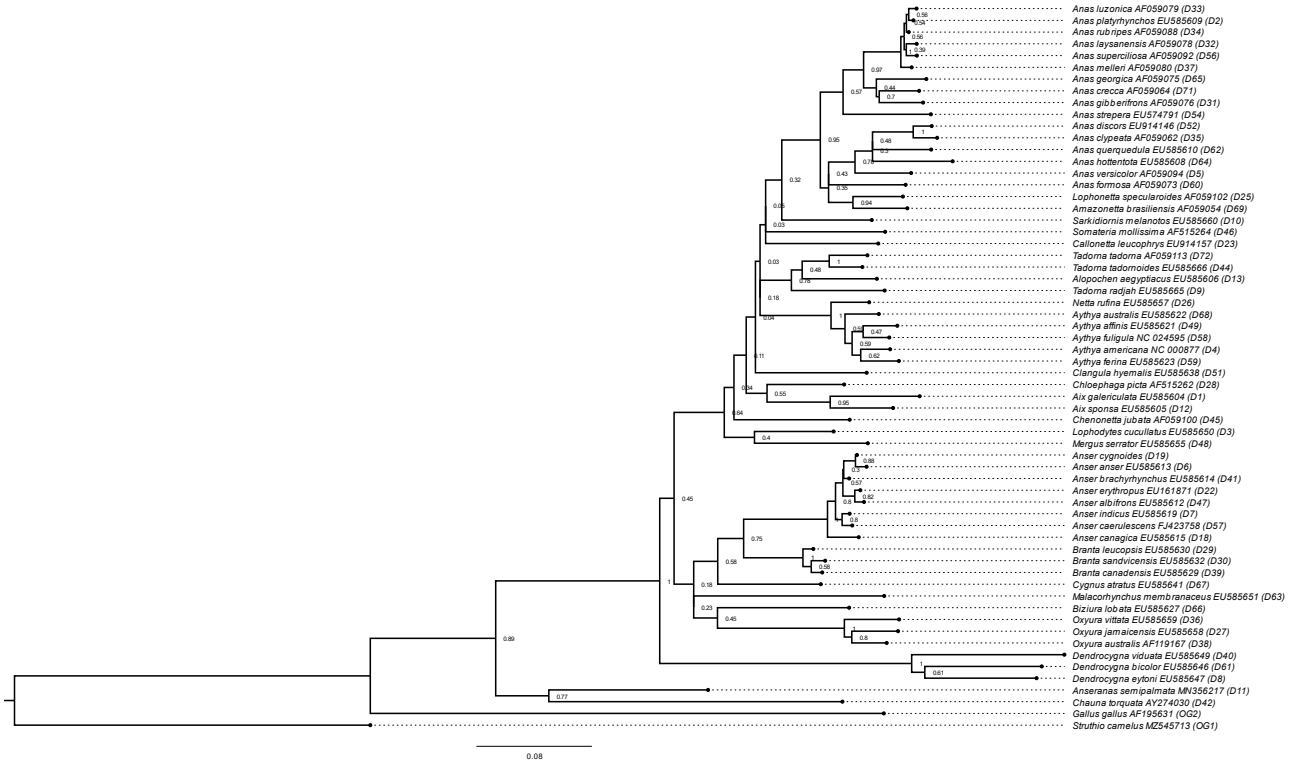


Figure S3.1. The evolutionary history was inferred by using the Maximum Likelihood method and Tamura-Nei model. The tree with the highest log likelihood (-11550.63) is shown. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Tamura-Nei model, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+*G*, parameter = 0.4430)). The rate variation model allowed for some sites to be evolutionarily invariable ([+*I*], 27.02% sites). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis involved 62 nucleotide sequences. There were a total of 992 positions in the final dataset. Evolutionary analyses were conducted in MEGA11.

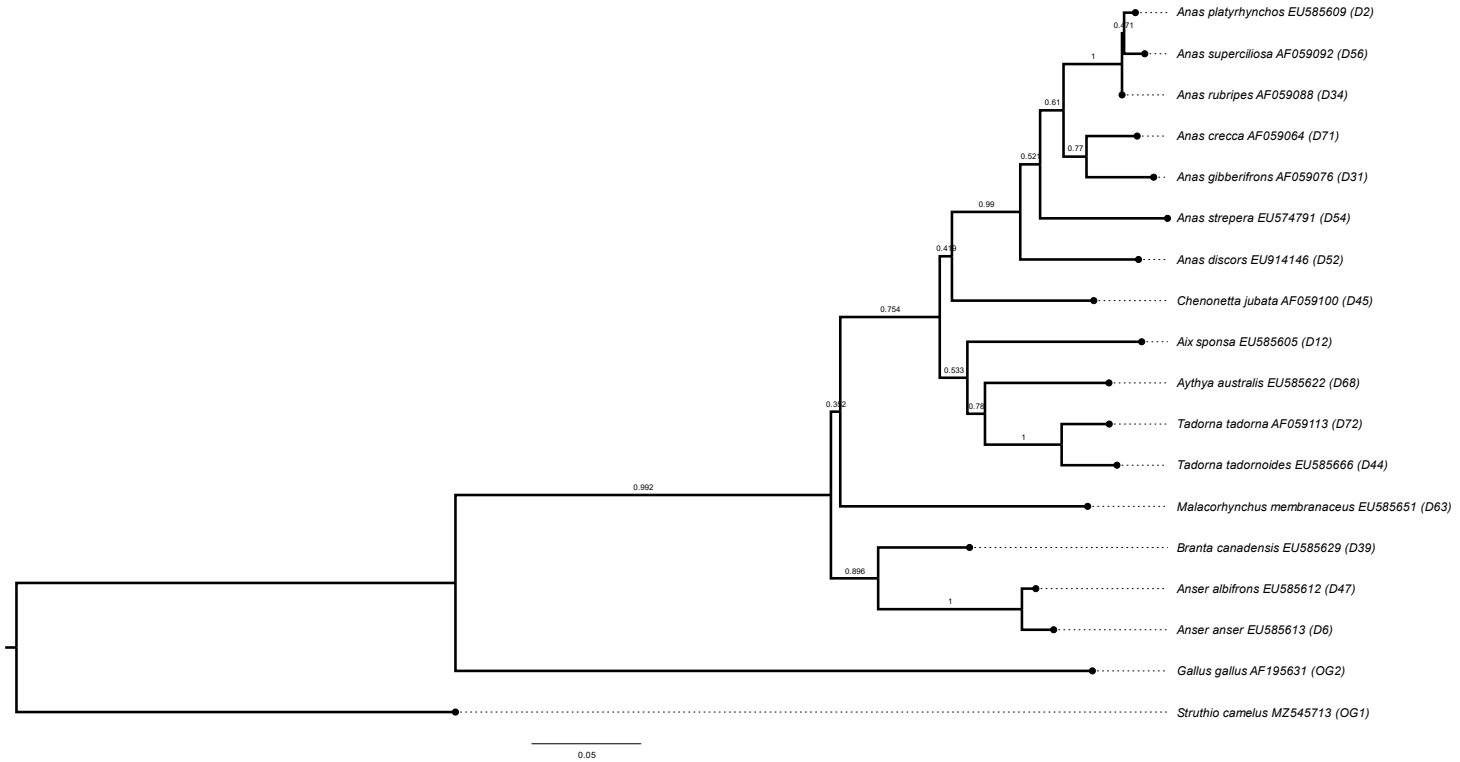


Figure S3.2. The evolutionary history was inferred by using the Maximum Likelihood method and Tamura-Nei model. The tree with the highest log likelihood (-5786.80) is shown. The percentage of trees in which the associated taxa clustered together is shown below the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Tamura-Nei model, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+*G*, parameter = 0.3642)). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis involved 18 nucleotide sequences. There were a total of 1082 positions in the final dataset. Evolutionary analyses were conducted in MEGA11.

Chapter 4 Concluding Remarks

This study addresses the challenges of analyzing sulfated *N*- and *O*-glycans in complex biological samples. We present a comprehensive workflow that integrates the strengths of Glycoblotting as a complementary purification, enrichment, methylation, and labeling technique for MALDI-TOF MS-based sulphoglycomics. Glycoblotting demonstrates its efficiency as a glycan enrichment platform, while the on-bead methyl esterification using MTT successfully overcomes the obstacles related to trace abundance, sample loss, and the presence of sialic acid. Notably, the on-bead methyl esterification step of Glycoblotting facilitates the discrimination between sulfated glycans and sialylated glycans and the differentiation of isomeric glycans containing sulfate or phosphate groups. This streamlined workflow enables efficient enrichment and detection of trace sulfated and phosphorylated *N*-glycans, offering a simplified approach to MALDI-TOF MS-based sulphoglycomics.

Moreover, employing the Glycoblotting-based sulphoglycomics approach, we uncovered a diverse array of sulfated and phosphorylated *N*-glycans in waterfowl egg whites, providing valuable insights into the differential expressions of acidic *N*-glycans in egg whites. We observed distinct variations in the expressions of acidic *N*-glycans among the four families (Anhimidae, Anseranatidae, Dendrocygnidae, and Anatinae) within the order Anseriformes. By examining sulfated trans-Gal(+) and trans-Gal(-) *N*-glycan structures and phosphorylated *N*-glycans, we successfully differentiate waterfowl species. Remarkably, waterfowl species with a high virus prevalence exhibit elevated phosphorylated hybrid and high-mannose *N*-glycans expression. These findings emphasize the significance of phosphorylated and sulfated *N*-glycans in comprehending the transmission and evolution of IAV within avian populations.

Further studies can be conducted to expand the application of the Glycoblotting-based sulphoglycomics workflow to other biological samples and species. In addition, investigations into the functional roles of specific sulfated and phosphorylated *N*-glycans in the interaction between waterfowl as a host and IAV can provide deeper insights into the mechanisms underlying viral infection and transmission. The integration of other analytical techniques, such as lectin binding assays and glycan microarrays, can enhance the understanding of the binding specificity and host-virus interactions mediated by sulfated and phosphorylated *N*-glycans, overall, continued advancements in sulphoglycomics research hold promise for elucidating the intricate relationship between glycans and viral infections, paving the way for novel preventive and therapeutic strategies against influenza and other related diseases.