Deciphering biological significance of dynamic glycosylation on glycan variants by mass spectrometry

質量分析法による動的で多様な糖鎖修飾の生物学的機能解明

Jurgen Teodosio Sanes

Submitted for the Degree of
Doctor of Life Science in

the Graduate School of Life Science

HOKKAIDO UNIVERSITY

September 2019

I dedicate this work to my parents and siblings...

Acknowledgments

My sincerest thanks to:

- Prof. Shin-Ichiro Nishimura for guiding me towards the accomplishment of this PhD. I am in awe of the brilliant ideas you have for the glycobiology field. The opportunity to work and learn under Nishimura-sensei is treasured,

- Prof. Hiroshi Hinou for valuable discussions regarding my various research projects, helping me with instrumentation troubles, and guidance for my work,

- Prof. Kenji Monde and Prof. Tomoyasu Aizawa for sharing their valuable time and meaningful insights,

- Previous and current Nishimura lab members, thank you for making this journey worthwhile. I will always remember the lab parties and gatherings we had. To Nakagawara, Miyoshi, Kambe, Miyajima, Yamada, Shiratori, Tanaka, Kobayashi, Masaki, Yasue, Abe, Kato, Ookado, Hayashi, Hammura, Koide, Yokoi, Hirane and Wakui-san, thank you for your omotenashi minna-san,

- Laboratory office administrators, Ms. Maki Morita and Ms. Tomoko Takahashi for the assistance from the start of my PhD program until the very end. Your wonderful smiles and thoughtfulness are always remembered,

- International Graduate Program (IGP) of Hokkaido University and Ministry of Education, Culture, Sports, Science and Technology (MEXT) Japan for providing me with a PhD scholarship,

Finally, to everyone who have helped, supported, and encouraged me in the accomplishment of this endeavor, ありがとうございま in Japanese and daghang salamat in my dialect.

Abstract

With the advent of chemical glycobiology, the use of homogeneous glycoconjugate structures is crucial to correctly evaluate glycan roles and biological functions. My work involved the use of different methodologies to study dynamic glycosylation on glycan variants. Glycan profiling from Galliformes quail egg whites was first conducted to assess diversity of structures using glycoblotting methodology. With this procedure, distinct structures among species were identified. The actual interest was for the isolation of the conserved $Man_3GlcNAc_2$ glycan structure found in quail egg whites useful in glycopeptide synthesis. The truncation of the conserved *N*-glycan was conducted, dehydrated to an oxazoline intermediate, and directly conjugated to two synthesized glycosyl acceptors, Fmoc-Asn(GlcNAc)-OH and IgG1-tryptic glycopeptide, using endoglycosidase-M-N175Q. The formation of the transglycosylated product was achieved with less purification steps. The procedure provides a direct isolation strategy of a glycan or a certain glycan class from a natural source depending on the endoglycosidase used. Starting with a conserved *N*-glycan structure is relevant for sugar elongation using glycosyltransferases. The goal for chemoenzymatic synthesis of glycopeptides is to use it for various assays. With this, different glycoforms of IgG1 and IgG2 tryptic glycopeptides were utilized to serve as calibration standards and optimized for multiplexing quantitations in human serum. Since most of the reactions performed in this work is based on enzymatic transformations, reusing the endoglycosidase enzyme was necessary. As such, an endoglycosidase immobilization strategy on chitosan nanoparticles was conducted in the last part of the work wherein applications on glycan release and conjugation was shown. Overall, the methodologies mentioned is of utmost importance in the preparation of homogeneous glycococonjugate structures that can be used for numerous applications.

Table of Contents

List of Tables

List of Figures

Abbreviation List

ABC : ammonium bicarbonate,

BSQ : blue-scaled quail,

BOA : O-benzyloxyamine hydrochloride,

BWQ : bobwhite quail, northern bobwhite quail;

CDMBI : 2-chloro-1,3-dimethyl-1H-benzimidazol-3-ium,

DIEA : *N, N*-diisopropylethylamine,

DHB : 2,5-dihydroxybenzoic acid,

DTT : dithiothreitol,

DMF : *N, N*-dimethylformamide,

Endo(s) : Endoglycosidase(s),

Endo-D : Endoglycosidase-D,

Endo-M : Endoglycosidase-M,

EW : Egg White,

Fmoc-Asn(Ac_3β-GlcNAc)-OH

:N-α-Fmoc-N-γ-(2-acetamido-2-deoxy-3,4,6-tri-O-acetyl-β-D-glucopyranosyl)-L-asparagine,

G0 : $Hex_3HexNAc_4$,

G2 : $Hex_5HexNAc_4$,

G2S2 : $Hex_5HexNAc_4NeuAc_2$,

GCC : Graphite Column Chromatography,

GlcNAc : *N*-Acetylglucosamine,

HBTU :1-[Bis(dimethylamino)methyliumyl]-1H-benzotriazole-3-oxide Hexafluorophosphate,

HILIC : Hydrophilic Interaction Chromatography,

HPLC : High-Pressure Liquid Chromatography,

HOBt : 1 hydroxybenzotriazole monohydrate,

IAA : Iodoacetamide,

IgG : Immunoglobulin,

IgG1-GlcNAc : H-Glu-Glu-Gln-Tyr-Asn(GlcNAc)-Ser-Thr-Tyr-Arg-OH,

IgG1-$Man_3GlcNAc_2$: H-Glu-Glu-Gln-Tyr-Asn($Man_3GlcNAc_2$)-Ser-Thr-Tyr-Arg-OH,

IgG4-GlcNAc : H-Glu-Glu-Gln-Phe-Asn(GlcNAc)-Ser-Thr-Tyr-Arg-OH,

IS : Internal standard,

JQ : Japanese quail,

MALDI-TOF : Matrix-Assisted Laser Desorption Ionization-Time of Flight

$Man_3GlcNAc$: Trimannosyl truncated chitobiose core,

$Man_3GlcNAc_2$: Trimannosyl chitobiose core,

MRM : Multiple Reaction Monitoring,

MWCO : Molecular Weight Cut-Off,

MQ : Mountain quail,

MS : Mass Spectrometry,

MTT : 3-methyl-1-p-tolyltriazene,

NHS : *N*-hydroxysuccinimide,

PBS : Phosphate-buffered saline,

PHM :1-propanesulfonic acid, 2-hydroxyl-3-myristamido,

PGC : Porous Graphitic Carbon,

PNGaseF : Peptide-*N*-Glycosidase-F,

PyBOP : 1 H-benzotriazol-1-yloxy-tri(pyrrolidino)phosphonium hexafluorophosphate,

RP-HPLC : Reversed Phase- High Performance Liquid Chromatography,

SRM : Single Reaction Monitoring,

SGP : Sialyglycopeptide, KVAN($Hex_5HexNAc_4NeuAc_2$)KT ,

TFA : Trifluoroacetic acid

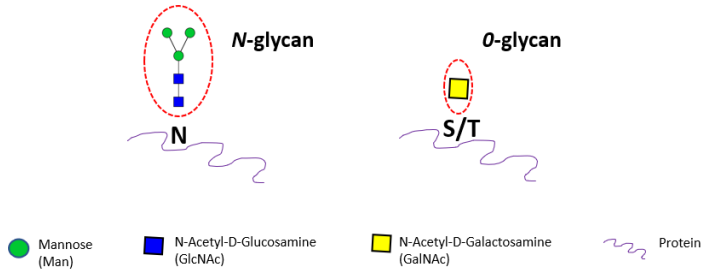
Chapter 1

General Introduction

Glycosylation is an important post-translational modification involved in many biological processes. Unlike DNA, there is no template for glycosylation. Adding to this complexity, a network of enzymes is involved with the formation of different glycans.[1,2] Glycan structures consists of numerous configurations and combinations which maybe due to different biological conditions wherein mechanisms are still unknown. The heterogeneity of glycan structures thus makes it a challenging biological process.

The sugar structures on proteins can be classified as either an *N*- or *O*-glycan depending on the amino acid residue it is attached to. It is an *N*-glycan when the sugar is conjugated to an asparagine (N) residue, while considered as an O-glycan if attached to a serine(S)/threonine(T). *N*-glycans share the same pentasaccharide conserved structure highlighted in Figure 1-1, while, O glycan starts with a galactosamine structure. .[1] The focus of this work is on *N*-glycans.

Figure 1-1. *N* and *O*-Glycan classification



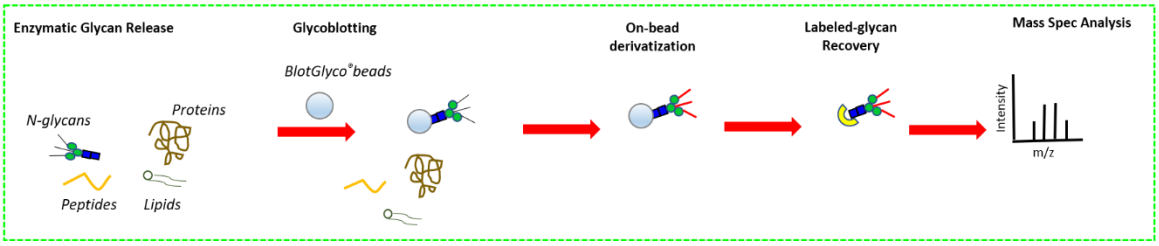
Different methodologies have been used to rationalize the function of sugar structures typically consisting of using either cleaved glycans or studied together with its peptide component. [3-7] In this dissertation, both cases were given emphasis. Glycans were cleaved using different enzymes that removes the structures from asparagine-linkage in the glycoprotein and quantitation of *N*-glycopeptides in samples were performed to decipher dynamic glycosylation on different variants. In both cases, mass spectrometry played a significant role.

There are only a few methodologies to synthesize glycan and glycoprotein structures owing to the difficulty of the stereochemistry and configurations of such glycans.[8-10] Other strategies rely on natural sources for glycan or glycopeptide isolation[11-13] or the derivatization of other saccharides[14-16]. Glycans can be released from protein/peptide attachment through an enzymatic and a chemical methodology involving different purification strategies to profile glycans in different samples, as well as comparison between healthy and diseased states. In this work, the glycoblotting methodology was used to quantitate the expression of glycans in the samples studied.

Glycoblotting methodology is a strategy for glycan enrichment.[17,18] The procedure starts with glycan removal through enzymatic digestion followed by purification with hydrazine beads. The enzymatic digests contain a lot of other biomolecules and salts that would interfere in the detection of glycans. The enrichment with the use of glycoblotting makes use of the stability of sugar hemiacetals under acidic medium to form hydrazone bonds with the bead, thereby, differentiating it from other biomolecules. The enriched glycans could be easily washed, derivatized, and labeled for mass spec analysis making the analysis faster. [17] The glycoblotting methodology was developed in this laboratory to serve as a high-throughput procedure that shortens glycan analysis and have been applied to various samples. Studying glycan profiles among different animal species would give an insight into glycan diversity.

Figure 1-2. Glyan enrichment through glycoblotting methodology. A. Glycoblotting procedure for glycan release, enrichment, derivatization, recovery, and mass spec analysis. B. Glycan exist as hemiacetals in acetic medium wherein can be selectively enriched through hydrazine attachment on beads

A.



B.

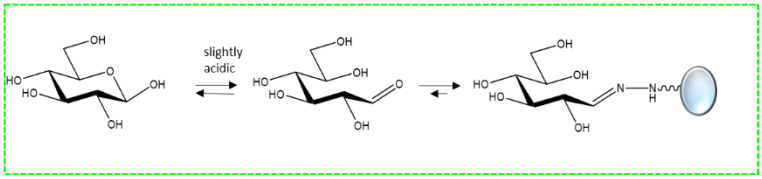
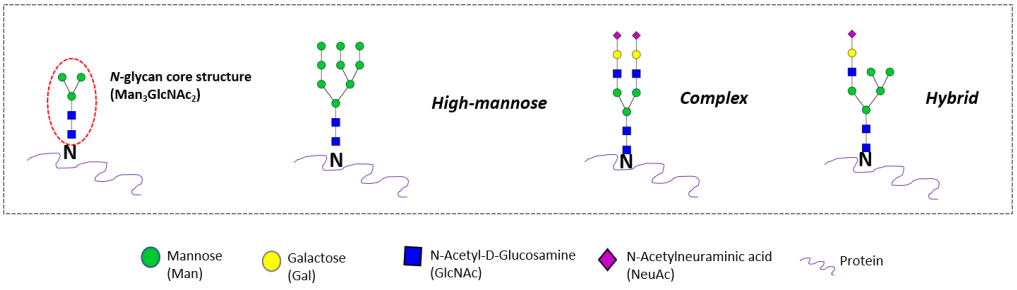


Figure 1-3. Different classes of *N*-glycans. The $Man_3GlcNAc_2$, pentasaccharide structure is conserved among *N*-glycans.



Of interest is the $Man_3GlcNAc_2$ structure, as it is conserved among *N*-glycans. Different methodologies have been used to prepare this conserved structure. [13,19-22] Truncated version of glycans released by cleavage in between the chitobiose core is needed for oxazoline conversion. [22,23] Glycans need to be converted to a dehydrated derivative called a sugar-oxazoline for it to be useful for glycopeptide synthesis. Sugar-oxazolines are intermediate structures requiring an endoglycosidase enzyme to catalyze the glycan attachment to a glycosyl acceptor. Forms of sugar-oxazolines have been synthesized through chemical means either from a direct glycan obtained from a natural source or

laboratory-synthesized glycans. Oxazolines can be prepared readily using dehydrating agents, e.g 2-chloro-1-3-dimethylimidazolinium chloride[24] and 2-chloro-1,3-dimethyl-1H-benzimidazole-3-ium chloride (CDMBI)[25].

Confirmation of the monosaccharide components in overall glycan and glycopeptide structure require high-resolution mass spectrometry.[26,27]Aside from exoglycosidase digestion, fragmentation patterns are usually deciphered to identify structures. When glycans are fragmented, a variety of product ion types arise within carbohydrate moieties. The designation of the ions is critical. A systematic nomenclature for carbohydrate fragmentations was first introduced by Domon and Costello (1988) [28] which can apply to ions produced from different MS/MS spectra recorded in either positive or negative ion mode. The glycan structure can be rationalized basing from fragmentation patterns. Mass spectrometry fragmentation for glycan and glycopeptide structures were conducted in MALDI-TOF/TOF and LC-MS/MS as shown in this work. The use of LC-MS/MS operating in Multiple Reaction Monitoring (MRM) modes have widely been used for the targeted quantitation and screening of different compounds because of known sensitivity and selectivity. [27,29] The use of MRM is shown in this work for the quantitation of glycopeptides in human serum by using chemoenzymatically synthesized standards. This is important because in nature, glycoproteins usually have the same protein components but different forms of the glycans. With this, studying a glycoprotein with correctly defined glycan components is relevant to understand its biological function. The dissertation theme can be dissected into 4 parts, namely, glycan profiling to source out natural glycans, isolation of glycans, glycoconjugate synthesis, and enzyme immobilization. The strategies presented is relevant in the glycobiology field.

References

1 Varki A, Sharon N. Historical Background and Overview. In: Varki A, Cummings RD, Esko JD, et al., editors. Essentials of Glycobiology. 2nd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2009. Chapter 1. Available from: https://www.ncbi.nlm.nih.gov/books/NBK1931/

2 Rini J, Esko J, Varki A. Glycosyltransferases and Glycan-processing Enzymes. In: Varki A, Cummings RD, Esko JD, et al., editors. Essentials of Glycobiology. 2nd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2009. Chapter 5. Available from: https://www.ncbi.nlm.nih.gov/books/NBK1921/

3 Nwosu, C.C., Seipert, R.R., Strum, J.S., Hua, S.S., An, H.J., Zivkovic, A.M., German, B.J. and Lebrilla, C.B (2011). Simultaneous and extensive site-specific N-and O-glycosylation analysis in protein mixtures. *Journal of proteome research*, *10*(5), 2612-2624.

4 An, H. J., Peavy, T. R., Hedrick, J. L., & Lebrilla, C. B. (2003). Determination of N-glycosylation sites and site heterogeneity in glycoproteins. *Analytical chemistry*, *75*(20), 5628-5637.

5 Pabst, M., & Altmann, F. (2011). Glycan analysis by modern instrumental methods. *Proteomics*, *11*(4), 631-643.

6 Nilsson, J., Rüetschi, U., Halim, A., Hesse, C., Carlsohn, E., Brinkmalm, G., & Larson, G. (2009). Enrichment of glycopeptides for glycan structure and attachment site identification. *Nature methods*, *6*(11), 809.

7 Goetz, J. A., Novotny, M. V., & Mechref, Y. (2009). Enzymatic/chemical release of O-glycans allowing MS analysis at high sensitivity. *Analytical chemistry*, *81*(23), 9546-9552.

8 Bernardes, G. J., B. Castagner and P. H. Seeberger (2009). "Combined approaches to the synthesis and study of glycoproteins." *ACS Chem Biol* 4(9): 703-713.

9 Lepenies, B., J. Yin and P. H. Seeberger (2010). "Applications of synthetic carbohydrates to chemical biology." *Curr Opin Chem Biol* 14(3): 404-411.

10 Wang, Z., Z. S. Chinoy, S. G. Ambre, W. Peng, R. McBride, R. P. de Vries, J. Glushka, J. C. Paulson and G. J. Boons (2013). "A general strategy for the chemoenzymatic synthesis of asymmetrically branched N-glycans." Science 341(6144): 379-383.

11 Seko, A., Koketsu, M., Nishizono, M., Enoki, Y., Ibrahim, H.R., Juneja, L.R., Kim, M. and Yamamoto, T (1997). Occurrence of a sialylglycopeptide and free sialylglycans in hen's egg yolk. *Biochim. Biophys. Acta (BBA)-General Subjects*, *1335*(1-2), 23-32.

12 Lis, H., & Sharon, N. (1978). Soybean agglutinin--a plant glycoprotein. Structure of the carboxydrate unit. *J. Biol. Chem*, *253*(10), 3468-3476.

13 Oda, Y., Nakayama, K., Abdul-Rahman, B., Kinoshita, M., Hashimoto, O., Kawasaki, N., Hayakawa, T., Kakehi, K., Tomiya, N. and Lee, Y.C (2000). Crocus sativus lectin recognizes Man3GlcNAc in the N-glycan core structure. *J. Biol. Chem. 275*(35), 26772-26779

14 Kumar H.V., R., Naruchi, K., Miyoshi, R., Hinou, H., Nishimura, S.-I. (2013) A new approach for the synthesis of hyperbranched N-glycan core structure from locust bean gum. *Org. Lett. 15*, 6268-6281.

15 Hammura, K., Ishikawa, A., Miyoshi, R., Yokoi, Y., Tanaka, M., Hinou, H., & Nishimura, S. I. (2018). Synthetic Glycopeptides Allow for the Quantitation of Scarce Nonfucosylated IgG Fc N-Glycans of Therapeutic Antibody. *ACS medicinal chemistry letters*, *9*(9), 889-894.

16 Yogesh, K.V., Kamiyama, T., Ohyama, C., Yoneyama, T., Nouso, K., Kimura, S., Hinou, H. and Nishimura, S.I. (2018). Synthetic glycopeptides as a designated standard in focused glycoproteomics to discover serum cancer biomarkers. *MedChemComm*, *9*(8), 1351-1358.

17 Miura, Y.,Hato, M., Shinohara, Y., Kuramoto,H., Furukawa, J.-I.,Kurogochi, M., Shimaoka, H., Tada, M., Nakanishi, K., Ozaki, M., Todo, S., and Nishimura, S.-I. (2008)

BlotGlycoABC™, an integrated glycoblotting technique for rapid and large scale clinical glycomics. *Mol. Cell. Proteomics. 7(2)*, 370–377.

18 Nishimura, S.-I., Niikura, K., Kurogochi, M., Matsushita, T., Fumoto, M., Hinou, H., Kamitani, R., Nakagawa, H., Deguchi, K., Miura,N., Monde, K., and Kondo, H. (2004) High-throughput protein glycomics: Combined use of chemoselective glycoblotting and MALDI-TOF/TOF mass spectrometry. *Angew. Chem., Int. Ed. 44*, 91–96.

19 Ratner, D. M., Swanson, E. R., & Seeberger, P. H. (2003). Automated synthesis of a protected N-linked glycoprotein core pentasaccharide. *Organic letters*, *5*(24), 4717-4720.

20 Zeng, Y., & Lehrman, M. A. (1991). Isolation of Chinese hamster ovary cell lines producing Man3GlcNAc2 asparagine-linked glycans. *Analytical biochemistry*, *193*(2), 266-271.

21 De Pourcq, K., Tiels, P., Van Hecke, A., Geysens, S., Vervecken, W., & Callewaert, N. (2012). Engineering Yarrowia lipolytica to produce glycoproteins homogeneously modified with the universal Man3GlcNAc2 N-glycan core. *PLoS One*, *7*(6), e39976.

22 Sun, B., Bao, W., Tian, X., Li, M., Liu, H., Dong, J., & Huang, W. (2014) A simplified procedure for gram-scale production of sialylglycopeptide (SGP) from egg yolks and subsequent semi-synthesis of Man3GlcNAc oxazoline. *Carbohydr. Res. 396*, 62-69

23 Li, B., Zeng, Y., Hauser, S., Song, H., & Wang, L. X. (2005). Highly efficient endoglycosidase-catalyzed synthesis of glycopeptides using oligosaccharide oxazolines as donor substrates. *Journal of the American Chemical Society*, *127*(27), 9692-9693.

24 Noguchi, M., Tanaka, T., Gyakushi, H., Kobayashi, A., & Shoda, S. I. (2009). Efficient synthesis of sugar oxazolines from unprotected N-acetyl-2-amino sugars by using chloroformamidinium reagent in water. *The Journal of organic chemistry*, *74*(5), 2210-2212.

25 Noguchi, M., Fujieda, T., Huang, W. C., Ishihara, M., Kobayashi, A., & Shoda, S. I. (2012). A Practical One-Step Synthesis of 1, 2-Oxazoline Derivatives from Unprotected Sugars and

Its Application to Chemoenzymatic β-N-Acetylglucosaminidation of Disialo-oligosaccharide. *Helvetica Chimica Acta*, *95*(10), 1928-1936.

26 Kailemia, M. J., Ruhaak, L. R., Lebrilla, C. B., & Amster, I. J. (2013). Oligosaccharide analysis by mass spectrometry: a review of recent developments. *Analytical chemistry*, *86*(1), 196-212.

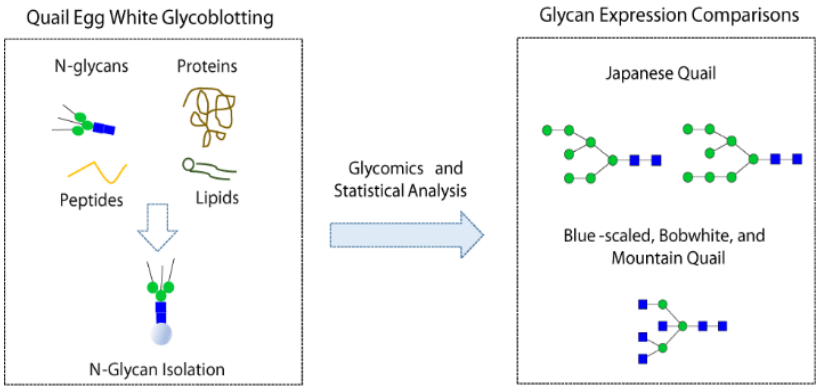
27 Kurogochi, M., Matsushista, T., Amano, M., Furukawa, J. I., Shinohara, Y., Aoshima, M., & Nishimura, S. I. (2010). Sialic acid-focused quantitative mouse serum glycoproteomics by multiple reaction monitoring assay. *Molecular & Cellular Proteomics*, *9*(11), 2354-2368.

28 Domon, B., & Costello, C. E. (1988) A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconj. J. 5*(4), 397-409.

29 Huang, J., Kailemia, M.J., Goonatilleke, E., Parker, E.A., Hong, Q., Sabia, R., Smilowitz, J.T., German, J.B. and Lebrilla, C.B. (2017). Quantitation of human milk proteins and their glycoforms using multiple reaction monitoring (MRM). *Analytical and bioanalytical chemistry*, *409*(2), 589-606.

Chapter 2

Glycoblotting of Egg White Reveals the Presence of the Core Pentasaccharide *N*-glycan Structure and Diversity in *N*-Glycan Expression among Quail Species



2.1 Chapter Summary

The glycan part of glycoproteins is known to be involved in the structure and modulatory functions of glycoproteins, serving as ligands for cell to cell interactions, and as specific ligands for cell to microbe interactions. It is believed that intraspecies and interspecies variations in glycosylation exist. As an approach to better understand glycan diversity, egg whites (EW) from four different quail species are studied by the well-established glycoblotting procedure, a glycan enrichment and analysis method. *N*-glycans were classified and the profiles were established for quail egg white samples which showed 21 relevant glycan peaks; 18 peaks were expressed significantly, and 10 glycan peaks are found to be abundant in certain species. The result establishes glycan profiles for Blue-scaled, Bobwhite, Japanese, and Mountain quail egg whites and shows unique difference among glycan expressions, particularly, high-mannose in Japanese quail and tetra-antennary glycan structure for other quail species. Of equal importance, the *N*-glycan core pentasaccharide ($Man_3GlcNAc_2$) structure is present in all the quail species.

Keywords: *N*-Glycans; Glycan Diversity; Inter-species Glycan Comparison; Glycoblotting; Egg White; Galliformes; Quail

2.2 Introduction

Glycans in glycoproteins are formed by post-translational modification of peptides. Intra-species and inter-species glycan variation in glycoproteins is known to exist, but less information about glycan structural diversity is available [1]. We chose to study EWs, because they are a nutrient source for the embryo, and most importantly, it serves as a barrier to harmful microbes that invade the developing embryo [2]. Thus, it is a very ideal material for examining the inter-species diversity of glycans [3-5].

A comprehensive profiling of *N*-glycans of the EWs from different avian species was conducted previously [6] which showcased that members of Galloanserae can be classified into 2 major groups and 5 submajor clusters based on MALDI-TOF/MS analysis. It was rationalized that glycan expression patterns are influenced by features such as body size, and the influence of lifestyle and diet of the birds. The comprehensive glycan analysis was achieved very efficiently using the *Glycoblotting/SweetBlot* technology developed in this laboratory [7]. The above-mentioned work compares between two avian orders, e.g. Galliformes and Anseriformes, and not between within the same order of birds [6]. The current work was conducted to compare glycan expression within one avian order.

In this work, the *N*-glycan expression from EW of four quail species, Blue-Scaled Quail (BSQ, *Callipepla squamata*), Northern Bobwhite Quail (BWQ, *Colinus virginianus*), Japanese Quail (JQ, *Coturnix japonica*), and Mountain Quail (MQ, *Oreortyx pictus*) [8,9] were analyzed. These quail species belong to the order Galliformes of ground-feeding birds. The glycan analyses were conducted by the glycoblotting enrichment procedure for glycans. Glycoblotting glycan enrichment makes use of the stability of hemiacetals under acidic medium to form hydrazone bonds with a hydrazide functionalized polymer that would differentiate the glycans from other biomolecules. The enriched glycans could be easily washed, derivatized, and labeled for mass spectrometry analysis [10]. At the time of writing, this is the first paper for *N*-glycan inter-species comparison among quails in the order Galliformes.

2.3 Materials and Methods

The procedure for digestion and glycoblotting was based on the published article [6] with few modifications and was performed as follows:

2.3.1 Enzymatic Release of N-Glycans from Quail Egg White

Japanese quail eggs were bought from the supermarket and the egg white separated from the yolk. The BSQEW, BWQEW, and MQEW were taken from an in-house egg white collection. Egg whites from 10 eggs provided ~50 mL of sample which was homogenized and then dried. A 1.0 mg lyophilized quail EW (BSQ, BWQ, JQ, and MQ) was weighed and placed in Eppendorf tubes. Three trials were prepared for each quail type. The egg white powder was dissolved in 20 µL 200 mM ammonium bicarbonate. Addition of a 26 µL 100 µM disialyloctasaccharide (Tokyo Chemical Industry Co., LTD.) internal standard followed. A 54 µL mixture containing 0.06% 1-propanesulfonic acid, 2-hydroxyl-3-myristamido (PHM) with 12 mM dithiothreitol (DTT) in 105 mM NH_4HCO_3 was added and incubated at 60°C for 90 min. After incubation, a 10 µL freshly prepared 123 mM iodoacetamide was added and incubated in the dark for 1 h at room temperature. Digestion with 10 µL of 40 U/ µL sequence grade trypsin (Sigma Aldrich) in 1mM HCl was performed and incubated overnight at 37°C. The trypsin enzyme was heat inactivated in a heat block at 90°C for 10 min. The trypsin-digested solution was cooled at room temperature before adding 2 units of PNGase F (Roche) and incubated overnight at 37°C, followed by addition of a 10 µL of 0.5 units/µL Proteinase K at 37°C and incubation for 3 hours. The Proteinase-K (Roche) was heat inactivated at 90°C for 10 min. The enzyme digested samples were then placed in a SpeedVac to dry.

2.3.2 Glycoblotting Methodology for Enrichment of N-Glycans

Five hundred µL BlotGlycoH bead suspension (10 mg/mL) was added to each of the 96-well multiScreen Solvinert filter plate. The 96-well filter plate was then attached to a vacuum manifold and the water removed. The dried digest containing released *N*-glycans was reconstituted with 20 µL Milli Q water and 180 µL of 2% acetic acid/acetonitrile. The 96-well plate was incubated at 80°C for 45 min or until dry. Washing with 200 µL of 2 M guanidine-HCl in NH_4HCO_3, H_2O, and 1%

triethylamine in methanol followed. Each solvent washing was performed twice and vacuumed after every solvent addition. The unreacted hydrazide functional beads were then capped by addition of freshly prepared 100 µL 10 % acetic anhydride in methanol and incubated at RT for 30 min. The capping solvent was removed in vacuum. Washing with 200 µL of 10 mM HCl, methanol, dioxane was then conducted. Washing with each solvent was done twice and vacuumed after each solvent addition. Addition of 100 µL 100 mM 3-methyl-1-*p*-tolyltriazene (MTT) in dioxane was conducted and the 96-well plate incubated at 60°C until dry. The 96-well plate was washed twice with 200 µL of dioxane, water, methanol, and water, with vacuuming between washings. A 20 µL of 50 mM *O*-benzyloxyamine hydrochloride (BOA) and 180 µL of 2 % acetic acid/acetonitrile were added to the wells. The 96-well plate was then incubated at 80°C until the wells are dry. Elution of the glycans was then performed by addition of 100 µL of MilliQ H_2O. Elution was conducted twice. The eluted labeled-glycans were pooled and dried in SpeedVac.

2.3.3 MALDI-TOF MS and MALDI-TOF/TOF Analysis

The dried labeled-glycans were reconstituted with 20 µL of ultrapure H_2O. The mass spectral analysis was with an Autoflex III (Bruker Daltonics) in reflector, positive ion mode, typically totaling 200x10 shots with acceleration voltage, reflector voltage, and pulsed ion extraction settings of 25.3 kV, 26.4 kV, and 100 ns, respectively. The matrix recommended by the Bruker Dalton company was prepared by dissolving a mixture of 20 mg/mL 2,5-dihydroxybenzoic acid (DHB) and 1 mM NaCl in 30% acetonitrile containing 0.1% trifluoroacetic acid in water. A 1 µL of the prepared DHB matrix was spotted on the Anchorchip Target Plate MTP 384 Target Plate (polished steel TF, Bruker), dried, followed by spotting with 1 µL of the reconstituted solution containing the labeled glycans. The experimental masses were obtained using the *FlexAnalysis* 3.0 software (Bruker Daltonics) and the structures annotated using the *GlycoMod* Database (https://web.expasy.org/glycomod/)[11,12] and the *Functional Glycomics Gateway* (http://www.functionalglycomics.org/glycomics/molecule/jsp/carbohydrate/searchByComposition.jsp)[13] websites. In MALDI-TOF/TOF Mode, precursor ions were accelerated to 8 kV and the fragments further accelerated to 20.1 kV in positive ion mode.

MS/MS analysis was conducted using Ultraflex III (Bruker Daltonics) to further confirm the identity of important *N*-glycan structures.

2.3.4 Glycotyping of N-glycans in Quail Egg Whites and Statistical Analysis

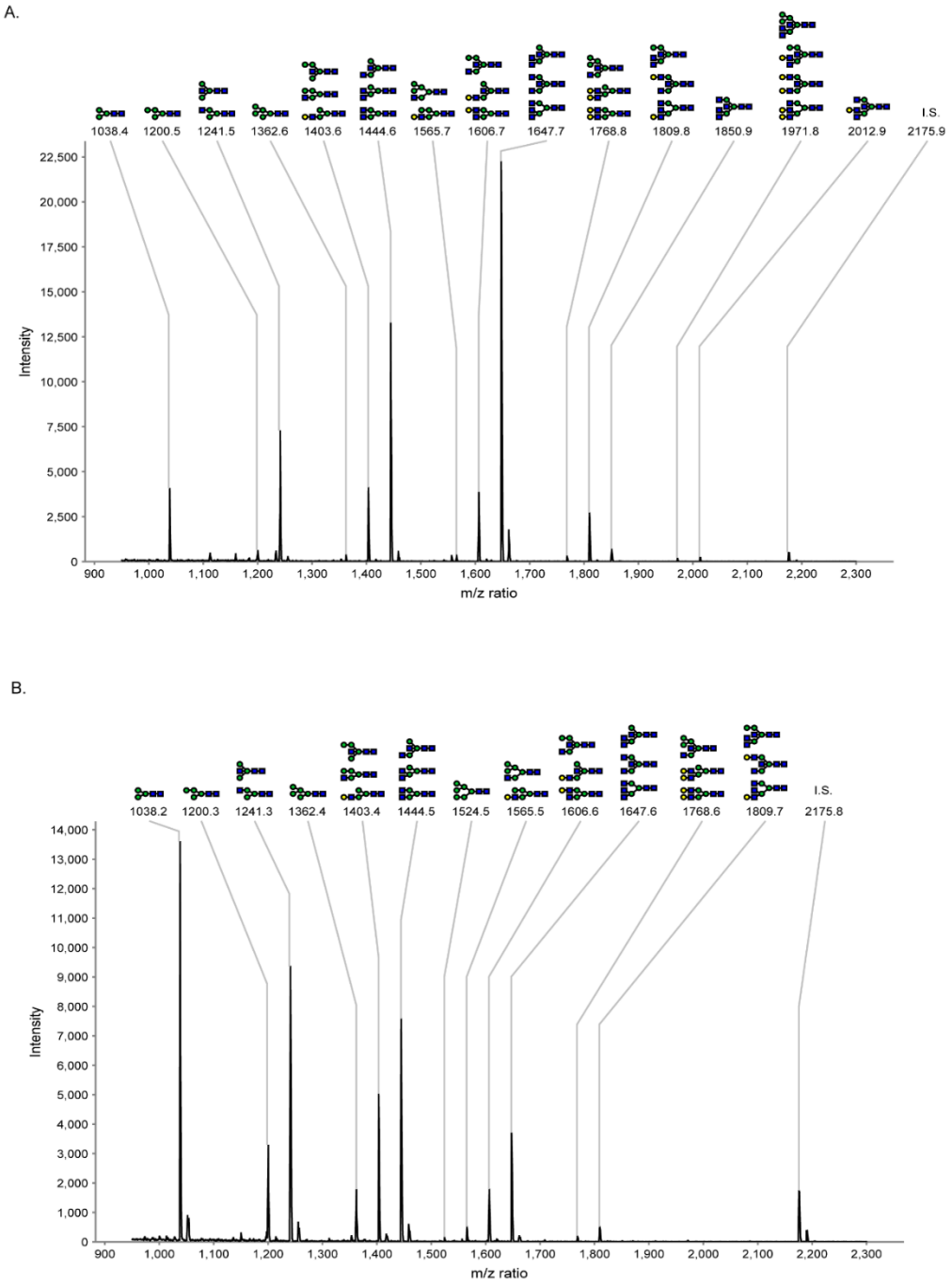
The area of the curve was taken and normalized with the internal standard. From the results obtained as total amount of glycans in pmols, the *N*-glycans could be classified into hybrid, complex, high-mannose, and paucimannose structures. The hybrid and complex glycans were classified into bisecting and non-bisecting structures. Furthermore, hybrid and complex glycans were classified based on antennary types. Estimated mole percentage of each glycan type was tabulated to further distinguish the difference of the glycan structures expressed in quail EWs and identify which of the glycan types were expressed differently. Distinct, abundant, and trace *N*-glycan structures were identified. Single-way Analysis of Variance (ANOVA) and Tukey's multiple comparison post-test were used for statistical analysis.

2.4 Results and Discussion

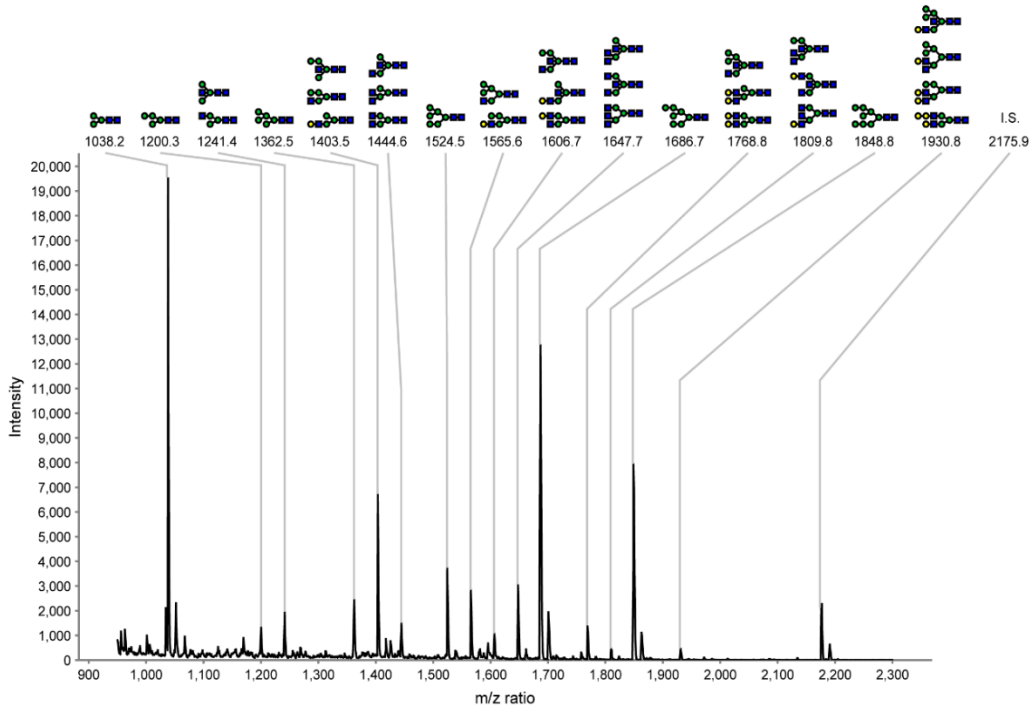
2.4.1 N-Glycans in Blue-Scaled Quail Egg White (BSQEW)

Figure 2-1 shows all major *N*-glycans for a single trial (see also Figure S2-1) and the structural compositions is displayed in Table 2-1. Nineteen mass/charge peaks corresponding to *N*-glycans were found in the analysis of BSQEW shown in Table 2-2. A description on how to calculate the theoretical *m/z* is presented in the supporting information. The peak at 2175 *m/z* pertains to the disialyloctasaccharide internal standard (2600 pmol) to which the relevant glycan *m/z* values were normalized with. Highest and lowest glycan amounts were seen for 1647 *and* 1972 *m/z* values, respectively. Statistical significant difference was found among ions under 99.9% confidence level using one-way analysis of variance as test statistic.

Figure 2-1. Comparison of mass spectra for major *N*-Glycans found in quail egg white samples. The spectra represent major glycan peaks for A) BSQ, B) BWQ, C) JQ, and D) MQ, respectively. Spectra is representative of 1 trial only. Other trials are also shown in Figure S2-1. I.S. signifies internal standard.



C.


1038.2
1200.3
1241.4
1362.5
1403.5
1444.6
1524.5
1565.6
1606.7
1647.7
1686.7
1768.8
1809.8
1848.8
1930.8
I.S.
2175.9
Intensity
m/z ratio

D.

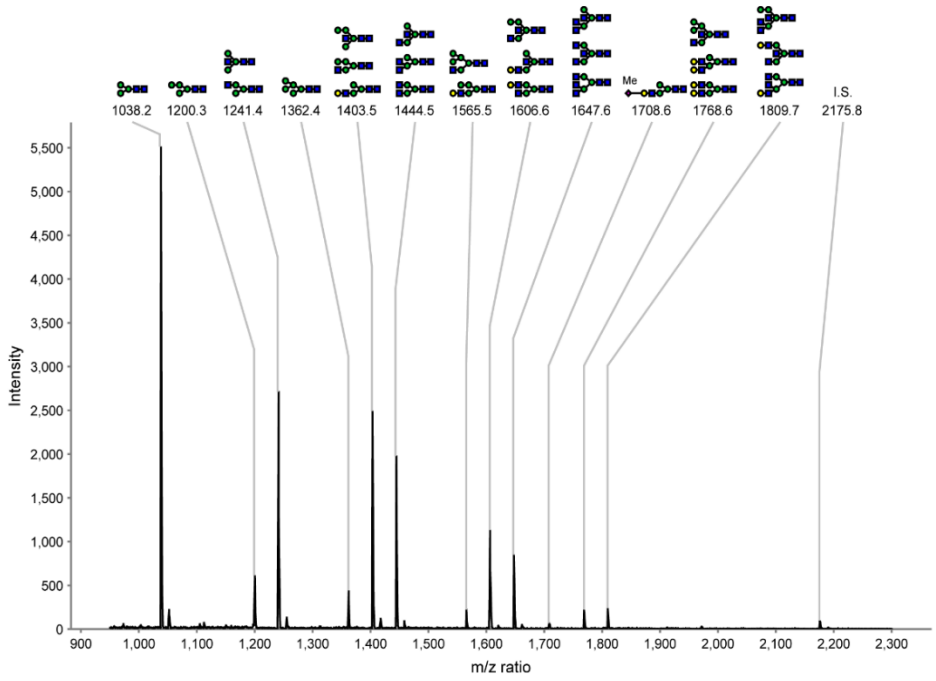
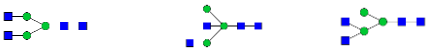
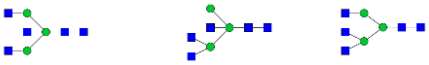
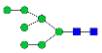
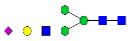
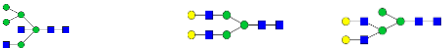
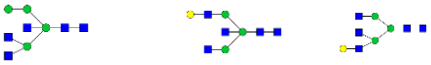
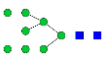
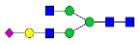
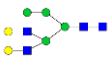
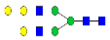
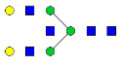
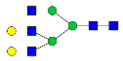
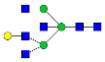

1038.2
1200.3
1241.4
1362.4
1403.5
1444.5
1565.5
1606.6
1647.6
Me
1708.6
1768.6
1809.7
I.S.
2175.8
Intensity
m/z ratio

Table 2-1. *N*-Glycans Expressed in Quail Egg White Galliformes

Peak No.	*m/z* values	Composition	*N*-Glycan Putative Structures	Ref. No.
1	1038	$(Man)_3(GlcNAc)_2$		14-19
2	1200	$(Man)_4(GlcNAc)_2$		20-21
3	1241	$(Hex)_3(HexNAc)_3$		14,15,21,22
4[a]	1362	$(Man)_5(GlcNAc)_2$		14,15,21
5	1403	$(Hex)_4(HexNAc)_3$		14,15,21,22
6	1444	$(Hex)_3(HexNAc)_4$		14,15, 20,21,22
7[a]	1524	$(Man)_6(GlcNAc)_2$		14,15,21
8	1565	$(Hex)_5(HexNAc)_3$		14,15,21
9	1606	$(Hex)_4(HexNAc)_4$		14,15,21
10	1647	$(Hex)_3(HexNAc)_5$		14,15,21,22
11[a]	1686	$(Man)_7(GlcNAc)_2$		14,15,21,23
12[a]	1708	$(Hex)_4(HexNAc)_3(NeuAc)_1$		6
13	1768	$(Hex)_5(HexNAc)_4$		14,15,17,21,23
14	1809	$(Hex)_4(HexNAc)_5$		15, 21
15[a]	1848	$(Man)_8(GlcNAc)_2$		23,24

16[a]	1850	$(Hex)_3(HexNAc)_6$					14,15,22,25
17	1911	$(Hex)_4(HexNAc)_4(NeuAc)_1$					15
18	1930	$(Hex)_6(HexNAc)_4$					17
19	1972	$(Hex)_5(HexNAc)_5$					14,15,17,21
20	2012	$(Hex)_4(HexNAc)_6$					14,15,21,22
21	2133	$(Hex)_6(HexNAc)_5$					16,17,21
22	2175	Internal Standard					

[a] MS/MS was conducted for the glycan structures.

GlcNAc; *N*-acetylglucosamine, Hex; hexose, HexNAc; *N*-acetylhexosamine, Man; mannose, NeuAc; *N*-acetylneuraminic acid. All *m/z* values are in terms of $[M+Na]^+$ adduct with the O-benzyloxyamine label accounting 105 *m/z*. For both 1708 and 1911 *m/z* values, 14 *m/z* was accounted for single methylation. It is worth noting that total number of glycan structures may exceed the total number of peaks observed herein as certain isomeric *N*-glycan structures may exist. The possible structures were derived from the previous references on avian egg glycosylation found in *GlycoMod*[11,12] and Functional Glycomics Gateway[13] databases. Glycan structures were prepared using GlycoWorkBench [26]

Table 2-2. Statistical Comparison among *m/z* Values for *N*-Glycan Amounts of Quail Egg White

Peak No.	*m/z* values	BSQ		BWQ		JQ		MQ		S/NS[2]
		pmol	SD	pmol	SD	pmol	SD	pmol	SD	
1	1038	10003	4250	18256	8774	14237	3005	66202	41623	*
2	1200	1875	701	4451	2210	801	157	7485	3578	*
3	1241	20856	7661	13316	6201	1472	183	35263	19474	*
4	1362	1101	340	2276	1195	1699	218	5808	2895	*
5	1403	12112	3934	6937	3079	5794	770	33762	17611	*
6	1444	47404	15665	11243	4839	1065	141	27666	14356	**
7	1524	---		156	155	3213	402	---		***
8	1565	1253	340	577	216	2496	213	3280	1289	**
9	1606	13214	3380	2191	801	800	186	16865	7986	**
10	1647	106534	32198	5276	2079	2805	343	13231	6391	***
11	1686	ND		ND		16164	2215	ND		***
12	1708	---		---		---		1557	471	***
13	1768	1051	161	163	153	1230	208	4156	1767	**
14	1809	12012	2369	639	209	158	274	3990	1782	***
15	1848	ND		ND		11007	1180	ND		***
16	1850	3009	294	---		ND		---		***
17	1911	---		---		---		---		CC
18	1930	---		---		162	280	---		NS
19	1972	678	26	---		---		---		***
20	2012	1056	55	---		---		---		***
21	2133	---		---		---		---		CC
22	2175 (I.S.)	2600		2600		2600		2600		
Total *N*-Glycan (pmol/mg)		232159		65482		63104		219265		
S_{pooled}			21542		7285		2353		38651	
S/NS[1]		***		***		***		***		

S- significant, NS- not significant, SD calculated for n=3, S_{pooled}- pooled standard deviation.[1]Statistical comparison among *m/z* values for each quail type, [2]Statistical comparison for each *m/z* value among quails* significant at 95%, ** significant at 99%, *** significant at 99.9% CI, CC-cannot be calculated and not statistically significant. ND-not detected, ----trace; Not Detected and Trace were both taken as 0 pmole for statistical calculations

Based from the total *N*-glycan amount of 232±22 nmols, glycotyping reveals the glycans in BSQEW were estimated to be <1% high-mannose-type, 88% complex-type, 6% hybrid-type, and 5% paucimannose structures. The hybrid and complex structures were subdivided into bisecting and non-bisecting glycan structures. Hybrid-bisecting and hybrid-non-bisecting *N*-glycans were 4% and 2%, respectively. The complex-bisecting and complex-non-bisecting glycan structures were 49% and 39%, respectively. The total monoantennary, biantennary, triantennary, and tetraantennary structures coming from the complex and hybrid *N*-glycans were 15%, 26%, 51%, and 2%, respectively. Furthermore, classification based on relationship between antennary structures and hybrid/complex types was made. There were 2% hybrid monoantennary, 2% hybrid biantennary, 2% hybrid triantennary, 12% complex monoantennary, 24% complex biantennary, 50% complex triantennary, and 2% complex tetraantennary. The paucimannose structures, $(Man)_3(GlcNAc)_2$ and $(Man)_4(GlcNAc)_2$, were 4%, and <1%, respectively. Glycotyping results are shown in Figure 2-2 and tabulated in Tables 2-3 to 2-5.

Figure 2-2. % Composition of *N*-Glycan amount in quail egg white samples based on glycan type. A) BSQEW, B) BWQEW, C) JQEW D) MQEW. All total *N*-glycan values are in pmol glycan/mg of egg white sample. The difference in circle size signifies the total amount of glycans present in each species. Abundant high-mannose expression is seen in JQEW.

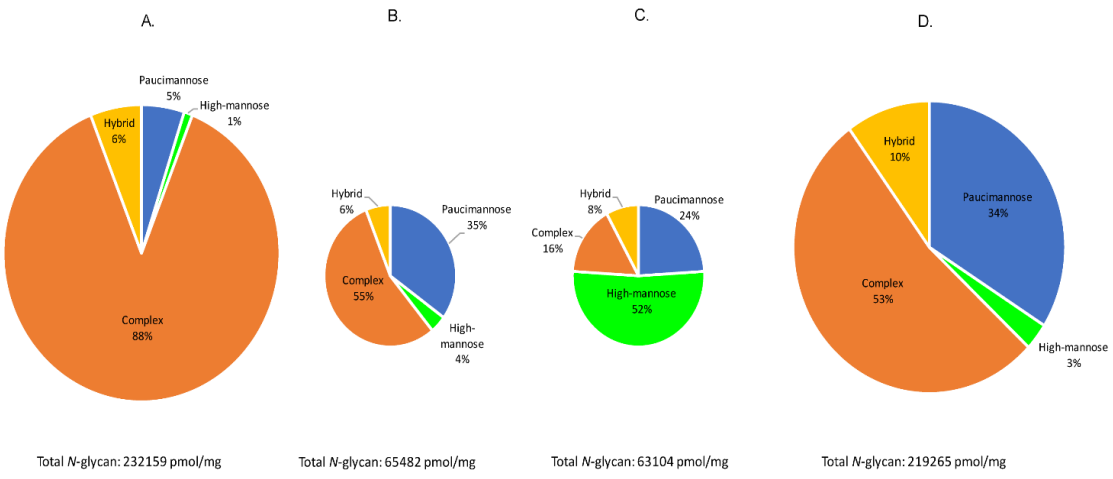


Table 2-3. % Composition of Bisecting and Non-Bisecting *N*-Glycan in Hybrid and Complex Glycan Types

Glycan Type	Bisecting Classification	% Composition from Total *N*-Glycan			
		BSQ	BWQ	JQ	MQ
Hybrid	Bisecting	4	2	1	4
	Non-Bisecting	2	4	6	7
Complex	Bisecting	49	26	8	24
	Non-Bisecting	39	29	8	29

Table 2-4. % Composition of Total Antennary *N*-Glycans from Hybrid and Complex Glycan Types

Antennary Classification	% Composition from Total *N*-Glycan			
	BSQ	BWQ	JQ	MQ
Monoantennary	15	32	13	34
Biantennary	26	20	5	22
Triantennary	51	9	5	8
Tetraantennary	2	---	---	---

--- trace

Table 2-5. % Composition of Antennary *N*-Glycans in Hybrid and Complex Glycan Types

Glycan Type	Antennary Classification	% Composition from Total *N*-Glycan			
		BSQ	BWQ	JQ	MQ
Hybrid	Monoantennary	2	4	6	7
	Biantennary	2	1	1	3
	Triantennary	2	<1	<1	<1
	Tetraantennary	---	---	---	---
Complex	Monoantennary	12	27	7	27
	Biantennary	24	20	4	19
	Triantennary	50	9	5	7
	Tetraantennary	2	---	---	---

--- trace

2.4.2 N-Glycans in Bobwhite Quail Egg White (BWQEW)

Nineteen mass/charge peaks corresponding to *N*-glycans were found in BWQEW. Highest glycan amount was seen at 1038 *m/z*, while low glycan amounts were seen at 1524 and 1768 *m/z*. Statistical comparison resulted in a significant difference among quantifiable *N*-glycan peaks at 99.9% confidence level.

The total amount of BWQEW *N*-glycans was 65±7 nmol based on normalization with the internal standard content. The glycans in BWQEW were estimated to be composed of 4% high-mannose type, 55% complex, 6% hybrid, and 35% paucimannose structures. From the total glycan content, hybrid-bisecting and hybrid-non-bisecting *N*-glycans were 2% and 4%, respectively. The complex-bisecting and complex-non-bisecting glycan structures were 26% and 29%, respectively. The monoantennary, biantennary, and triantennary structures from the complex and hybrid *N*-glycans were 32%, 20%, and 9%, respectively. There were 4% hybrid monoantennary, 1% hybrid biantennary, <1% hybrid triantennary, 27% complex monoantennary, 20% complex biantennary, and 9% complex triantennary. The paucimannose structures, $(Man)_3(GlcNAc)_2$ and $(Man)_4(GlcNAc)_2$, were 28%, and 7%, respectively.

2.4.3 N-Glycans in Japanese Quail Egg White (JQEW)

The analysis of JQEW revealed 20 mass/charge peaks corresponding to *N*-glycans. High glycan amounts were seen for 1038 and 1686 *m/z*, while low glycan amounts were seen at 1809 and 1930 *m/z* values. Intensity of the peaks does not at all correlate with area as seen in some *m/z* values, i.e. 1038 and 1686. As such glycan amounts are higher because of a wider area in 1686 *m/z*. Statistical comparison resulted in a significant difference among all quantifiable peaks at 99.9% confidence level.

Based on the total amount of 63±2 nmol *N*-glycans, glycotyping reveals the glycans in JQEW were estimated to be composed of 52% high-mannose type, 16% complex, 8% hybrid, and 24% paucimannose structures. From the total *N*-glycan content, hybrid-bisecting and hybrid-non-bisecting

N-glycans were 1% and 6%, respectively. The complex-bisecting and complex-non-bisecting glycan structures comprises both 8% of the total glycan amount. The monoantennary, biantennary, and triantennary structures from the complex and hybrid *N*-glycans were 13%, 5%, and 5%, respectively, based on the whole glycan content. There were 6% hybrid monoantennary, 1% hybrid biantennary, <1% hybrid triantennary, 7% complex monoantennary, 4% complex biantennary, and 5% complex triantennary. The paucimannose structures, $(Man)_3(GlcNAc)_2$ and $(Man)_4(GlcNAc)_2$, were 23%, and 1%, respectively.

2.4.4 N-Glycans in Mountain Quail Egg White (MQEW)

Nineteen mass/charge peaks corresponding to *N*-glycans were found in the analysis of MQEW. Highest glycan amount was accounted to $(Man)_3(GlcNAc)_2$ at 1038 *m/z*, while the lowest was a monosialylated glycan structure at 1708 *m/z* value. A statistically significant difference among all quantifiable *N*-glycan peaks were at 99.9% confidence level. Based on the 219±39 nmol total amount of *N*-glycans, glycotyping reveals the glycans in MQEW were estimated to be 3% high-mannose type, 53% complex, 10% hybrid, and 34% paucimannose structures. Hybrid-bisecting and hybrid-non-bisecting *N*-glycans from the total glycan content were 4% and 7%, respectively. The complex-bisecting and complex-non-bisecting glycan structures were 24% and 29 %, respectively. The monoantennary, biantennary, and triantennary structures from the complex and hybrid *N*-glycans were 34%, 22%, and 8%, respectively. There were 7% hybrid monoantennary, 3% hybrid biantennary, <1% hybrid triantennary, 27% complex monoantennary, 19% complex biantennary, and 7% complex triantennary. The paucimannose structures, $(Man)_3(GlcNAc)_2$ and $(Man)_{4(}GlcNAc)_2$, were 30%, and 3%, respectively.

2.4.5 Glycotyping Reveals Expression of N-glycan Structures Varies from Species to Species

BSQEW, BWQEW, and MQEW were highly abundant in complex type *N*-glycans, while having the least expression of high-mannose type. JQEW on the other hand contains predominantly high-mannose type glycan, while the hybrid-structure was low.

Bisecting structures, complex-non-bisecting structures were highest for BWQEW and MQEW, while hybrid-bisecting structures were low for 2 of the quail species. For BSQEW, complex-bisecting structures was highest and hybrid-non-bisecting structures was lowest. It is around the same amount for JQEW.

For antennary types, BWQEW, JQEW, and MQEW were predominantly complex-monoantennary structures while BSQEW were largely complex-triantennary expressed. Only BSQEW has a quantifiable amount of complex-tetraantennary structure, but only trace for other quail species. The complex-tetraantennary structure in BSQEW was the lowly expressed antennary type, while the hybrid-triantennary structures were lowest for BWQEW, JQEW, and MQEW.

2.4.6 Expression of Distinct and/or Abundant N-glycans is Found in Specific Quail Species

From 21 relevant glycan peaks, 18 peaks were statistically significant using single-way Analysis of Variance (ANOVA) at 95% confidence limit. Some of the glycan peak areas were higher in specific species and some peaks are distinctly found in certain species. Trace amounts were also found, and notations are specified in Table 2. The abundant ions were defined as a peak belonging to one quail type that is statistically significant at 99% confidence interval for ANOVA and statistically significant among the 3 other quail types using post-test Tukey's multiple comparison test at 95% confidence interval. The values which are statistically significant, but do not meet the criteria, were not considered abundant. The Tukey's multiple comparison test details are presented in Table S1. Abundant ions were found at *m/z* values 1647, 1809, 1850, 1972, and 2012 for BSQEW, at 1524, 1686, and 1848 *m/z* values for JQEW, and at 1708 and 1768 *m/z* values for MQEW.

Small peaks were considered as trace, if present but not quantifiable by the *FlexAnalysis* 3.0 software. Figures were added in Figures S2 (i to vii) to show that these ions were too small to be quantified but were present, e.g. at *m/z* 1524, 1708, 1911, 1930, 1972, 2012, and 2133. Specific comparison of the *m/z* regions is discussed in detail as follows.

In Figure 2-3, high mannose structures were found in JQEW at *m/z* 1362, 1524, 1686, and 1848 corresponding to $(Man)_5(GlcNAc)_2$, $(Man)_6(GlcNAc)_2$, $(Man)_7(GlcNAc)_2$, and $(Man)_8(GlcNAc)_2$, respectively. Species variations in the expression levels of individual structures were seen. $(Man)_5(GlcNAc)_2$ is in all four species, $(Man)_6(GlcNAc)_2$ at detectable levels in 2 species (BWQEW and JQEW), and $(Man)_7(GlcNAc)_2$ and $(Man)_8(GlcNAc)_2$ only in JQEW (see Figure S2-2 viii).

The abundant glycans found for BSQEW and MQEW corresponds to complex and hybrid structures. The peaks at *m/z* 1647, 1809, and 1972 are triantennary, while 1850 and 2012 *m/z* values have tetraantennary structures which are all highest for BSQEW. For MQEW abundant structures, *m/z* at 1708 and 1768 corresponds to a monosialylated and biantennary glycan, respectively. The peaks at *m/z* 1911 and 2133, corresponding to a monosialylated and tetraantennary structure, respectively, were trace for all quail species and is not statistically significant.

2.4.7 N-Glycan Expression in Quail Egg Whites

Previous glycan structure analysis was conducted for some of the quail species studied herein but comparison of glycan content was not emphasized, and comprehensive inter-species glycan comparison among quail species was never presented. All the peaks specified in Figure 2-1 were also found in the previous published work on avian species[6], except for the peak at *m/z* ~1848 which could be seen for the JQEW sample studied herein. Hirose *et. al* [6] did not include JQEW in the previous paper for inter-species glycan comparison.

The total number of glycan structures may exceed the total number of peaks observed herein as certain isomeric *N*-glycan structures may exist. The putative structures were derived from the previous references found in the *GlycoMod* Database and Functional Glycomics Gateway similar to the previous report[6]. Nevertheless, all these structures were accounted upon glycotyping calculations. Important structures were also confirmed through MALDI-MS/MS fragmentation analysis.

No fucosylated glycan structures were found in all quail types studied, in agreement with the previous report[3,6,27]. The previous report[6] showed only low abundance of these glycan types in other Galliformes avian species, i.e. not in quails. The confirmation of the fucosylated structure by MALDI-MS/MS was impossible as it was only in trace quantity[6]. The sialylglycan structures presented herein were also only in trace amounts for some quail types, but, one of the monosialylated structure was quantifiable in MQEW. The peak at *m/z* 1708 monosialylated glycan shown in Figure S2-3 was fragmented through MALDI-MS/MS which confirms the previous report of Hirose *et. al* regarding this structure.

Some *N*-glycans occurred only in trace amounts and maybe interpreted due to the difference in areas of the relevant glycan peaks. The amount of glycans found in this paper range from 100 to 100000 pmol values (per mg of lyophilized egg white) which are higher than what was previously reported[6]. Some intensities were so high that areas of the curve for the lowly expressed glycans were not quantified but peaks were still observable.

Although *N*-glycans were presented in this work, sulfated[28,29] and fucosylated[30] glycans were previously found in cells from quail embryo. The presence of O-glycans in salivary glands[30] from other avian species, as with the Galβ1-4Gal epitope[24,31,32] in egg whites were also reported.

2.4.8 The Distinct High-Mannose Structures in Japanese Quail

The total *N*-glycan contents for each quail type is specified in Table 2-1. BSQEW and MQEW have higher *N*- glycan content as compared with BWQEW and JQEW. BSQEW, BWQEW, and MQEW were highly abundant with complex type *N*-glycans while having the least expression of high-mannose type glycans. On the other hand, JQEW contains the highest expression of high-mannose structure. The difference was due mostly to the 1848 and 1850 *m/z* peak values which were very close to each other. *FlexAnalysis* 3.0 software accounts the 1848-1852 *m/z* range peaks for JQEW as originating from the parent ion peak at *m/z* 1848 (Figure S2-4) whereas no peak at *m/z* 1848 were found for BSQEW, BWQEW, and MQEW. Distinction between this peak could be seen clearly from the MALDI-MS peaks shown in Figure 2-4 even with trace amounts. Furthermore, the parent ions and

structural information were confirmed using MALDI-MS/MS for the 1848 and 1850 *m/z* peaks as shown in Figure 2-3D and Figure 2-5, respectively. The peak at *m/z* 1848 corresponds to a high mannose structure, $(Man)_8(GlcNAc)_2$ while the peak at *m/z* 1850 was for a bisecting tetraantennary structure, $(Hex)_3(HexNAc)_6$, that was highest for BSQEW. The peak at *m/z* 1848 glycan structure could also be found in pigeon egg white[24]. The high mannose structures, $(Man)_7(GlcNAc)_2$ and $(Man)_8GlcNAc_2$, were only found in JQEW. Both structures were present on a previous glycan analysis on ovalbumin[23]. Presence of lesser oligomannoside (5-6 mannose) structures were found in quail types studied in this work.

Figure 2-3. High-mannose structures found in JQEW. The spectra represent MS/MS fragmentation patterns of precursor ions from A) *m/z* 1362, B) *m/z* 1524, C) *m/z* 1686, and D) *m/z* 1848, respectively, shown as $[M+Na]^+$. The glycans, $(Man)_7(GlcNAc)_2$ and $(Man)_8(GlcNAc)_2$, shown in C and D, respectively, were only found in JQEW. Parent ions and the 552.2 *m/z* (in C) peak shown denotes structures attached to a BOA label at the reducing end. Additional annotation is shown in Tables S2-2 to S2-5.

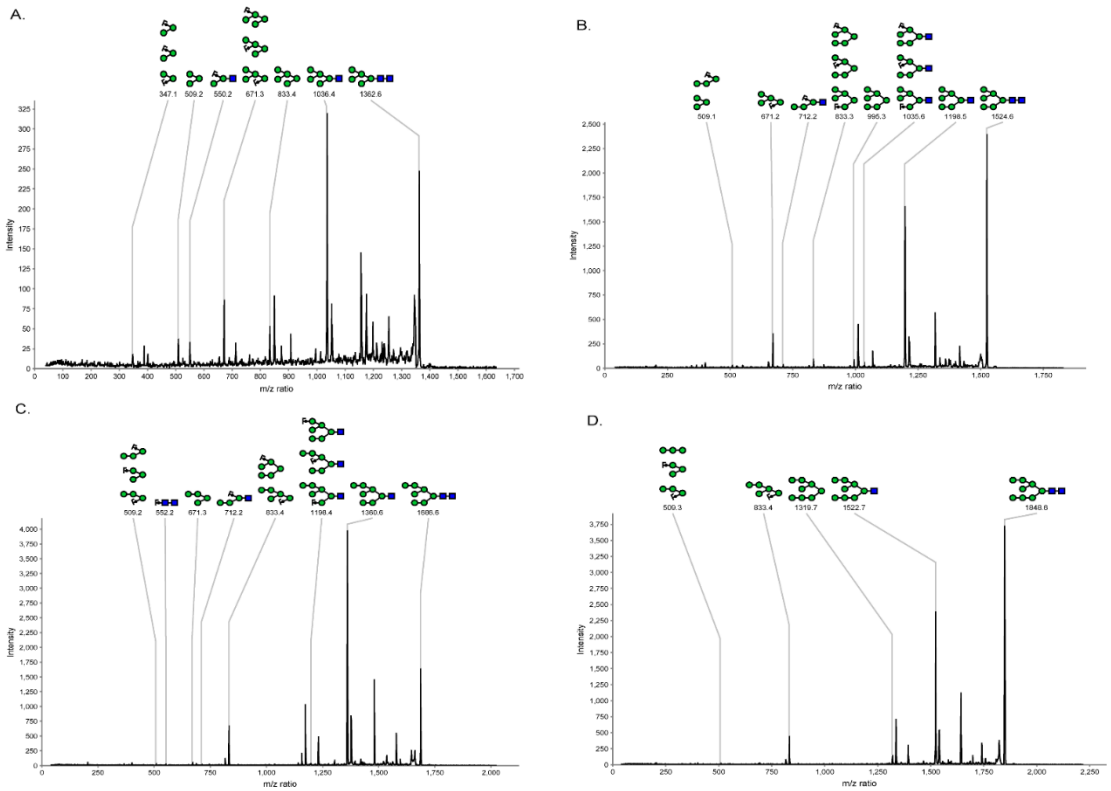


Figure 2-4. Comparison among MALDI-TOF/MS peaks at 1848 and 1850 *m/z* values for quail egg white samples shown as $[M+Na]^+$. The spectra from top to bottom represents A) BSQEW, B) BWQEW, C) JQEW, and D) MQEW, respectively. The glycan, $(Man)_8(GlcNAc)_2$, at *m/z* 1848 value can only be seen in JQEW while, $(Hex)_3(HexNAc)_6$, was highly abundant in BSQEW.

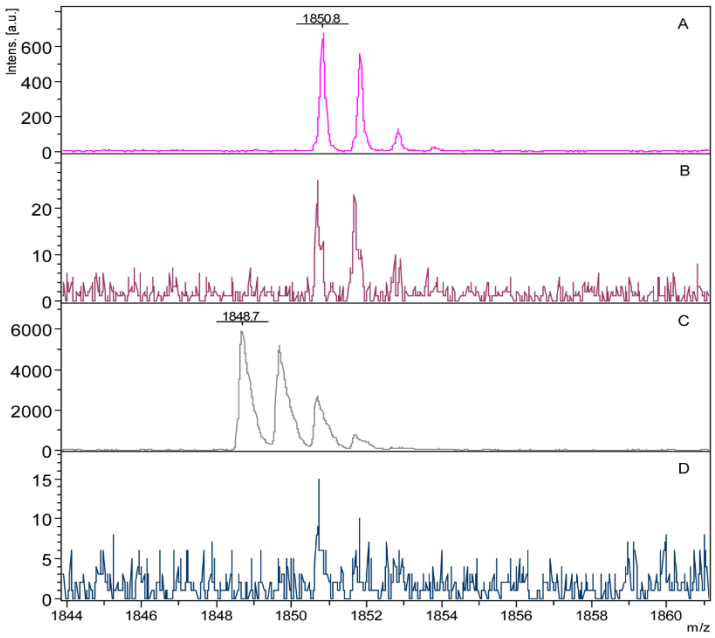
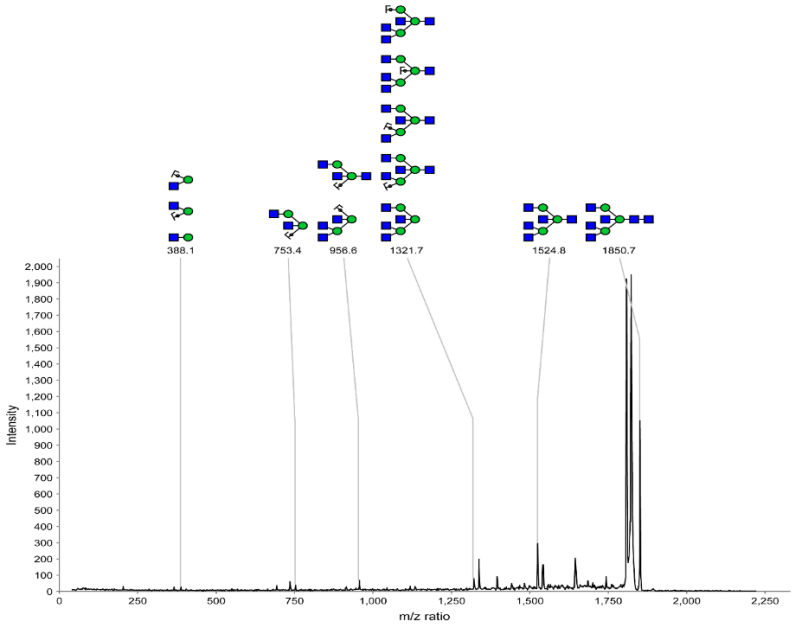


Figure 2-5. MALDI-TOF MS/MS spectra of 1850 *m/z* that confirms the identity of the $(Hex)_3(HexNAc)_6$ tetra-antennary structure for BSQEW shown as $[M+Na]^+$.



Mechanism for the Diversity of Glycan Structures

Reasons explaining diversity of glycans suggest that diversity may be due to the differences in morphology and function, may reflect differing selection pressures resulting from exposure to different pathogens, or significant intra-species polymorphism in glycan structure can exist without an obvious functional value[1].

Glycan diversity may also be due to the condition of the egg at the time of sampling. Maturity of the egg and physical state of the mother hen before egg laying would be crucial as epigenetic mechanisms [33,34], e.g. regulations of glycogenes, may generate a glycome pliability in response to different microorganisms [35,36]. Egg whites were used as samples as it serves as a barrier for microbes protecting the developing embryo.

The varying amounts of glycoproteins present in the egg whites may also affect the overall *N*-glycan amount. Recently, glycoproteomic data on *N*-glycosites for both egg white [37] and yolk [38] were conducted. Varying amounts and distinct glycan structures identified in quail egg whites may be dependent on enzymatic networks that assembles or trims the products of prior stages of synthesis to form the final glycan structure [39], of which, the presumed biosynthetic pathway was described elsewhere[40]. Existing monosaccharide structures were possibly utilized by enzymes to yield the final glycan[41].

As the quail species are from different regions of the world, BSQ, BWQ, and MQ are mostly found in the Northern Americas and in a few parts of Europe while JQ is in Asia[8,9], it may be related to the selective evolutionary characteristic needed for the specific quail to adapt to its environment. Accordingly, the quality and composition of farmed JQ eggs may be affected by different factors, e.g., age, feed composition, stocking density, and storage time[42]. The distinct glycan structures in JQ as compared to the mostly similar glycan structures for BSQ, BWQ, MQ may be related to the environment.

An alternative to glycan synthetic methodology[43,44] is to use glycans from natural sources[45,46] that is important for glycan-based drug discovery research as demonstrated in our previous

communications[47-49]. The abundant and varied supply of glycans in egg whites can serve as sources of carbohydrate structures. Particularly, the $(Man)_3(GlcNAc)_2$ glycan structure can be isolated as a template for *N*-glycopeptide production as it can be easily extended using different types of glycosyltransferases. Isolation and utilization of this conserved structure is currently being investigated and the results will be communicated as soon as possible.

The results revealed different expression of *N*-glycan structures among four quail egg white studied, i.e. glycan structural types, glycan amounts, distinct, and abundant structures. Remarkably, high-mannose glycans, $(Man)_7(GlcNAc)_2$ and $(Man)_8(GlcNAc)_2$, were only found in JQEW while $(Hex)_3(HexNAc)_6$ was present in other species. The results would suggest that a difference in glycosylation maybe likely for quail egg white samples as highlighted by glycan expressions in the Order Galliformes of birds.

2.5 Conclusion

The results revealed different expression of *N*-glycan structures among four quail egg white studied, i.e. glycan structural types, glycan amounts, distinct, and abundant structures. Remarkably, high-mannose glycans, $(Man)_7(GlcNAc)_2$ and $(Man)_8(GlcNAc)_2$, were only found in JQEW while $(Hex)_3(HexNAc)_6$ was present in other species. The results would suggest that a difference in glycosylation maybe likely for quail egg white samples as highlighted by glycan expressions in the Order Galliformes of birds. The mitochondrial DNA sequence (*see also Supporting Information 2.8.2*) comparison of the quail species reveal a difference in numbers while, glycan difference was clearly defined using glycoblotting. Of equal importance, the *N*-glycan core pentasaccharide ($Man_3GlcNAc_2$) structure is present in all the quail species. Isolation of the core pentasaccharide structure is relevant and Chapter 3 of this dissertation provides different methodologies to obtain the glycan. The goal of this chapter is to explore glycan diversity as well as searching for relevant glycan structures that can be isolated from natural sources. Both goals were achieved using the glycoblotting methodology.

2. 6 References

1 Varki A, Freeze HH, Gagneux P. Evolution of Glycan Diversity. In: Varki A, Cummings RD, Esko JD, et al., editors. Essentials of Glycobiology. 2nd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2009. Chapter 19. Available from: https://www.ncbi.nlm.nih.gov/books/NBK1942/.

2 Kovacs-Nolan, J., Phillips, M., & Mine, Y. (2005). Advances in the value of eggs and egg components for human health. *J. Agric. Food Chem. 53*(22), 8421-8431.

3 Suzuki, N., Khoo, K. H., Chen, H. C., Johnson, J. R., and Lee, Y. C. (2001) Isolation and characterization of major glycoproteins of pigeon egg white: Ubiquitous presence of unique *N*-glycans containing Galα1-4Gal. *J. Biol. Chem. 276*, 23221–23229.

4 Suzuki, N., Laskowski, M., Jr., and Lee, Y. C. (2004) Phylogenetic expression of Galα1-4Gal on avian glycoproteins: Glycan differentiation inscribed in the early history of modern birds. *Proc. Natl. Acad. Sci. U.S.A. 101(24)*, 9023–9028.

5 Suzuki, N., Laskowski, M., Jr., and Lee, Y. C. (2006) Tracing the history of Galα1-4Gal on glycoproteins in modern birds. *Biochim.Biophys. Acta-General Subjects. 1760*, 538–546.

6 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*(21), 4757-4774.

7 Miura, Y.,Hato, M., Shinohara, Y., Kuramoto,H., Furukawa, J.-I.,Kurogochi, M., Shimaoka, H., Tada, M., Nakanishi, K., Ozaki, M., Todo, S., and Nishimura, S.-I. (2008) BlotGlycoABC™, an integrated glycoblotting technique for rapid and large scale clinical glycomics. *Mol. Cell. Proteomics. 7(2)*, 370–377.

8 Lepage, D. (2014). Avibase–the world bird database. *Published online at http://avibase. bsc-eoc. org*.

9 Sibley, D. A. (2000). *The Sibley Guide to Birds*. New York: Alfred A. Knopf, Inc.

10 Nishimura, S.-I., Niikura, K., Kurogochi, M., Matsushita, T., Fumoto, M., Hinou, H., Kamitani, R., Nakagawa, H., Deguchi, K., Miura,N., Monde, K., and Kondo, H. (2004) High-throughput protein glycomics: Combined use of chemoselective glycoblotting and MALDI-TOF/TOF mass spectrometry. *Angew. Chem., Int. Ed. 44*, 91–96.

11 Cooper, C. A., Gasteiger, E., & Packer, N. H. (2001). GlycoMod–a software tool for determining glycosylation compositions from mass spectrometric data. *Proteomics*, *1*(2), 340-349.

12 Cooper C.A., Gasteiger E., Packer N. (2003) *Predicting Glycan Composition from Experimental Mass Using GlycoMod* (in) Handbook of Proteomic Methods (Conn P.M. ed.), pp. 225-232, Humana Press, Totowa, NJ.

13 Raman, R., Venkataraman, M., Ramakrishnan, S., Lang, W., Raguram, S., & Sasisekharan, R. (2006) Advancing glycomics: implementation strategies at the consortium for functional glycomics. *Glycobiology*, *16*(5), 82R-90R.

14 Lattova, E., Perreault, H., and 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(5)*, 725–735.

15 Sumiyoshi, W., Nakakita, S., Hasehira, K., Miyanishi, N., Kubo, Y., Kita, T., and Hirabayashi, J. (2010) Comprehensive analysis of N linked oligosaccharides from eggs of the family Phasianidae. *Biosci., Biotechnol., Biochem. 74*, 606–613.

16 Takahashi, N., Khoo, K. H., Suzuki, N., Johnson, J. R., and Lee, Y. C. (2001) N-Glycan structures from the major glycoproteins of pigeon egg white: Predominance of terminal Galα (1-4) Gal. *J. Biol. Chem. 276,* 23230–23239.

17 Suzuki, N., Su, T. H., Wu, S. W., Yamamoto, K., Khoo, K. H., and Lee, Y. C. (2009) Structural analysis of N-glycans from gull egg white glycoproteins and egg yolk IgG. *Glycobiology 19*, 693–706.

18 Risley, J. M., and Van Etten, R. L. (1986) Structures of the carbohydrate moiety attached to one site in the first domain of turkey ovomucoid: Elucidation by 1H NMR spectroscopy. *Carbohydr. Res.147*, 21–29.

19 Hase, S., Okawa, K., & Ikenaka, T. (1982) Identification of the trimannosyl-chitobiose structure in sugar moieties of Japanese quail ovomucoid. *J. Biol. Chem. 91*(2), 735-737.

20 Nomoto, H., Endo, T., and Inoue, Y. (1982) Preparation and characterization of fragment glycoasparagines from ovalbumin glycopeptides: Reference compounds for structural and biochemical studies of the oligo-mannose and hybrid types of carbohydrate chains of glycoproteins. *Carbohydr. Res. 107*(1), 91–101.

21 Harvey, D. J., Wing, D. R., Küster, B., and Wilson, I. B. (2000). Composition of N-linked carbohydrates from ovalbumin and co-purified glycoproteins. *J. Am. Soc. Mass Spectrom. 11(6)*, 564–571.

22 Yamashita, K., Kamerling, J. P., and 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. *J. Biol. Chem. 258*, 3099–3106.

23 Mutsaers, J. H., van Halbeek, H., Vliegenthart, J. F., Iwase, H., Kato, Y., & Hotta, K. (1985) Structural analysis of dansyl glyco-asparagines from quail ovalbumin. *Biochim. Biophys. Acta (BBA)-General Subjects*, *843*(1-2), 123-127.

24 Suzuki, N., Khoo, K. H., Chen, C. M., Chen, H. C., & Lee, Y. C. (2003) N-Glycan Structures of Pigeon IgG A Major Serum Glycoprotein Containing Termini Galα1–4Gal. *J. Biol. Chem. 278*(47), 46293-46306.

25 Hase, S., Sugimoto, T., Takemoto, H., Ikenaka, T., & Schmid, K. (1986) The structure of sugar chains of Japanese quail ovomucoid. The occurrence of oligosaccharides not expected from the classical biosynthetic pathway for N-glycans; a method for the assessment of the structure of glycans present in picomolar amounts. *J. Biol. Chem. 99*(6), 1725-1733.

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*(4), 1650-1659.

27 Oda, Y., Senaha, T., Matsuno, Y., Nakajima, K., Naka, R., Kinoshita, M., Honda, E., Furuta, I. and Kakehi, K. (2003). A new fungal lectin recognizing α (1–6)-linked fucose in the N-glycan. *J. Biol. Chem.*, *278*(34), 32439-32447.

28 Rossi, S. G., & Rotundo, R. L. (1996) Transient interactions between collagen-tailed acetylcholinesterase and sulfated proteoglycans prior to immobilization on the extracellular matrix. *J. Biol. Chem.*, *271*(4), 1979-1987.

29 Hennig, A. K., & Maxwell, G. D. (1996) Persistent correlation between expression of a sulfated carbohydrate antigen and adrenergic differentiation in cultures of quail trunk neural crest cells. *Differentiation*, *59*(5), 299-306.

30 Sriwilaijaroen, N., Kondo, S., Yagi, H., Wilairat, P., Hiramatsu, H., Ito, M., Ito, Y., Kato, K. and Suzuki, Y., (2009) Analysis of N-glycans in embryonated chicken egg chorioallantoic and amniotic cells responsible for binding and adaptation of human and avian influenza viruses. *Glycoconj J.*, *26*(4), 433-443.

31 Wieruszeski, J.M., Michalski, J.C., Montreuil, J., Strecker, G., Peter-Katalinic, J., Egge, H., Van Halbeek, H., Mutsaers, J.H. and Vliegenthart, J.F. (1987) Structure of the monosialyl oligosaccharides derived from salivary gland mucin glycoproteins of the Chinese swiftlet (genus Collocalia). Characterization of novel types of extended core structure, Gal beta (1----3)[GlcNAc beta (1----6)] GalNAc alpha (1----3) GalNAc (-ol), and of chain termination,[Gal alpha (1----4)] 0-1 [Gal beta (1----4)] 2GlcNAc beta (1----.). *J. Biol. Chem.*, *262*(14), 6650-6657.

32 Suzuki, N., Nawa, D., Su, T. H., Lin, C. W., Khoo, K. H., & Yamamoto, K. (2013) Distribution of the Galβ1-4Gal epitope among birds: species-specific loss of the glycan structure in chicken and its relatives. *PloS one*, *8*(3), e59291.

33 Kizuka, Y., Nakano, M., Miura, Y., & Taniguchi, N. (2016) Epigenetic regulation of neural N-glycomics. *Proteomics*, *16*(22), 2854-2863.

34 Dall'Olio, F., & Trinchera, M. (2017) Epigenetic Bases of Aberrant Glycosylation in Cancer. *Int. J. Mol. Sci. 18*(5), 998.

35 Lauc, G., & Zoldoš, V. (2009). Epigenetic regulation of glycosylation could be a mechanism used by complex organisms to compete with microbes on an evolutionary scale. *Med. Hypotheses*. *73*(4), 510-512.

36 Lauc, G., Vojta, A., & Zoldoš, V. (2014) Epigenetic regulation of glycosylation is the quantum mechanics of biology. *Biochim. Biophys. Acta (BBA)-General Subjects*. *1840*(1), 65-70.

37 Geng, F., Wang, J., Liu, D., Jin, Y., & Ma, M. (2017) Identification of N-Glycosites in Chicken Egg White Proteins Using an Omics Strategy. *J. Agric. Food Chem*., *65*(26), 5357-5364.

38 Geng, F., Xie, Y., Wang, J., Majumder, K., Qiu, N., & Ma, M. (2018) *N*-glycoproteomic Analysis of Chicken Egg Yolk. *J. Agric. Food Chem., 66* (43), 11510–11516.

39 Rini J., Esko J., Varki A. (2009) in *Essentials of Glycobiology* (Varki A., Cummings R.D., Esko J. D., Freeze H. H., Stanley P., Bertozzi C.R., Hart G.W., Etzler M. E., eds) 2nd Ed., pp. 63–74, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

40 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*(3), 543-551.

41 Springer, S. A., & Gagneux, P. (2013) Glycan evolution in response to collaboration, conflict, and constraint. *J. Biol. Chem. 288*(10), 6904-6911.

42 Douglas, T. A. (2013) Coturnix revolution: The success in keeping the versatile coturnix: Everything you need to know about the Japanese quail. Create Space Independent Publishing Platform, USA.

43 Lepenies, B., J. Yin and P. H. Seeberger (2010). "Applications of synthetic carbohydrates to chemical biology." *Curr Opin Chem Biol* 14(3): 404-411.

44 Wang, Z., Z. S. Chinoy, S. G. Ambre, W. Peng, R. McBride, R. P. de Vries, J. Glushka, J. C. Paulson and G. J. Boons (2013). "A general strategy for the chemoenzymatic synthesis of asymmetrically branched N-glycans." Science 341(6144): 379-383.

45 Seko, A., Koketsu, M., Nishizono, M., Enoki, Y., Ibrahim, H.R., Juneja, L.R., Kim, M. and Yamamoto, T (1997). Occurrence of a sialylglycopeptide and free sialylglycans in hen's egg yolk. *Biochim. Biophys. Acta (BBA)-General Subjects*, *1335*(1-2), 23-32.

46 Lis, H., & Sharon, N. (1978). Soybean agglutinin--a plant glycoprotein. Structure of the carboxydrate unit. *J. Biol. Chem*, *253*(10), 3468-3476.

47 Kumar H.V., R., Naruchi, K., Miyoshi, R., Hinou, H., Nishimura, S.-I. (2013) A new approach for the synthesis of hyperbranched N-glycan core structure from locust bean gum. *Org. Lett. 15*, 6268-6281.

48 Hammura, K., Ishikawa, A., Miyoshi, R., Yokoi, Y., Tanaka, M., Hinou, H., & Nishimura, S. I. (2018). Synthetic Glycopeptides Allow for the Quantitation of Scarce Nonfucosylated IgG Fc N-Glycans of Therapeutic Antibody. *ACS Medicinal Chemistry Letters*, *9*(9), 889-894.

49 Yogesh, K.V., Kamiyama, T., Ohyama, C., Yoneyama, T., Nouso, K., Kimura, S., Hinou, H. and Nishimura, S.I. (2018). Synthetic glycopeptides as a designated standard in focused glycoproteomics to discover serum cancer biomarkers. *MedChemComm*, *9*(8), 1351-1358.

2.7 Supporting Information

2.7.1 Calculations for Theoretical Glycan Mass

The calculations for the theoretical *N*-glycan mass was taken from the underivatized glycan structural mass obtained from the GlycoMod database, the mass accounted to the O-benzyloxyamine (BOA) hydrochloride label, and/or methylation of sialic-acid containing glycans for which the structure registers as a Na adduct in MS. Calculations were accounted as follows:

a. *Neutral glycan m/z peak detected*

= Underivatized glycan mass obtained from Glycomod + ~105 from BOA label + ~23 from Na

b. *Acidic glycan m/z peak detected*

= Underivatized glycan mass obtained from Glycomod + ~105 from BOA label + ~14 for methylation (per sialic acid) + ~23 from Na

The schematic of the reaction was presented elsewhere [6].

[6] Hirose, K., Amano, M., Hashimoto, R., Lee, Y. C., & Nishimura, S. I. Insight into glycan diversity and evolutionary lineage based on comparative avio-N-glycomics and sialic acid analysis of 88 egg whites of Galloanserae. *Biochemistry* ***2011****, 50*(21), 4757-4774.

2.7.2 Mitochondrial DNA Sequence Similarity

Before glycan analysis, alignment of mitochondrial sequence from the DNA accession numbers of the quail samples was performed through BLAST using the National Center for Biotechnology Information (NCBI) website (https://blast.ncbi.nlm.nih.gov/) [S-2] comparing the 3 quails to the sequence of JQ. The megablast algorithm was used. Sequence similarity of BWQ to BSQ, MQ to BSQ, and MQ to BWQ were also performed. BLAST was conducted to identify whether there is similarity within the mitochondrial sequence.

The 3 quail samples BSQ (KT722338.1), BWQ (KJ914548.1), and MQ (AY952749.1), are 84%, 84%, and 80% similar in mitochondrial DNA sequence to that of JQ (AP003195.2), respectively. Sequence similarity of BWQ to BSQ, MQ to BSQ, and MQ to BWQ, are 99%, 87%, and 86% similar, respectively. It is important to note that there is no complete mitochondrial genome sequence of MQ as compared to the other quail types. The accession number used for MQ pertains to the mitochondrial NADH dehydrogenase 2 gene. The mitochondrial DNA sequence of BWQ to BSQ, MQ to BSQ, and MQ to BWQ are slightly higher as compared when the sequence of the quail types was aligned to JQ. The complete mitochondrial DNA sequence for Pigeon (Accession Number: GU908131.1) is 92% similar to JQ, 81% to BSQ, 89% to BWQ, and no significant similarity to MQ using the megablast algorithm. However, it shows 92 and 99% similarity for Pigeon to MQ if the discontiguous megablast and blastn algortihms were used, respectively. The highly similar sequence for Pigeon to JQ may explain the expression of the $(Hex)_5$ +$(Man)_3(GlcNAc)_2$ glycan structure in both avian species.

Supplementary Tables

Table S2-1. Post-Test Statistical Comparison using Tukey's Multiple Comparison Test

Peak No.	*m/z* values	Tukey's Multiple Comparison Test (S/NS)					Abundant in
		Between	q	95%	99%	99.9%	
1	1038	BSQ VS BWQ	0.6672	NS	NS	NS	
		BSQ VS JQ	0.3423	NS	NS	NS	
		BSQ VS MQ	4.543	S	NS	NS	
		BWQ VS JQ	0.3249	NS	NS	NS	
		BWQ VS MQ	3.879	NS	NS	NS	
		JQ VS MQ	4.201	NS	NS	NS	
2	1200	BSQ VS BWQ	0.2140	NS	NS	NS	
		BSQ VS JQ	0.08924	NS	NS	NS	
		BSQ VS MQ	5.345	S	NS	NS	
		BWQ VS JQ	0.3033	NS	NS	NS	
		BWQ VS MQ	5.131	S	NS	NS	
		JQ VS MQ	5.435	S	NS	NS	
3	1241	BSQ VS BWQ	1.197	NS	NS	NS	
		BSQ VS JQ	3.076	NS	NS	NS	
		BSQ VS MQ	2.287	NS	NS	NS	
		BWQ VS JQ	1.880	NS	NS	NS	
		BWQ VS MQ	3.483	NS	NS	NS	
		JQ VS MQ	5.363	S	NS	NS	
4	1362	BSQ VS BWQ	1.289	NS	NS	NS	
		BSQ VS JQ	0.6557	NS	NS	NS	
		BSQ VS MQ	5.162	S	NS	NS	
		BWQ VS JQ	0.6329	NS	NS	NS	
		BWQ VS MQ	3.874	NS	NS	NS	
		JQ VS MQ	4.507	NS	NS	NS	
5	1403	BSQ VS BWQ	0.9784	NS	NS	NS	
		BSQ VS JQ	1.195	NS	NS	NS	
		BSQ VS MQ	4.094	NS	NS	NS	
		BWQ VS JQ	0.2162	NS	NS	NS	
		BWQ VS MQ	5.072	S	NS	NS	
		JQ VS MQ	5.288	S	NS	NS	
6	1444	BSQ VS BWQ	5.748	S	NS	NS	

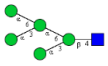
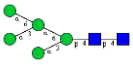
		BSQ VS JQ	7.366	S	S	NS	
		BSQ VS MQ	3.138	NS	NS	NS	
		BWQ VS JQ	1.618	NS	NS	NS	
		BWQ VS MQ	2.610	NS	NS	NS	
		JQ VS MQ	4.228	NS	NS	NS	
7	**1524**	BSQ VS BWQ	1.253	NS	NS	NS	**JQ**
		BSQ VS JQ	**25.85**	**S**	**S**	**S**	
		BSQ VS MQ	0.0	NS	NS	NS	
		BWQ VS JQ	**24.60**	**S**	**S**	**S**	
		BWQ VS MQ	1.253	NS	NS	NS	
		JQ VS MQ	**25.85**	**S**	**S**	**S**	
8	1565	BSQ VS BWQ	1.713	NS	NS	NS	
		BSQ VS JQ	3.148	NS	NS	NS	
		BSQ VS MQ	5.136	S	NS	NS	
		BWQ VS JQ	4.861	S	NS	NS	
		BWQ VS MQ	6.849	S	S	NS	
		JQ VS MQ	1.988	NS	NS	NS	
9	1606	BSQ VS BWQ	4.384	NS	NS	NS	
		BSQ VS JQ	4.937	S	NS	NS	
		BSQ VS MQ	1.452	NS	NS	NS	
		BWQ VS JQ	0.5534	NS	NS	NS	
		BWQ VS MQ	5.835	S	NS	NS	
		JQ VS MQ	6.389	S	S	NS	
10	**1647**	**BSQ VS BWQ**	**10.66**	**S**	**S**	**S**	**BSQ**
		BSQ VS JQ	**10.92**	**S**	**S**	**S**	
		BSQ VS MQ	**9.826**	**S**	**S**	**S**	
		BWQ VS JQ	0.2602	NS	NS	NS	
		BWQ VS MQ	0.8378	NS	NS	NS	
		JQ VS MQ	1.098	NS	NS	NS	
11	**1686**	BSQ VS BWQ	0.0	NS	NS	NS	**JQ**
		BSQ VS JQ	**25.28**	**S**	**S**	**S**	
		BSQ VS MQ	0.0	NS	NS	NS	
		BWQ VS JQ	**25.28**	**S**	**S**	**S**	
		BWQ VS MQ	0.0	NS	NS	NS	
		JQ VS MQ	**25.28**	**S**	**S**	**S**	
12	**1708**	BSQ VS BWQ	0.0	NS	NS	NS	**MQ**
		BSQ VS JQ	0.0	NS	NS	NS	
		BSQ VS MQ	**11.44**	**S**	**S**	**S**	
		BWQ VS JQ	0.0	NS	NS	NS	
		BWQ VS MQ	**11.44**	**S**	**S**	**S**	
		JQ VS MQ	**11.44**	**S**	**S**	**S**	
13	**1768**	BSQ VS BWQ	1.714	NS	NS	NS	**MQ**

		BSQ VS JQ	0.3471	NS	NS	NS	
		BSQ VS MQ	**6.001**	**S**	NS	NS	
		BWQ VS JQ	2.061	NS	NS	NS	
		BWQ VS MQ	**7.715**	**S**	S	NS	
		JQ VS MQ	**5.654**	**S**	NS	NS	
14	**1809**	**BSQ VS BWQ**	**13.20**	**S**	**S**	**S**	**BSQ**
		BSQ VS JQ	**13.76**	**S**	**S**	**S**	
		BSQ VS MQ	**9.312**	**S**	**S**	**S**	
		BWQ VS JQ	0.5587	NS	NS	NS	
		BWQ VS MQ	3.889	NS	NS	NS	
		JQ VS MQ	4.448	NS	NS	NS	
15	**1848**	BSQ VS BWQ	0.0	NS	NS	NS	**JQ**
		BSQ VS JQ	**32.33**	**S**	**S**	**S**	
		BSQ VS MQ	0.0	NS	NS	NS	
		BWQ VS JQ	**32.33**	**S**	**S**	**S**	
		BWQ VS MQ	0.0	NS	NS	NS	
		JQ VS MQ	**32.33**	**S**	**S**	**S**	
16	**1850**	**BSQ VS BWQ**	**35.54**	**S**	**S**	**S**	**BSQ**
		BSQ VS JQ	**35.54**	**S**	**S**	**S**	
		BSQ VS MQ	**35.54**	**S**	**S**	**S**	
		BWQ VS JQ	0.0	NS	NS	NS	
		BWQ VS MQ	0.0	NS	NS	NS	
		JQ VS MQ	0.0	NS	NS	NS	
17	1911			CC	CC	CC	
18	1930	BSQ VS BWQ	0.0	NS	NS	NS	
		BSQ VS JQ	2.000	NS	NS	NS	
		BSQ VS MQ	0.0	NS	NS	NS	
		BWQ VS JQ	2.000	NS	NS	NS	
		BWQ VS MQ	0.0	NS	NS	NS	
		JQ VS MQ	2.000	NS	NS	NS	
19	**1971**	**BSQ VS BWQ**	**88.73**	**S**	**S**	**S**	**BSQ**
		BSQ VS JQ	**88.73**	**S**	**S**	**S**	
		BSQ VS MQ	**88.73**	**S**	**S**	**S**	
		BWQ VS JQ	0.0	NS	NS	NS	
		BWQ VS MQ	0.0	NS	NS	NS	
		JQ VS MQ	0.0	NS	NS	NS	
20	**2012**	**BSQ VS BWQ**	**67.0**	**S**	**S**	**S**	**BSQ**
		BSQ VS JQ	**67.0**	**S**	**S**	**S**	
		BSQ VS MQ	**67.0**	**S**	**S**	**S**	
		BWQ VS JQ	0.0	NS	NS	NS	
		BWQ VS MQ	0.0	NS	NS	NS	
		JQ VS MQ	0.0	NS	NS	NS	
21	2133			CC	CC	CC	

Critical Values for Tukey Multiple Comparison Test for 4 quail types at each m/z value for df=3 is 4.50 for α=0.05, 8.26 for α=0.01, and 18.28 for α=0.001, S- Significant, NS- Not Significant, CC -Cannot be calculated.

The abundant peaks were defined as a peak belonging to one quail type that is statistically significant at 99% confidence interval for ANOVA and statistically significant among the 3 other quail types using post-test Tukey's multiple comparison at 95% confidence interval. The abundant peaks are highlighted. The values which are statistically significant, but, do not meet the criteria in both ANOVA and Tukey's test were not considered abundant.

Table S2-2. MALDI-TOF/TOF Fragmentation Patterns from 1362 *m/z* Parent Ion from JQEW

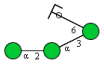
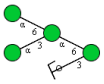
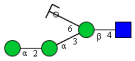
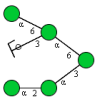
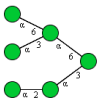
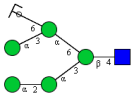
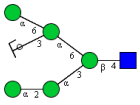
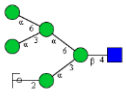
m/z	Theoretical Fragmentation *m/z*	Ion Type	Structure(s)		
347.1	347.0949	BY			
388.2 401.2					
509.2	509.1477	B			
550.2	550.1742	BY			
671.2	671.2005	BY			
712.2	713.2111	$B^{0,2}X_{Man}$			
833.4	833.2533	B			
849.3 907.4					
1036.4	1036.3327	B			
1052.4 1156.5 1175.5 1198.5 1254.7 1344.8 1348.3					
1362.6	1362.4226	Parent Ion w/ BOA			

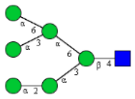
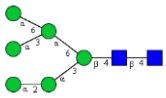
w/BOA-with O-benzyloxyamine hydrochloride, $[M+Na]^+$

Glycan structures were prepared and annotated using GlycoWorkBench [26]

Ion types are based on the fragmentation patterns by Domon & Costello, 1988 [S-2]

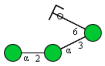
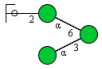
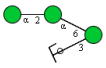
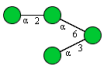
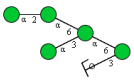
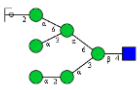
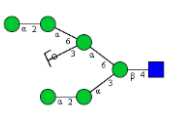
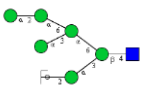
Table S2-3. MALDI-TOF/TOF Fragmentation Patterns from 1524 *m/z* Parent Ion from JQEW

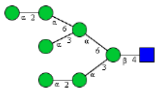
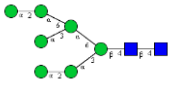
m/z	Theoretical Fragmentation *m/z*	Ion Type	Structure(s)
204.1*	203.0794	BY	
366.0			
388.1			
401.1			
509.2	509.1477	BY	
		B	
550.2			
653.2			
655.2			
671.2	671.2005	BY	
712.2	712.2271	BY	
833.3	833.2533	BY	 
995.3	995.3062	B	
1011.3			
1035.7	1036.3327	BY	  

1069.4			
1198.5	1198.3855	B	
1214.5			
1318.6			
1337.5			
1359.6			
1373.7			
1416.6			
1499.6			
1503.2			
1524.6	1524.4755	Parent Ion	 w/ BOA

w/BOA-with O-benzyloxyamine hydrochloride, $[M+Na]^+$, * signifies $[M+H]^+$
Glycan structures were prepared and annotated using GlycoWorkBench [26]
Ion types are based on the fragmentation patterns by Domon & Costello, 1988 [S-2]

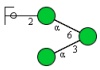
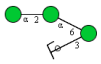
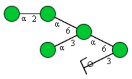
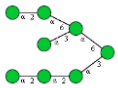
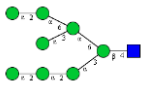
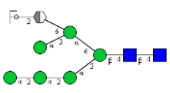
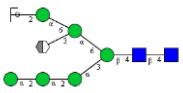
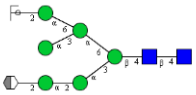
Table S2-4. MALDI-TOF/TOF Fragmentation Patterns from 1686 *m/z* Parent Ion from JQEW

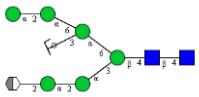
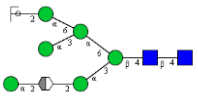
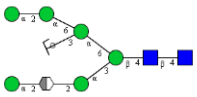
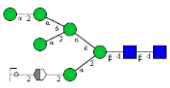
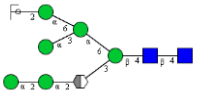
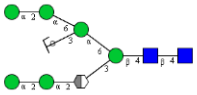
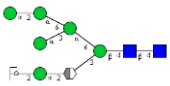
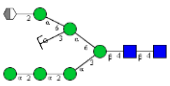
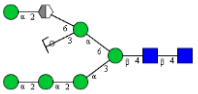
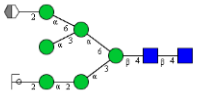
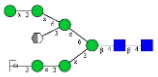
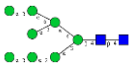
m/z	Theoretical Fragmentation *m/z*	Ion Type	Structure(s)		
204.1*	203.0794	BY			
366.1 401.2					
509.2	509.1477	BY			
552.2	552.1585	Y	 w/BOA		
671.3	671.2005	B			
712.3	712.2271	BY			
815.4					
833.4	833.2533	BY			
1141.5 1157.5 1173.5					
1198.4	1198.3855	BY			
1231.5 1303.6					

1360.6	1360.4384	B	
1376.6			
1378.6			
1480.7			
1535.7			
1578.7			
1645.8			
1660.6			
1686.6	1686.5283	Parent Ion w/BOA	

w/BOA-with O-benzyloxyamine hydrochloride, $[M+Na]^+$, * signifies $[M+H]^+$
Glycan structures were prepared and annotated using GlycoWorkBench [26]
Ion types are based on the fragmentation patterns by Domon & Costello, 1988[S-2]

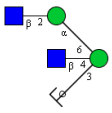
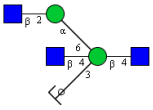
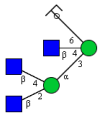
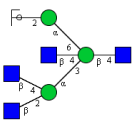
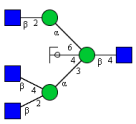
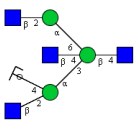
Table S2-5. MALDI-TOF/TOF Fragmentation Patterns from 1848 *m/z* Parent Ion from JQEW

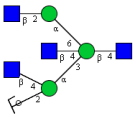
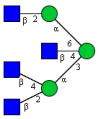
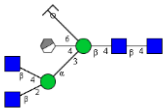
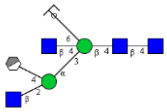
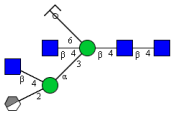
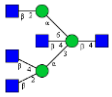
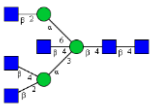
m/z	Theoretical Fragmentation *m/z*	Ion Type	Structure(s)		
204.2*	203.0794	BY			
509.3	509.1477	BY			
		B			
815.5 818.2					
833.4	833.2533	BY			
1319.7	1319.4118	B			
1335.6 1393.7 1465.7					
1522.7	1522.4912	B			
1538.7					
1582.8	1582.4810	$^{2,5}X_{Man}Y$	 w/BOA	 w/BOA	 w/BOA

			 w/BOA	 w/BOA	 w/BOA
			 w/BOA	 w/BOA	 w/BOA
			 w/BOA		
		$Y^{2,5}X_{Man}$	 w/BOA	 w/BOA	 w/BOA
			 w/BOA	 w/BOA	
1642.7					
1697.8					
1740.8					
1757.7					
1807.9					
1814.4					
1821.5					
1826.2					
1848.6	1848.811	Parent Ion	 w/BOA		

w/BOA-with O-benzyloxyamine hydrochloride, $[M+Na]^+$, * signifies $[M+H]^+$
Glycan structures were prepared and annotated using GlycoWorkBench [26]
Ion types are based on the fragmentation patterns by Domon & Costello, 1988 [S-2]

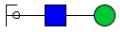
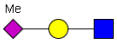
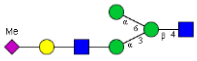
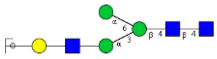
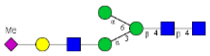
Table S2-6. MALDI-TOF/TOF Fragmentation Patterns from 1850 *m/z* Parent Ion from BSQEW

m/z	Theoretical Fragmentation *m/z*	Ion Type	Structure(s)		
204.1*	203.0794	BY			
		B			
366.1*	365.1322	BY			
		B			
388.1	388.1214	BY			
		B			
693.4	693.2325	$B^{2,4}X_{Man}$			
735.4					
753.4	753.2536	BY			
956.6	956.3330	BY			
1321.7	1321.4652	BY			

		BY			
		B			
1337.7					
1395.7	1395.4706	$^{0,3}X_{GlcNAc}Y$	 w/BOA	 w/BOA	 w/BOA
1524.7	1524.5446	B			
1540.8					
1644.8					
1807.7					
1814.3					
1818.6					
1822.7					
1826.0					
1829.2					
1850.7	1850.6345	Parent Ion	 w/BOA		

w/BOA-with O-benzyloxyamine hydrochloride, $[M+Na]^+$, * signifies $[M+H]^+$
Glycan structures were prepared and annotated using GlycoWorkBench [26]
Ion types are based on the fragmentation patterns by Domon & Costello, 1988[S-2]

Table S2-7. MALDI-TOF/TOF Fragmentation Patterns from 1708 *m/z* Parent Ion from MQEW

m/z	Theoretical Fragmentation *m/z*	Ion Type	Structure(s)
204.2*	203.0794	BY	
388.2	388.1214	BY	 
693.4	693.2325	B	
1195.7			
1382.7	1382.4703	B	
1403.1	1403.4492	Y	 w/BOA
1674.1 1676.3 1678.6 1680.9 1683.0 1685.2			
1708.7	1708.5603	Parent Ion	 w/BOA

Me-methylated sialyl glycan structure, w/BOA-with O-benzyloxyamine hydrochloride, $[M+Na]^+$,* signifies $[M+H]^+$
Glycan structures were prepared and annotated using GlycoWorkBench [26]
Ion types are based on the fragmentation patterns by Domon & Costello, 1988 [S-2]

Reference:

[S-1] Zhang, Z., Schwartz, S., Wagner, L., & Miller, W. (2000). A greedy algorithm for aligning DNA sequences. *Journal of Computational biology*, *7*(1-2), 203-214.

[S-2] Domon, B., & Costello, C. E. A systematic nomenclature for carbohydrate fragmentations in FAB- MS/MS spectra of glycoconjugates. *Glycoconjugate. J.* **1988**, *5* (4), 397-409.

Supplementary Figures

Figure S2-1. Mass spectra for quail egg white *N*-glycans showing 3 trials for A.) BSQ, B.) BWQ, C.) JQ, and D.) MQ.

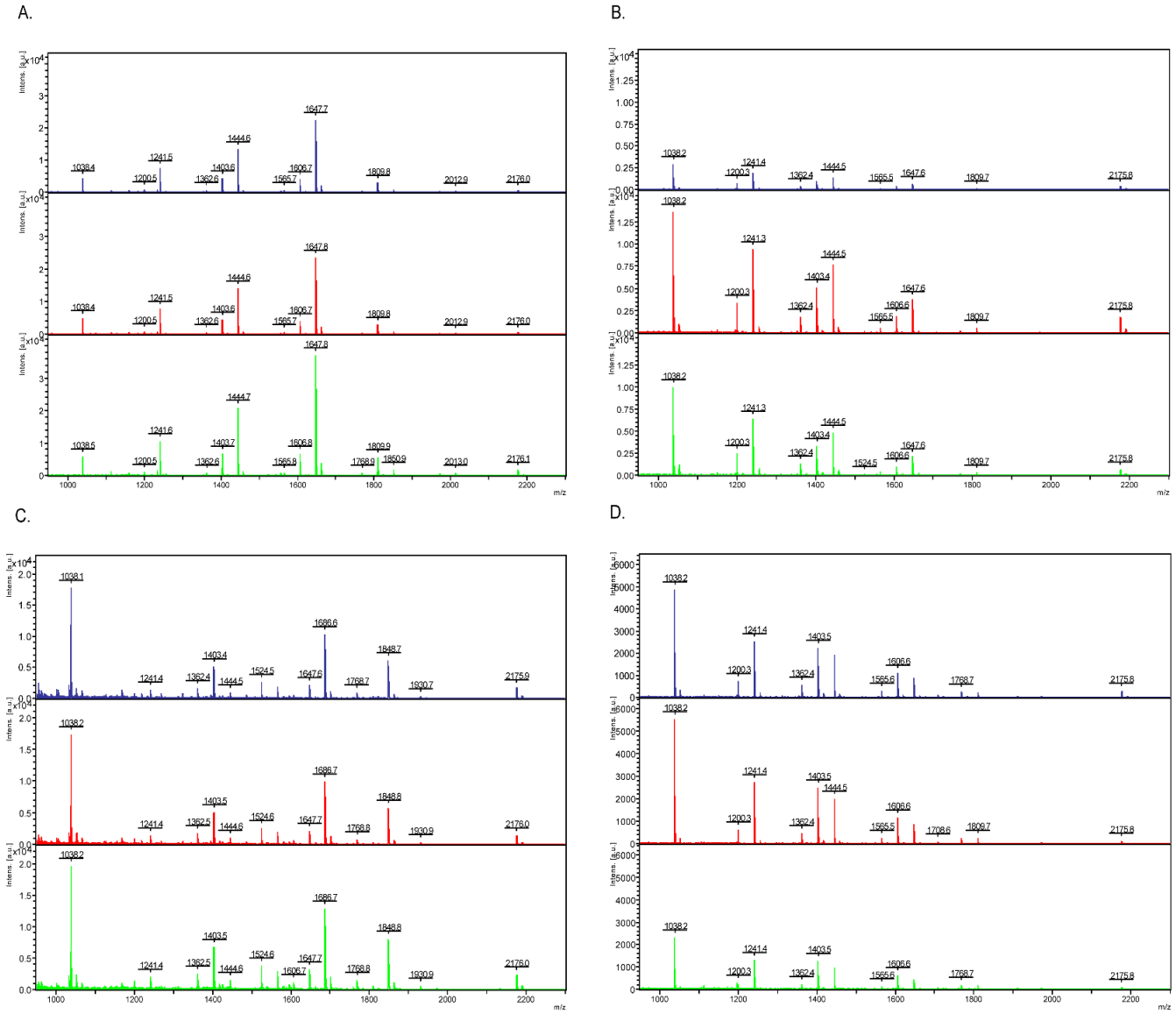


Figure S2-2. Comparison of *N*-glycan MALDI-TOF peaks for egg white quail samples. The spectra from top to bottom for each item represents A.) BSQ, B.) BWQ, C.) JQ, and D.) MQ.

i. At 1524 *m/z*. $(Man)_6(GlcNAc)_2$ was highly expressed in JQ, seen in quantifiable levels for BWQ, and considered trace for BSQ and MQ.

ii. At 1708 *m/z*. $(Hex)_4$ $(HexNAc)_3(NeuAc)_1$, a monosialylated structure, was only quantifiable for MQ.

iii. At 1911 *m/z*. $(Hex)_4(HexNAc)_4(NeuAc)_1$ glycan was considered trace for all quail types.

iv. At 1930 *m/z*. $(Hex)_6$ $(HexNAc)_4$ glycan was only quantifiable for JQ but was not statistically significant when compared among the trace quantities for all other quail types.

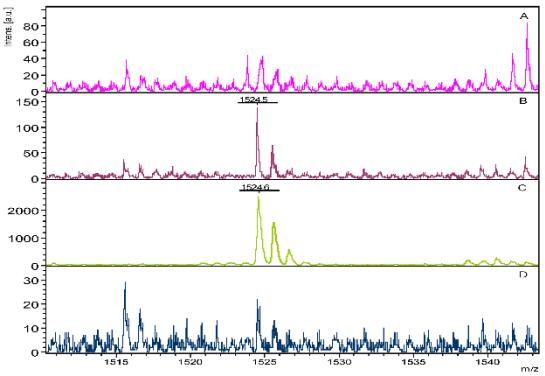
v. At 1972 *m/z*. $(Hex)_5$ $(HexNAc)_5$ glycan was highly abundant in BSQ but were only trace for other quail types.

vi. At 2012 *m/z*. $(Hex)_4$ $(HexNAc)_6$ glycan was highly abundant for BSQ but was trace in all other quail types.

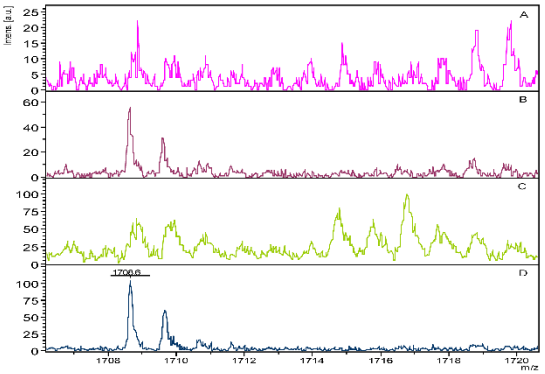
vii. At 2133 *m/z*. $(Hex)_6$ $(HexNAc)_5$ glycan was trace in all quail types.

viii. At 1686 *m/z*. $(Man)_7(GlcNAc)_2$ glycan was highly abundant in JQ but was not detected in other quail types.

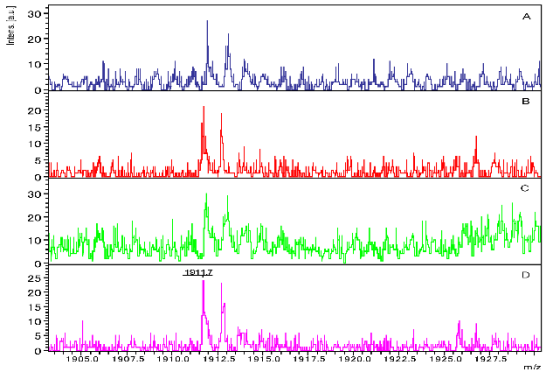
i.



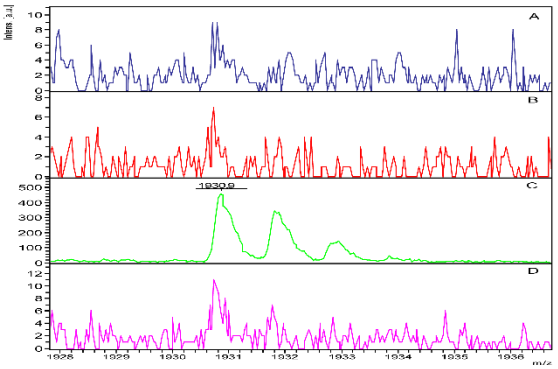
ii.



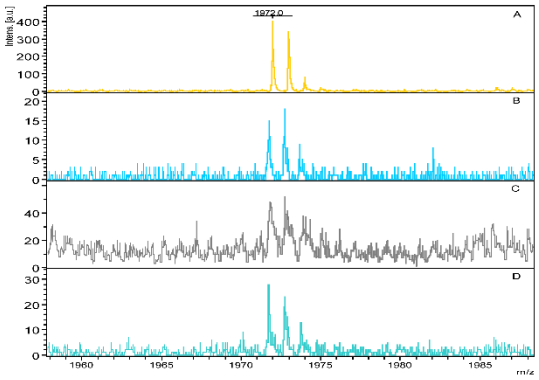
iii.



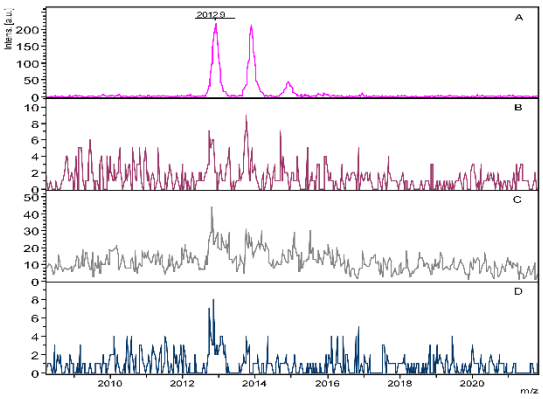
iv.



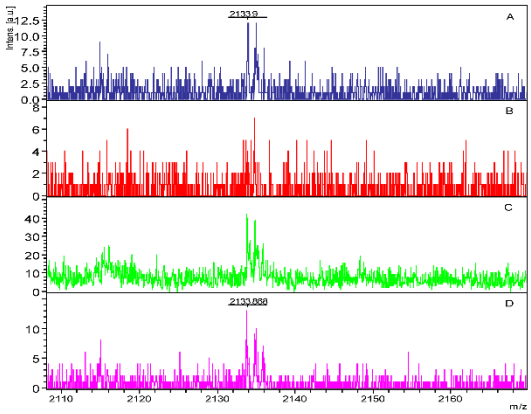
v.



vi.



vii.



viii.

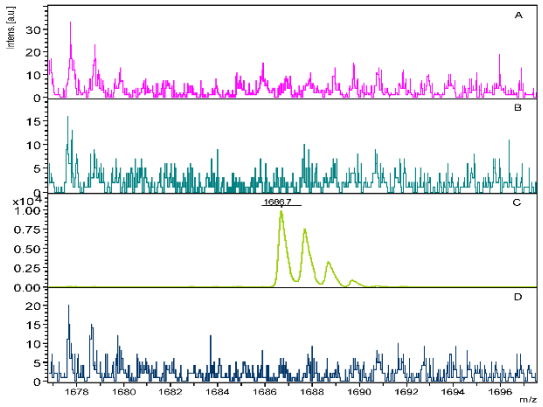


Figure S2-3. MALDI-TOF/TOF parent ion at 1708 *m/z* for MQEW. The fragmentation patterns pertain to a monosialylated glycan structure, $(Hex)_4$ $(HexNAc)_3(NeuAc)_1$ similar to the fragmentation patterns in a previous report (6). The parent ion shown denotes a structure attached to a BOA label at the reducing end. Annotation of this peak is also placed on Table S2-7 for clarity.

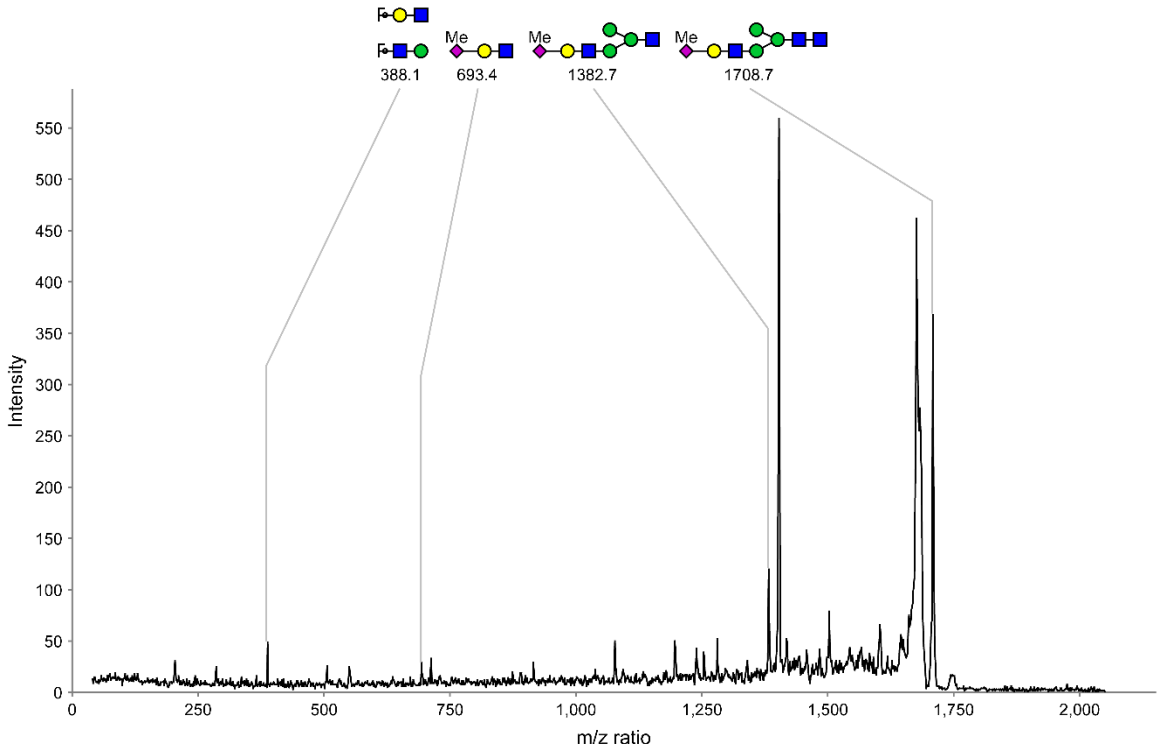
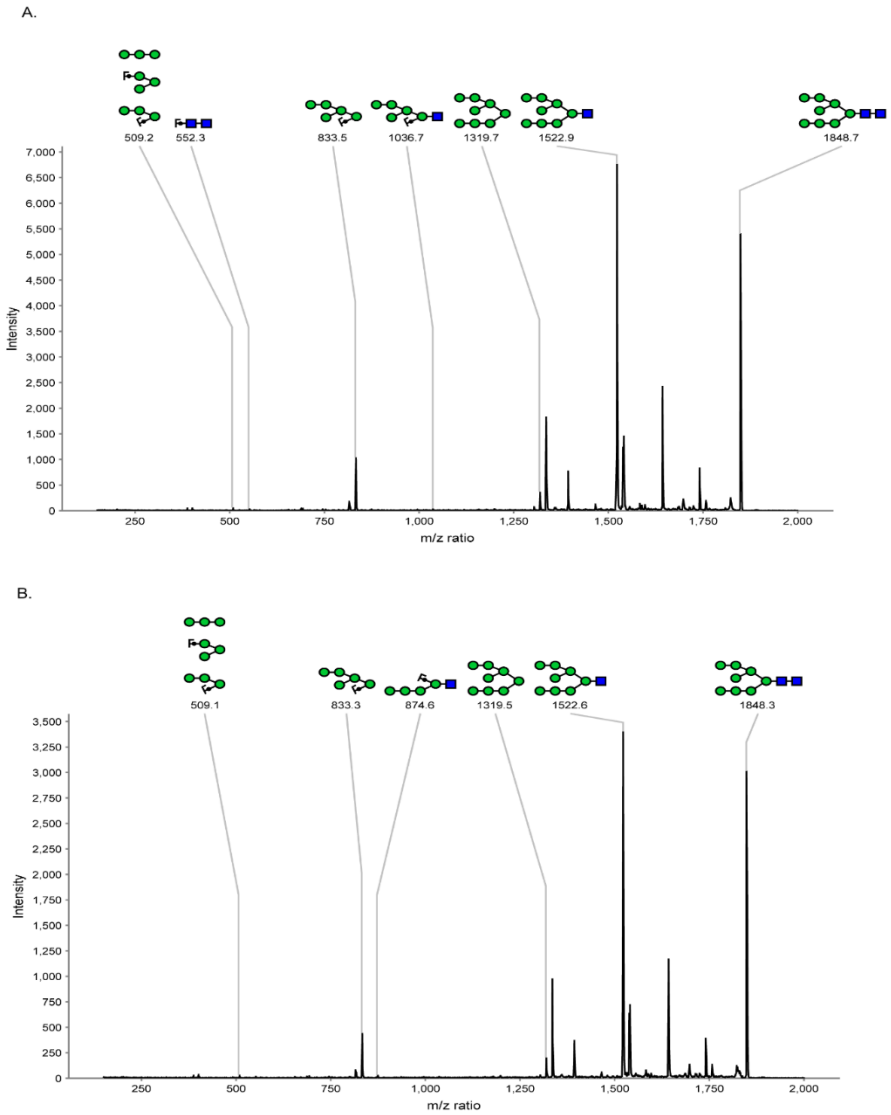
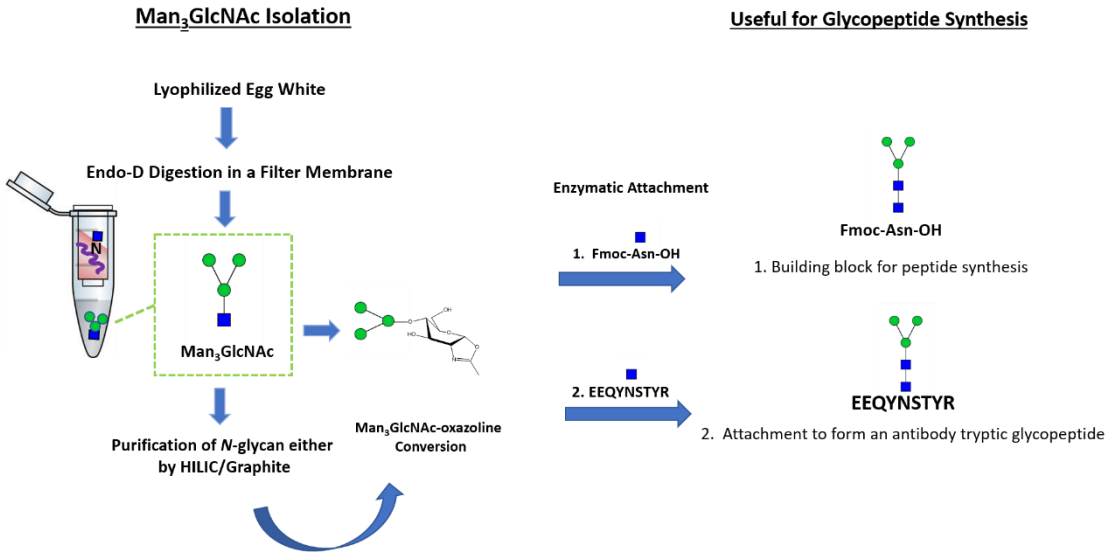


Figure S4. JQEW has a high-mannose structure at m/z 1848 rather than a complex *m/z* 1850 structure. Peak picking was conducted at the 1850 region to check whether it is the parent ion. The A.) 1848 parent ion and B.) 1850 peak obtained from peak picking for JQEW revealed the same precursor ion, both at 1848, as same parent and daughter ions can be seen from fragmentation patterns. The parent ion and chitobiose structure (at 552.3 *m/z* for A) shown in the MS/MS spectra denotes a structure attached to a BOA label at the reducing end.



Chapter 3

$Man_3GlcNAc$ Cleavage from Japanese Quail Egg White: A Strategy for Direct Glycoconjugate Preparation.



A version of this work will be submitted for journal publication.

3.1 Chapter Summary

Japanese quail egg white contains an abundant amount of the trimannosyl chitobiose core glycan, $Man_3GlcNAc_2$, which is a conserved structure among *N*-glycans. Isolating an *N*-glycan that would serve as a starting structural template for different glycoconjugate preparations is highly relevant. As such, two strategies were designed. The first strategy presented is to digest lyophilized egg white using an endoglycosidase enzyme that would cleave in between the chitobiose core, yielding $Man_3GlcNAc$. Performing digestion on a filter membrane isolates the released truncated glycans from the glycoproteins with a remaining GlcNAc attachment that would cause interference in subsequent reactions. Simple glycan purification methods, i.e. Hydrophilic Interaction Chromatography (HILIC) and graphite column chromatography (GCC), would then desalt the glycans which provide for a faster methodology to obtain $Man_3GlcNAc$. The second strategy involves tryptic digestion in regular Eppendorf tubes followed by purification using hydrazine beads to remove non-carbohydrate biomolecules. Desalting of glycans was then conducted by GCC. Incorporating the methods presented in this work is of great importance for an easy, non-tedious, and direct isolation strategy to obtain natural glycans from a complex heterogeneous starting material. As proof-of-concept, direct sugar-oxazoline conversion and transglycosylation yielded an Fmoc-Asn-OH and immunoglobulin G (IgG) glycopeptide with a core *N*-glycan attached.

Keywords: *trimannosyl chitobiose core, $Man_3GlcNAc$, sugar-oxazoline, endoglycosidase-D, direct glycan isolation, IgG1-$Man_3GlcNAc_2$*

3.2 Introduction

Quail egg whites contain a variety of *N*-glycans as shown in a previous glycan diversity study among different avian species and quails.[1-4] Particularly present in egg white is the trimannosyl chitobiose core structure, $Man_3GlcNAc_2$. To have a viable glycan structure useful in glycopeptide synthesis, sugars need to be dehydrated to form sugar-oxazolines.[5,6] In this case, *N*-glycans must be released in between the chitobiose yielding a structure with an *N*-acetyl-D-glucosamine (GlcNAc) sugar less.

Obtaining naturally sourced out *N*-glycan structures are usually tedious. Previous isolation of a natural $Man_3GlcNAc$ involves digestion of an ovomucoid glycoprotein from the egg white of Japanese quail, followed by column purification, identification of the glycan containing fraction, $Man_3GlcNAc$-release using endoglycosidase-D (Endo-D)[7], and purification of the glycan. The use of ovomucoid would mean isolation of the glycoprotein from egg white that involves a long process. Furthermore, co-elution of other glycoproteins is common when precipitation and separation strategies[8-10] are conducted, especially from a complex and viscous sample like egg white. Another preparation strategy involves exoglycosidase trimming[11] from a purified chicken egg yolk sialylglycopeptide[12]. Again, the sialylglycopeptide must be isolated from egg yolk, followed by purification protocols, glycan trimming, then $Man_3GlcNAc$-release. The previous techniques involved numerous purification steps on glycoprotein and glycopeptide isolation before actual glycan release. As such, this work details direct release and isolation of glycans, particularly, $Man_3GlcNAc$ from a heterogeneous mixture of glycoproteins with two strategic methodologies using simple chromatographic techniques.

Glycan purification usually involves the use of Hydrophilic Interaction Chromatography (HILIC) and Porous Graphitic Carbon (PGC) cartridge. When used alone, the purification strategies also involve co-elution of smaller molecular weight peptides and glycopeptides together with the glycans.[13] Some of these biomolecules are abundant and may mask glycan ionization. In addition, non-glycan biomolecules are difficult to separate because of the limited elution solvents used for

glycan isolation. One strategy in this work eliminates abundant glycoproteins by using a filter membrane. Desalting was then conducted by either HILIC or GCC, thereby, having native glycan structures.

The use of native and natural glycan structures, not labeled and not derivatized, is important for sugar-oxazoline conversion needed for glycopeptide synthesis. Another glycan isolation technique is glycoblotting, wherein, hydrazide-functionalized polymer called glycobeads are used. Glycobeads (BlotGlycoH™, Sumitomo Bakelite Co., LTD, Japan) was developed in this laboratory for the specific isolation of labeled *N*- and *O*-glycans from different biomolecules.[14-16] Owing to the reversible nature of the hydrazide-functionalized polymer[16-18], another strategy of this work uses glycobeads for the recovery of unlabeled *N*-glycans together with desalting strategies.

With the advent in glycobiology, this work presents the isolation of $Man_3GlcNAc$ from Japanese quail egg white using two strategies and by first releasing glycans using an appropriate endoglycosidase specific for truncated trimmanosyl chitobiose structures. As proof-of-concept, sugar-oxazoline conversion and transglycosylation reactions were conducted. Furthermore, this paper marks the first strategy of directly using an isolated $Man_3GlcNAc$ glycan from a natural source to sequentially form a transglycosylation product, an IgG-glycopeptide having a core *N*-glycan attachment.

3.3 Materials and Methods

3.3.1 Endoglycosidase-D(Endo-D) Release of N-Glycans from Japanese Quail Egg White

To check and confirm the truncated *N*-glycans released in Japanese quail egg white (JQEW), enzymatic digestion was conducted. Supermarket bought Japanese quail eggs were cracked and the egg white separated from the yolk and lyophilized. 0.5 mg of lyophilized egg white sample was added in a YM-10 centrifugal unit (Merck, Millipore) with 20 µL 200 mM ammonium bicarbonate (ABC), followed by the addition of a 26 µL 100 µM internal standard disialyloctasaccharide (Tokyo Chemical Industry Co., LTD.) and homogenized. Addition of 54 µL mixture of 0.06% 1-propanesulfonic acid, 2-hydroxyl-3-myristamido with 12 mM dithiothreitol in 105 mM ABC was added. Incubation at 60° C for 90 min followed. A 10 µL 123 mM iodoacetamide (IAA) was added and incubated in the dark for 1 h at RT. Addition of 2 µL of 1 U/µL units of Endoglycosidase-D (Endo-D) (New England Biolabs) followed and was incubated at 37° C for 24 h. The tube was placed in a centrifuge at 10,000 rpm for 10 minutes at RT. Further elution with addition of 100 µL MilliQ water was conducted and repeated once. The procedure stated is for Strategy 1 Endo-D digestion. The filtrates were pooled and dried in a SpeedVac for glycoblotting.

For digestion without YM-10 filter for Strategy 2, 0.5 mg egg white samples were placed in a 1.5 mL Eppendorf tube, added with the same reagents as above, together with a 10 µL of 40 U/µL Trypsin (Sigma-Aldrich) after IAA addition, then incubated for 24 h at 37° C, and heat deactivated for 10 min at 90° C. Addition of Endo-D followed, similar to the procedure above but not placed in a filter membrane.

3.3.2 Glycoblotting Methodology

The procedure was similar to a previous report on *N*-glycan analysis on egg whites.[1,2] Briefly, 500 µL BlotGlycoH™ bead suspension (10 mg/mL) was added in a 96-well multiScreen Solvinert filter plate and attached in a vacuum manifold. Water was aspirated. Addition of 20 µL ultrapure water reconstituted released *N*-glycans was mixed with 180 µL of 2% acetic acid/acetonitrile and incubated

at 80° C for 45 min or until dry. 200 µL of 2M guanidine-HCl in 16.6 mM ammonium bicarbonate, ultrapure water, and 1% triethylamine in methanol washings followed. Solvent washing was performed. Unreacted glycobeads were then capped by addition of 100 µL 10 % acetic anhydride in methanol, incubated at RT for 30 minutes, and aspirated. Washing with 200 µL of 10 mM HCl, methanol, dioxane was then conducted. Washing for each solvent was done twice. 100 µL of 100 mM 3-methyl-1-*p*-tolyltriazene (MTT) in dioxane was added to the solvinert plate and incubated at 60° C until dry. Washing twice with 200 µL of dioxane, water, methanol, and water was then conducted. Labeling with 20 µL 50 mM O-benzyloxyamine hydrochloride with 180 µL of 2 % acetic acid/acetonitrile and the solvinert plate incubated at 80 °C until dry. *N*-glycans containing truncated trimannosyl chitobiose core structures were eluted twice with 100 µL of MilliQ water, pooled and dried in SpeedVac. The labeled glycans were reconstituted with 20 µL of MilliQ water. A 1 µL of the reconstituted sample was spotted and dried on an Anchorchip MTP 384 Target Plate (polished steel TF, Bruker) right after spotting and drying 1 µL mixture of matrix (containing 20 mg/mL 2,5-DHB and 1mM NaCl in 30:70[v/v] acetonitrile: 0.1% Trifluoroacetic acid (TFA) in water) on the target plate. Experimental masses were obtained using the *FlexAnalysis* 3.0 software (Bruker Daltonics) and calculated for the unlabeled equivalent. The structures were annotated using the *GlycoMod* Database (https://web.expasy.org/glycomod/).[19-20] Glycan structures were prepared using GlycoWorkBench.[21] MALDI-TOF and MS/MS analysis was conducted using Ultraflex III (Bruker Daltonics) to further confirm the identity of truncated trimannosyl chitobiose core *N*-glycan for both labeled and unlabeled structures. Daughter ions from precursors were annotated using the carbohydrate fragmentation format from Domon & Costello (1988).[22] The mass spectral analysis was with an Ultraflex III (Bruker Daltonics) in reflector, positive ion mode, typically totaling 100x20 shots with acceleration voltage, reflector voltage, and pulsed ion extraction settings of 25.3 kV, 26.4 kV, and 100 ns, respectively. In MALDI-TOF/TOF Mode, parent ions were accelerated to 8 kV which was further accelerated to 20.1 kV in positive ion mode.

3.3.3 Isolation Strategies of Glycans from Japanese Quail Egg White

Different isolation strategies for unlabeled glycans were performed to determine the best method. The strategies include a hydrazone-linkage method using the glycoblotting procedure, HILIC and GCC. Combinations of isolation strategies were also conducted. Briefly, for *N*-glycan release on a filter membrane, either HILIC or GCC could be directly used for glycan desalting; while, for tryptic digested glycoproteins and *N*-glycans released in a normal Eppendorf tube, glycobeads were utilized followed by GCC to desalt the glycans.

3.3.3.1 Recovery of Glycans Using Hydrophilic Liquid Interaction Chromatography

GlycoWorks™ (Waters) HILIC µL elution plate was placed on a vacuum manifold, conditioned 6x with 700 µL of MilliQ H_2O and aspirated. The column was activated with 100 µL of 1% acetic acid in H_2O and aspirated. Washing twice with 200 µL of 95% acetonitrile in water containing 1% acetic acid followed. The lyophilized sample was reconstituted in 750 µL 1% acetic acid in acetonitrile and added in the GlycoWorks™ HILIC column. The sample was incubated for 60 minutes at room temperature and aspirated slowly. Washing twice with 200 µL 95% acetonitrile in water containing 1% acetic acid followed. Elution with 100 µL of 5% acetonitrile/H_2O containing 1% acetic acid followed and performed thrice. The collected filtrate was lyophilized, reconstituted in 20 µL ultrapure water, and analyzed in MALDI-TOF/MS.

3.3.3.2 Recovery of Glycans Using Graphite Column Chromatography (GCC)

A 1.0 g graphite (Wako Chemicals) was added to 10 mL MilliQ water and mixed well. 200 µL of graphite and water mixture was placed in 96-well solvinert filter plates until 1200 µL total volume was added. The solvent was removed by vacuuming for 10 minutes. The graphite mini-column was activated with 200 µL of 0.1 % TFA in 50 % acetonitrile and the solvent was then aspirated. Another round of activation was then performed. Washing with 200 µL of 0.1% TFA in MilliQ water followed and the solvent aspirated. Then, addition of 100 µL reconstituted sample in 0.1% TFA in water was conducted and the solvinert plate incubated for 10 minutes before solvent aspiration.

Washing twice with 200 µL 0.1% TFA in water was then performed. The glycans were eluted with 100 µL 0.1% formic acid in 50% acetonitrile. Elution was repeated twice, lyophilized, and analyzed in MALDI-TOF MS.

3.3.3.3 Recovery of Glycans Using Glycoblotting Methodology

The method is a deviation of the procedure presented in the previous section on glycoblotting (GB), wherein, no sialic acid- MTT methylation and no glycan labeling were performed as both reactions were not needed. Briefly, the procedure involved using 500 µL of glycobeads, aspirated, and 20 µL of water reconstituted sample together with 180 µL of 2% acetic acid in acetonitrile was added. The solvinert filter plate was incubated at 80° C for 45 min or until dry. Washing twice with 200 µL of 2M guanidine-HCl in ammonium bicarbonate, ultrapure water, and 1% triethylamine in methanol followed. Capping with the addition of 100 µL 10 % acetic anhydride in methanol followed, the solvinert plate was incubated at RT for 30 minutes, and then the solvent aspirated. Washing with 200 µL methanol to remove impurities was conducted. 200 µL of 2% acetic acid in acetonitrile was added, and then incubated at 80° C for 45 minutes. After the incubation period, elution with 100 µL ultrapure water followed and conducted twice. Additional elution with 100 µL of 5% acetic acid in water was performed and the eluents pooled. Eluted sample was lyophilized, reconstituted in 20 µL MilliQ water, and analyzed in MALDI-TOF MS. Another sample that was not lyophilized was used directly for GCC recovery.

3.3.4 Sugar-Oxazoline Conversion

From a 1 mg egg white, digested and purified with HILIC using Strategy 1 then lyophilized, sugar-oxazoline conversion was performed by adding 10 µL of 0.5 M 2-chloro-1,3-dimethyl-1H-benzimidazol-3-ium (CDMBI) and 10 µL of 1.5 M Na_3PO_4. Mixture was stirred for 2 h at 4° C and centrifuged at 4° C (13.2 rpm) for 10 minutes. Supernatant was removed and filtered. For mass spectrometry analysis and checking, a 1 µL of supernatant, cleaned with a C18 10 µL pipette tip (Merck Millipore) following manufacturer's guidelines, was spotted in a target plate with DHB as a matrix. MALDI-TOF and MALDI-TOF/TOF were conducted to confirm the formation of

$Man_3GlcNAc$-oxazoline. MetFrag (https://msbi.ipb-halle.de/MetFragBeta/)[23] and ChemDraw (Cambridgesoft) fragmentation tool was used to compare and annotate daughter ions in MALDI-TOF/TOF.

3.3.5 Transglycosylation Reaction

Further transglycosylation reactions were conducted to utilize the $Man_3GlcNAc$ from Japanese quail egg white. Oxazoline conversion was conducted first for the transglycosylation to a glycoamino acid and a glycopeptide acceptor. The details are stated below.

3.3.5.1 For Fmoc-Derivatization

The N- α – Fmoc – N – γ - (2-acetamido-2-deoxy-3,4,6-tri-O-acetyl-β-D-glucopyranosyl)-L-asparagine (Fmoc-Asn(Ac_3β-GlcNAc)-OH) (Medicinal Chemistry Pharmaceuticals, Co., LTD., Japan) was first deacetylated to serve as the glycosyl acceptor for $Man_3GlcNAc$-oxazoline transglycosylation with Endo-M-N175Q. Deacetylation was conducted by dissolving 11.18 mg of Fmoc-Asn(Ac_3β-GlcNAc)-OH in 500 µL methanol and then the pH was adjusted to 12.5 with 1 N NaOH. The solution was incubated for 1h with constant shaking and checked for complete deacetylation. The deacetylated compound was then neutralized with 1 N acetic acid and lyophilized. Purification was performed under HPLC with solvents, A. water B. acetonitrile, both containing 0.1% TFA under a gradient, 0%B (0 min)→95%B (50min)→95%B (60min) in a preparative Inertsil-ODS 3 column (20x250 mm) at 265 nm. Additional confirmation of the purified Fmoc-Asn(GlcNAc)-OH product was performed under analytical HPLC containing the same solvent with 0%B (0 min)→60%B (50min)→95%B (50.1 min)→95%B (60min) in an analytical Inertsil-ODS 3 column (4.5x250 mm) at 265 nm. Both preparative and analytical HPLC contain only 1 peak and was checked in MALDI-TOF/MS (see Figure S3).

Strategy 1 digestion was performed and purified for sugar-oxazoline conversion. Briefly, 100 mg lyophilized egg white was added onto an Amicon-Ultra 4-mL centrifugal unit with 10kDa MWCO (Merck, Millipore). The lyophilized egg white was added to 500 µL 200 mM ABC and 800 µL 0.06%

PHM with 12 mM DTT in 105 mM ABC. The solution was incubated at 60° C for 90 min. A 200 µL 123 mM IAA was added and incubated in the dark for 1 h at RT. Addition of 50 units of Endo-D (New England Biolabs) followed and the solution was incubated at 37° C for 8 days. The tube was centrifuged at 10,000 rpm for 10 minutes at RT. The filtrate was then dried in a SpeedVac before sugar-oxazoline conversion.

Sugar-oxazoline conversion was performed by adding 20 µL of 0.5 M CDMBI and 20 µL of 1.5 M Na_3PO_4. After stirring for 2 h at 4° C, 2 µL of 2 M HCl was added to neutralize the solution. The mixture was then centrifuged at 4° C (13.2 rpm) for 10 minutes. The supernatant from the previous reaction was separated and added to 5 µL 20 mM Fmoc-Asn(GlcNAc)-OH, 10 µL 50 mM phosphate buffer (pH 7.0), and 2.5 mU Endo-M-N175Q (Tokyo Chemical Industry Co. LTD, Japan) in order to effect transglycosylation. The mixture was incubated at 30° C with constant shaking and the formation of the transglycosylated product was monitored. The Endo-M-N175Q enzyme was heat-inactivated by placing the Eppendorf tube in a heat-block at 90º C for 10 min and the solution was directly analyzed for MS and MS/MS or purified under HPLC. HPLC solvents include A. water B. acetonitrile, both containing 0.1% TFA with a multi-step gradient, 0%B (0 min)→60%B (50 min)→95%B (50.1 min)→95%B (60min) in an analytical Inertsil-ODS 3 column (4.6x250 mm) at 265 nm. The wavelength is for the absorbance of the Fmoc chromophore. The transglycosylated product was found at 32 mins and analyzed for MS and MS/MS.

3.3.5.2 For Tryptic Immunoglobulin H-Glu-Glu-Gln-Tyr-Asn(GlcNAc)-Ser-Thr-Tyr-Arg-OH (IgG1-GlcNAc) Derivatization

Since the previous derivatization is only for an Fmoc-asparagine glycoamino acid, the procedure was extended with the transglycosylation reaction to a glycopeptide containing a GlcNAc attachment. The solid-phase peptide synthesis of IgG1-GlcNAc, the acceptor of the $Man_3GlcNAc$-oxazoline, is synthesized following a procedure based on a previous work from our laboratory[24]. Slight deviations from the previous procedure were based solely on the use of microwave-assisted solid phase peptide synthesis strategy in this work for faster glycopeptide synthesis of H-Glu-Glu-Gln-Tyr-Asn(GlcNAc)-Ser-Thr-Tyr-Arg-OH (IgG1-GlcNAc). Briefly, trityl-ChemMatrix resin (0.3 mmol/g,

50 mg, 15 µmol), Fmoc-amino acids (60 µmol, 4.0 eq), and Fmoc-Asn(Ac_3β-GlcNAc)-OH (18µmol, 1.2 eq) glycoamino acid were used. The identity of the Fmoc-amino acids used for the synthesis are Fmoc-Arg(Pbf)-OH, Fmoc-Tyr(tBu), Fmoc-Thr(tBu)-OH, Fmoc-Ser(tBu)-OH, Fmoc-Gln(Trt)-OH, and Fmoc-Glu(OtBu)-OH. For all microwave irradiation, the equipment used is Green Motif I microwave synthesis reactor (IDX Corp., Tochigi, Japan) set at 2,450 MHz and 50ºC.

The weighed trityl-ChemMatrix was placed in a 10 mL Libra Tube (Hipep Laboratories). The matrix was chlorinated first by addition of 2% $SOCl_2$ in 2 mL of dry CH_2Cl_2 and was agitated in a vortex mixer overnight. The solution was then removed and washing with 6x of CH_2Cl_2 followed. The first amino acid, Fmoc-Arg(Pbf)-OH (60µmol, 4.0 eq), was coupled onto the resin by first dissolving it in 210 µL CH_2Cl_2 and 22 µL of *N, N*-diisopropylethylamine (DIEA)(8.0 eq). The Libra tube was placed under microwave irradiation for 15 min with continuous shaking. After the solution was removed, washing with 85% CH_2Cl_2: 10% methanol: 5 % DIEA solution, CH_2Cl_2, *N, N*-dimethylformamide (DMF) was conducted. Each washing solvent was shaken in a vortex mixer before removal. Solvent washing was conducted thrice. Fmoc group was removed with the addition of 20% piperidine in DMF under microwave irradiation for 5 minutes. From the second Fmoc-amino acids except the glycosylated Fmoc-Asn(Ac_3β-GlcNAc)-OH amino acid, coupling was conducted by first dissolving with 150 µL 0.4 M 1- [Bis(dimethylamino)methyliumyl]-1H-benzotriazole-3-oxide hexafluorophosphate (HBTU), 1- hydroxybenzotriazole monohydrate (HOBt)/DMF (60 µmol, 4 eq.) and 16 µL DIEA (90 µmol, 6 eq.) in DMF. The glycosylated amino acid was dissolved in 8 µL of 0.4 M 1H-benzotriazol-1-yloxy-tri(pyrrolidino)phosphonium hexafluorophosphate (PyBOP), HOBt/DMF (18µmol, 1.2 eq.) and 8µL of DIEA (45µmol, 3 eq.) in DMF. Coupling of the Fmoc-amino acid and Fmoc-glycosylated amino acid was conducted under microwave irradiation for 15 mins. Another round of coupling for the Fmoc-glycosylated amino acid was performed by addition of 45 µL 0.4M PyBOP, HOBt/DMF to the Libra tube and then the solution irradiated back for 15 mins. After coupling the corresponding amino acid, washing 3x of DMF, CH_2Cl_2, and DMF followed. Capping with 85% DMF:10% acetic anhydride:5% DIEA was then conducted for 5 minutes with constant shaking at room temperature. Washing 3x of DMF, CH_2Cl_2, and DMF was again performed.

Deprotection by adding 20% piperidine in DMF then followed. A series of coupling, acetyl capping, and Fmoc-removal was repeated until the addition of the final Fmoc-amino acid. After completion of synthesis, the acetyl capped IgG1-GlcNAc was cleaved by addition of 2 mL of 95% TFA: 2.5% H_2O: 2.5% triisopropylsilane to the resin and shaken for 2h. The solution was then recovered. An additional washing with 1 mL cleavage cocktail was performed and repeated once to completely cleave the glycopeptide. The filtrate was pooled and concentrated by streaming air until around 2 mL of the solution remains. The glycopeptide was then precipitated by addition of 20 mL cold *tert*-butylmethylether and centrifuged for 10 mins. The supernatant was discarded, and the residue left was dissolved in 20 mL 50% acetonitrile in H_2O and then lyophilized. The precipitate was dissolved in 5 mL methanol and the pH adjusted to 12.5 with 1 N NaOH to deacetylate and obtain the glycosyl acceptor, IgG1-GlcNAc. The mixture was shaken for 2 h at room temperature. After which, the solution was neutralized with 1N CH_3COOH and lyophilized. The crude glycopeptide was then purified by RP-HPLC using solvents, A. water B.) acetonitrile, both containing 0.1% TFA with a linear gradient, 0%B (0 min)→20%B (60min) in a preparative Inertsil-ODS 3 column (20x250 mm) at 220 nm with 4mL/min flow rate. The purified glycopeptide was further checked in analytical HPLC Inertsil-ODS 3 column (4.5x250 mm) at 220 nm with 1mL/min flow rate using the same linear gradient. The purified IgG1-GlcNAc was then used for transglycosylation reaction.

Importantly, to determine whether skipping the purification step, i.e., HILIC, GCC, and GB after $Man_3GlcNAc$-release, can be directly used for sugar-oxazoline and transglycosylation reactions at ease with IgG1-GlcNAc, the procedure below was performed (see also Figure 5).

A similar oxazoline-conversion procedure as with Fmoc-derivatization was conducted. After drying the solution containing $Man_3GlcNAc$, conversion to oxazoline was done by adding 30 µL of CDMBI and 30 µL of 1.5 M Na_3PO_4. The mixture was stirred for 2 h at 4° C and 4 µL of 2 M HCl was added to neutralize the solution. The mixture was then centrifuged at 4° C (13.2 rpm) for 10 minutes. Supernatant was removed and used for transglycosylation reaction with IgG1-GlcNAc. Simultaneously, another 100-mg lyophilized egg white following Strategy 1 digestion without any purification and directly converted to $Man_3GlcNAc$-oxazoline skipping the HILIC/GCC purification

step was performed. Addition of reagents for oxazoline conversion was conducted similar to the procedure above and the supernatant was recovered.

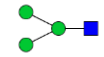
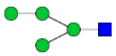
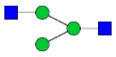
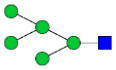
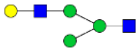
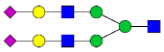
The supernatant from the sugar-oxazoline conversion steps (either with or without HILIC purification after $Man_3GlcNAc$ release) was added to 5 µL of 10 mM IgG1-GlcNAc, 10 µL of 50 mM sodium phosphate buffer (pH 7), and 2.5 mU (5 µL of 500 mU/mL) of Endo-M-N175Q. The reaction mixture was incubated at 30° C with constant shaking and the formation of the transglycosylated product monitored under MALDI-TOF MS and HPLC. Chosen time intervals were 5 min, 10 min, 30 min, and 2 h. From the incubated mixture, 1 µL was taken and diluted to 100 µL with ultrapure water containing 0.1% TFA per time interval. The Endo-M-N175Q enzyme was heat-inactivated by placing the Eppendorf tube in a heat-block at 90º C for 10 min and the solution was purified under HPLC. HPLC monitoring using solvents, A. water B. acetonitrile, both containing 0.1% TFA with a linear gradient, 0%B (0 min)→20%B (60min) in an analytical Inertsil-ODS 3 column (4.6x250 mm) with a flow rate of 1mL/min at 220 nm was used. The fractions were checked in MALDI-TOF MS and MALDI-TOF/TOF.

3.4 Results and Discussion

3.4.1 Glycans Released from Endo-D Digestion

Digesting the Japanese quail egg white released five truncated *N*-glycans, namely, $(Man)_3(GlcNAc)_1$, $(Man)_4(GlcNAc)_1$, $(HexNAc)_1 + (Man)_3(GlcNAc)_1$, $(Man)_5(GlcNAc)_1$, and $(Hex)_1(HexNAc)_1 + (Man)_3(GlcNAc)_1$. The *N*-glycans were normalized with the internal standard found at 2175 *m/z*. Figure 1 shows the mass spectra from the released labeled *N*-glycans using the glycoblotting methodology. Strategy 1 (Figure 3-1A) involves digestion of Japanese quail egg whites on a filter membrane wherein glycans were isolated by centrifugation. Strategy 2 digestion (Figure 3-1B) was conducted in an Eppendorf tube. The intensities seen in Figure 1 were further normalized with the internal standard to quantitate the glycans. Comparing the glycan structures released by Endo-D from the Peptide-*N*-Glycosidase-F (PNGaseF) released glycans[1-4,25-32] confirms that the non-truncated chitobiose versions were also found in egg white as shown in Table S1. The highest glycan found in Figure 3-1 and Table 3-1 is the $Man_3GlcNAc$ structure with 1080±71 and 1095±196, with and without the filter membrane, respectively. Confirmation of the BOA-labeled $Man_3GlcNAc$ glycan structure by MALDI-TOF/TOF MS is presented in Table S3-2. The other 4 lowly expressed glycans are shown clearly in Figure S3-1. The data between glycans digested under filtered and non-filtered conditions showed no statistical difference using t-test at 99.9% confidence interval for Peaks 1, 3, 4, and 5, while a significant difference at 95% confidence interval was seen for Peak 2. The difference in significance of the results may be accounted to the difference in the ionization of the labeled glycans. Nevertheless, with or without trypsin digestion did not affect the Endo-D release of $Man_3GlcNAc$. Analysis and quantification using the glycoblotting methodology establishes that abundant $Man_3GlcNAc$ was released by Endo-D digestion. The next step was to determine the best desalting strategy and effective isolation of glycans from other biomolecules. Two strategies were presented.

Table 3-1. Truncated chitobiose N-glycan structures in Japanese quail egg white released by using Endoglycosidase-D (Endo-D)

Peak No.	Glycan *m/z* values		Composition	*N*-Glycan Putative Structures	Amount, pmol per 0.5 mg sample	
	Unlabeled	BOA-labeled			With Filter	Without Filter
1	730	835	$(Man)_3(GlcNAc)_1$		1080±71	1095±196
2	892	997	$(Man)_4(GlcNAc)_1$		88±3	50±6
3	933	1038	$(HexNAc)_1$ + $(Man)_3(GlcNAc)_1$		28±2	108±110
4	1054	1159	$(Man)_5(GlcNAc)_1$		33±4	33±11
5	1095	1200	$(Hex)_1$ $(HexNAc)_1$ + $(Man)_3(GlcNAc)_2$		39±3	14±24
6	---	2175	Internal Standard		---	---

All *m/z* values were sodium adducts, $[M+Na]^+$. BOA label mass is approximately ~105 *m/z*.
Glycan structures were prepared using GlycoWorkBench [21]
All BOA-labeled glycans were normalized with the internal standard to obtain the glycan amounts in pmol.
Japanese quail egg white samples, n=3. Internal standard was only added for BOA-labeled samples.
--- means not applicable

Figure 3-1. Comparison of mass spectra for labeled-*N*-glycans under different reaction tubes and analyzed by glycoblotting methodology. A. Strategy 1 Endo-D digestion in a filter tube (YM-10, 10kDa cut-off). B. Strategy 2 Trypsin and Endo-D digestion in an Eppendorf tube

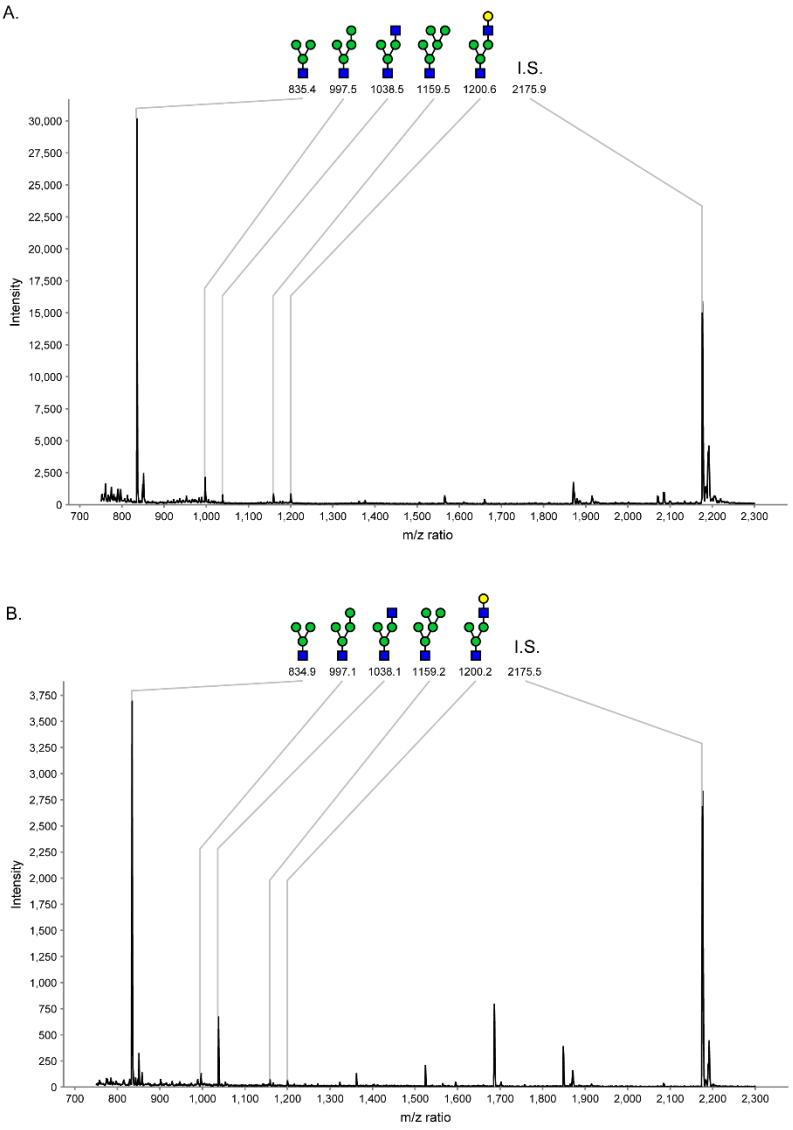
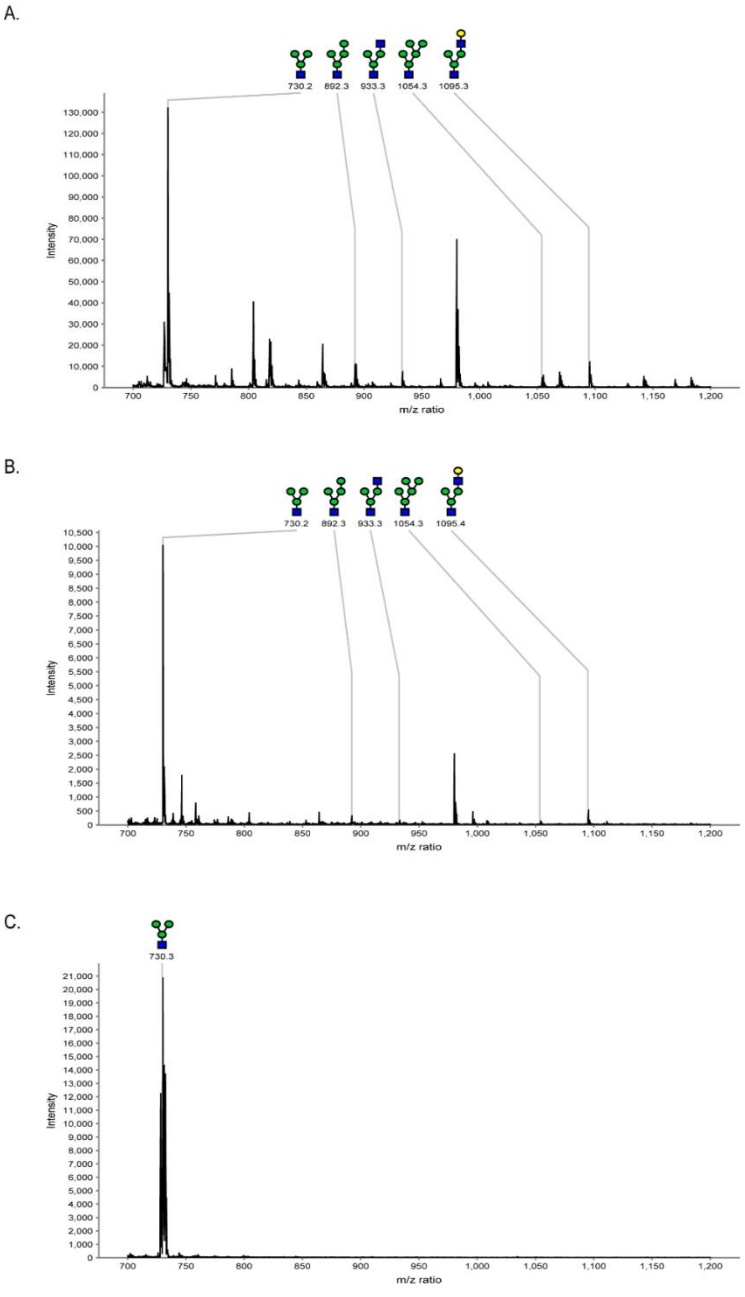


Figure 3-2. Unlabeled *N*-glycan recovery using Strategy 1 and 2. A. Strategy 1 digestion then HILIC purified. B. Strategy 1 digestion then GCC purified. C. Strategy 2 digestion then GB and GCC purified. Only one intense peak at 730 *m/z* was seen and identified to be the recovered $Man_3GlcNAc$ glycan for C.



3.4.2 First Strategy: JQEW Endo-D Digestion on a filter membrane followed by glycan desalting using HILIC or GCC

The first strategy involved conducting Endo-D digestion on a filter membrane with incubation up to 24 h followed by centrifugation to obtain the released truncated *N*-glycans. The results in Figure 3-2 show the difference among desalting strategies of the released glycans. Trypsin digestion was not conducted for egg white incubated in the filter membrane to prevent the cleavage of abundant glycoproteins in glycopeptides, which could mask off glycan ionization and contaminate subsequent reactions. In that way, glycans could be easily isolated using the filter. Whichever glycan desalting is used, HILIC or graphite, would work. Other unknown peaks in the spectra were also seen after purification. Comparing relative intensities (Table S3-3) in either HILIC or GCC, $Man_3GlcNAc$ showed higher values compared to other 4 glycan structures. Tables S3-4 to S3-8 confirm the structures for unlabeled truncated glycans.

Glycoblotting purification for Strategy 1 digestion was also conducted. But, the methodology makes use of salts for washing, i.e., guanidine-HCl and ammonium bicarbonate, that may render less ionization of the low abundant glycans in mass spectrometry (MS). In addition, acetic acid followed by water was used to elute the glycans. Further dilution was necessary since less crystals were found upon spotting the MS target plates; the implication would be that an additional desalting procedure is needed after glycoblotting. Same was true when no glycan purification was conducted for the Endo-D digested egg white. Since the purpose was to have a direct desalting procedure right after glycan filter membrane isolation, further purification after glycoblotting was no longer necessary as either HILIC or GCC would already suffice for Strategy 1. As such, the glycoblotting purification was reserved for the second strategy after glycan release.

3.4.3. Second Strategy: JQEW Endo-D Digestion followed by Isolation of $Man_3GlcNAc$ Using Hydrazone Beads and GCC

The second strategy involved Endo-D digestion of Japanese quail egg white in a regular Eppendorf tube followed by glycoblotting and GCC. The unlabeled $Man_3GlcNAc$ could not be seen in the mass spectra when glycoblotting was conducted alone as explained above. The theoretical unlabeled $Man_3GlcNAc$ mass at 730 *m/z* could not be seen in glycoblotting alone, which would suggest that ionization of the glycan structure was masked with salts present, thereby, a desalting procedure was needed. Typically, PGC is used for desalting glycans. The graphite carbon which was used in this strategy also provided clear desalting of $Man_3GlcNAc$ as could be seen in Figure 3-2C. Even when tryptic digestion was performed, small peptides were not observed in the mass spectra. Having a basic solution like ammonium bicarbonate was necessary to remove other biomolecules and at the same time maintaining $Man_3GlcNAc$ attachment on the glycobeads. In this case, desalting was necessary. The hydrazide functionalized polymer would just be specific for aldehydes and ketones, which would be the form a reduced glycan (hemiacetal in acetic condition) [16,18] would take after Endo-D digestion. The procedure for hydrazine and graphite recovery for *N*-glycan expression analysis in serum was also performed by Yang and Zhang (2012) [18] for relative quantitation but not for the purpose presented in this paper. Since glycans on peaks 2 to 5 (as seen in Table 1) are in low abundance, the $Man_3GlcNAc$ was recovered well after glycoblotting and GCC. Starting with HILIC or GCC with a diluted solution after glycan release for this strategy gave convoluted spectra that featured $Man_3GlcNAc$ (seen at 731 *m/z*) but with lower intensities because of highly abundant salts and tryptic peptides (see Figure S3-2) as expected. As such, the GB and GCC strategy 2 is beneficial for samples which have undergone tryptic digestion.

3.4.4 Relevance of $Man_3GlcNAc$ for Sugar-Oxazoline Formation

Both the first and second strategies provide for routes on obtaining the truncated *N*-glycan template. The first strategy would also isolate small glycan amounts other than $Man_3GlcNAc$; this does not limit the potential of this strategy in anyway as the other glycans are less abundant. If in large scale, further separation would be easily conducted through High Performance Anion-Exchange

Chromatography (HPAEC) or through HPLC conditions with an appropriate column using a refractive index detector or charged aerosol detection. Otherwise, direct sugar-oxazoline conversion[33-34] could be further conducted for $Man_3GlcNAc$ obtained from this work. As proof-of-concept, $Man_3GlcNAc$ was fully converted to $Man_3GlcNAc$-oxazoline clearly seen at 712 *m/z* $[M+Na]^+$ (Figure 3-3A). No peak at 730 *m/z*, $[M+Na]^+$ was observed, indicative of complete dehydration of $Man_3GlcNAc$ to the sugar-oxazoline counterpart. The peak was fragmented (Figure 3-3B) to confirm the 712 *m/z* precursor ion. All experimental m/z values for the daughter ions correspond to the *in silico* fragment data using combined MetFrag and ChemDraw fragmentation tool, which was annotated in Table S3-9. Furthermore, no sugar-oxazolines that may be formed from the other 4 low abundant glycans were seen from the mass spectra. The $Man_3GlcNAc$-oxazoline may be used for a variety of applications in chemoenzymatic synthesis of various glycoconjugates. [11,35-36]

Figure 3-3. Complete $Man_3GlcNAc$-oxazoline conversion. A. MALDI-TOF-MS spectrum at 712 *m/z*, $[M+Na]^+$. $Man_3GlcNAc$-oxazoline was purified by using C18 Ziptip and no $Man_3GlcNAc$ was seen at 730 *m/z*. B. MALDI-TOF/TOF MS spectra confirmation for $Man_3GlcNAc$-oxazoline, $[M+Na]^+$. The peak with * signifies an $[M+H]^+$ ion while all other annotated structures are $[M+Na]^+$ *m/z* ions.

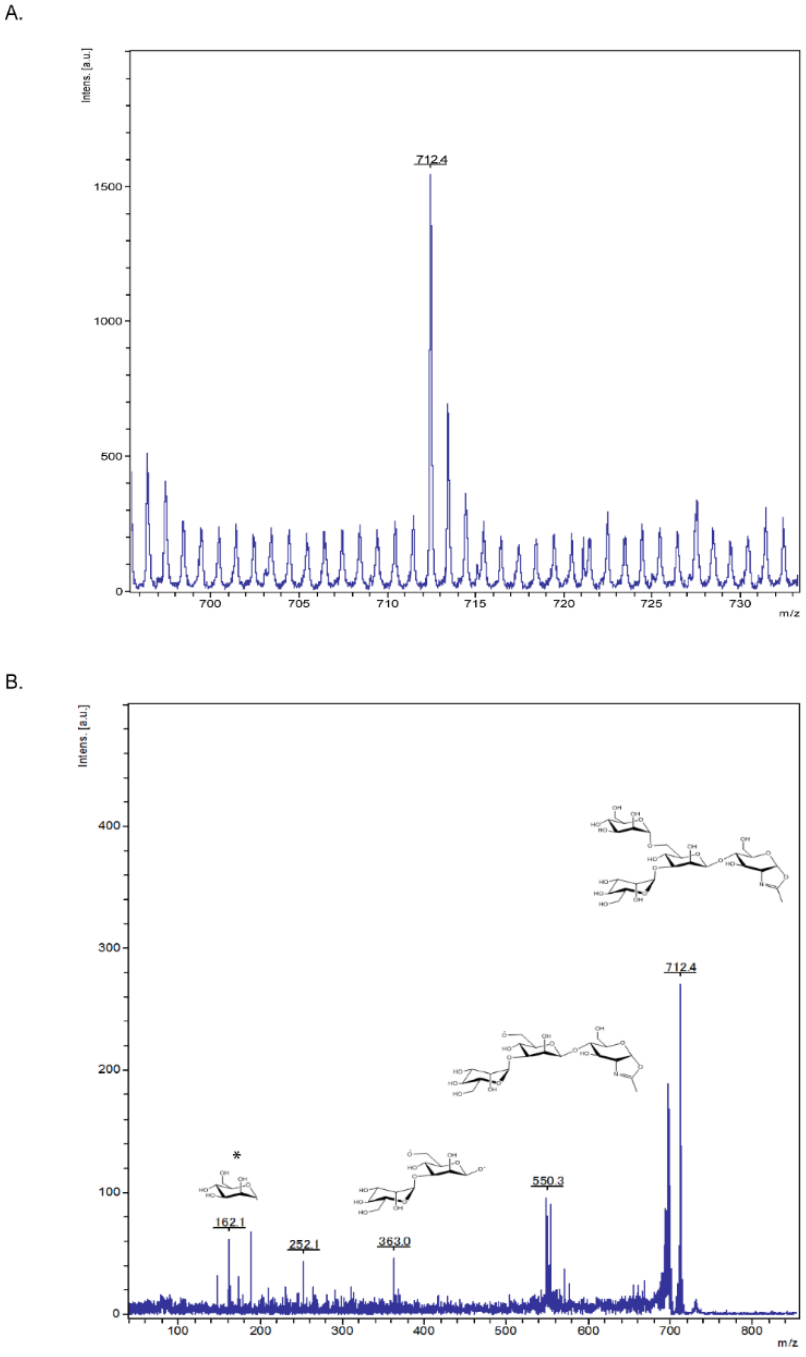
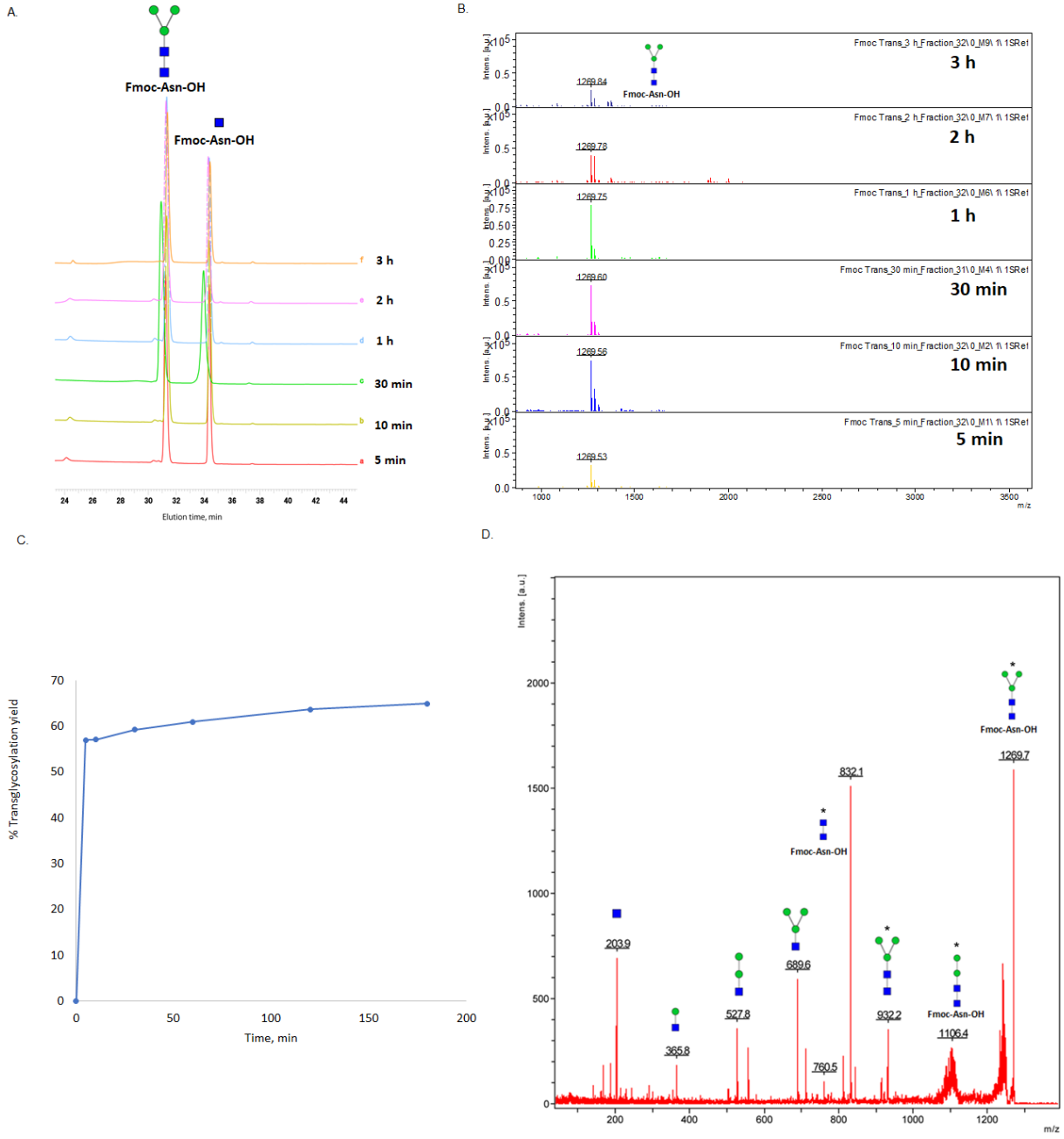


Figure 3-4. Formation of Fmoc-Asn($Man_3GlcNAc_2$)-OH Transglycosylation Product. A.) HPLC Spectra. B.) MALDI-TOF/MS Spectra for Fraction ~32 showing peak at 1269 m/z at different incubation time. C. % Transglycosylation Yield for different incubation time. D. MALDI-TOF/TOF MS spectra confirmation Fmoc-Asn($Man_3GlcNAc_2$)-OH at 1269 *m/z*, $[M+Na]^+$. The peaks with * signifies $[M+Na]^+$ ions while all other annotated structures are $[M+H]^+$ *m/z* ions. Table S3-9 shows detailed fragment annotations for the transglycosylation product.



3.4.5 Transglycosylation Reaction Fmoc-Asn(Ac_3β-GlcNAc)-OH and IgG1-GlcNAc Acceptors

In this work, transglycosylation reaction from $Man_3GlcNAc$-oxazoline was performed using Endo-M-N175Q to two purified glycosyl acceptors, Fmoc-Asn(GlcNAc)-OH and IgG1-GlcNAc, with 8.68 mg (78% yield) and 1.9 mg (9.5% yield), respectively.

For the transglycosylation reaction with Fmoc-Asn(GlcNAc)-OH, a transglycosylated product was already observed for 5 minutes of incubation. HPLC purification showed the Fmoc-Asn($Man_3GlcNAc_2$)-OH product at 1269.53 *m/z* (Calculated $[M+Na]^+$, 1269.43) with ~60% transglycosylation yield across the incubation period. HPLC, MS, and confirmatory MALDI-TOF/TOF spectrum is shown in Figure 3-4, which reinforces the presence of the Fmoc-Asn($Man_3GlcNAc$)-OH transglycosylated product. The applicability of the method was further shown by attaching the converted natural truncated core *N*- glycan-oxazoline to an IgG1-GlcNAc.

From the synthesized IgG1-GlcNAc glycopeptide, a transglycosylated reaction was conducted right after oxazoline formation. The HPLC and mass spectra of the purified IgG1-GlcNAc are shown in Figure S3-4. Either a HILIC purified $Man_3GlcNAc$ after $Man_3GlcNAc$ cleavage or a continuous sugar oxazoline conversion after glycan release as shown in Figure 3-5 was performed. After oxazoline conversion, no purification was conducted other than centrifugation and separation of precipitate as CDMBI was used. [33] Both procedures yielded a transglycosylated product even for 5-min incubation (see Figure S3-5) with Endo-M-N175Q enzyme under HPLC monitoring. Figure 3-6 shows the IgG1-$Man_3GlcNAc_2$ product and the comparison between the transglycosylation yields up to 2 hours. Transglycosylation yield was higher in purified $Man_3GlcNAc$ as compared with the glycan which was not purified for all the time-course reaction monitored. The procedure was almost a one-pot transglycosylation reaction involving release of $Man_3GlcNAc$ from Japanese quail egg white, then oxazoline conversion, and transglycosylation reaction with IgG1-GlcNAc. Nevertheless, eliminating purification strategies also gave out a transglycosylated product which would be important for a faster methodology in large-scale synthesis. Even though it is not fully fragmented, MALDI-TOF/TOF confirms the identity of the transglycosylated product as shown in Figure S3-6 as the cleaved component as well as the parent peak was seen. Aside from Endo-M-N175Q, transglycosylation can

be accomplished using other endoglycosidases, i.e., Endo-A[11,37] and Endo F1[35] to possibly enhance the yield especially for IgG1-GlcNAc transglycosylation. Another possible way is to use high amounts of the donor glycan. With this, an abundant amount of the endoglycosidase enzyme is important for the strategy mentioned.

Even when the procedures for purification, isolation strategies, and the glycan specificity of the endoglycosidase enzyme used in this work have been around for many years, incorporating all the methods for a direct isolation strategy involving $Man_3GlcNAc$ in a heterogeneous mixture of glycoproteins from Japanese quail egg whites has not been conducted as to our knowledge. Separating glycoproteins from a natural source and obtaining a homogeneous glycopeptide involves a tedious process. The methodology presented in this work eliminates the use of different expensive enzymes necessary for glycan trimmings to achieve a $Man_3GlcNAc$ *N*-glycan structure and skips laborious column purification, fraction checking, and extraction procedures otherwise obtained starting from a glycopeptide or purified glycoprotein. Depending on the *N*-glycan of interest which is abundantly expressed from a complicated sample, fishing out the glycan using a specific endoglycosidase enzyme involving the strategy presented herein, would make glycan preparations and further reactions, i.e. oxazoline conversion for transglycosylation reaction as simple as possible.

Figure 3-5. Scheme for direct IgG1-$Man_3GlcNAc_2$ synthesis using $Man_3GlcNAc$ obtained from Japanese quail egg white. Both purified and unpurified $Man_3GlcNAc$ were converted to an oxazoline by addition of CDMBI and Na_3PO_4. $Man_3GlcNAc$-oxazoline was utilized directly and mixed with the synthesized IgG1-GlcNAc glycosyl acceptor, together with Endo-M-N175Q and sodium phosphate buffer (pH 7). The mixture was incubated at 30ºC to obtain IgG1-$Man_3GlcNAc_2$. The final reaction mixture was then purified under HPLC.

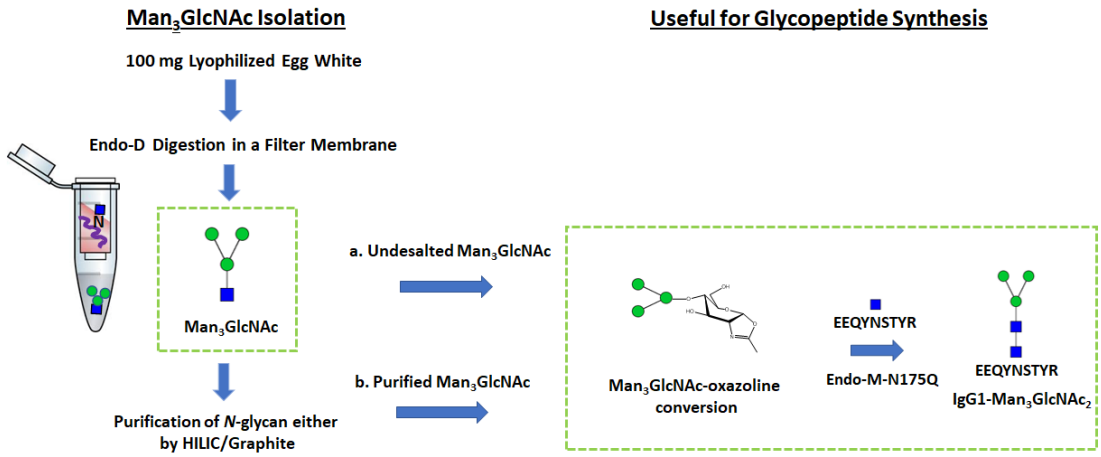
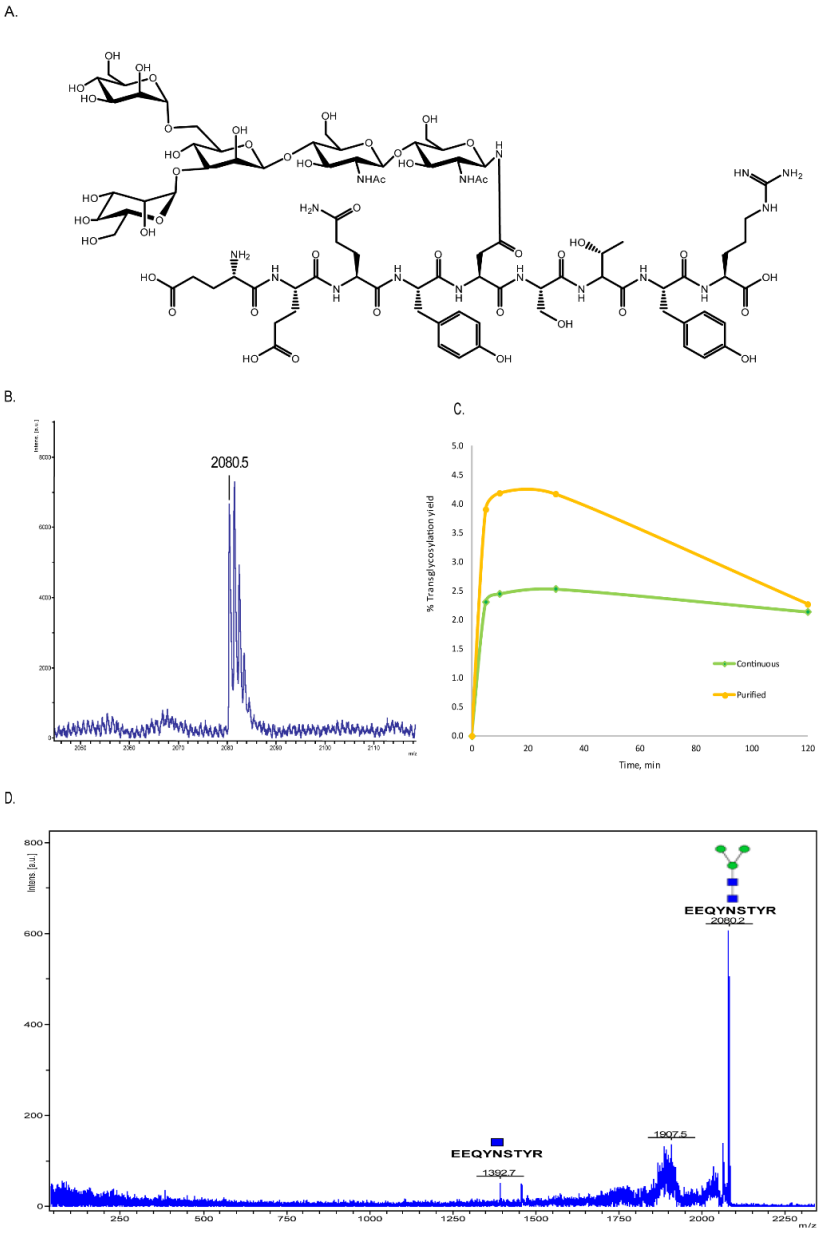
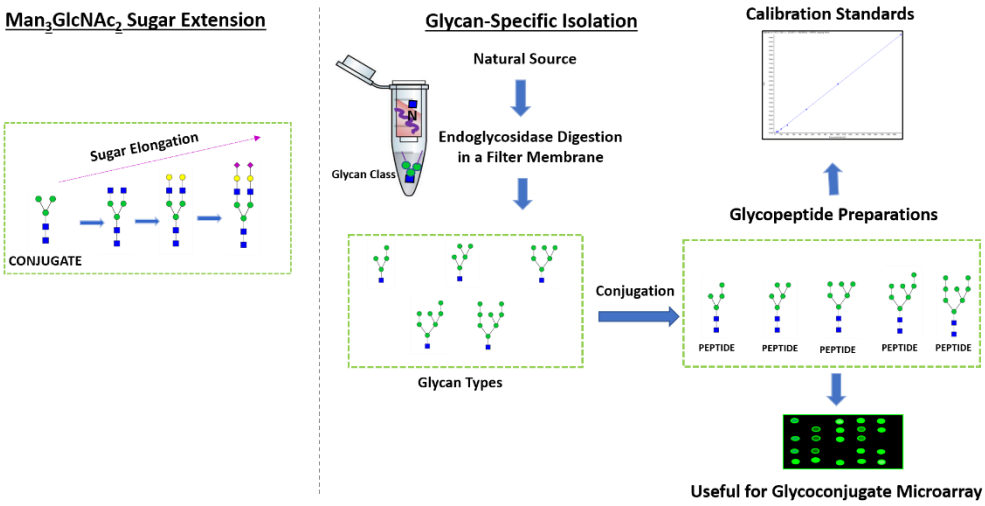


Figure 3-6. Transglycosylated IgG1-$Man_3GlcNAc_2$ glycopeptide product. A. Chemical structure of IgG1-$Man_3GlcNAc_2$ B. MALDI-TOF/MS Spectrum at 2080 *m/z*, $[M+H]^+$. C. Comparison between yields for continuous(undesalted) and purified $Man_3GlcNAc$ on the transglycosylation of $Man_3GlcNAc$-oxazoline to IgG1-GlcNAc acceptor. D. MALDI-TOF/TOF MS confirmation of transglycosylated product, IgG1-$Man_3GlcNAc_2$, which shows the IgG1-GlcNAc fragment.



3.5 Chapter Conclusion

Figure 3-7. Methodology application. Methodology can be used to acquire a starting glycan for sugar elongation of a $Man_3GlcNAc_2$-glycoconjugate (left) and can also be relevant for a glycan class-specific isolation (right) utilized for different assays.



Direct isolation using filter membranes separating the glycoproteins from released *N*-glycans and purification procedures by either HILIC or GCC, is very important for the recovery of *N*-glycan structures. Equally important, the use of glycoblotting and GCC for $Man_3GlcNAc$ isolation was presented in this work. The release of $Man_3GlcNAc$ from glycoproteins in Japanese quail egg white using Endo-D provides for a direct route in obtaining the trimannosyl truncated chitobiose core glycan. $Man_3GlcNAc$ can serve as a starting *N*-glycan template for *N*-glycopeptide synthesis or any glycoconjugate procedures, i.e., Fmoc-Asn-OH and IgG1 glycopeptide with a core *N*-glycan attached, wherein different glycosyltransferases can be used to extend the structure (Figure 3-7). The procedure stated herein is an improvement of the methodology of previous works (*see Supporting Information*) of obtaining the conserved *N*-glycan structure. The purpose of preparing different glycopeptides is discussed in Chapter 4. The results in this chapter detail a necessary procedure that is of great significance for the large-scale production of $Man_3GlcNAc$ obtained from a natural source, and synthesis of Fmoc-Asn($Man_3GlcNAc_2$)-OH and IgG1-$Man_3GlcNAc_2$.

3.6 References

1 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*(21), 4757-4774.

2 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*(1), 531-540.

3 Mutsaers, J. H., van Halbeek, H., Vliegenthart, J. F., Iwase, H., Kato, Y., & Hotta, K. (1985) Structural analysis of dansyl glyco-asparagines from quail ovalbumin. *Biochim. Biophy. Acta (BBA)-General Subjects*. *843*(1-2), 123-127.

4 Hase, S., Okawa, K., & Ikenaka, T. (1982) Identification of the trimannosyl-chitobiose structure in sugar moieties of Japanese quail ovomucoid. *J. Biochem*. *91*(2), 735-737.

5 Fujita, M., Shoda, S. I., Haneda, K., Inazu, T., Takegawa, K., & Yamamoto, K. (2001) A novel disaccharide substrate having 1, 2-oxazoline moiety for detection of transglycosylating activity of endoglycosidases. *Biochim. Biophys. Acta (BBA)-General Subjects*. *1528*(1), 9-14.

6 Fairbanks, A. J. (2017) The ENGases: versatile biocatalysts for the production of homogeneous N-linked glycopeptides and glycoproteins. *Chem. Soc. Rev*. *46*(16), 5128-5146.

7 Oda, Y., Nakayama, K., Abdul-Rahman, B., Kinoshita, M., Hashimoto, O., Kawasaki, N., Hayakawa, T., Kakehi, K., Tomiya, N. and Lee, Y.C (2000). Crocus sativus lectin recognizes Man3GlcNAc in the N-glycan core structure. *J. Biol. Chem. 275*(35), 26772-26779.

8 Robinson, D. S., & Monsey, J. B. (1971) Studies on the composition of egg-white ovomucin. *Biochem. J. 121*(3), 537-547.

9 Donovan, J. W., Davis, J. G., & White, L. M. (1970) Chemical and physical characterization of ovomucin, a sulfated glycoprotein complex from chicken eggs. *Biochim. Biophys. Acta (BBA)-Protein Structure*. *207*(1), 190-201.

10 Paredes, B., Gonzalez, S., Rendueles, M., & Diaz, M. (2006) Chromatographic separation at a preparative scale of egg white ovalbumin and its application in the elaboration of yogurt mousse. *J. Food Process Eng. 29*(1), 36-52.

11 Sun, B., Bao, W., Tian, X., Li, M., Liu, H., Dong, J., & Huang, W. (2014) A simplified procedure for gram-scale production of sialylglycopeptide (SGP) from egg yolks and subsequent semi-synthesis of Man3GlcNAc oxazoline. *Carbohydr. Res. 396*, 62-69.

12 Seko, A., Koketsu, M., Nishizono, M., Enoki, Y., Ibrahim, H.R., Juneja, L.R., Kim, M. and Yamamoto, T. (1997) Occurrence of a sialylglycopeptide and free sialylglycans in hen's egg yolk. *Biochim. Biophys. Acta (BBA)-General Subjects*. *1335*(1-2), 23-32.

13 Packer, N. H., Lawson, M. A., Jardine, D. R., & Redmond, J. W. (1998) A general approach to desalting oligosaccharides released from glycoproteins. *Glycoconj. J. 15*(8), 737-747.

14 Nishimura, S.-I., Niikura, K., Kurogochi, M., Matsushita, T., Fumoto, M., Hinou, H., Kamitani, R., Nakagawa, H., Deguchi, K., Miura,N., Monde, K., and Kondo, H. (2004) High-throughput protein glycomics: Combined use of chemoselective glycoblotting and MALDI-TOF/TOF mass spectrometry. *Angew. Chem., Int. Ed. 44*, 91–96.

15 Miura, Y., Kato, K., Takegawa, Y., Kurogochi, M., Furukawa, J.I., Shinohara, Y., Nagahori, N., Amano, M., Hinou, H. and Nishimura, S.I. (2010) Glycoblotting-assisted O-

glycomics: ammonium carbamate allows for highly efficient O-glycan release from glycoproteins. *Anal. Chem. 82*(24), 10021-10029.

16 Nishimura, S. I. (2011) Toward automated glycan analysis. In *Adv. Carbohydr. Chem. Biochem.* (Vol. 65, pp. 219-271). Academic Press.

17 Miura, Y., Hato, M., Shinohara, Y., Kuramoto, H., Furukawa, J.I., Kurogochi, M., Shimaoka, H., Tada, M., Nakanishi, K., Ozaki, M. and Todo, S. (2008) BlotGlycoABC™, an integrated glycoblotting technique for rapid and large scale clinical glycomics. *Mol. Cell. Prot. 7*(2), 370-377.

18 Yang, S. J., & Zhang, H. (2012) Glycan analysis by reversible reaction to hydrazide beads and mass spectrometry. *Anal. Chem. 84*(5), 2232-2238.

19 Cooper, C. A., Gasteiger, E., & Packer, N. H. (2001) GlycoMod–a software tool for determining glycosylation compositions from mass spectrometric data. *Proteomics*. *1*(2), 340-349.

20 Cooper C.A., Gasteiger E., Packer N. *Predicting Glycan Composition from Experimental Mass Using GlycoMod* (in) Handbook of Proteomic Methods (Conn P.M. ed.), pp. 225-232, Humana Press, Totowa, NJ (2003).

21 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*(4), 1650-1659.

22 Domon, B., & Costello, C. E. (1988) A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconj. J. 5*(4), 397-409.

23 Ruttkies, C., Schymanski, E. L., Wolf, S., Hollender, J., & Neumann, S. (2016). MetFrag relaunched: incorporating strategies beyond in silico fragmentation. *J. Cheminformatics*. *8*(1), 3.

24 Hammura, K., Ishikawa, A., Miyoshi, R., Yokoi, Y., Tanaka, M., Hinou, H., & Nishimura, S. I. (2018) Synthetic glycopeptides allow for the quantitation of scarce non-fucosylated IgG Fc N-glycans of therapeutic antibody. *ACS Med. Chem. Lett.* 9 (9), 889–894.

25 Lattova, E., Perreault, H., and 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.

26 Sumiyoshi, W., Nakakita, S., Hasehira, K., Miyanishi, N., Kubo, Y., Kita, T., and Hirabayashi, J. (2010) Comprehensive analysis of N linked oligosaccharides from eggs of the family Phasianidae. *Biosci., Biotechnol., Biochem. 74*, 606–613.

27 Takahashi, N., Khoo, K. H., Suzuki, N., Johnson, J. R., and Lee, Y. C. (2001) N-Glycan structures from the major glycoproteins of pigeon egg white: Predominance of terminal Galα (1-4) Gal. *J. Biol. Chem. 276*, 23230–23239.

28 Suzuki, N., Su, T. H., Wu, S. W., Yamamoto, K., Khoo, K. H., and Lee, Y. C. (2009) Structural analysis of N-glycans from gull egg white glycoproteins and egg yolk IgG. *Glycobiology. 19*, 693–706.

29 Risley, J. M., and Van Etten, R. L. (1986) Structures of the carbohydrate moiety attached to one site in the first domain of turkey ovomucoid: Elucidation by 1H NMR spectroscopy. *Carbohydr. Res.147*, 21–29.

30 Nomoto, H., Endo, T., and Inoue, Y. (1982) Preparation and characterization of fragment glycoasparagines from ovalbumin glycopeptides: Reference compounds for structural and biochemical studies of the oligo-mannose and hybrid types of carbohydrate chains of glycoproteins. *Carbohydr. Res. 107*, 91–101.

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

32 Yamashita, K., Kamerling, J. P., and 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. *J. Biol. Chem. 258*, 3099–3106.

33 Noguchi, M., Tanaka, T., Gyakushi, H., Kobayashi, A., & Shoda, S. I. (2009) Efficient synthesis of sugar oxazolines from unprotected N-acetyl-2-amino sugars by using chloroformamidinium reagent in water. *J. Org. Chem. 74*(5), 2210-2212.

34 Noguchi, M., Fujieda, T., Huang, W. C., Ishihara, M., Kobayashi, A., & Shoda, S. I. (2012) A Practical One-Step Synthesis of 1, 2-Oxazoline Derivatives from Unprotected Sugars and Its Application to Chemoenzymatic β-N-Acetylglucosaminidation of Disialo-oligosaccharide. *Helv. Chim. Acta*. *95*(10), 1928-1936.

35 Huang, W., Li, J., & Wang, L. X. (2011) Unusual Transglycosylation Activity of Flavobacterium meningosepticum Endoglycosidases Enables Convergent Chemoenzymatic Synthesis of Core Fucosylated Complex N-Glycopeptides. *ChemBioChem*. *12*(6), 932-941.

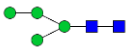
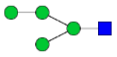
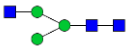
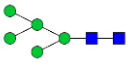
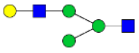
36 Yin, J., Li, L., Shaw, N., Li, Y., Song, J.K., Zhang, W., Xia, C., Zhang, R., Joachimiak, A., Zhang, H.C. and Wang, L.X. (2009) Structural basis and catalytic mechanism for the dual functional endo-β-N-acetylglucosaminidase A. *PLoS One*. *4*(3), e4658.

37 Li, B., Zeng, Y., Hauser, S., Song, H., & Wang, L. X. (2005). Highly efficient endoglycosidase-catalyzed synthesis of glycopeptides using oligosaccharide oxazolines as donor substrates. *J. Am. Chem. Soc*. *127*(27), 9692-9693.

3.7 Supplementary Information

Supplementary Tables

Table S3-1. Comparison of *N*-glycan structures released by PNGase-F and Endo-D

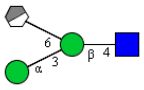
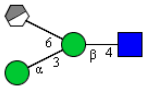
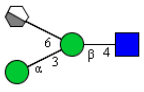
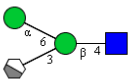
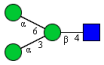
Peak No.	Endo-D *m/z* values	PNGase-F *m/z* values	Non-truncated Composition	*N*-Glycan Structures		References No.
				PNGase-F-released	Endo-D-released	
1	835	1038	$(Man)_3(GlcNAc)_2$			1-4,25-29
2	997	1200	$(Man)_4(GlcNAc)_2$			1,2,30, 31
3	1038	1241	$(HexNAc)_1$ + $(Man)_3(GlcNAc)_2$			2,25,26,31,32
4	1159	1362	$(Man)_5(GlcNAc)_2$			2, 25,26,31
5	1200	1403	$(Hex)_1$ $(HexNAc)_1$ + $(Man)_3(GlcNAc)_2$			2,25,26,31,32

All *m/z* values include the BOA label which has a mass of ~105.

Other *N*-glycans are released by PNGase-F. The table only shows comparison of truncated *N*-glycans released by Endo-D that can also be found as non-truncated equivalents released by PNGase-F.

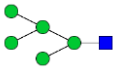
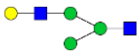
Glycan structures were prepared using GlycoWorkBench [21]

Table S3-2. MALDI-TOF/TOF Fragmentation Patterns from Labeled Parent Ion, $(Man)_3(GlcNAc)_1$ at 835 *m/z* from JQEW

m/z	Theoretical Fragmentation, *m/z*	Ion Type	Putative Structure(s)	
200.95				
240.95				
315.01				
331.04				
347.02	347.0949	BY		
388.05				
493.11				
509.05	509.1477	B		
522.09				
525.09				
629.18				
676.81				
678.84				
681.93				
684.22				
727.25				
744.27	745.2059	$^{0,3}X_{Man}$	 w/BOA	 w/BOA
		$^{1,4}X_{Man}$	 w/BOA	 w/BOA
835.34	835.2376	Parent Ion	 w/BOA	

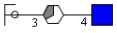
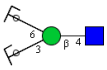
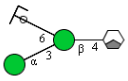
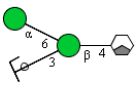
w/BOA-with O-benzyloxyamine hydrochloride, $[M+Na]^+$. Glycan structures were prepared using GlycoWorkBench [21]Ion types are based on the fragmentation patterns by Domon & Costello, 1988 [22]

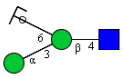
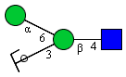
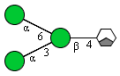
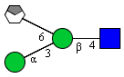
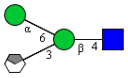
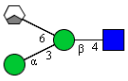
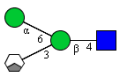
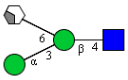
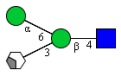
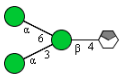
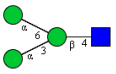
Table S3-3. Relative Intensities of Unlabeled Truncated *N*-Glycans Recovered Using Different Strategies

Peak No.	Glycan *m/z* values	Composition	*N*-Glycan Putative Structures	Relative Intensity per 0.5 mg sample		
				1st Strategy		2nd Strategy
				HILIC	GCC	GB then GCC
1	730	$(Man)_3(GlcNAc)_1$		1	1	1
2	892	$(Hex)_1$ +$(Man)_3(GlcNAc)_1$		0.12±0.08	0.05±0.01	ND
3	933	$(HexNAc)_1$ + $(Man)_3(GlcNAc)_1$		0.07±0.03	0.03±0.01	ND
4	1054	$(Hex)_2$ + $(Man)_3(GlcNAc)_1$		0.05±0.03	0.03±0.01	ND
5	1095	$(Hex)_1$ $(HexNAc)_1$ + $(Man)_3(GlcNAc)_2$		0.08±0.03	0.07±0.01	ND

HILIC- Hydrophilic Interaction Chromatography, GCC- Graphite Column Chromatography, GB- Glycoblotting
Relative Intensities were only a comparison among the 5 glycan structures with the highest intensity glycan (at 730) considered as 1, n=3. Non-glycan peaks were not considered.
ND- Not detected

Table S3-4. MALDI-TOF/TOF Fragmentation Patterns from Unlabeled Parent Ion, $(Man)_3(GlcNAc)_1$ at 730 *m/z* from JQEW (HILIC Purified) Strategy 1 Digestion

m/z	Theoretical Fragmentation, *m/z*	Ion Type	Putative Structure(s)	
60.04				
160.01				
178.04				
209.98				
226.07				
244.09	244.0792	Y		
274.05				
292.03				
332.07	332.0952	$^{3,5}X_{Man}Y$		
347.14	347.0949	BY		
365.13				
390.05				
406.17	406.1320	YY		
457.92				
467.16	467.1371	$^{1,3}X_{GlcNAc}Y$		
509.18	509.1477	B		
527.20				
548.92				
552.96				

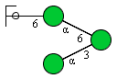
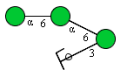
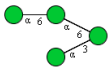
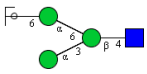
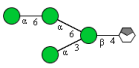
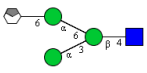
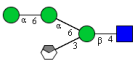
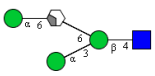
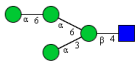
568.25	568.1848	Y		
572.84				
574.94				
588.77				
598.95				
629.22	629.1900	$^{1,3}X_{GlcNAc}$		
670.24	670.2165	$^{0,4}X_{Man}$		
		$^{1,3}X_{Man}$		
		$^{2,4}X_{Man}$		
		$^{0,4}X_{GlcNAc}$		
684.99				
712.62				
730.22	730.2376	Parent Ion		

Mass values, $[M+Na]^+$

Glycan structures were prepared using GlycoWorkBench [21]

Ion types are based on the fragmentation patterns by Domon & Costello, 1988 [22]

Table S3-5. MALDI-TOF/TOF Fragmentation Patterns from Unlabeled Parent Ion, $(Hex)_1$ + $(Man)_3(GlcNAc)_1$ at 892 *m/z* from JQEW (HILIC Purified) Strategy 1 Digestion

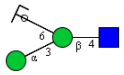
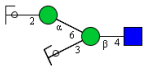
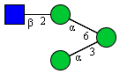
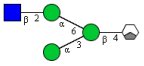
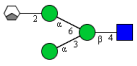
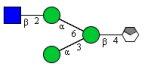
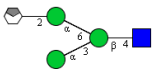
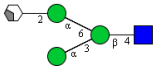
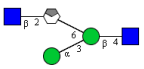
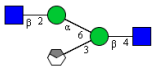
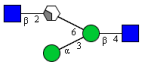
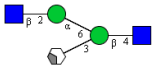
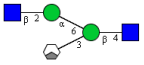
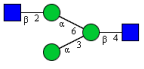
m/z	Theoretical Fragmentation, *m/z*	Ion Type		Putative Structure(s)
509.21	509.1477	BY		
550.16 569.24 598.22				
671.26	671.2005	B		
728.64 730.46	730.2376	Y		
734.88 736.95 739.04				
791.33	791.2428	$^{1,3}X_{GlcNAc}$		
832.34	832.2693	$^{0,4}X_{GlcNAc}$		
		$^{0,4}X_{Man}$		
		$^{1,3}X_{Man}$		
		$^{2,4}X_{Man}$		
892.27	892.2905	Parent Ion		

Mass values, $[M+Na]^+$

Glycan structures were prepared using GlycoWorkBench [21]

Ion types are based on the fragmentation patterns by Domon & Costello, 1988 [22]

Table S3-6. MALDI-TOF/TOF Fragmentation Patterns from Unlabeled Parent Ion, $(HexNAc)_1$ + $(Man)_3(GlcNAc)_1$ at 933 *m/z* from JQEW (HILIC Purified) Strategy 1 Digestion

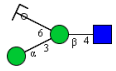
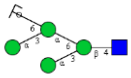
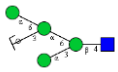
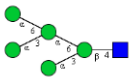
m/z	Theoretical Fragmentation, *m/z*	Ion Type	Putative Structure(s)	
509.12				
568.23	568.1848	Y		
		YY		
730.31	730.2376	Y		
774.75 776.87 781.07				
832.37	832.2693	$^{1,3}X_{GlcNAc}$		
873.36	873.2959	$^{0,4}X_{GlcNAc}$		
		$^{2,4}X_{GlcNAc}$		
		$^{0,4}X_{Man}$		
		$^{2,4}X_{Man}$		
		$^{1,3}X_{Man}$		
933.30	933.3170	Parent Ion		

Mass values, $[M+Na]^+$

Glycan structures were prepared using GlycoWorkBench [21]

Ion types are based on the fragmentation patterns by Domon & Costello, 1988 [22]

Table S3-7. MALDI-TOF/TOF Fragmentation Patterns from Unlabeled Parent Ion, $(Hex)_2$ + $(Man)_3(GlcNAc)_1$ at 1054 *m/z* from JQEW (HILIC Purified) Strategy 1 Digestion

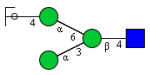
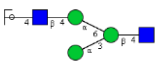
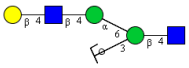
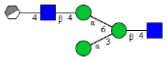
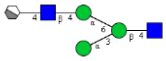
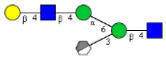
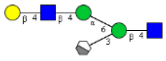
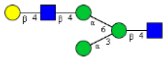
m/z	Theoretical Fragmentation , *m/z*	Ion Type	Putative Structure(s)
569.29	568.1848	Y	
890.57			
892.60	892.2905	Y	  
894.77			
896.90			
900.02			
1030.90			
1055.42	1054.3433	Parent Ion	

Mass values, $[M+Na]^+$

Glycan structures were prepared using GlycoWorkBench [21]

Ion types are based on the fragmentation patterns by Domon & Costello, 1988 [22]

Table S3-8. MALDI-TOF/TOF Fragmentation Patterns from Unlabeled Parent Ion, $(Hex)_1(HexNAc)_1$ + $(Man)_3(GlcNAc)_2$ at 1095 *m/z* from JQEW (HILIC Purified) Strategy 1 Digestion

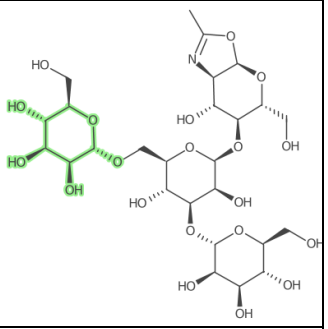
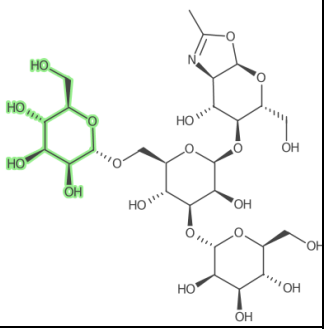
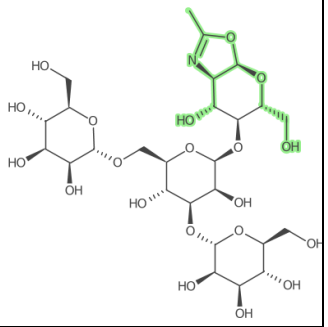
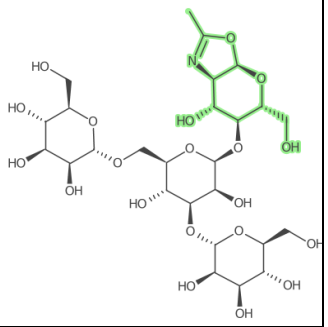
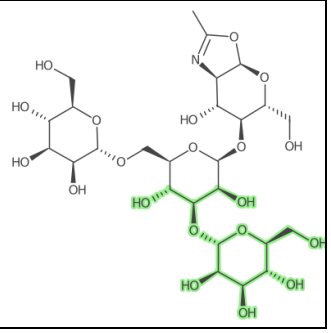
m/z	Theoretical Fragmentation, *m/z*	Ion Type	Putative Structure(s)
730.36	730.2376	Y	
926.35 928.42			
932.60	933.3170	Y	 
934.67 936.78 938.86 940.93 943.04 948.48 950.53 952.70			
1004.44	1005.3381	$^{0,3}X_{Gal}$	
		$^{1,4}X_{Gal}$	
		$^{0,3}X_{Man}$	
		$^{1,4}X_{Man}$	
1095.43		Parent Ion	

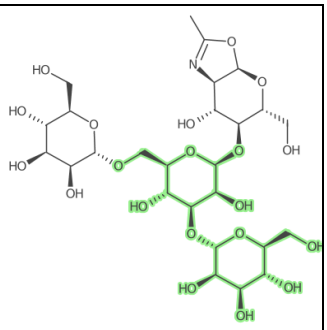
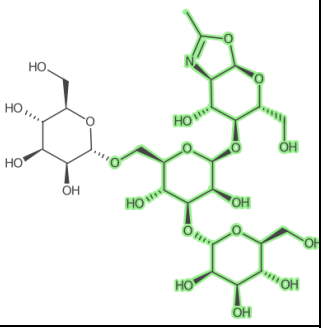
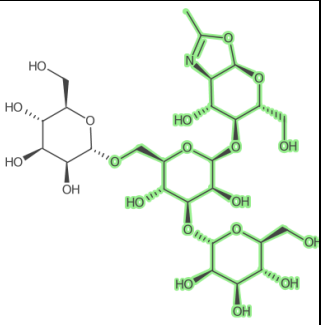
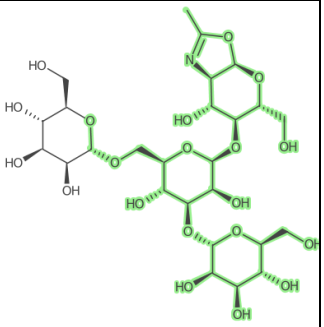
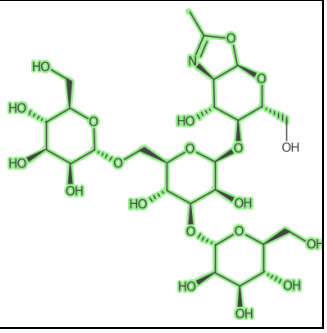
Mass values, $[M+Na]^+$

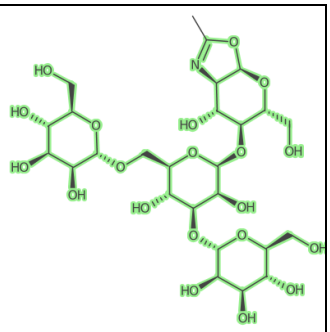
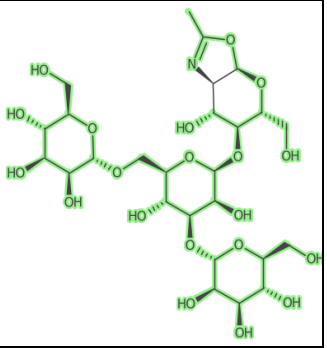
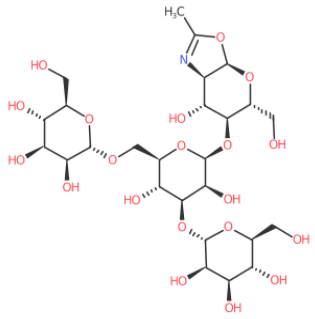
Glycan structures were prepared using GlycoWorkBench [21]

Ion types are based on the fragmentation patterns by Domon & Costello, 1988 [22]

Table S3-9. MALDI-TOF/TOF Fragmentation Patterns from $Man_3GlcNAc$-oxazoline at 712 *m/z*

m/z	Theoretical Fragmentation, *m/z*	Formula	Putative Structure
147.0683	148.03664	$[C_5H_8O_5]^+$	
162.0867	163.06013	$[C_6H_{11}O_5]^+$	
174.0768	174.07612	$[C_7H_{11}NO_4+H]^+$	
188.0723	187.08395	$[C_8H_{12}NO_4+H]^+$	
252.1312	252.08401	$[C_9H_{16}O_8]^+$	

363.0209	363.08983	$[C_{12}H_{20}O_{11}]+Na^+$	
548.6281	549.1665	$[C_{20}H_{32}NO_{15}]+Na^+$	
550.3306	550.17433	$[C_{20}H_{32}NO_{15}+H]+Na^+$	
554.8057	555.17947	$[C_{21}H_{33}NO_{16}]^+$	
694.4253	695.22444	$[C_{26}H_{42}NO_{19}]+Na^+$	

697.7062	697.20369	$[C_{25}H_{40}NO_{20}]+Na^+$	
700.0170	699.21935	$[C_{25}H_{42}NO_{20}]+Na^+$	
712.4406	712.2379	$[C_{26}H_{43}NO_{20}+Na]^+$	

MetFrag was used for *in silico* fragmentation, comparison of peaks to MALDI-TOF/TOF, and annotation [23]
For the parent ion at 712 m/z: red highlighted marks represent O or OH and blue represents N
For daughter ions: green highlighted structures pertain to the fragment at that *m/z* value
$Man_3GlcNAc$-oxazoline, ChemSpider Identifier: 57620467, InChIKeyBlock1 = CFJYTGJZCRGCLC
Name :(3aR,5R,6S,7R,7aR)-7-Hydroxy-5-(hydroxymethyl)-2-methyl-5,6,7,7a-tetrahydro-3aH-pyrano[3,2-d] [1,3]oxazol-6-yl alpha-D-mannopyranosyl-(1->3)-[alpha-D-mannopyranosyl-(1->6)]-beta-D-mannopyranoside

Supplementary Figures.

Figure S3-1. Spectra comparison of labeled released glycans A.) Endo-D digestion in an Eppendorf Tube and B.) Endo-D digestion on a filter membrane. Comparison of mass spectra for labeled-*N*-glycans under different reaction tubes and analyzed by glycoblotting methodology. A. Strategy 1 Endo-D digestion in a filter tube (YM-10, 10kDa cut-off). B. Strategy 2 Trypsin and Endo-D digestion in an Eppendorf tube

i. $(Hex)_1 + (Man)_3(GlcNAc)_1$ at 997 *m/z*.
ii. $(HexNAc)_1 + (Man)_3(GlcNAc)_1$ at 1038 *m/z*
iii. $(Hex)_2 + (Man)_3(GlcNAc)_1$ at 1159 *m/z*
iv. $(Hex)_1(HexNAc)_1 + (Man)_3(GlcNAc)_1$ at 1200 *m/z*

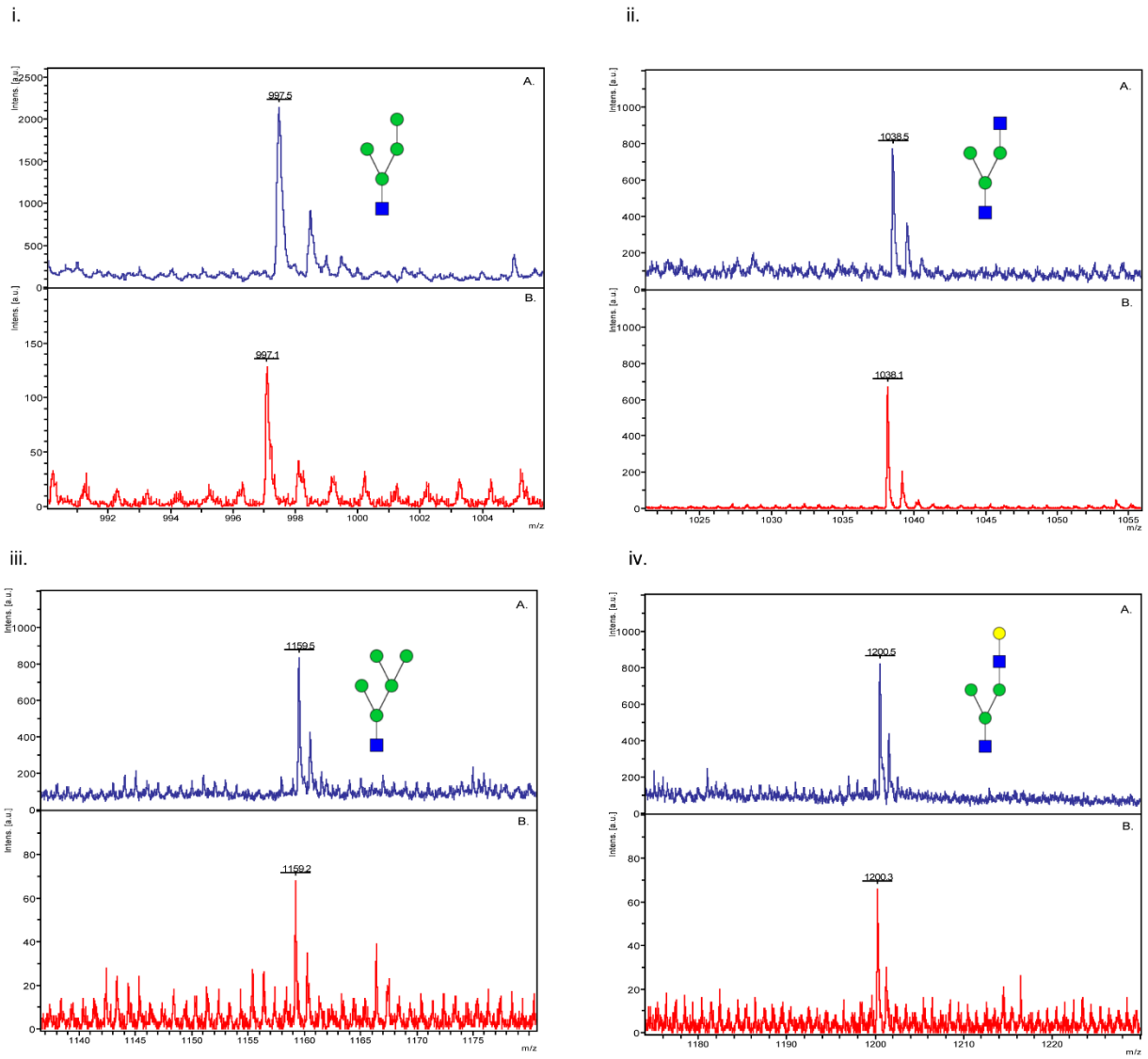


Figure S3-2. Comparison of HILIC and GCC Purified JQEW using Strategy 2 digestion (Trypsin and Endo-D) in Eppendorf tubes followed by either. A. HILIC only or B. GCC only.

i. Whole spectra of $Man_3GlcNAc$ found at 731 *m/z* instead of the 730 *m/z* which may be due to difference in ionization in the presence of salt and peptide contaminants. The spectra shows convoluted peaks that may be accounted for co-eluted tryptic glycopeptides and low molecular weight peptides.
ii. Expanded region near 731 m/z. $Man_3GlcNAc$ was present but co-elution of other tryptic glycopeptides and peptides were also present.

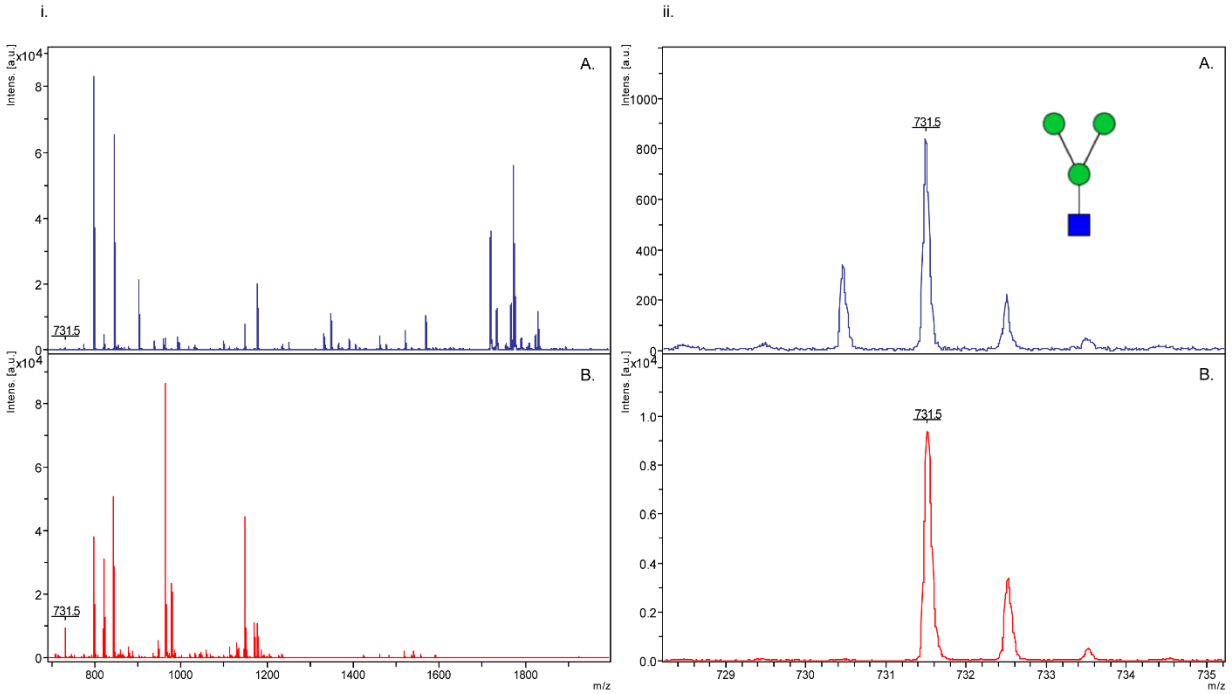


Figure S3-3. Purification of Fmoc-Asn(GlcNAc)-OH. A. Preparative HPLC B. Analytical HPLC C. MALDI-TOF/MS. Complete conversion to the deacetylated structure was accomplished.

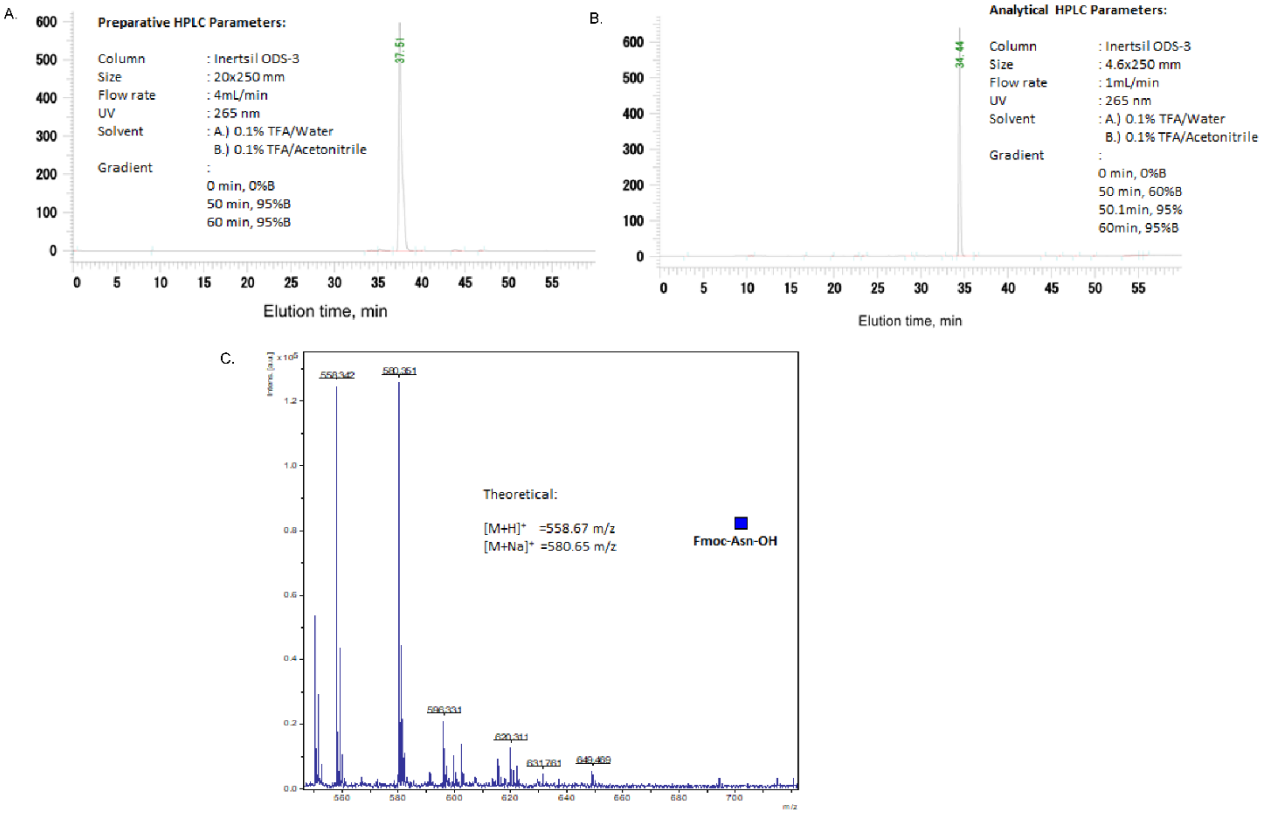


Figure S3-4. Purified IgG1-GlcNAc as glycosyl acceptor for $Man_3GlcNAc$-oxazoline. A. HPLC spectrum B. MALDI-TOF/MS spectrum of purified fraction found at 1393 *m/z*. The purified IgG1-GlcNAc glycopeptide was checked in analytical RP-HPLC using water and acetonitrile as solvents, both containing 0.1% TFA with a linear gradient, 0 min (100% water/0% acetonitrile) → 60min (80% water/20% acetonitrile) in an Inertsil-ODS 3 column (4.5x250 mm) at 220 nm with 1mL/min flow rate.

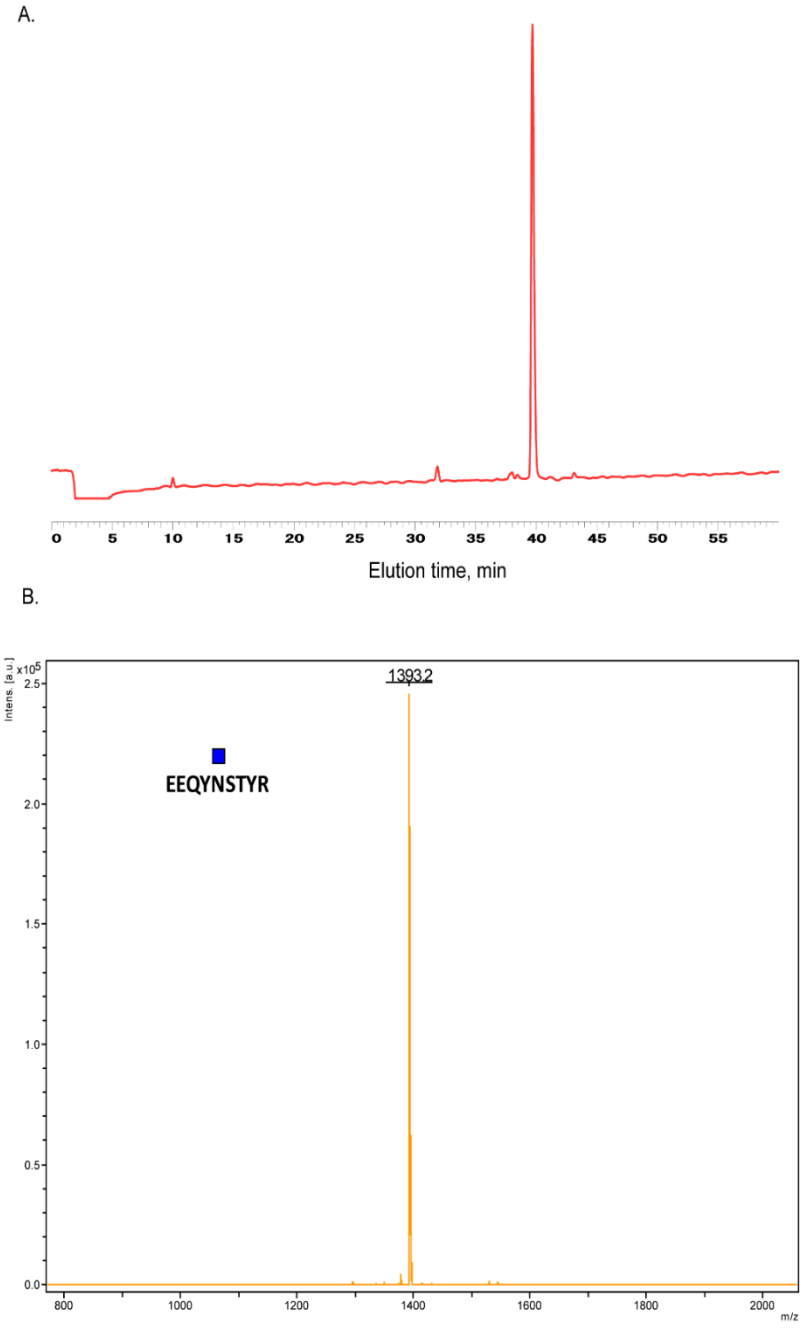


Figure S3-5. HPLC spectra of a 5-min transglycosylation reaction comparing with or without purification before oxazoline conversion and directly used as donor for IgG1-GlcNAc transglycosylation reaction. A. Unpurified $Man_3GlcNAc$ B. Purified $Man_3GlcNAc$ before oxazoline conversion. Conditions for analytical RP-HPLC involved using water and acetonitrile as solvents, both containing 0.1% TFA with a linear gradient, 0 min (100% water/0% acetonitrile) → 60min (80% water/20% acetonitrile) in an Inertsil-ODS 3 column (4.5x250 mm) at 220 nm with 1mL/min flow rate. IgG1-$Man_3GlcNAc_2$ transglycosylation yield are 3.9% and 2.3% for purified and unpurified $Man_3GlcNAc$, respectively.

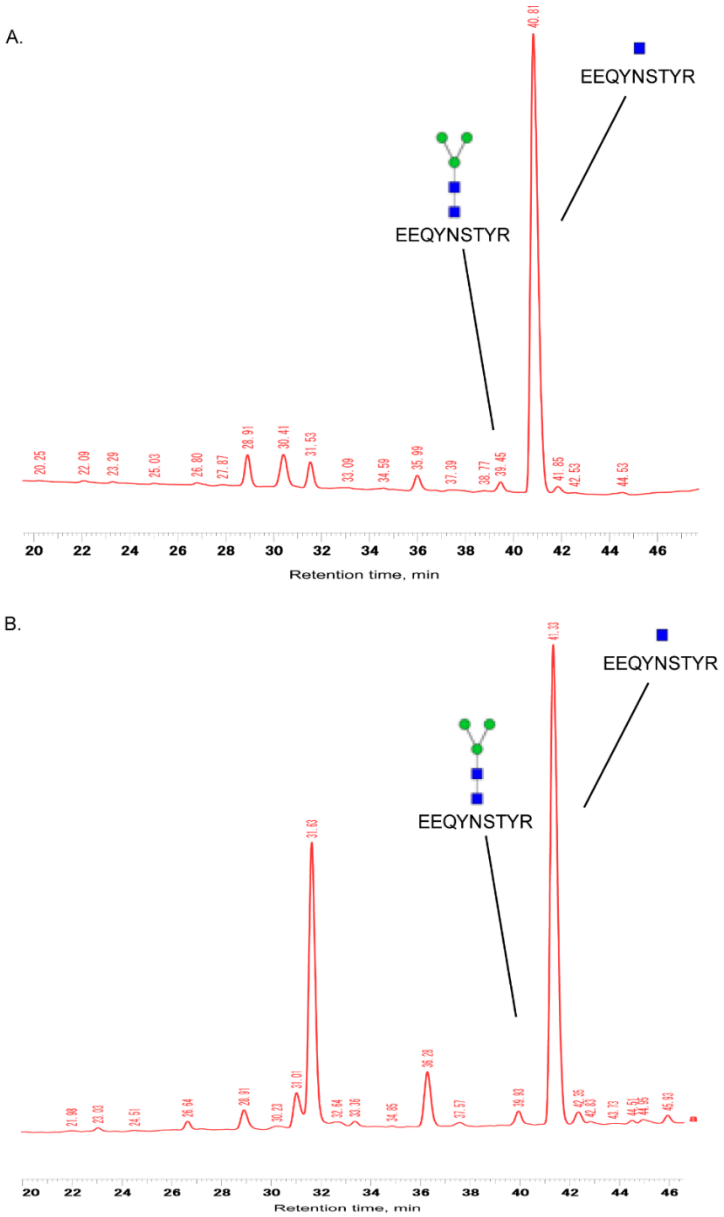
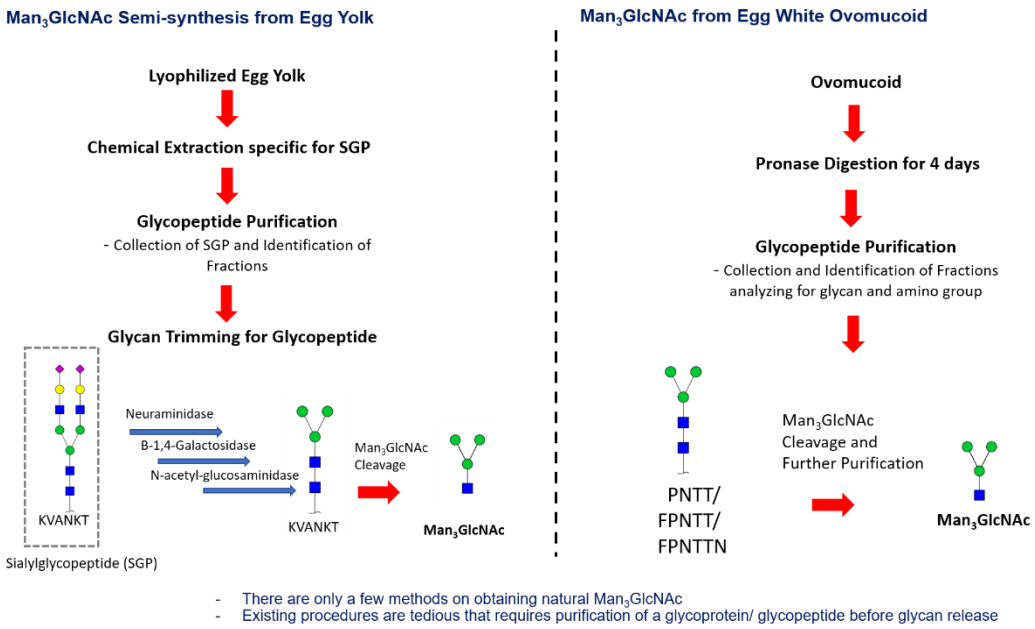


Figure S3-6. Previous methodologies on acquiring the conserved $Man_3GlcNAc_2$ glycan.



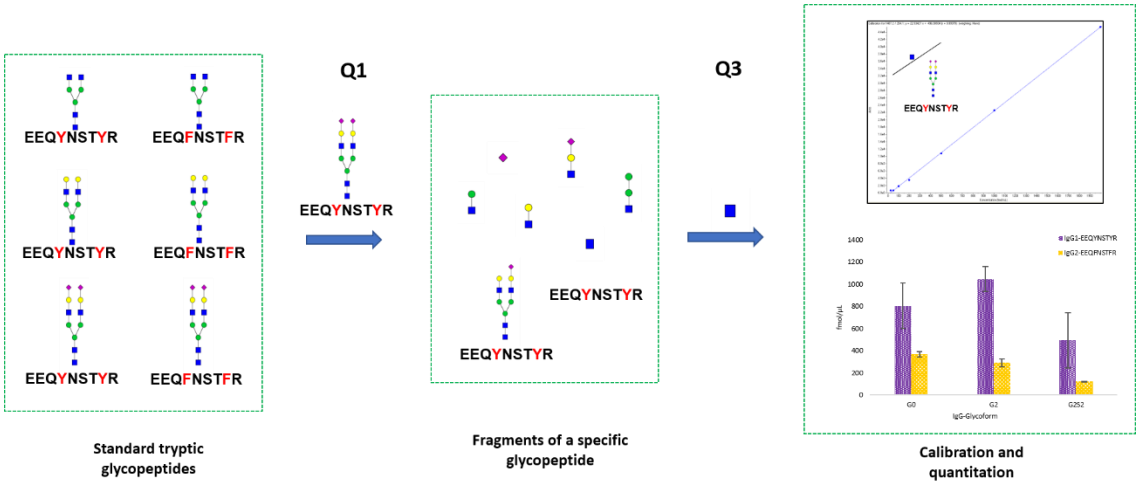
Reference:

Oda, Y., Nakayama, K., Abdul-Rahman, B., Kinoshita, M., Hashimoto, O., Kawasaki, N., Hayakawa, T., Kakehi, K., Tomiya, N. and Lee, Y.C (2000). Crocus sativus lectin recognizes Man3GlcNAc in the N-glycan core structure. *J. Biol. Chem. 275*(35), 26772-26779.

Sun, B., Bao, W., Tian, X., Li, M., Liu, H., Dong, J., & Huang, W. (2014) A simplified procedure for gram-scale production of sialylglycopeptide (SGP) from egg yolks and subsequent semi-synthesis of Man3GlcNAc oxazoline. *Carbohydr. Res. 396*, 62-69.

Chapter 4

Chemoenzymatic synthesized glycopeptide standards allow for the glycoform-specific absolute quantitation of IgG in human serum analyzed using multiple reaction monitoring



A version of this work will be submitted for journal publication

4.1 Chapter Summary

The synthesis of homogeneous glycoprotein or glycopeptide structures as calibration standards is necessary for possible biomarker quantitations. Only a few studies involving IgG glycopeptide quantitation was conducted and mostly performed on purified IgG standards isolated from human serum. The IgG standards from serum would have different component glycoforms per production batch and lot-to-lot variation would exist. As a result, only relative quantitations is achieved due to the heterogeneity of glycoform in such complex mixtures. For absolute quantitation of glycopeptides, a correctly defined glycan and peptide component is essential and can be synthesized by chemoenzymatic means. In this chapter, quantitation of IgG tryptic glycopeptides in human serum from synthesized calibration standard using a sensitive Multiple Reaction Monitoring which is an LC-MS/MS strategy was performed. The strategy directly measured fragments from the parent glycopeptides and were used to quantitate the presence of the said biomolecules in human serum. The importance of the methodology is for multiplexing analysis wherein multiple analytes are screened at a single sample injection. By having correctly defined glycan components attached to peptides, performing quality control checks for therapeutics and monitoring of essential glycosylated biomarkers from various samples is accomplished.

Keywords: *multiple reaction monitoring, glycopeptide, IgG glycopeptide, IgG quantitation, glycoform quantitation, serum, glycopeptide synthesis, absolute quantitation, multiplexing, biomarkers*

4.2 Introduction

Differences in serum *N*-glycan profiles between benign and diseased states have been shown in many profiling studies.[1-3] Particularly interesting is the aberrant *N*-glycosylation profile of serum immunoglobulins (IgG).[4-6] Assessment and quantitation of glycopeptide biomarkers are highly needed because there are only few glycopeptide standards available commercially. Associating glycans as coming from IgG are commonly scrutinized because purification of IgGs from serum most likely contain impurities from other glycoproteins.[7] The incomplete purification is obvious upon subjecting the IgG enriched fractions to sodium dodecyl sulfate (SDS)-gel with bands of other proteins still noticeable. As such, cut-outs of the gel bands are still needed before performing an in-gel tryptic digestion and glycan cleavage. Consequently, Fc and Fab regions both contain glycans.[8] Verification of the associated IgG glycans is necessary to pinpoint actual alterations in glycosylation. The affinity purifications of IgG using proteins A or G is never absolute and contaminating proteins still exist. Using a relevant amount of IgG is also needed for correct assessment.

Direct IgG quantitation from serum was conducted previously using commercial IgG standards.[9,10] Multiple reaction monitoring (MRM) was performed on human sera but absolute quantitation was based on Total IgG, not per glycoform. The glycoform concentration was expressed based on relative amounts since normalization was conducted on Total IgG.[9] Another work on glycopeptide quantification was conducted through a MALDI-TOF/TOF procedure. The quantification was based on the IgG Fc tryptic GlcNAc-containing glycopeptide residue since an endoglycosidase digestion was used first-hand to cleave the chiotobiose core separating the glycan.[10] The procedure does not measure per glycoform concentration but total glycoform content. Although the previous reports are relevant, the analysis at the individual glycopeptide level is the best approach for absolute quantitation.

Due to glycan-site heterogeneity, the need for clearly defined glycopeptide standards are important for absolute quantitation. The availability of glycopeptide standards is beneficial for assessing current therapeutic antibodies and confirmation of possible glycopeptide biomarkers. In this

work, two glycopeptide structures were chemoenymatically synthesized. Moreover, glycopeptide calibration standards were prepared for IgG1 and IgG2 glycoforms leading to the quantitation of these trace biomolecules in human serum without the need for IgG enrichment. The use of MRM made multiplexing analysis faster and the achievement of the goals for this Chapter possible.

4.3 Materials and Methods

4.3.1 Glycosyl Acceptor Synthesis

Two glycosyl acceptors, IgG1-GlcNAc and IgG4-GlcNAc, were prepared using solid-phase peptide synthesis. The synthesis of the IgG1-GlcNAc was similar to the procedure presented 3.3.5.2 of the previous chapter. The solid-phase peptide synthesis of IgG4-GlcNAc, H-Glu-Glu-Gln-Phe-Asn(GlcNAc)-Ser-Thr-Tyr-Arg-OH (IgG1-GlcNAc) is presented in this part. Both synthesized glycosyls, IgG1-GlcNAc and IgG4-GlcNAc, are the acceptor of the sialic acid-containing glycan from Sialylglycopeptide (Tokyo Chemical Industry Co. LTD, Japan) from chicken egg yolk for the transglycosylation reaction with Endoglycosidase-M (Tokyo Chemical Industry Co. LTD, Japan).

Synthesis of H-Glu-Glu-Gln-Phe-Asn(GlcNAc)-Ser-Thr-Tyr-Arg-OH (IgG4-GlcNAc). Briefly, trityl-ChemMatrix resin (0.3 mmol/g, 50 mg, 15 µmol), Fmoc-amino acids (60 µmol, 4.0 eq), and Fmoc-Asn(Ac_3β-GlcNAc)-OH (18µmol, 1.2 eq) glycoamino acid were used. The identity of the Fmoc-amino acids used for the synthesis are Fmoc-Arg(Pbf)-OH, Fmoc-Tyr(tBu), Fmoc-Thr(tBu)-OH, Fmoc-Ser(tBu)-OH, Fmoc-Phe-OH, Fmoc-Gln(Trt)-OH, and Fmoc-Glu(OtBu)-OH. For all microwave irradiation, the equipment used is Green Motif I microwave synthesis reactor (IDX Corp., Tochigi, Japan) set at 2450 MHz and 50ºC. A similar Fmoc-removal, coupling, and capping procedures as in 3.3.5.2 of the previous chapter was conducted. The crude glycopeptide was then purified by RP-HPLC using solvents A.) water B.) acetonitrile, both containing 0.1% TFA with a linear gradient, 0%B (0 min)→40%B (60min) in a preparative Inertsil-ODS 3 column (20x250 mm) at 220 nm with 4mL/min flowrate. The purified glycopeptide was further checked in analytical HPLC Inertsil-ODS 3 column (4.5x250 mm) at 220 nm with 1mL/min flowrate using the same linear gradient. The purified IgG4-GlcNAc was then used for transglycosylation reaction.

4.3.2 Transglycosylation Reaction

The transglycosylation using IgG1-GlcNAc as glycosyl acceptor was performed by using a mixture containing 5 µL 10 mM IgG1-GlcNAc tryptic peptide, 5 µL 50 mM SGP (Tokyo Chemical Industry Co. Ltd), 18 µL 60 mM potassium phosphate buffer (pH 6.25), and 2 µL of 1 mU/µL free Endo-M in an Eppendorf tube and the solution incubated at 30° C with constant shaking. Aliquots of 1 µL were taken at different time intervals and added to a 99 µL water. Aliquot samples were taken after 1, 4, 6, 8, and 24h. After taking aliquot at each time interval, the tube was place in 90ºC to heat-inactivate Endo-M. After cooling to room temperature, HPLC analysis was performed. Analytical HPLC monitoring of the transglycosylation product conducted under solvents, A. water B. acetonitrile, both containing 0.1% TFA. Solvent separation was conducted under a gradient of 0%B (0 min)→20%B (60min) in an Inertsil-ODS 3 column (4.5x250 mm) set at a 220 nm wavelength. The transglycosylation yield was estimated based on the HPLC peak areas of the transglycosylated product, EEQYN($Hex_5HexNAc_4NeuAc_2$) STYR, IgG1-G2S2, as compared to the total peak areas coming from both IgG1-G2S2 and IgG1-GlcNAc

The transglycosylation using IgG4-GlcNAc as glycosyl acceptor was performed by using a mixture containing 5 µL 10 mM IgG4-GlcNAc tryptic peptide, 5 µL 50 mM SGP (Tokyo Chemical Industry Co. Ltd), 18 µL 60 mM potassium phosphate buffer (pH 6.25), and 2 µL of 1 mU/µL free Endo-M in an Eppendorf tube and the solution incubated at 30° C with constant shaking. Aliquots of 1 µL were taken at different time intervals and added to a 99 µL water. Aliquot samples were taken after 1, 2, 4, 8, and 16h. Analytical HPLC monitoring of the transglycosylation was then conducted with solvents, A. water B. acetonitrile, both containing 0.1% TFA. For transglycosylated product monitoring, solvent separation was conducted under a gradient of 0%B (0 min)→40%B (60min) in an Inertsil-ODS 3 column (4.5x250 mm) set at a 220 nm wavelength. The transglycosylation yield was estimated based on the HPLC peak areas of the transglycosylated product, EEQFN($Hex_5HexNAc_4NeuAc_2$) STYR, IgG4-G2S2, as compared to the total peak areas coming from both IgG4-G2S2 and IgG4-GlcNAc

4.3.3 High-Performance Liquid Chromatography Checking of Tryptic IgG Glycopeptides

Synthesized tryptic IgG glycopeptides were purchased from Medicinal Chemistry Pharmaceuticals, Co., Ltd. and was checked for purity. The peptide sequence is EEQYNSTYR and EEQFNSTFR for IgG1 and IgG2, respectively. The IgG1 and IgG2 tryptic glycopeptides contain $Hex_5HexNAc_4NeuAc_2$, $Hex_5HexNAc_4$, and $Hex_3HexNAc_4$ which was packed separately. Analytical HPLC purity checking was conducted with solvents, A. water B. acetonitrile, both containing 0.1% TFA under a gradient of 0%B (0 min)→20%B (60min) in an Inertsil-ODS 3 column (4.5x250 mm) set at a 220 nm wavelength.

4.3.4 MALDI-TOF and MS/MS Analysis

The matrix consists of 20 mg/mL 2,5-dihydroxybenzoic acid dissolved in 30% acetonitrile containing 0.1% trifluoroacetic acid (TFA) solution. A 1 µL of the prepared matrix was spotted, dried, followed by spotting of the reconstituted sample into an Anchorchip MTP 384 Target Plate (polished steel TF, Bruker), and then dried. *FlexAnalysis* 3.0 software (Bruker Daltonics) was used to obtain spectra of experimental masses. MALDI-TOF and MALDI-TOF/TOF parameters using an Ultraflex III (Bruker Daltonics). Samples in MS were analyzed in reflector, positive ion mode, typically totaling 2000 shots with settings of acceleration voltage, reflector voltage, and pulsed ion extraction at 25.3 kV, 26.4 kV, and 100 ns, respectively. In MS/MS Mode, chosen parent peaks were initially accelerated to 8 kV and further accelerated to 20.1 kV. *N*-glycan structures were prepared by using GlycoWorkBench.

4.3.5 MRM Compound Optimization Parameter Settings

4.3.5.1 Mass spectrometry modes

Different mass spectrometry modes were conducted, namely, enhanced mass spectrometry (EMS), enhanced resolution (ER), and enhanced product ion (EPI) to assess the IgG tryptic glycopeptides response in MS and MS/MS collision. All 3 modes were measured using 10,000 fmol/µL of the glycopeptide at an injection flowrate of 5 µL/min. A 4000 QTRAP® triple quadrupole mass spectrometer (AB Sciex Pte. Ltd.) with UltiMate™ 3000 HPLC (Thermo Fisher Scientific Inc.) was used for MRM analysis.

4.3.5.1 Sample injection and HPLC separation

10 µL IgG glycopeptide calibration standards and tryptic digested human serum samples of the samples were injected using an WPS-3000 autosampler. Analytical HPLC separation was conducted with solvents, A. water B. acetonitrile, both containing 0.1% formic acid under a gradient of 0%B (0 min)→20%B (15min)→90%B (20 min)→2%B (20.1min)→2%B (30 min) with Dionex HPLC and AB Sciex 4000Q Trap® TurboIonSpray systems using an Inertsil-ODS 3 column 2.1x150 mm (GL Science) set at a 25ºC temperature,280 nm wavelength, and at a flowrate of 200 µL/min controlled using a Chromeleon™ 6.80 software. Data analysis were conducted using Analyst 1.5 and MultiQuant™ softwares.

4.3.6 Preparation of tryptic IgG Calibration Standards

Different calibration standard solution of IgG glycopeptides was prepared by serial dilution. The concentration of the glycopeptides is 0, 25, 50, 100, 200, 500, 1000, 2000, and 5000 fmol/µL. 10 µL of each standard was injected in the 4000 QTrap LC-MS/MS instrument for calibration.

4.3.7 Human Serum Tryptic Digestion

20 µL of human serum (human male AB plasma, Sigma, H4522) and Japanese human serum taken from pancreatic cancer and non-cancer patients were used for tryptic digestion. The human serum sample was added with 30 µL 0.33 M NH_4HCO_3 and 60 µL of 0.1% 1-propane sulfonic acid (PHM) in 10 mM NH_4HCO_3. The mixture was then incubated at 37 °C for 10 min. Then, 10 µL 120 mM dithiothreitol and incubated at 60° C. After 30 minutes, 20 µL 123 mM iodoacetamide was added and incubated at room temperature. After 1 h, 10 µL 40 U/ µL trypsin (Sigma Aldrich) in 1 mM HCl and incubate at 37° C for 24-h. The trypsin was heat-inactivated at 90°C for 10 min. The solution was then lyophilized.

4.3.8 Matrix effects: Calibration standard spiking

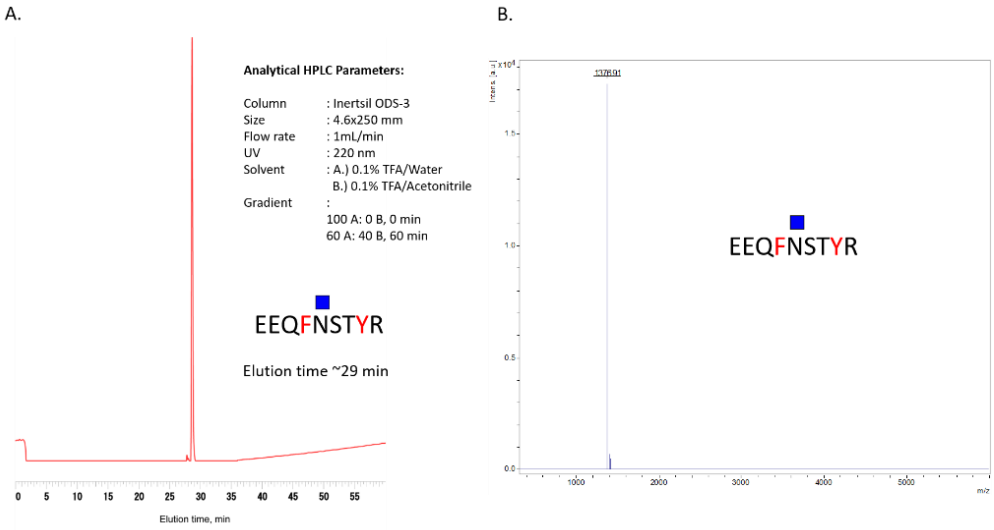
100 µL of different calibration standards (25, 100, 500, 1000, and 2000 fmol/µL) were added on lyophilized 20 µL of tryptic digested human serum to assess matrix effects. The same scheduled MRM parameters was used in the analysis of the spiked human serum samples and was analyzed in positive mode. The positive mode was for IgG1-G0, IgG1-G2, IgG1-G2S2, IgG2-G0, IgG2-G2, and IgG2-G2S2. A negative polarity mode was assessed with MRM parameters and analyzed for IgG1-G2S2 and IgG2-G2S2. Same 10 µL injection of sample was loaded for the analysis.

4.4 Results and Discussion

4.4.1 Transglycosylation reaction of IgG glycopeptide

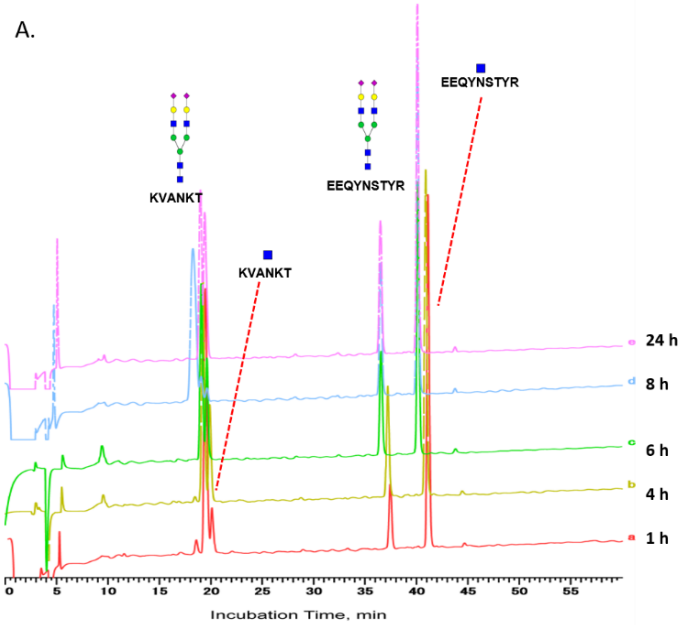
The IgG1-GlcNAc and IgG4-GlcNAc glycosyl acceptors were synthesized and 1.9 mg (9.5% yield), 4.7 mg (22% yield), respectively, were recovered. The purified IgG1-GlcNAc was shown in Figure S3-4 and the purified IgG4-GlcNAc is shown in Figure 4-1. Only a small-scale synthesis was conducted for the transglycosylation reaction and monitoring.

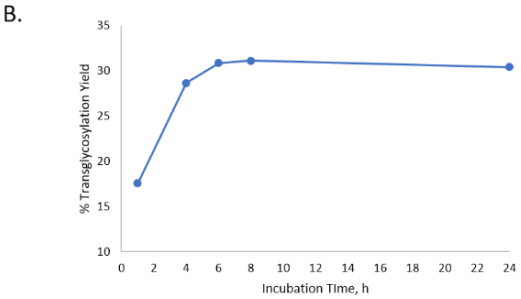
Figure 4-1. Purification of IgG4-GlcNAc. A. Analytical HPLC purification. B. MALDI-TOF MS spectra confirmation found at 1376 *m/z*. The eluted purified glycosyl acceptor is found ~29 min with the gradient used.



The transglycosylation reaction was accomplished by the aid of endoglycosidase-M wherein the sialic acid containing glycan from chicken egg yolk was cleaved and transferred to the glycosyl acceptor set at 30 ° C. Monitoring of transglycosylation reaction shown in Figure 4-2 was performed through analytical HPLC and the yield estimated based on the HPLC peak areas of the transglycosylated product, EEQYN($Hex_5HexNAc_4NeuAc_2$) STYR, IgG1-G2S2, as compared to the total peak areas coming from both IgG1-G2S2 and IgG1-GlcNAc. The same is true for IgG4-G2S2 shown in Figure 4-3.

Figure 4-2. Transglycosylation reaction monitoring of IgG1-G2S2 formation. A. HPLC monitoring B. % Transglycosylation curve C. % Transglycosylation yield showing the 6h gave the highest yield.

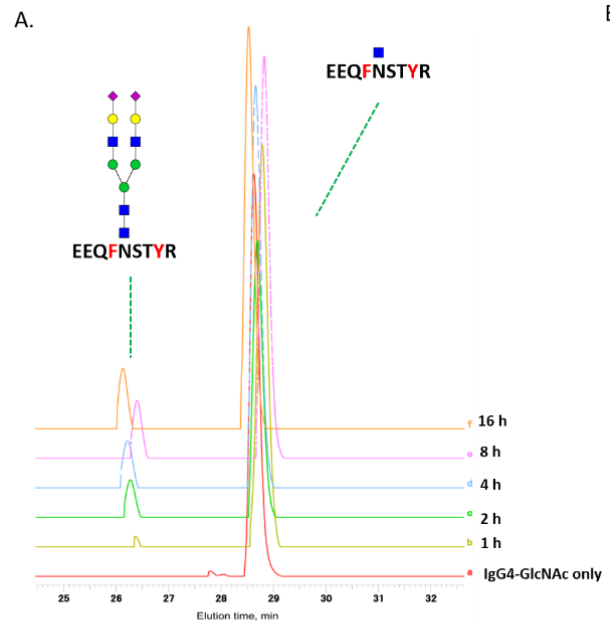


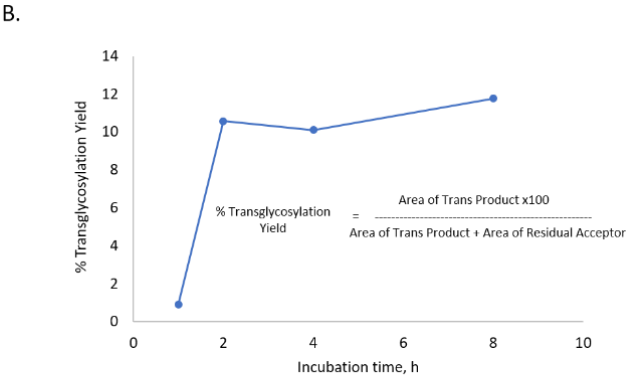


C.

Time (h)	% Transglycosylation Yield
1	17.5
4	28.6
6	30.9
8	31.1
24	30.4

Figure 4-3. Transglycosylation reaction monitoring of IgG4-G2S2 formation. A. HPLC monitoring B. % Transglycosylation curve C. % Transglycosylation yield showing the near amounts after 2h incubation.





C.

Time (h)	% Transglycosylation Yield
1	1
2	11
4	10
8	12
16	12

The eluted transglycosylated product fractions, IgG1-G2S2 and IgG4-G2S2, were checked in MALDI-TOF/MS (shown in Figures 4-4 and Figure 4-6), analyzed in both positive and negative mode, and further fragmented in MALDI-TOF/TOF for confirmation of the structure shown in Figures 4-5 and 4-7.

Figure 4-4. MALDI-TOF/MS Spectra of IgG1-G2S2 found at ~ 3396 *m/z*. The transglycosylation product was checked in both A. Reflector positive and B. Linear negative modes.

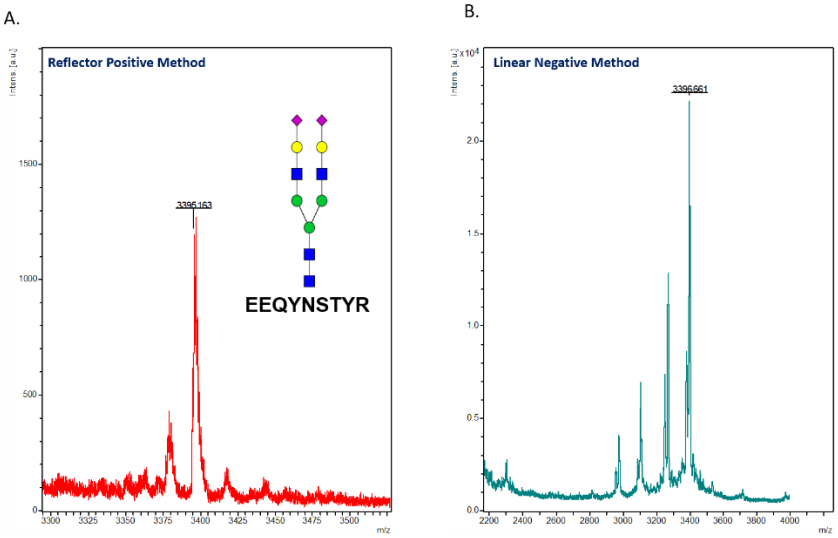
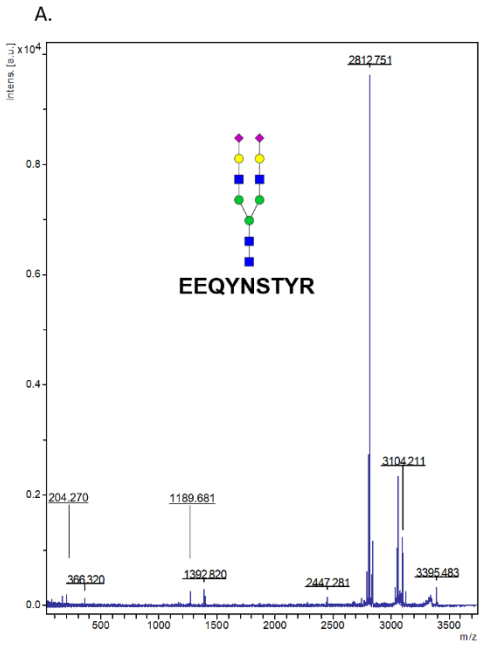


Figure 4-5. MALDI-TOF/TOF MS Spectra confirmation of IgG1-G2S2 structure. Fragments are annotated as can be seen in the plot. The annotation of the 204 and 1189 *m/z* was labeled manually since the figure needs to be expanded to show the label.



m/z	Structure
204.270	■
366.320	●■
1189.681	EEQYNSTYR (Pep)
1392.820	■ Pep
2447.281	●■■ Pep
2812.751	●■■ Pep
3104.211	◆●■■ Pep
3395.483	Parent Ion

Figure 4-6. MALDI-TOF/MS Spectra of IgG4-G2S2 found at ~3379 *m/z*. The transglycosylation product was checked in both A. Reflector positive and B. Linear negative modes.

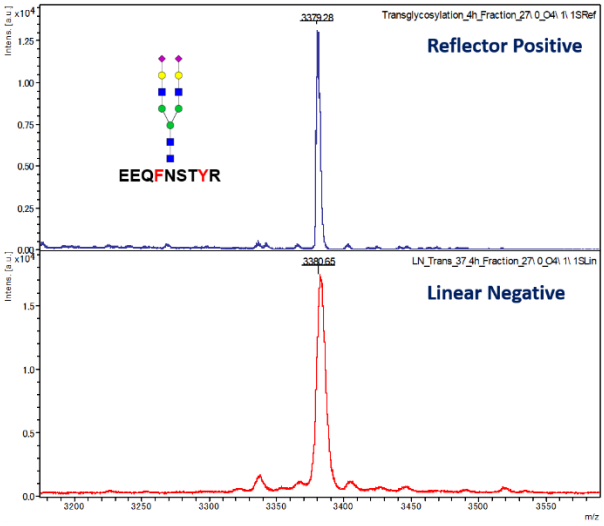
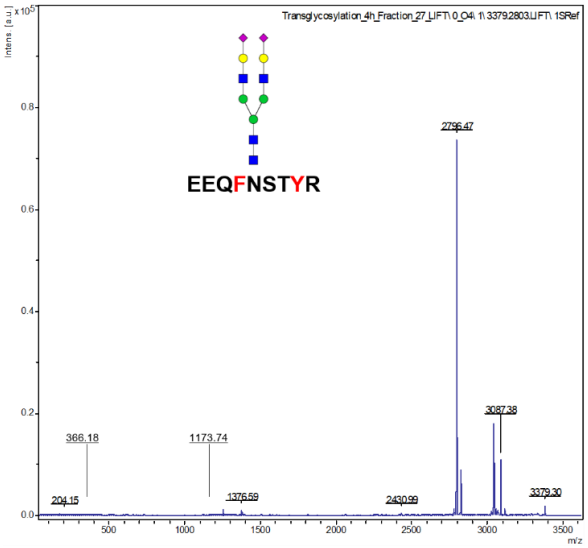


Figure 4-7. MALDI-TOF/TOF MS Spectra confirmation of IgG4-G2S2 structure. Fragments are annotated as can be seen in the plot. The annotation of the 366 and 1173 *m/z* was labeled manually since the figure needs to be expanded to show the label.

A.



B.

m/z	Structure
204.15	■
366.18	●—■ or ●—●
1173.74	EEQFNSTYR (Pep)
1376.59	■ Pep
2430.99	●—●(■)—■ Pep
2796.47	●—●(■)—■—■ Pep
3087.38	◆—●—●(■)—■—■ Pep
3379.30	Parent Ion

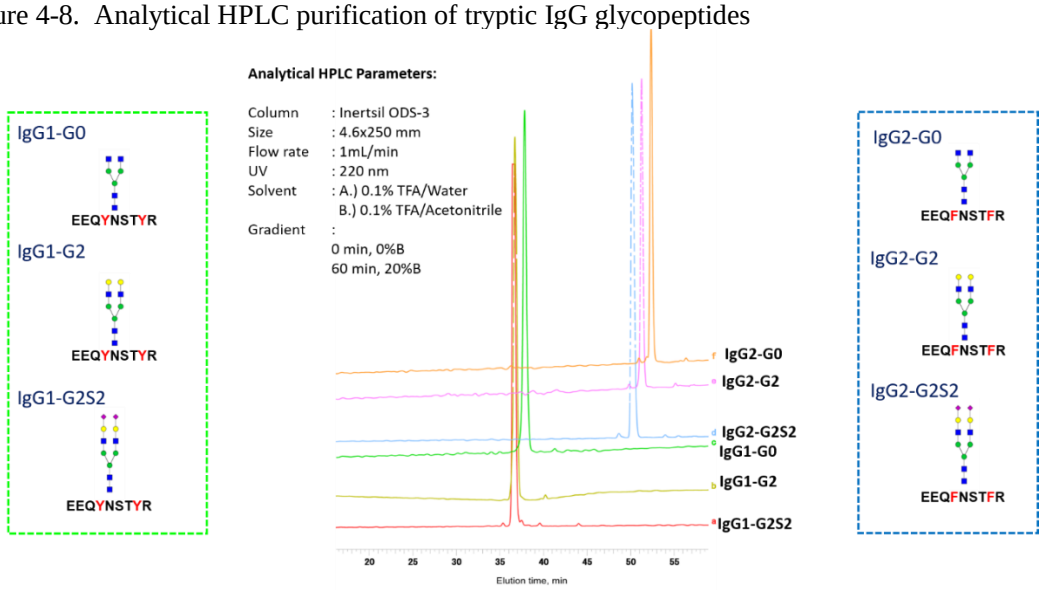
The transglycosylation reactions were only conducted on a small scale. The purpose was to show that natural glycopeptides, e.g., glycans from eggs can be used to synthesize glycoconjugate structures as shown in this work. One aspect of this dissertation is the use of natural glycans for chemoenzymatic preparations. The requirement to synthesize relevant amounts of glycopeptides is

the need of an abundant supply of endoglycosidases and starting materials to be able to do such reactions. This work provides the glycopeptide synthesis of IgG4 tryptic glycopeptide with the corresponding sialic acid glycoform component. The tryptic IgG4-G2S2 glycopeptide is relevant to serve as calibration standard for possible biomarker quantifications. The starting glycoform can be cleaved using exoglycosidases to create new glycopeptides for the same purpose. Addition of fucose on the synthesized glycopeptides can then be accomplished with glycosyltransferases, i.e., Fut8, to prepare fucosylated versions. For the next methods in this chapter, the glycopeptides used were that of the purchased IgG tryptic antibody glycopeptides from Medicinal Chemistry Pharmaceutical Co., Ltd.

HPLC Purification

The chemoenzymatically synthesized glycopeptides from Medicinal Chemistry Pharmaceutical Co., Ltd. were checked simultaneously in analytical HPLC and MALDI-TOF/MS. The HPLC spectra in Figure 4-8 showed high intense peaks of pure samples. Confirmation of the IgG tryptic glycopeptides structures were also conducted using MALDI-TOF/TOF analysis shown in Figures 4-9 and 4-10. Checking the IgG glycopeptides were conducted also in LC-MS/MS.

Figure 4-8. Analytical HPLC purification of tryptic IgG glycopeptides



The IgG2 glycopeptides eluted last while IgG1 glycopeptides eluted first due to polarity. The analytical HPLC is composed of a C18, nonpolar column. Thereby, phenylalanines present in IgG2 is

more non-polar which is attracted to the nonpolar column as compared to tyrosine present in IgG1 which is more polar. The glycoforms with the same peptide sequence elute closely because retention times rely mainly on the peptide moiety. [11] In the HPLC spectra, the glycan moiety for IgG2 glycoforms showed it has different retention times as compared to the IgG1 glycoforms wherein IgG1-G2S2 and IgG1-G2 are eluting close to each other. The analytical HPLC conducted here is for checking the purity of the purchased glycopeptide which is found to be pure.

Figure 4-9. MALDI-TOF MS spectra of tryptic IgG glycopeptides. Reflector positive mode was used for the neutral glycans and a linear negative mode for the sialic acid containing glycopeptide.

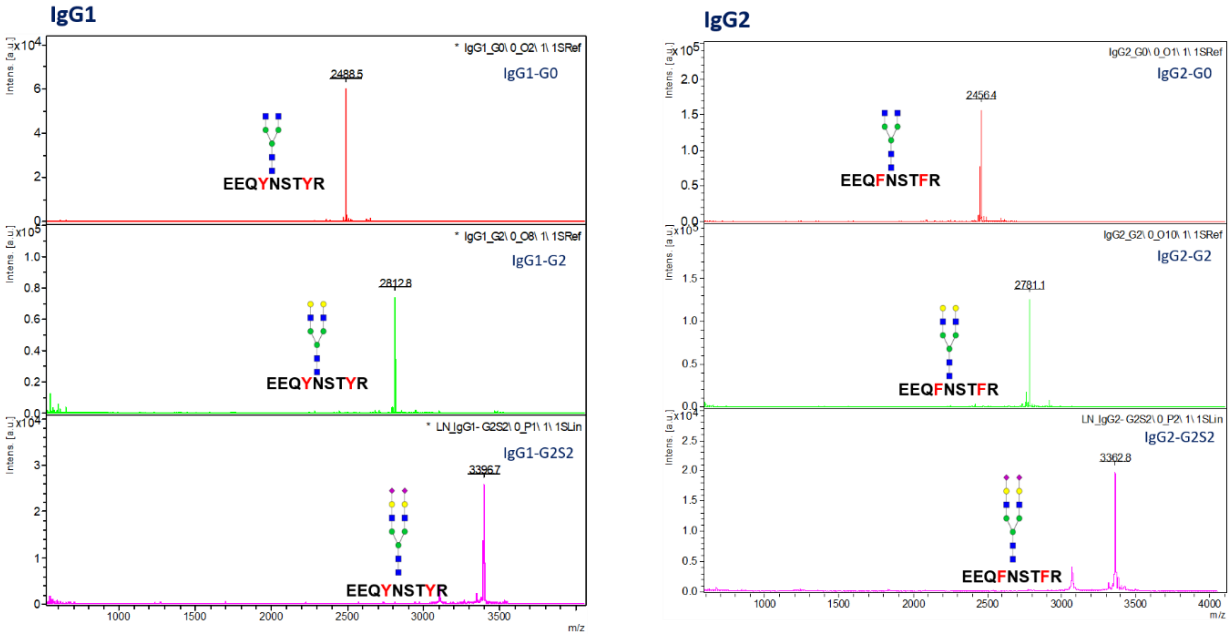
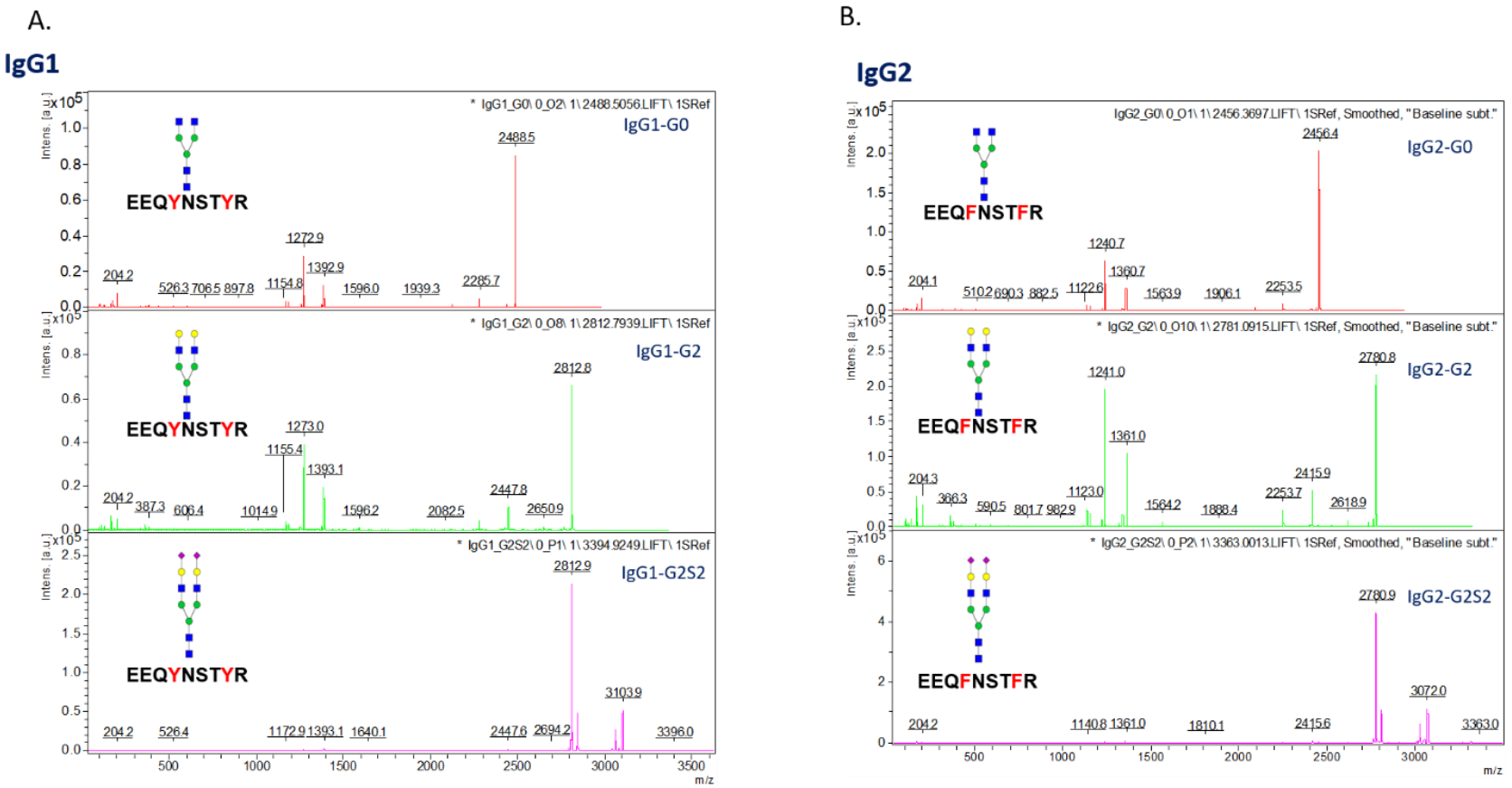


Figure 4-10. MALDI-TOF/TOF MS spectra of tryptic IgG glycopeptides.



Preliminary Screening of tryptic IgG glycopeptides in LC-MS/MS

Different mass spectrometry modes were conducted, namely, enhanced mass spectrometry (EMS), enhanced resolution (ER), and enhanced product ion (EPI) to assess the IgG tryptic glycopeptides response in MS and MS/MS collision. The EMS mode gave out either 2+ or 3+ adducts in positive mode and the ER mode was needed to highlight and expand the parent ion peaks. EPI on other hand have produced collision induced fragments relevant for assessing whether specific fragments can be used for calibration procedures. The corresponding spectra is shown in Figures 4-11 to 4-15. Parameters for the initial assessments are placed in the supplementary information.

Figure 4-11. Enhanced mass spectrometry mode mass spectra of IgG tryptic glycopeptides in positive mode. The different glycopeptides showed high intensity peaks at either $[M+2H]^{2+}$ or $[M+3H]^{3+}$ adducts for the parent ion peaks analyzed in 200-2800 m/z range.

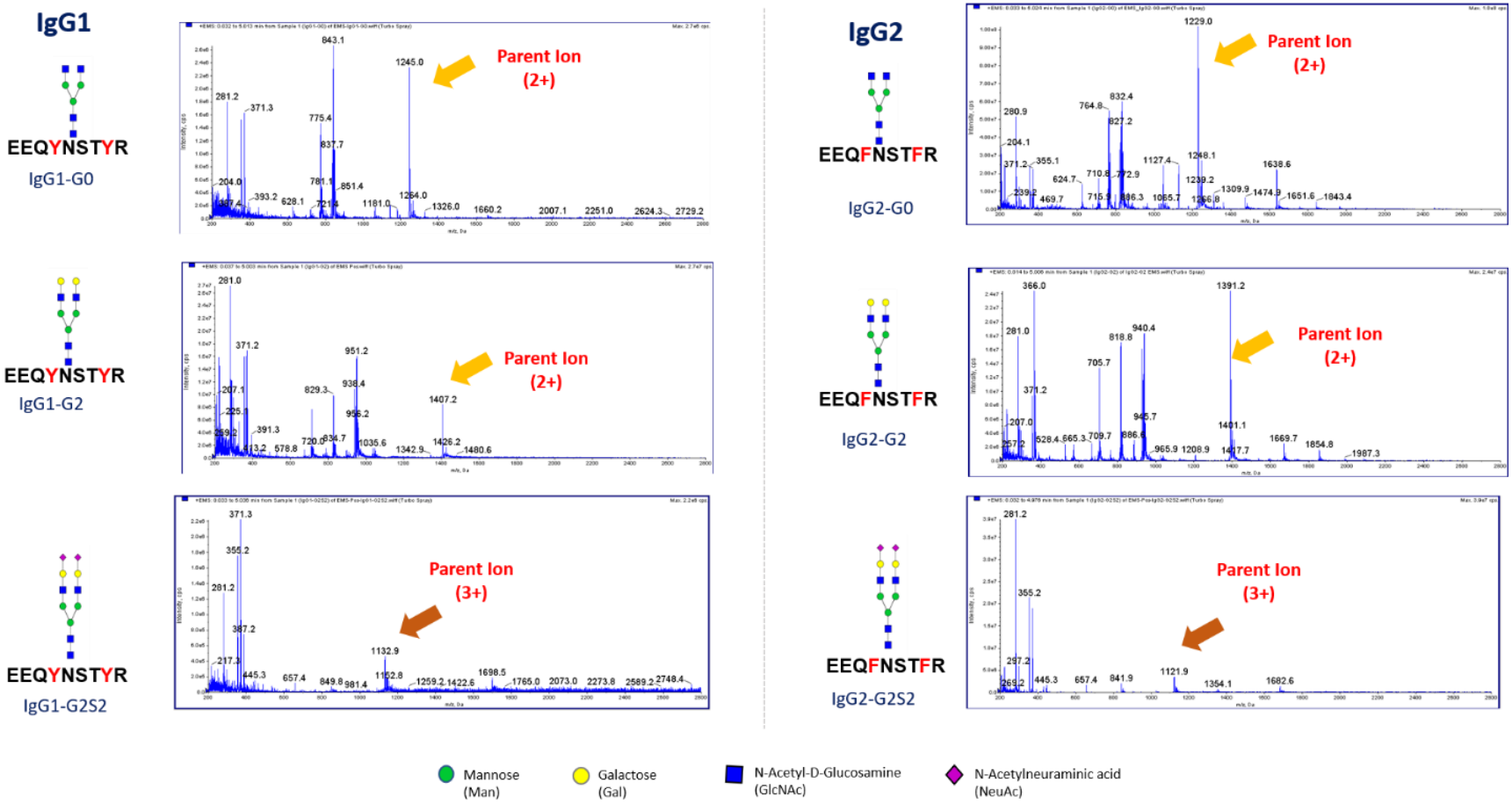


Figure 4-12. Enhanced mass spectrometry mode mass spectra of sialic acid containing IgG tryptic glycopeptides in negative mode. The different glycopeptides showed highest peaks at $[M-3H]^{3-}$ adducts for the parent ion peaks analyzed in 200-2800 m/z range.

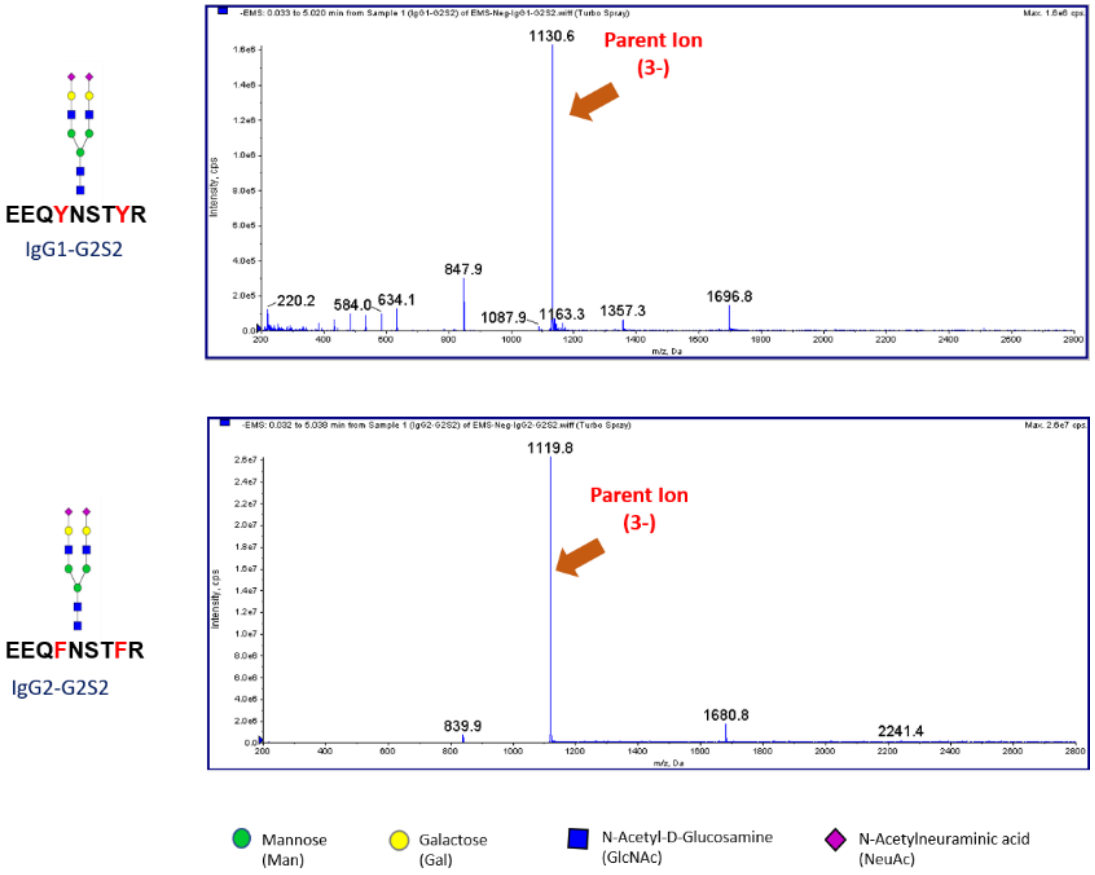


Figure 4-13. Enhanced resolution mode mass spectra of IgG tryptic glycopeptides in positive mode. The peak highlighted is for the parent peak at either $[M+2H]^{2+}$ or $[M+3H]^{3+}$ adducts.

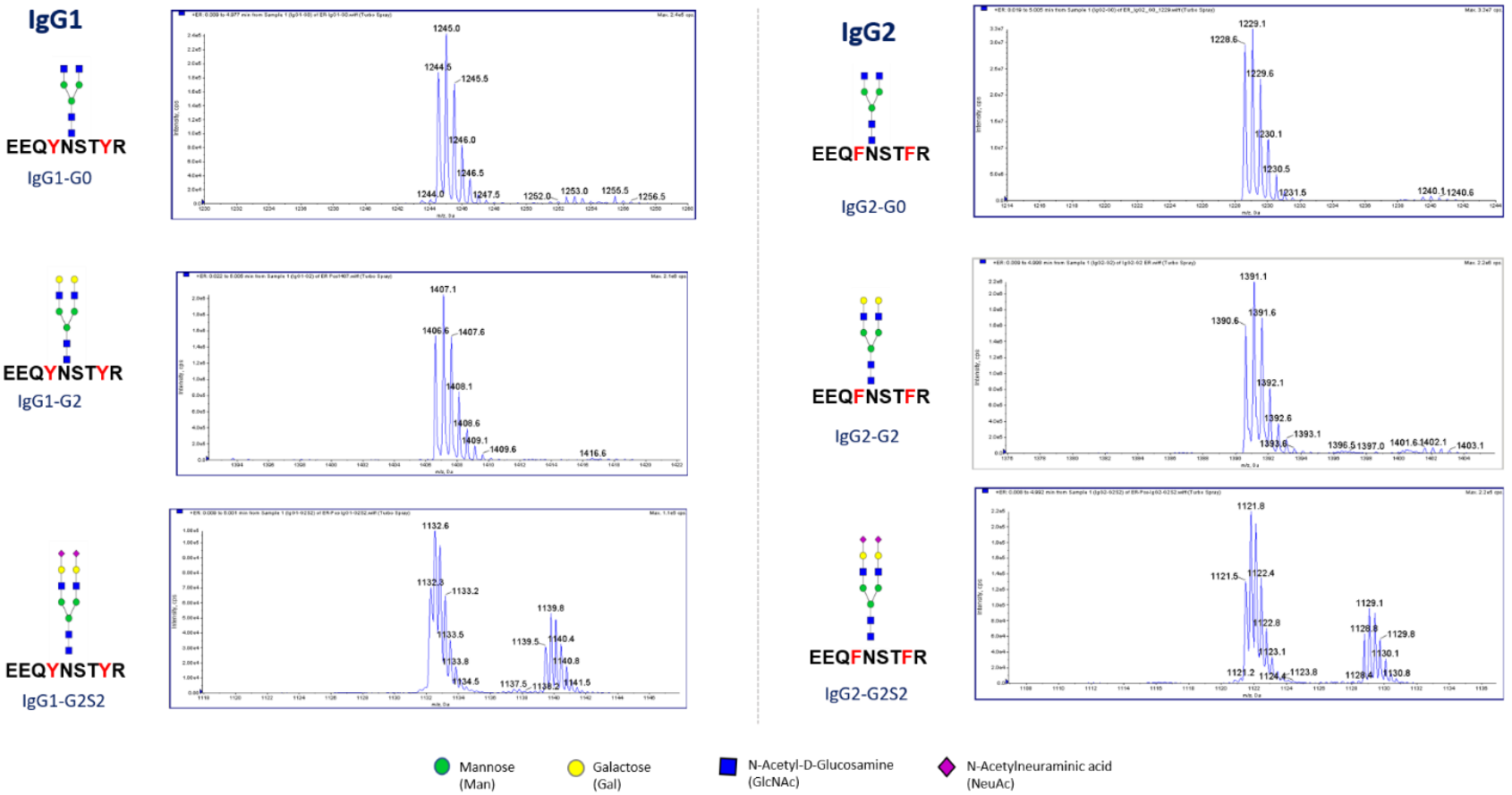
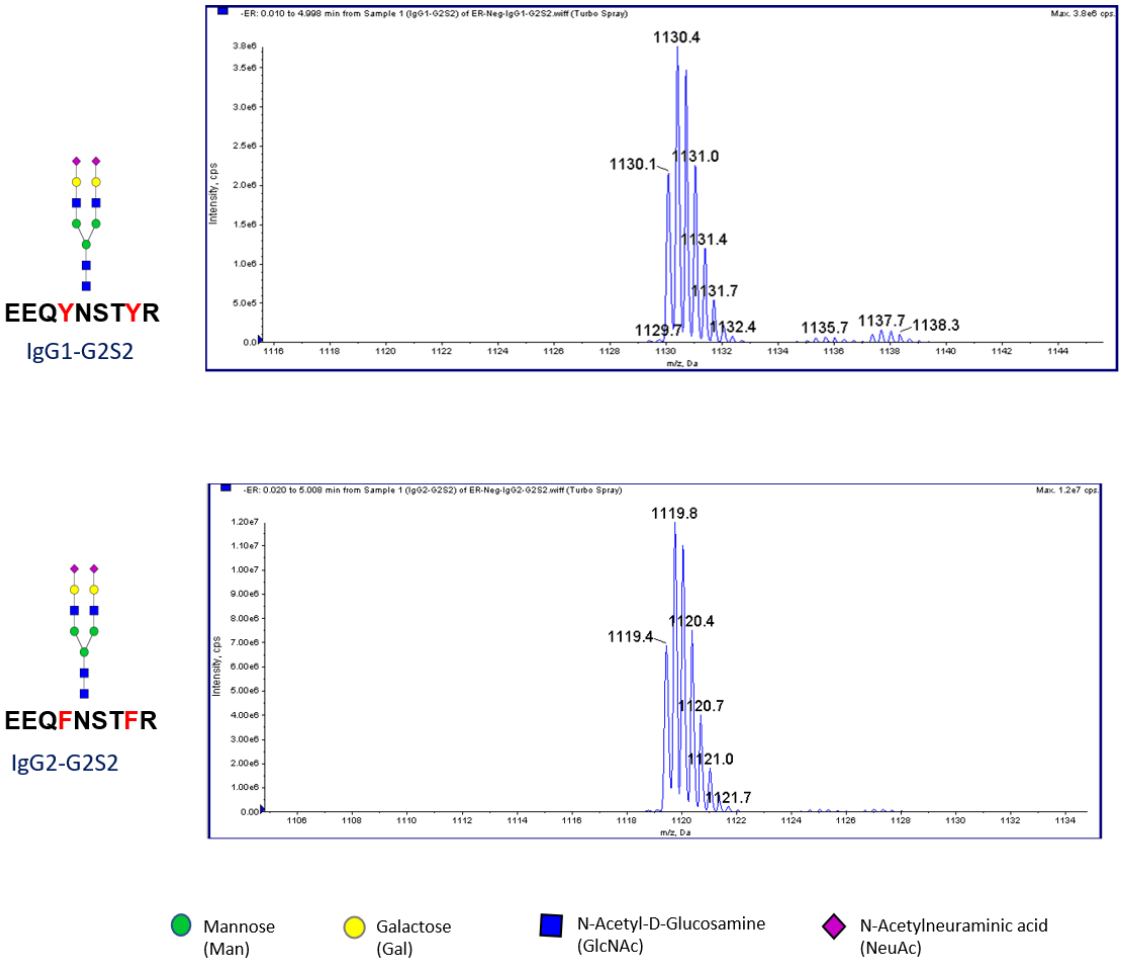
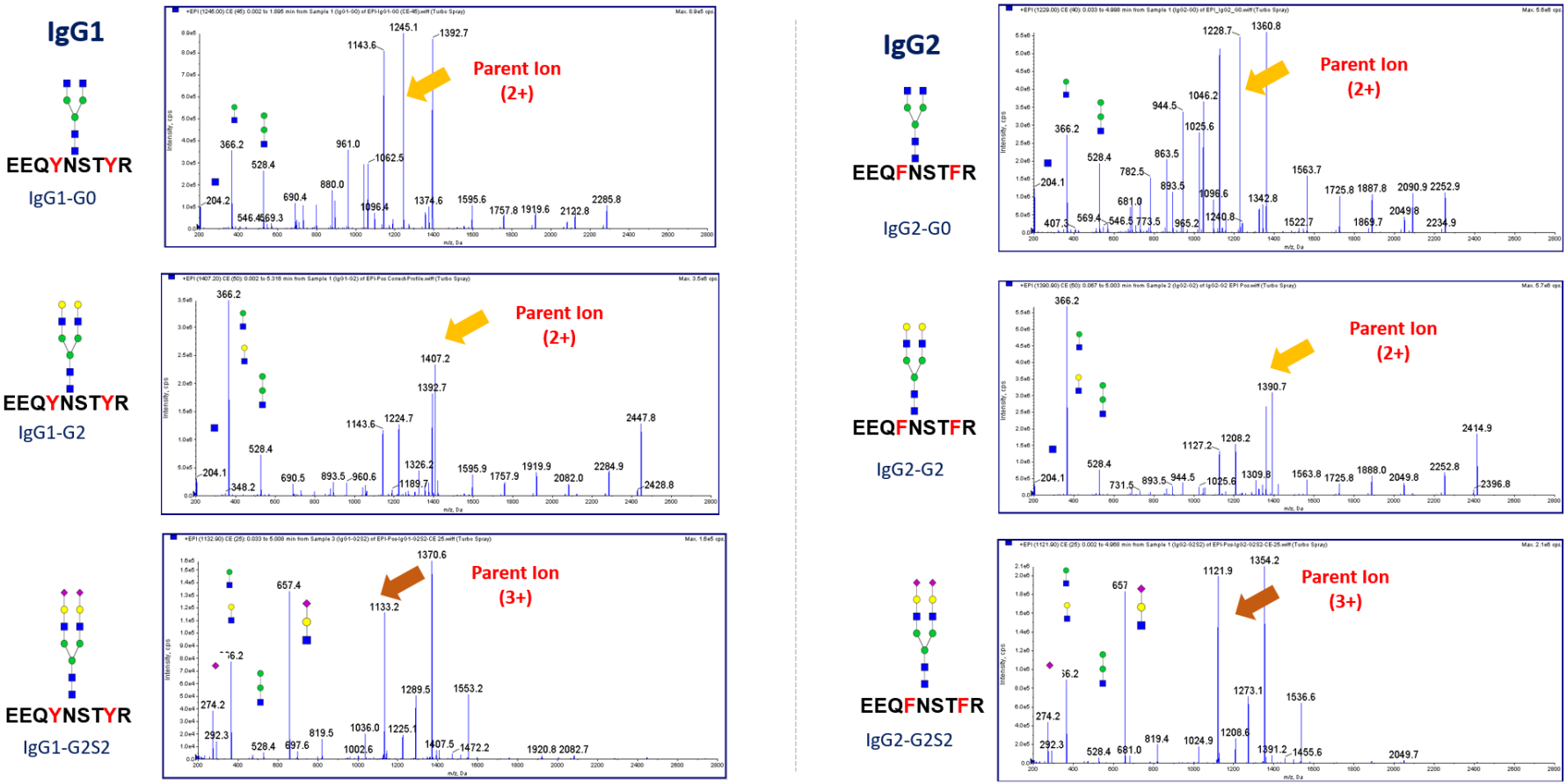


Figure 4-14. Enhanced resolution mode mass spectra of sialic acid containing IgG tryptic glycopeptides in negative mode. The peak highlighted is for the parent peak in $[M-3H]^{3-}$adduct.



The fragments from the parent ions were first identified before using for compound optimization parameters. Collision energies were optimized to get good intense peaks. A sample annotation of the oxonium fragments is presented in Figure 4-15.

Figure 4-15. Enhanced product ion mode of IgG tryptic glycopeptides in negative mode. The peak highlighted is for the parent ions with good S/N ratio. Gold and brown represent parent ions at either $[M+2H]^{2+}$ or $[M+3H]^{3+}$. Only a few annotations were presented here.



Assessment of Glycopeptide Structures

Looking at the MALDI-TOF/TOF and Enhanced Product Ion Mode in LC-MS/MS, sequential neutral losses of monosaccharides from the original parent glycopeptide, made possible the identification of possible oxonium ions for multiplex-monitoring. Carbohydrate oxonium fragment ions were used to indirectly associate it with the parent glycopeptide structure. These oxonium ions are HexNAc, $Hex_1HexNAc$, Neu5Ac, and $Hex_1HexNAc_1NeuAc_1$, corresponding to 204, 366, 292, 657 *m/z*, respectively, Neu5Ac can also be seen as a major product at 274 *m/z* pertaining to Neu5Ac subtracted with H_2O. In this work, B and Y type ions are common fragments showing high intense peaks in the spectra for all the IgG glycopeptides. Thus, Q3 fragments were chosen together with the respective Q1 value as a transition for compound optimization parameters and then LC-MS/MS monitoring.

Sialic acids are labile in nature and easily gets ionized in MALDI-TOF reflector positive mode. A less intense peak was seen for the IgG1-G2S2 and IgG2-G2S2 in positive mode. Switching to negative mode in MALDI-TOF/TOF, made the analysis easier. In LC-MS/MS, this was not seen as a problem as, we were able to optimize two glycopeptides, IgG1-G2S2 and IgG2-G2S2, in positive and negative modes. Furthermore, calibration standards were prepared, and a calibration of the line was calculated.

In the glycoform compound optimization, a series of standards were serially diluted to produce a range of calibration standards. A glycopeptide mixture of IgG1 and IgG2 tryptic glycoforms of known concentration were used as initial check for elution time. After which, a scheduled MRM was conducted after determining the elution of time for each of the glycopeptides and calibration line prepared.

Scheduled MRM Parameters for IgG glycopeptide Calibration and Quantitation

Table 4-1. Compound optimization parameters for scheduled multiple reaction monitoring

Q1	*Q3*	*Time (min)*	*Identity*	*DP*	*EP*	*CE*	*CXP*
1245.0	204.2	8.28	IgG1-G0	101	10	57	10
1407.2	204.1	8.20	IgG1-G2	91	10	71	14
1132.9	204.0	11.93	IgG1-G2S2	201	10	77	8
1229.0	204.1	11.11	IgG2-G0	101	10	65	14
1390.7	204.1	10.97	IgG2-G2	121	10	77	16
1121.9	204.1	15.98	IgG2-G2S2	116	10	65	8

DP; declustering potential, EP; entrance potential, CE; collision energy, CXP; collision cell exit potential

A scheduled MRM analysis was used for the calibration. The transitions presented in Table 4-1 gave out the best linearity and yielded high intensity peaks which are relevant in quantitation. High intensity peaks from the transitions is necessary because the glycopeptide analytes are in trace quantities.

Quantitation of IgG Glycopeptides in Human Serum

A scheduled MRM analysis with compound optimization parameters was used to analyze the presence of the scarce non-fucosylated IgG glycopeptides in human serum basing on the calibration standards prepared. The calibration of the IgG glycoforms yielded good linearity as shown by the r^2 values shown in Tables 4-2 and 4-3.

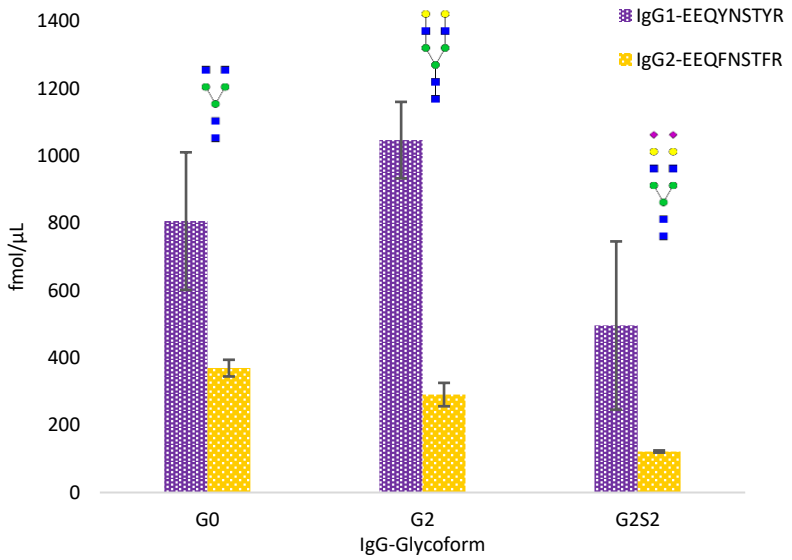
Table 4-2. Calibration standard for IgG1 glycoforms analyzed in positive polarity mode. The calibration standards used were 0, 25, 50, 100, 200, 500, 1000, 2000, and 5000 fmol/ µL.

IgG1	*Transition*	*Equation of the Line*	r^2
IgG1-G0	1245/204.2	y=28.94084x-1188.24883	0.99742
IgG1-G2	1407.2/204.1	y=7.49364x-634.28300	0.99317
IgG1-G2S2	1132.9/204	y=1.77922x-88.12468	0.99730

Table 4-3. Calibration standard for IgG2 glycoforms analyzed in positive polarity mode. The calibration standards used were 0, 25, 50, 100, 200, 500, 1000, 2000, and 5000 fmol/ µL.

IgG2	*Transition*	*Equation of the Line*	r^2
IgG2-G0	1229.0/204.1	y=118.47528x-4951.12629	0.99926
IgG2-G2	1390.7/204.1	y=19.15928x-624.74676	0.99994
IgG2-G2S2	1121.9/204.1	Y=157.20091x-4658.29151	0.99974

Figure 4-16. IgG glycoform amount present in human serum expressed in fmol/µL.



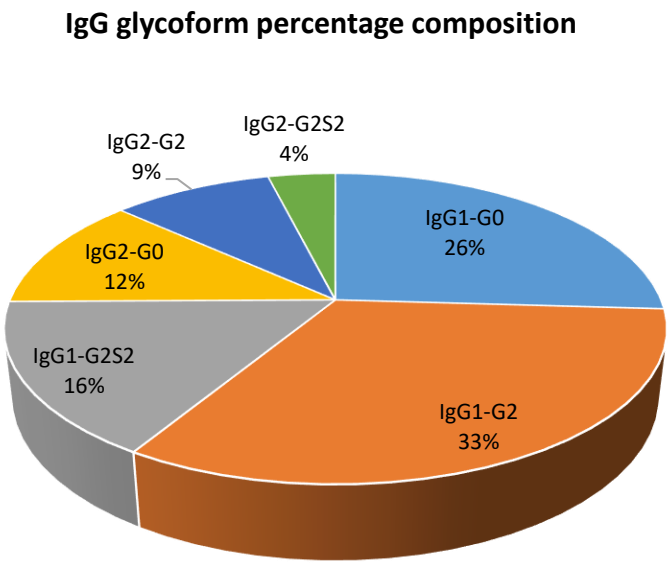
IgG1 glycoforms yielded higher amounts compared to IgG2 glycopeptides as shown in Figure 4-16 and Table 4-4. Values are expressed in fmol/µL. The total IgG tryptic glycopeptides are 2348 fmol/µL and 782 fmol/µL for IgG1 and IgG2 glycoforms, respectively. There could be other glycoforms present in IgG but in this work was only focused on G0, G2, and G2S2.

Table 4-4. Percentage composition based on total amount per IgG group and total IgG glycoform amount present in human serum

IgG glycoform	*IgG1-EEQYNSTYR*				*IgG2-EEQFNSTFR*			
	Average Concn., fmol/µL	SD (n=20)	*IgG1 % composition*	*Total IgG % composition*	*Average Concn., fmol/µL*	SD (n=20)	*IgG2 % composition*	*Total IgG % composition*
G0	806	204	35	26	369	25	47	12
G2	1047	114	45	33	291	35	37	9
G2S2	496	250	21	16	122	3	16	4
Total IgG per group, fmol/µL	2348	198*			782	25*		
Total IgG, fmol/µL	3130±345*							

*S_{pooled}

Figure 4-17. Percentage composition of each IgG glycopeptide based on the total IgG gycoform amount present in human serum.



IgG glycoform percentage comparison is shown in Figure 4-17. The % composition of IgG1 and IgG2 present in human serum is 60-70 and 20-25, respectively, according to a previous report. [12] The tryptic digestion was conducted for 24h. But it is important to note that there could be missed cleavages during digestion. IgG enrichment was no longer imperative with the targeted MRM strategy presented in this work.

Internal standard spiking of IgG glycopeptides in Human Serum

In both positive and negative polarities, the effect of internal standard spiking was conducted for additional confirmation of the peak believed to be the analyte in question. Theoretically, this is no longer necessary since MRM is a targeted method wherein the transitions and parameters inputted either in scheduled or unscheduled MRM makes it already selective for the analyte. As such, the peaks at that elution time is for the target analyte. A representative of the internal standard spiking for IgG2-G2S2 analyzed in both positive (Scheduled MRM) and negative (MRM) polarities is shown in Figures 4-18 and 4-19. Additional figures are placed in the supplementary information.

Figure 4-18. Spiking of IgG2-G2S2 tryptic glycopeptide in human serum analyzed with transition 1121.9/204.1 with equation of the line, y=121.94474x+461.94449 (r^2=0.99946) in positive polarity under MRM. The standards added are 25, 100, 500, 1000, and 2000 fmol/µL. The bottom spectra are the expanded versions of the top spectra.

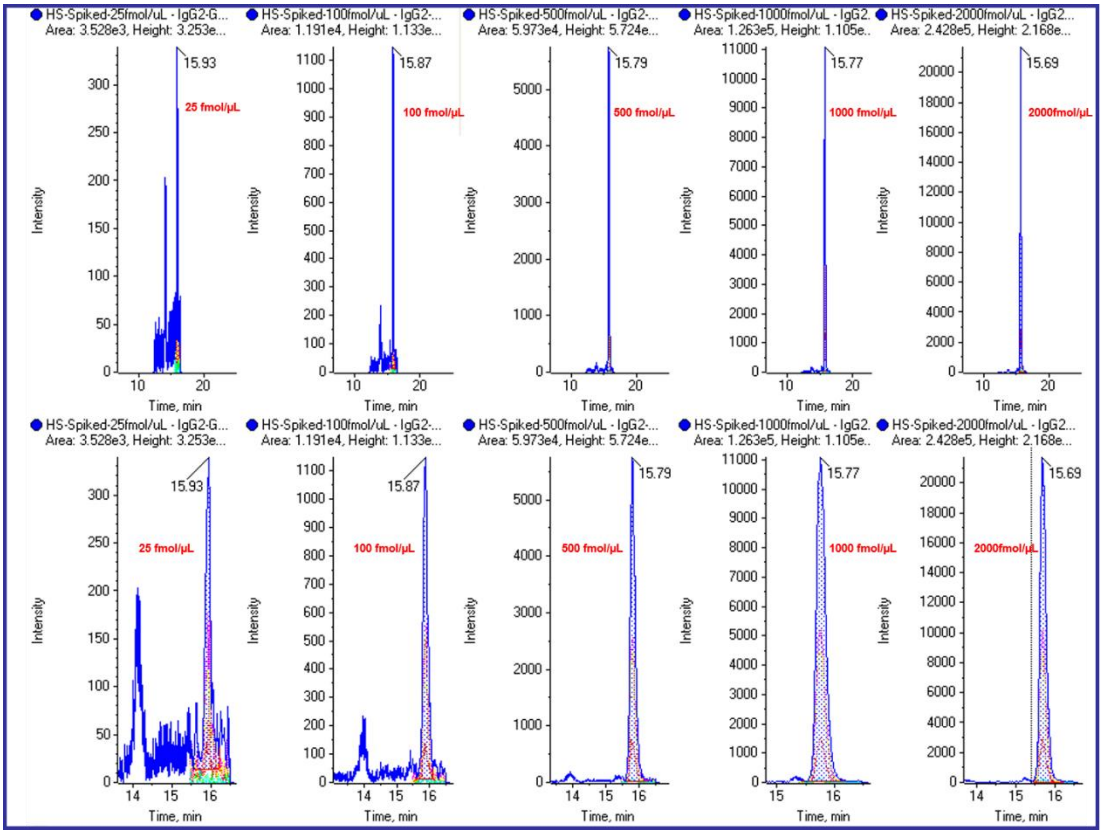
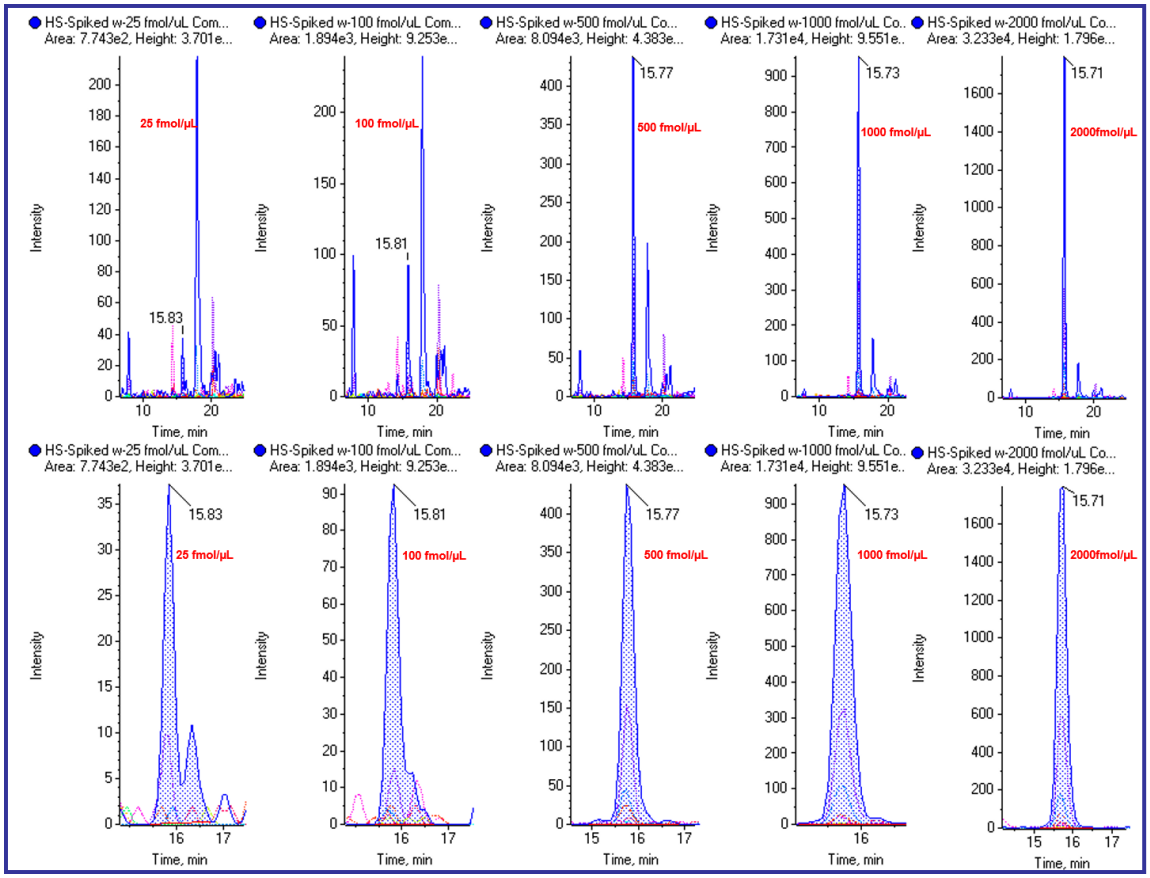


Figure 4-19. Spiking of IgG2-G2S2 tryptic glycopeptide in human serum analyzed with transition 1119.8/290.2 in negative polarity with equation of the line, y=37.06224x-2199.87043 (r^2=0.99654) under MRM. The standards added are 25, 100, 500, 1000, and 2000 fmol/µL. The bottom spectra are the expanded versions of the top spectra with seen elution at ~16 min.



The addition of the standard would reveal that the elution time checked for the specific analyte is confirmed to be in that position. It is also important to note that the different transitions were used for both polarities, 1390.7/204.1 and 1119.8/290.2, for positive and negative, respectively. Quantitation of the sample in negative polarity using the specified transition was not conducted because peak was not intense and only the sialic acid containing glycopeptide was analyzed on that polarity. As such, the final quantification was based on multiplexing in the positive polarity for the 6 glycopeptides.

Previously, synthesized glycopeptides[13-15] for single-reaction monitoring studies[14,15] was performed from our group.[13-14] The quantitation of scarce non-fucosylated tryptic IgG1 antibodies in Herceptin were conducted on a bi antennary and bisecting-glycan attached to EEQYNSTYR. Single reaction monitoring was conducted for the previous work. [14] Lot to lot variations of glycans are seen due to the inconsistent and unoptimized drug glycosylation. [15]In such a case, having quality control checks and synthesized calibration standards would correctly quantitate the presence of different IgG glycoforms in therapeutic antibodies. Currently, this chapter provides a multiplexing analysis of 6 glycopeptides at a single sample injection. The need for the quantification of immunoglobulin (IgG) glycopeptides are needed to confirm glycomics data for *N*-glycan profiling results. Important functions in immune homeostasis of IgG glycosylation were highlighted .[8]

IgG Glycopeptide Quantitation in Pancreatic Cancer Human Serum Samples

The previous sections were directed towards the quantitation of the presence of the IgG glycopeptides from the pooled human male serum of US origin (CHS). The current section details the quantitation in Japanese human serum samples in both non-cancer (JHS) and pancreatic cancer (PC) samples. Different comparison was conducted to assess the relevance of a few characteristics, e.g. gender, cancer and non-cancer, and based on TNM classification. The demographics of the samples are presented in Tables 4-5 to 4-8. Each of the comparison is important and was made possible using MRM.

Table 4-5. Demographic Characteristics of Human Serum Samples

Label	Description	Male Samples	Female Samples	Age Range	Mean Age
JHS	Japanese Human Serum Control	6	5	50's and 60's	NS
CHS	Human Serum from human male AB plasma, USA Origin, (H4522, Sigma)	pooled	0	NS	NS
Stage 1	Pancreatic Cancer Stage 1, Japanese patients	1	0	85	-
Stage 2	Pancreatic Cancer Stage 2, Japanese patients	0	1	73	-
Stage 4a	Pancreatic Cancer Stage 4a, Japanese patients	4	1	56-69	~64
Stage 4b	Pancreatic Cancer Stage 4b, Japanese patients	9	8	53-82	~68

NS- not specified. CHS ethnicity of US origin was not specified. Each sample was analyzed thrice except CHS wherein n=20.

Table 4-6. TNM Classification of Pancreatic Cancer Human Serum Samples

Label	Original Coding	Sex	Age	TNM Classification
PC-1	14PANS0449M6A	M	67	IVb:T4N0M
PC-2	14PANS0450M5A	M	53	IVb:T4N2M
PC-3	14PANSO451F7A	F	77	IVb:T4N2M
PC-4	14PANS0452F7A	F	73	II:T2N0M0
PC-5	14PANS0453M6A	M	69	IVb:T4N1M
PC-6	14PANS0454F7A	F	77	IVb:T4N2M
PC-7	14PANS0455F7A	F	77	IVb:T4N0M
PC-8	14PANS0456M8A	M	85	I: T1N0M0
PC-9	14PANS0457M8A	M	82	IVb: T4N0M
PC-10	14PANS0458F7A	F	76	IVb: T4N0M
PC-11	14PANS0459F6A	F	62	IVb:T4N2M
PC-12	14PANS0460M6A	M	65	IVa:T4N1M0
PC-13	14PANS0461M5A	M	56	IVa:T4N0M0
PC-14	14PANS0462M6A	M	63	IVb: T3N0M
PC-15	14PANS0463F7A	F	71	IVb: TXN3M
PC-16	14PANS0464M6A	M	62	IVb: T3N1M
PC-17	14PANS0465F5A	F	56	IVb: T4N0M
PC-18	14PANS0466M6A	M	68	IVb:T4N3M
PC-19	14PANS0467M6A	M	65	IVa:T4N0M
PC-20	14PANS0468M6A	M	69	IVa:T4N0M
PC-21	14PANS0469F6A	F	66	IVb: T3N0M
PC-22	14PANS0470F6A	F	64	IVA:T4N1M
PC-23	14PANS0471M6A	M	69	IVb: T4N2M
PC-24	14PANS0472M5A	M	59	IVb: T4N3M

PC-Pancreatic cancer

Table 4-7. Description of Human Serum Samples

Before Treatment						After Treatment	
Label	Original Coding	Sex	Age	TNM Classification		Label	Original Coding
PC-17	14PANS0465F5A	F	56	IVb: T4N0M	→	PC-25	14PANS0465F5B
PC-21	14PANS0469F6A	F	66	IVb: T3N0M	→	PC-26	14PANS0469F6B
PC-22	14PANS0470F6A	F	64	IVA: T4N1M	→	PC-27	14PANS0470F6B
PC-23	14PANS0471M6A	M	69	IVb: T4N2M	→	PC-28	14PANS0471M6B
PC-24	14PANS0472M5A	M	59	IVb: T4N3M	→	PC-29	14PANS0472M5B

Comparison of human serum samples from US origin and Japanese samples

It is important that the control should be closely related to the human serum pancreatic cancer. All the samples must closely match with age, sex, and ethnicity[17,18] as glycosylation are seen to be different with the aforementioned characteristics. The t-test statistical comparison of Japanese human serum non-pancreatic control samples (JHS) and human serum of US-origin purchased (Sigma, H4522) in Figure 4-20 showed all significant results at 99.9% confidence interval. With this, the results clearly emphasize that it is best suited to use the Japanese human serum to serve as control for the human serum obtained from cancer patients.

Comparison of human serum samples from non-pancreatic and pancreatic cancer patients

The human serum from non-pancreatic cancer donors (JHS) and the pancreatic cancer (PC) patients were obtained from collaborators of Showa and Okayama University, respectively. After running samples in MRM, all TNM classified human serum were classified as PC human serum and compared with the JHS samples. All IgG glycopeptide comparison for JHS and PC classified samples in Figure 4-21 showed statistical significance. Furthermore, a gender-based comparison of the two classifications in Figure 4-22 were performed. No statistically significant difference between JHS male VS JHS female and PC-male VS PC-female for all IgG glycopeptides were seen. JHS-male and

PC-female for IgG2-G0 showed no statistically significant difference. For IgG2-G2, only JHS-male and PC-male showed statistically significant difference at 95% confidence interval.

Comparison of human serum samples from non-pancreatic and pancreatic cancer patients based on TNM classification

Since Stage IVa and Stage IVb pancreatic cancer human serum samples were not statistically significant at 95% confidence interval (see Supporting Figure S4-21), it is collectively classified as Stage 4. MRM quantitation of various cancer patients reveal interesting results for disease progression. Although human serum samples were taken from different subjects and different number of trials were conducted for each stage, a general trend could be assessed for Figure 4-23. There is an increase in concentration of the IgG glycopeptides from non-cancer human serum to Stage I patients except for IgG1-G2S2. A decrease from Stage I to Stage II patients except for sialylated glycopeptides and IgG2-G2 was also seen. No significant change from Stage II to Stage IV for all glycopeptides was seen. The concentrations from Stage I to Stage IV patients is not significant for sialylated glycopeptides. Overall, this MRM quantitation gives the 1st absolute quantitation of 6 low abundant IgG glycopeptides in Japanese pancreatic cancer patients which was made possible by the use of pure chemo-enzymatically synthesized glycopeptide calibration standards.

The increase in concentration at Stage I would mean that the body is triggered and starting the immune response until such time that chronic inflammation would reach out a certain plateau corresponding to the non-significant change in concentrations of the glycopeptides. Some of the glycopeptides may exhibit an increase or decrease in concentration as a consequence of being used-up for synthesis of the next glycoform for certain glycosylation pathways which are unclear as it is an interplay between enzymatic networks and sugar donors. In order to fully comprehend a trend, more samples for all TNM stages and analysis with enzymatic networks is necessary. When total glycopeptides were compared at each state (see Figure S4-22) that will in a way estimate the total concentration per IgG class, no huge difference is seen. With this, glycoform-specific absolute quantitation is relevant.

Comparison of human serum samples from before and after treatment of pancreatic cancer patients

Some Stage IV pancreatic cancer patients listed in Table 3 were treated, and new serum samples were taken. A comparison of the concentrations of the IgG glycopeptides were conducted before and after treatment shown in Figure 4-24. Statistically significant difference was shown for all glycopeptides except IgG1-G0 and IgG1-G2. The decrease in concentration of 4 glycopeptides would mean a change in the glycosylation profile.

Figure 4-20. Comparison of the presence of IgG Glycopeptides in human serum from US origin and Japanese human serum controls. JHS; Japanese human serum, CHS; human serum of US origin. Concentrations are expressed in fmol/µL. The middle line for each group is expressed as mean of the results while the outer lines depicts standard deviation. All comparisons showed significant difference at 99.9% confidence interval.

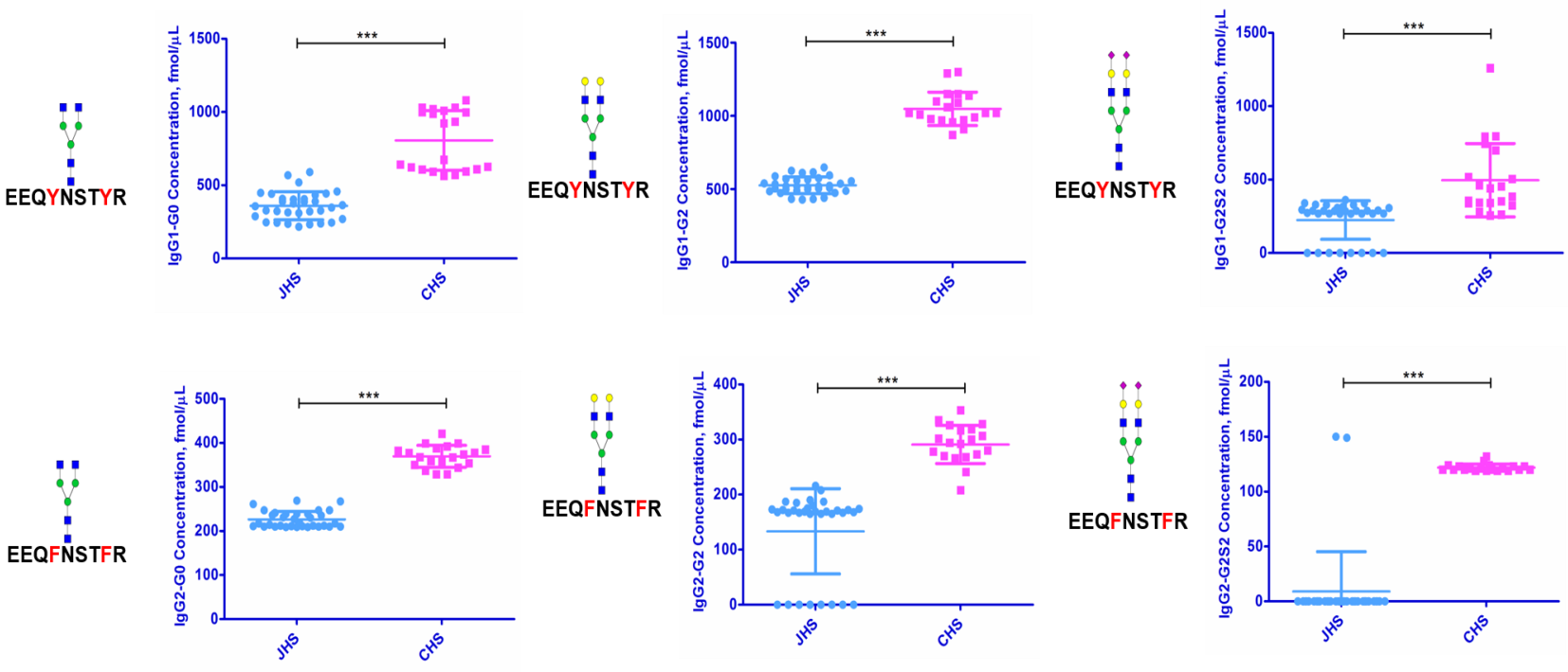


Figure 4-21. Comparison of the presence of IgG Glycopeptides in pancreatic and non-pancreatic control Japanese human serum samples. JHS; Japanese human serum, PC; human serum from pancreatic cancer patients. Concentrations are expressed in fmol/μL. The middle line for each group is expressed as mean of the results while the outer lines depicts standard deviation. All comparisons showed significant difference at 99.9% confidence interval, respectively.

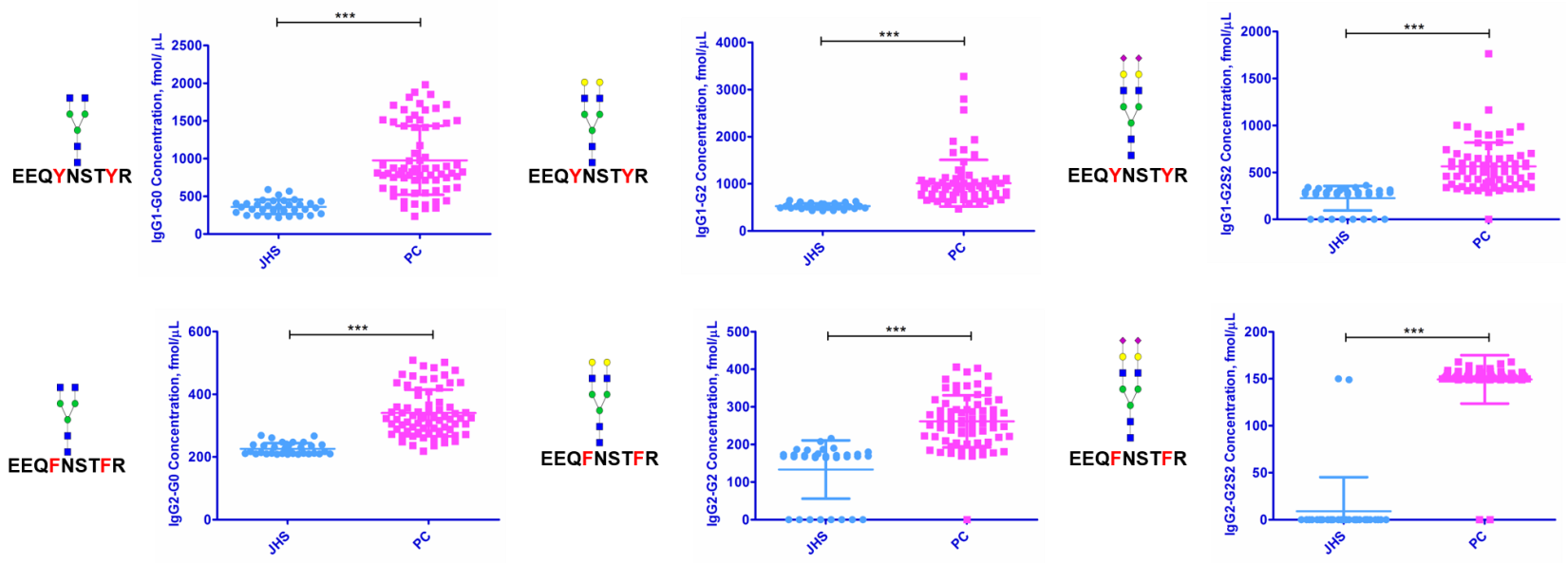


Figure 4-22. Gender comparison of the presence of IgG Glycopeptides in pancreatic and non-pancreatic control Japanese human serum samples. JHS; Japanese human serum, PC; human serum from pancreatic cancer patients. Concentrations are expressed in fmol/µL. The middle line for each group is expressed as mean of the results while the outer lines depicts standard deviation. All JHS-male VS JHS-female controls and PC-male VS PC-female does not show any significant difference. IgG2-G2 is not significantly different when JHS-female and PC-female samples are compared. *, **, and *** depicts significant difference at 95, 99, and 99.9% confidence interval, respectively.

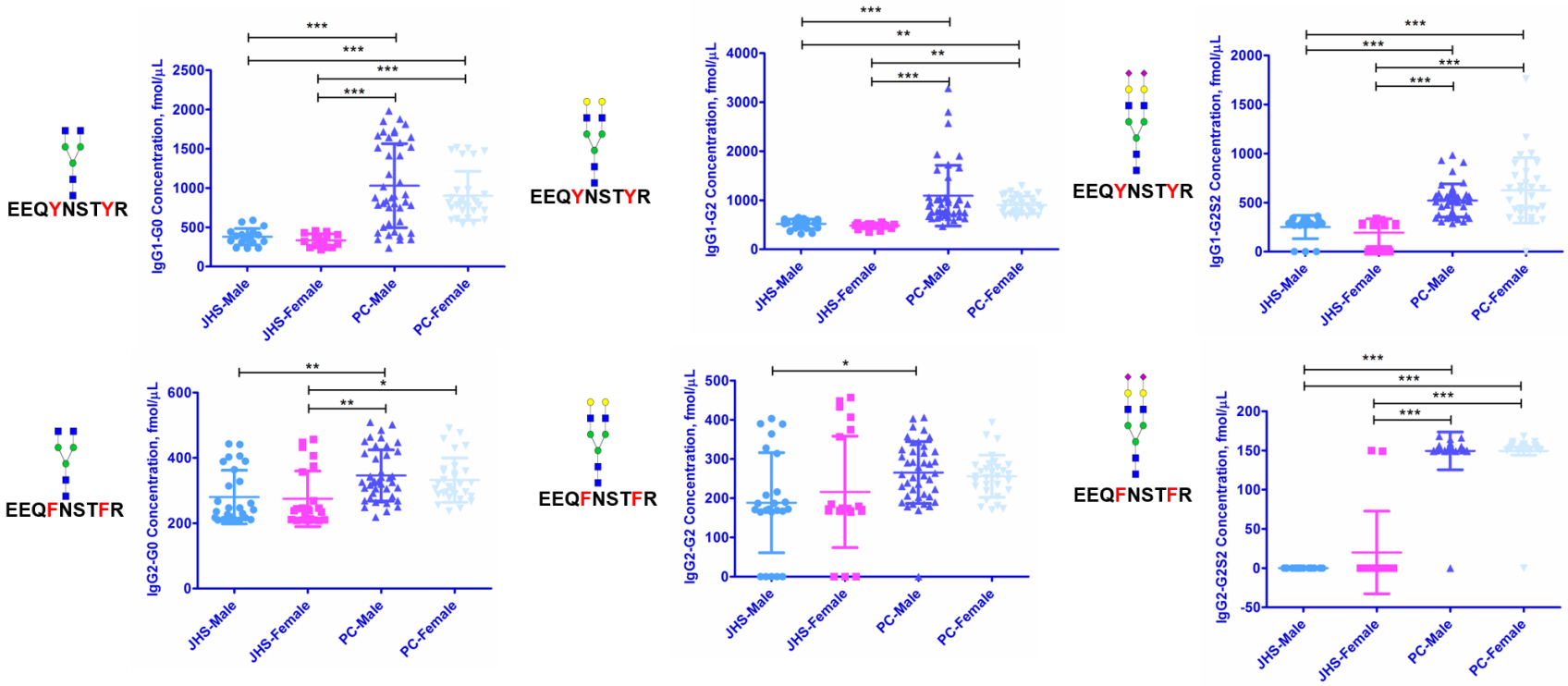


Figure 4-23. Comparison of the presence of IgG Glycopeptides at different TNM classification of pancreatic cancer patients. JHS; Japanese human serum non-cancer donors. Concentrations are expressed in fmol/µL. The middle line for each group is expressed as mean of the results while the outer lines depicts standard deviation. *, **, and *** depicts significant difference at 95, 99, and 99.9% confidence interval, respectively.

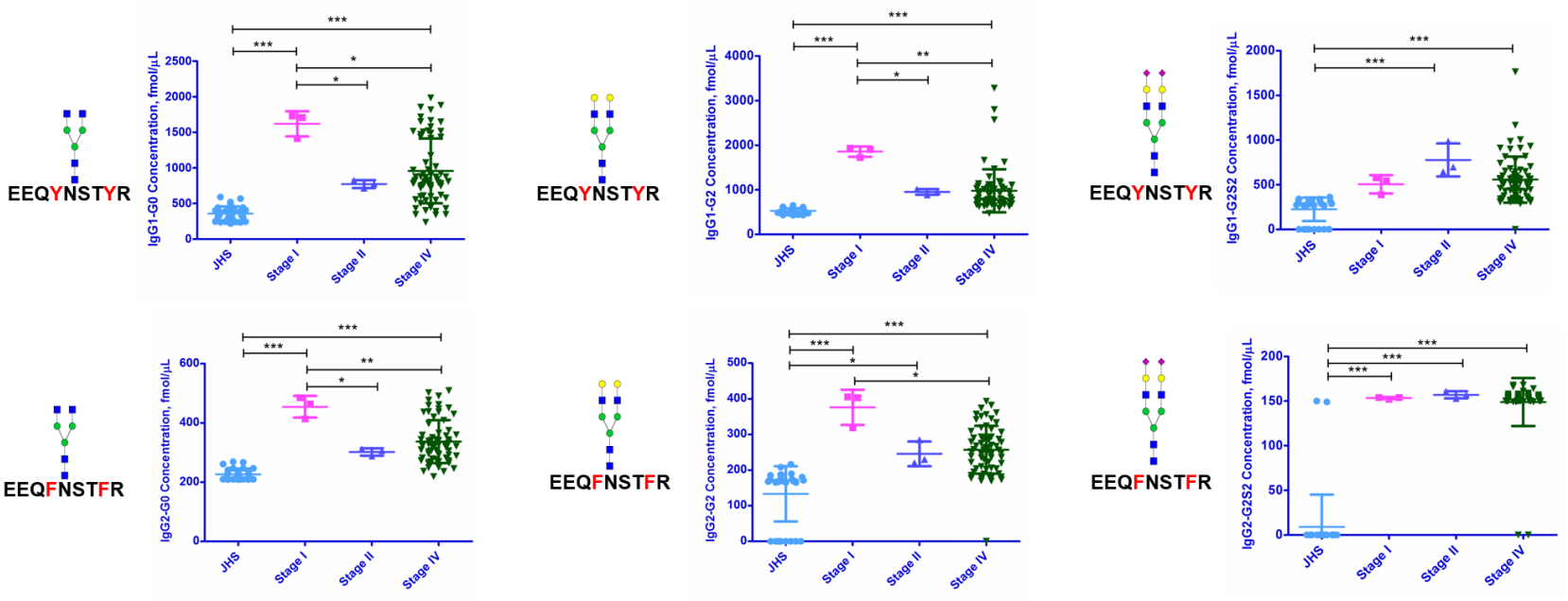
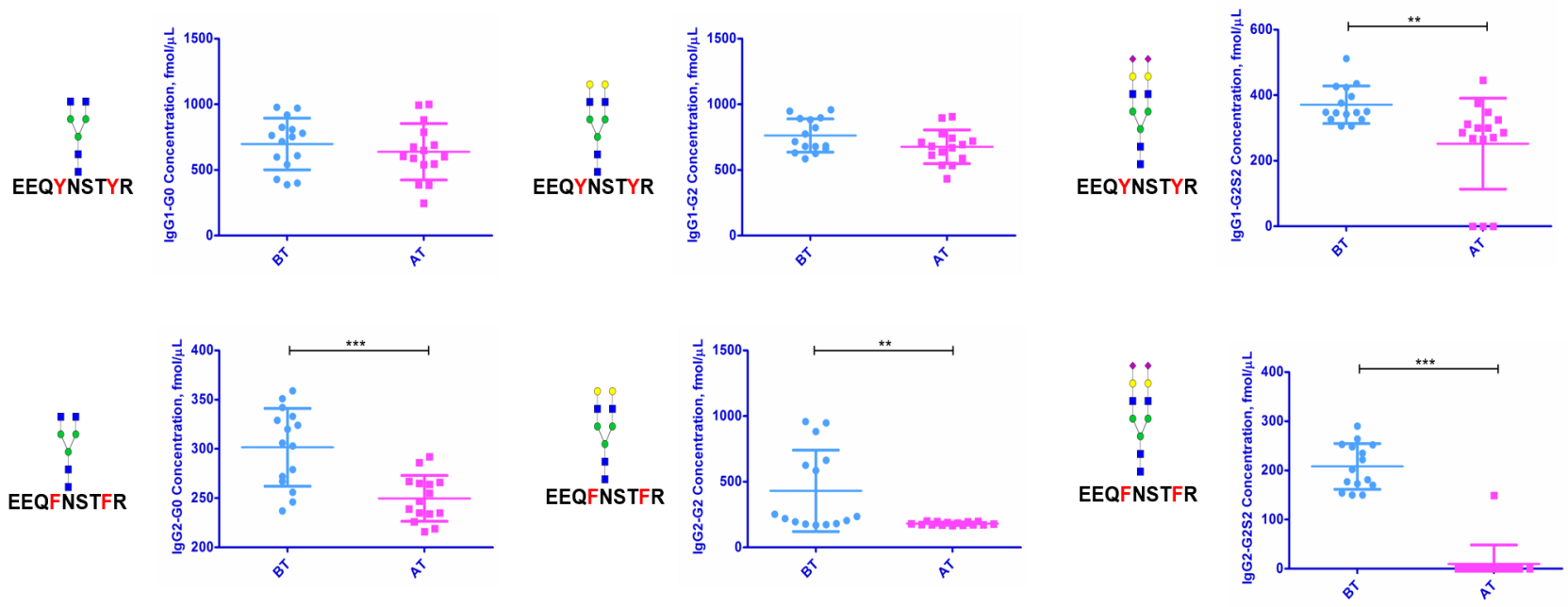


Figure 4-24. Comparison of the presence of IgG Glycopeptides for patients with Stage IV pancreatic cancer before and after treatment. The human serum is labeled as BT for before treatment and AT for after treatment. Concentrations are expressed in fmol/µL. The middle line for each group is expressed as mean of the results while the outer lines depicts standard deviation. 4 of the glycopeptides are statistically significant. The **, and *** depicts significant difference at 99, and 99.9% confidence interval, respectively.



4.5 Chapter Conclusion

IgG1-G2S2 and IgG4-G2S2 were chemoenzymatically prepared in this work using solid-phase peptide synthesis to obtain the glycosyl acceptor and the use of chicken egg yolk sialoglycopeptide as sugar donor. The transglycosylation reaction was accomplished using Endo-M. From this work, the need for endoglycosidase immobilization strategies to use for transglycosylation reaction is seen to be a necessity as these enzymes are only commercially available. Having no protein expression systems would make this a bottleneck for the chemoenzymatic preparation of glycopeptides. In the next chapter, Chapter 5, a procedure relating to an enzyme-immobilization support was performed.

The use of synthesized glycopeptides to serve as calibration standards has been shown in this chapter to quantitate the presence of tryptic IgG glycopeptides in human serum. Although scarce, the use of a sensitive and selective MRM equipment, made it possible to quantify IgG1 and IgG2 tryptic glycoforms. Applications of the work entails correct assessment of glycopeptide/glycoprotein components relevant in therapeutics. With a correctly defined glycan component, glycan-activity studies are better assayed. The use of multiplexing techniques in the current chapter such as MRM made quantification of the presence of tryptic glycopeptides faster without the need for IgG purification.

4.6 References

1 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. *Biochimica et Biophysica Acta (BBA)-General Subjects*, *1860*(8), 1716-1727.

2 Inafuku, S., Noda, K., Amano, M., Nishimura, S. I., & Ishida, S. (2016). increase of sialylated N-glycansin eyes with neovascular glaucoma secondary to proliferative diabetic retinopathy. *Current eye research*, *41*(5), 721-724.

3 Tomoda, T., Nouso, K., Kato, H., Miyahara, K., Dohi, C., Morimoto, Y., Kinugasa, H., Akimoto, Y., Matsumoto, K., Yamamoto, N. and Noma, Y. (2016). Alteration of serum N-glycan profile in patients with autoimmune pancreatitis. *Pancreatology*, *16*(1), 44-51.

4 Tanaka, T., Yoneyama, T., Noro, D., Imanishi, K., Kojima, Y., Hatakeyama, S., Tobisawa, Y., Mori, K., Yamamoto, H., Imai, A. and Yoneyama, T. (2017). Aberrant N-glycosylation profile of serum immunoglobulins is a diagnostic biomarker of urothelial carcinomas. *International journal of molecular sciences*, *18*(12), 2632.

5 Konno, N., Sugimoto, M., Takagi, T., Furuya, M., Asano, T., Sato, S., Kobayashi, H., Migita, K., Miura, Y., Aihara, T. and Komatsuda, A. (2018). Changes in N-glycans of IgG4 and its relationship with the existence of hypocomplementemia and individual organ involvement in patients with IgG4-related disease. *PloS one*, *13*(4), e0196163.

6 Zauner, G., Selman, M. H., Bondt, A., Rombouts, Y., Blank, D., Deelder, A. M., & Wuhrer, M. (2013). Glycoproteomic analysis of antibodies. *Molecular & Cellular Proteomics*, mcp-R112.

7 Lauc, G., Vučković, F., Bondt, A., Pezer, M., & Wuhrer, M. (2018). Trace N-glycans including sulphated species may originate from various plasma glycoproteins and not necessarily IgG. *Nature communications*, *9*(1), 2916.

8 Maverakis, E., Kim, K., Shimoda, M., Gershwin, M.E., Patel, F., Wilken, R., Raychaudhuri, S., Ruhaak, L.R. and Lebrilla, C.B. (2015). Glycans in the immune system and The Altered Glycan Theory of Autoimmunity: a critical review. *Journal of autoimmunity*, *57*, 1-13.

9 Hong, Q., Lebrilla, C. B., Miyamoto, S., & Ruhaak, L. R. (2013). Absolute quantitation of immunoglobulin G and its glycoforms using multiple reaction monitoring. *Analytical chemistry*, *85*(18), 8585-8593.

10 Roy, R., Ang, E., Komatsu, E., Domalaon, R., Bosseboeuf, A., Harb, J., Hermouet, S., Krokhin, O., Schweizer, F. and Perreault, H. (2018). Absolute quantitation of glycoforms of two human IgG subclasses using synthetic Fc peptides and glycopeptides. *Journal of The American Society for Mass Spectrometry*, *29*(6), 1086-1098.

11 Huang, J., Kailemia, M.J., Goonatilleke, E., Parker, E.A., Hong, Q., Sabia, R., Smilowitz, J.T., German, J.B. and Lebrilla, C.B. (2017). Quantitation of human milk proteins and their glycoforms using multiple reaction monitoring (MRM). *Analytical and bioanalytical chemistry*, *409*(2), 589-606.

12 Deshpande, V., Zen, Y., Chan, J.K., Eunhee, E.Y., Sato, Y., Yoshino, T., Klöppel, G., Heathcote, J.G., Khosroshahi, A., Ferry, J.A. and Aalberse, R.C. (2012). Consensus statement on the pathology of IgG4-related disease. *Modern pathology*, *25*(9), 1181.

13 Kumar H.V., R., Naruchi, K., Miyoshi, R., Hinou, H., Nishimura, S.-I. (2013) A new approach for the synthesis of hyperbranched N-glycan core structure from locust bean gum. *Org. Lett. 15*, 6268-6281.

14 Hammura, K., Ishikawa, A., Miyoshi, R., Yokoi, Y., Tanaka, M., Hinou, H., & Nishimura, S. I. (2018). Synthetic Glycopeptides Allow for the Quantitation of Scarce Nonfucosylated IgG Fc N-Glycans of Therapeutic Antibody. *ACS Medicinal Chemistry Letters*, *9*(9), 889-894.

15 Yogesh, K.V., Kamiyama, T., Ohyama, C., Yoneyama, T., Nouso, K., Kimura, S., Hinou, H. and Nishimura, S.I. (2018). Synthetic glycopeptides as a designated standard in focused glycoproteomics to discover serum cancer biomarkers. *MedChemComm*, *9*(8), 1351-1358.

16 Fernandes, D. L. in Biopharmaceutical Production Technology 545-583 (Wiley-VCH Verlag GmbH & Co. KGaA, 2012).

17 Gebrehiwot, A.G., Melka, D.S., Kassaye, Y.M., Rehan, I.F., Rangappa, S., Hinou, H., Kamiyama, T. & Nishimura, S. I. (2018). Healthy human serum N-glycan profiling reveals the influence of ethnic variation on the identified cancer-relevant glycan biomarkers. *PloS one*, *13*(12), e0209515.

18 Šimurina, M., de Haan, N., Vučković, F., Kennedy, N.A., Štambuk, J., Falck, D., Trbojević-Akmačić, I., Clerc, F., Razdorov, G., Khon, A. & Latiano, A. (2018). Glycosylation of immunoglobulin G associates with clinical features of inflammatory bowel diseases. *Gastroenterology*, *154*(5), 1320-1333.

4-7 Supplementary Information

Table S4-1. Settings for initial assessment of IgG glycopeptides for enhanced mass spectrometry, enhanced resolution, and enhanced product ion modes in either positive or negative polarity.

Parameters	*Positive Polarity*	*Negative Polarity*
Curtain Gas	10.0	10.0
Collision Gas	High	High
IonSpray Voltage	5500	-4500.0
Temperature (TEM)	0.0	0.0
Ion Source Gas 1	12.0	12.0
Ion Source Gas 2	0.0	0.0
Interface Heater	On	On
Declustering Potential	30	-130.0
Collision Energy	10.0	-10.0

Table S4-2. Compound optimization parameters for multiple reaction monitoring in negative polarity

Q1	*Q3*	*Time, msec*	*Identity*	*DP*	*EP*	*CE*	*CXP*
1130.6	290.0	200	IgG1-G2S2	-175	-10	-74	-17
1119.8	290.2	200	IgG2-G2S2	-155	-10	-72	-19

DP; declustering potential, EP; entrance potential, CE; collision energy, CXP; collision cell exit potential

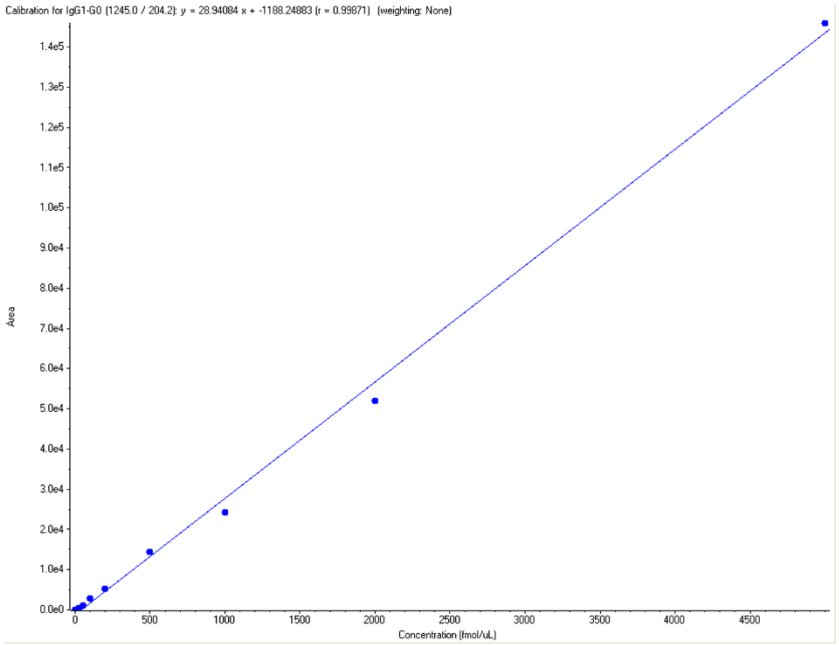
Table S4-3. Settings for multiple reaction monitoring

Parameters	*MRM (+)*	*MRM (-)*
Curtain Gas	10.0	10.0
Collision Gas	4	4
IonSpray Voltage	5500.0	-4500.0
Temperature (TEM)	0.0	0
Ion Source Gas 1	12.0	18.0
Ion Source Gas 2	0.0	0.0
Interface Heater	On	On
Entrance Potential	10.0	-10.0

Individual transition parameters are placed on Table 4-1 for positive polarity

Figure S4-1. Calibration of IgG1-G0 glycopeptide standard analyzed with transition 1245.0/204.2 in positive polarity with equation of the line, y=28.94084x-1188.24883 (r^2=0.99742) under MRM. A. Calibration curve. B. Elution time. The standards added are 0, 25, 50, 100, 200, 500, 1000, 2000 and 5000 fmol/µL dissolved in water.

A.



B.

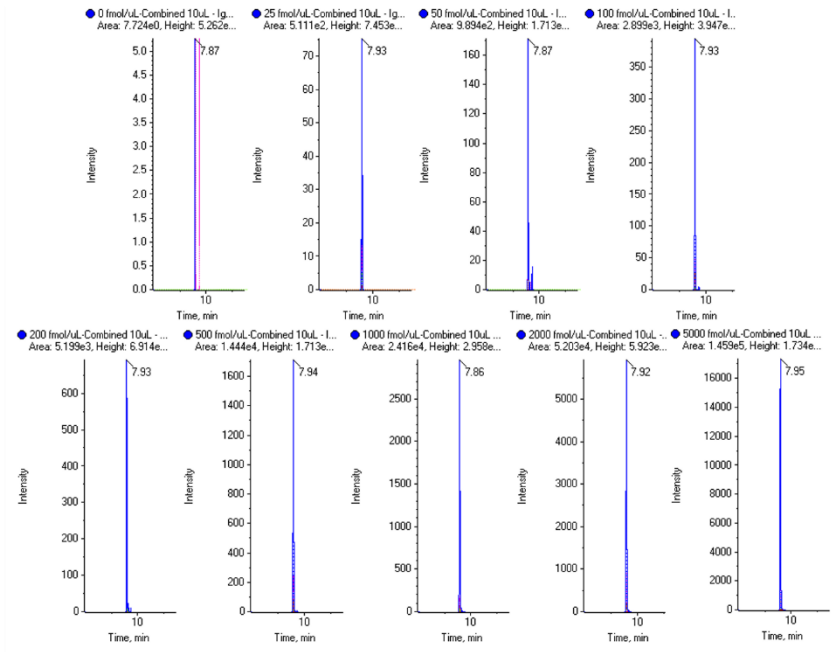
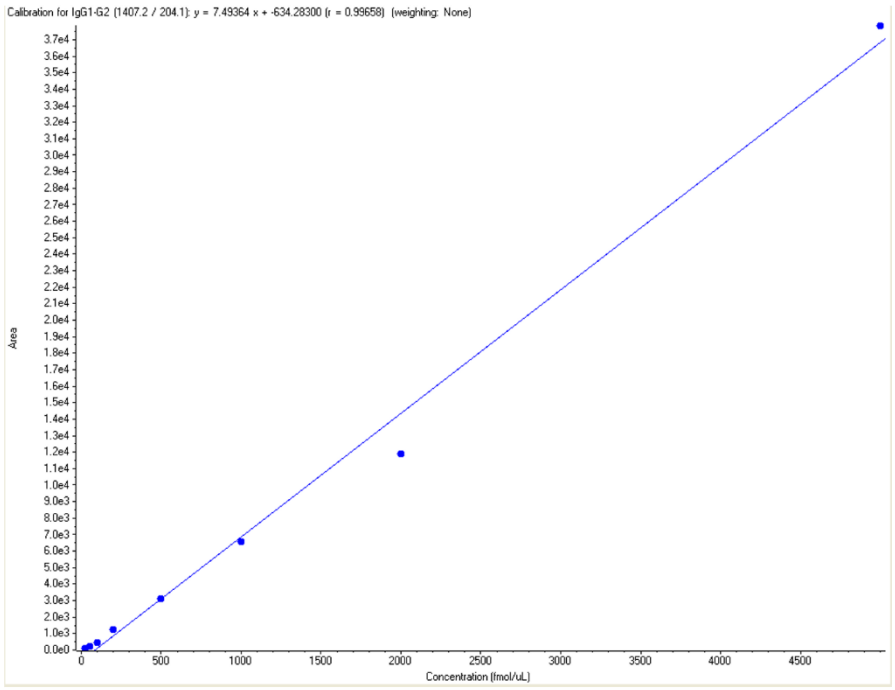


Figure S4-2. Calibration of IgG1-G2 glycopeptide standard analyzed with transition 1407.2/204.1 in positive polarity with equation of the line, y=7.49364x-634.28300 (r^2=0.99317) under MRM. . A. Calibration curve. B. Elution time. The standards added are 0, 25, 50, 100, 200, 500, 1000, 2000 and 5000 fmol/µL dissolved in water.

A.



B.

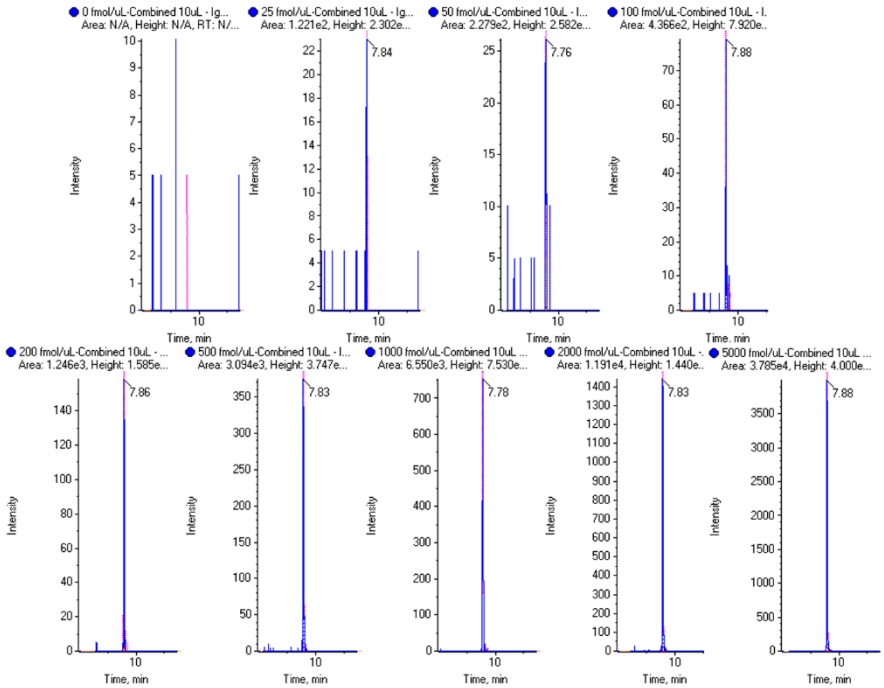
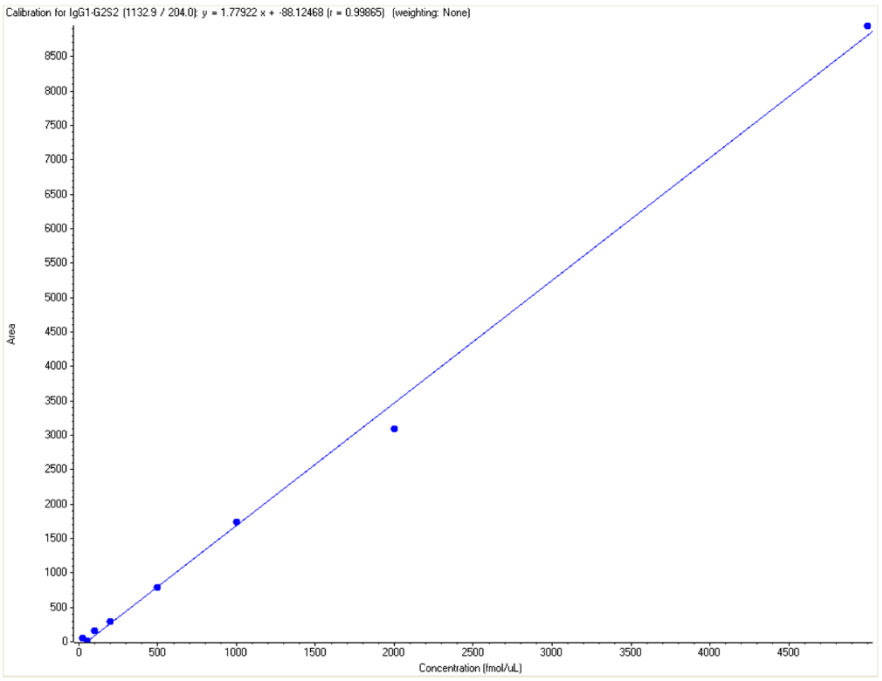


Figure S4-3. Calibration of IgG1-G2S2 glycopeptide standard analyzed with transition 1132.2/204.0 in positive polarity with equation of the line, y=1.77922x-88.12468 (r^2=0.99730) under MRM. A. Calibration curve. B. Elution time. The standards added are 0, 25, 50, 100, 200, 500, 1000, 2000 and 5000 fmol/µL dissolved in water.

A.



B.

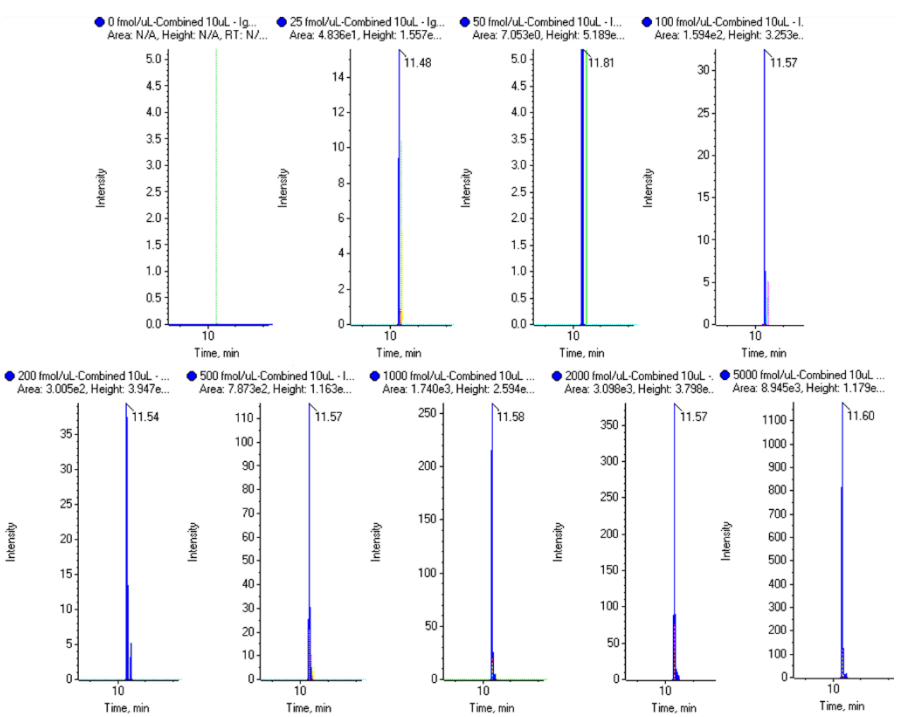


Figure S4-4. Calibration of IgG2-G0 glycopeptide standard analyzed with transition 1229.0/204.1 in positive polarity with equation of the line, y=118.47528x-4951.12629 (r^2=0.99928) under MRM. . A. Calibration curve. B. Elution time. The standards added are 0, 25, 50, 100, 200, 500, 1000, 2000 and 5000 fmol/µL dissolved in water.

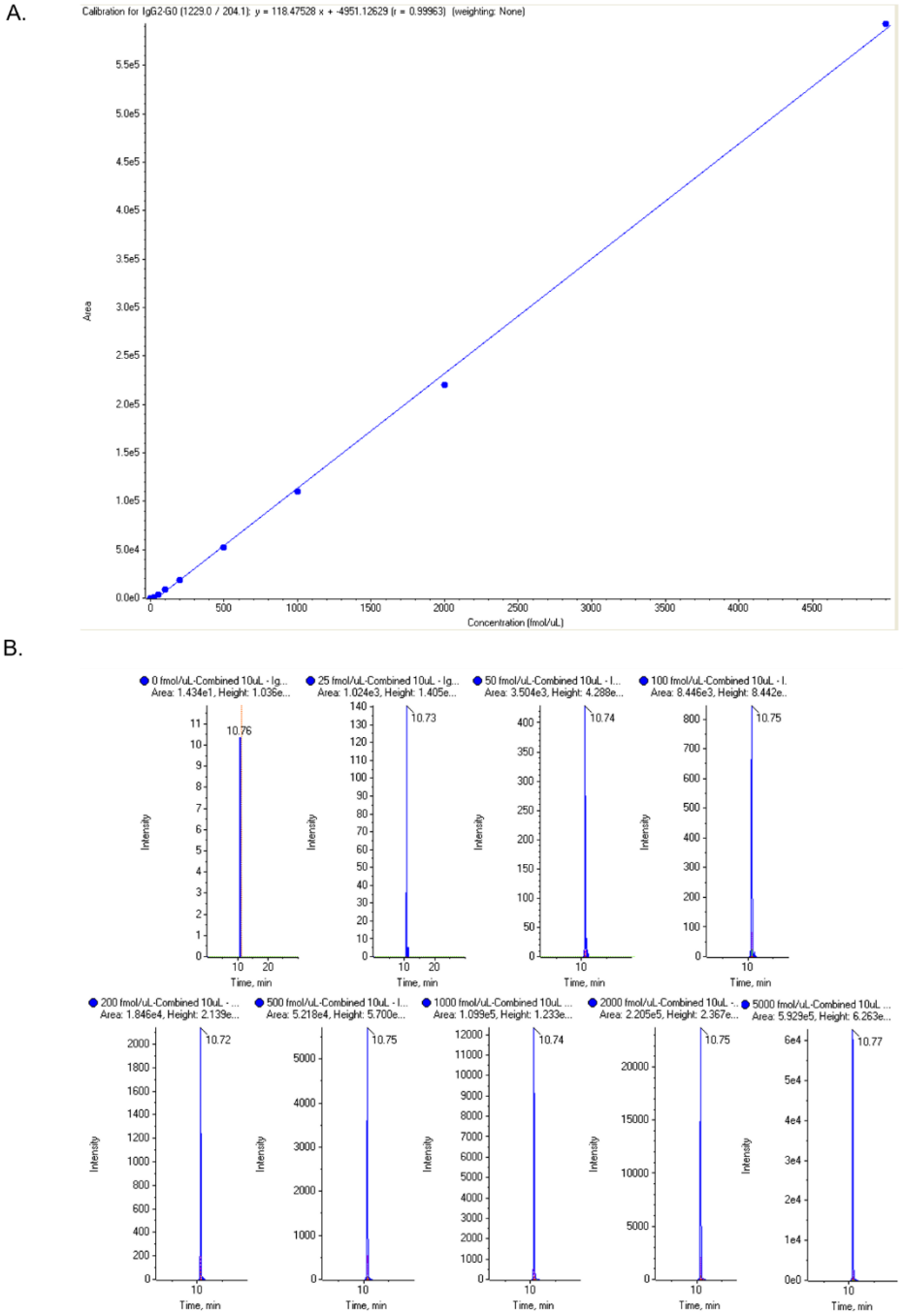


Figure S4-5. Calibration of IgG2-G2 glycopeptide standard analyzed with transition 1390.7/204.1 in positive polarity with equation of the line, y=19.15928x-624.74676 (r^2=0.99994) under MRM. . A. Calibration curve. B. Elution time. The standards added are 0, 25, 50, 100, 200, 500, 1000, 2000 and 5000 fmol/µL dissolved in water.

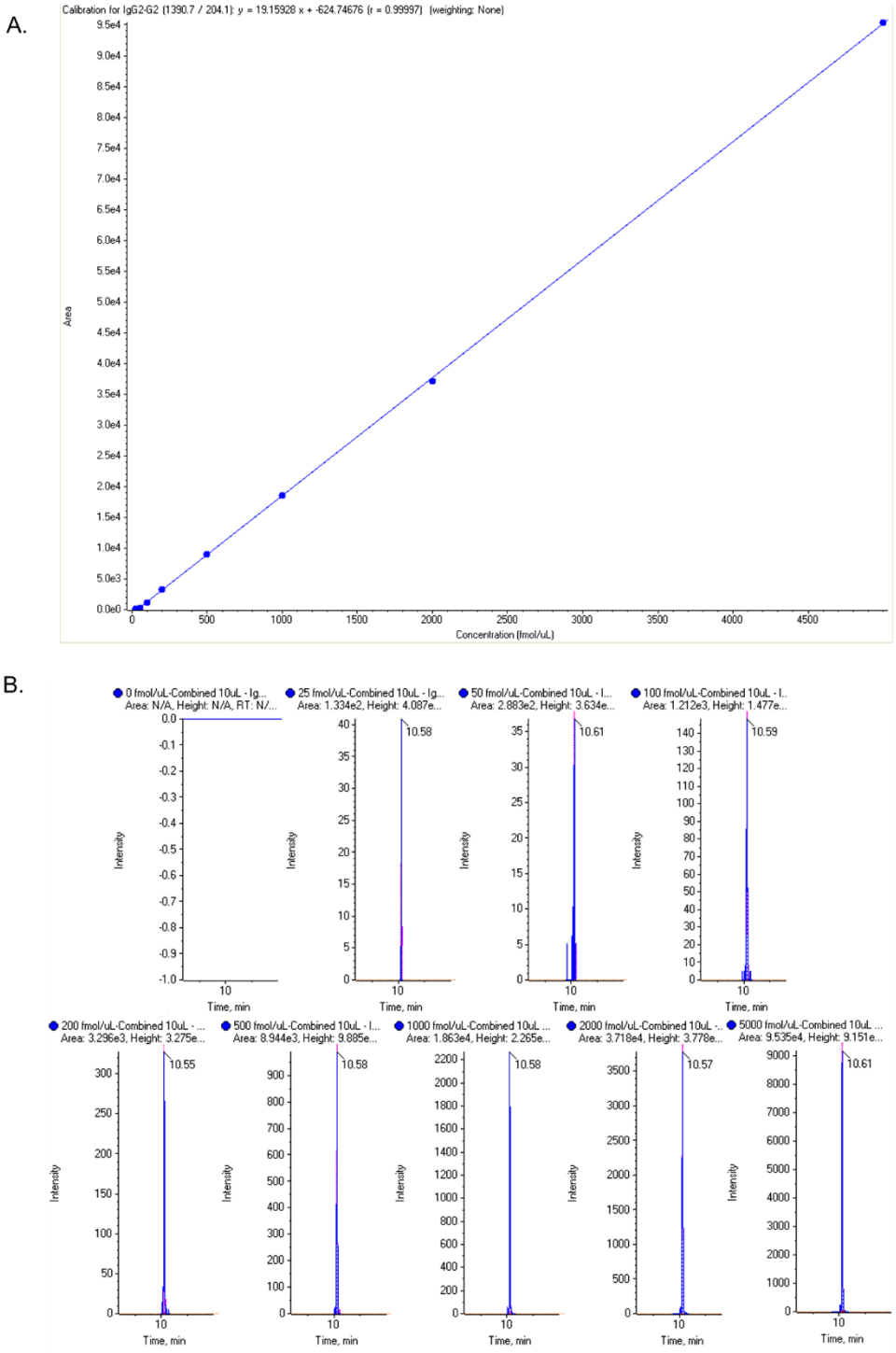
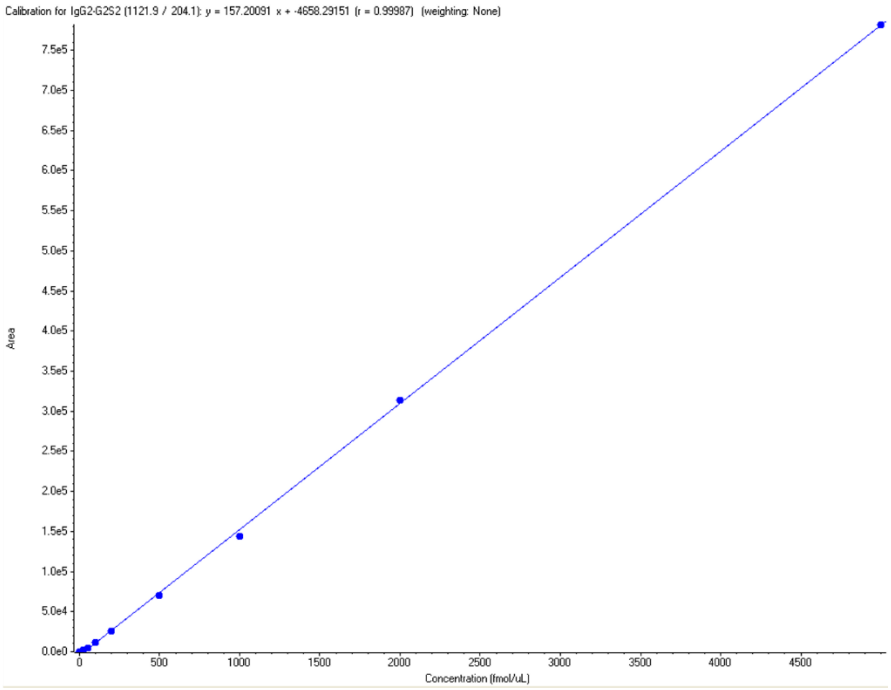


Figure S4-6. Calibration of IgG2-G2S2 glycopeptide standard analyzed with transition 1121.9/204.1 in positive polarity with equation of the line, y=157.20091x-4658.29151 (r^2=0.99974) under MRM. A. Calibration curve. B. Elution time. The standards added are 0, 25, 50, 100, 200, 500, 1000, 2000 and 5000 fmol/µL dissolved in water.

A.



B.

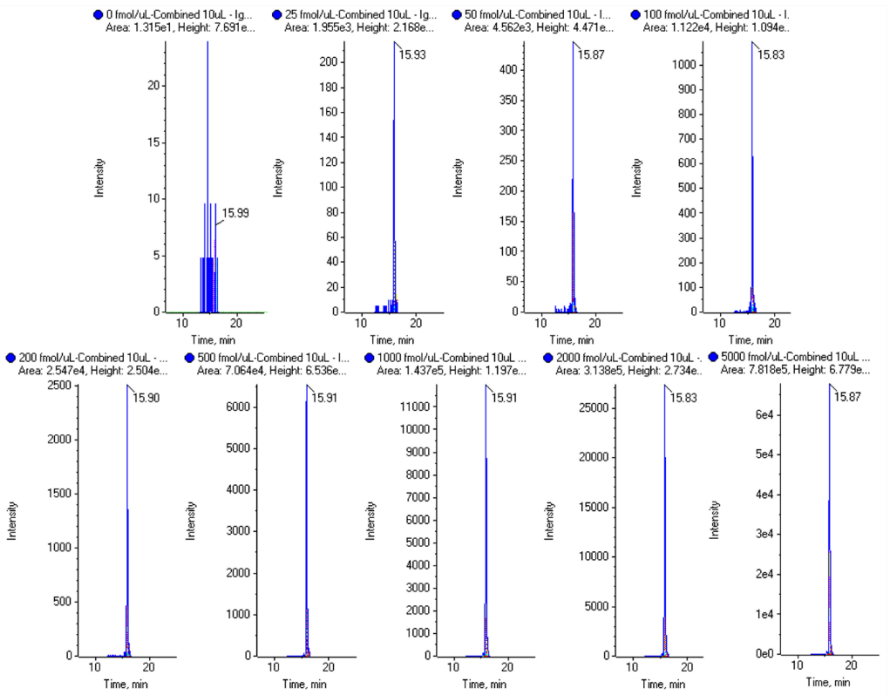


Figure S4-7. IgG1-G0 glycopeptide quantification analyzed with transition 1245.0/204.2 in positive polarity with equation of the line, y=28.94084x-1188.24883 (r^2=0.99742) under MRM. Shown below are 10 representative spectra from n=20.

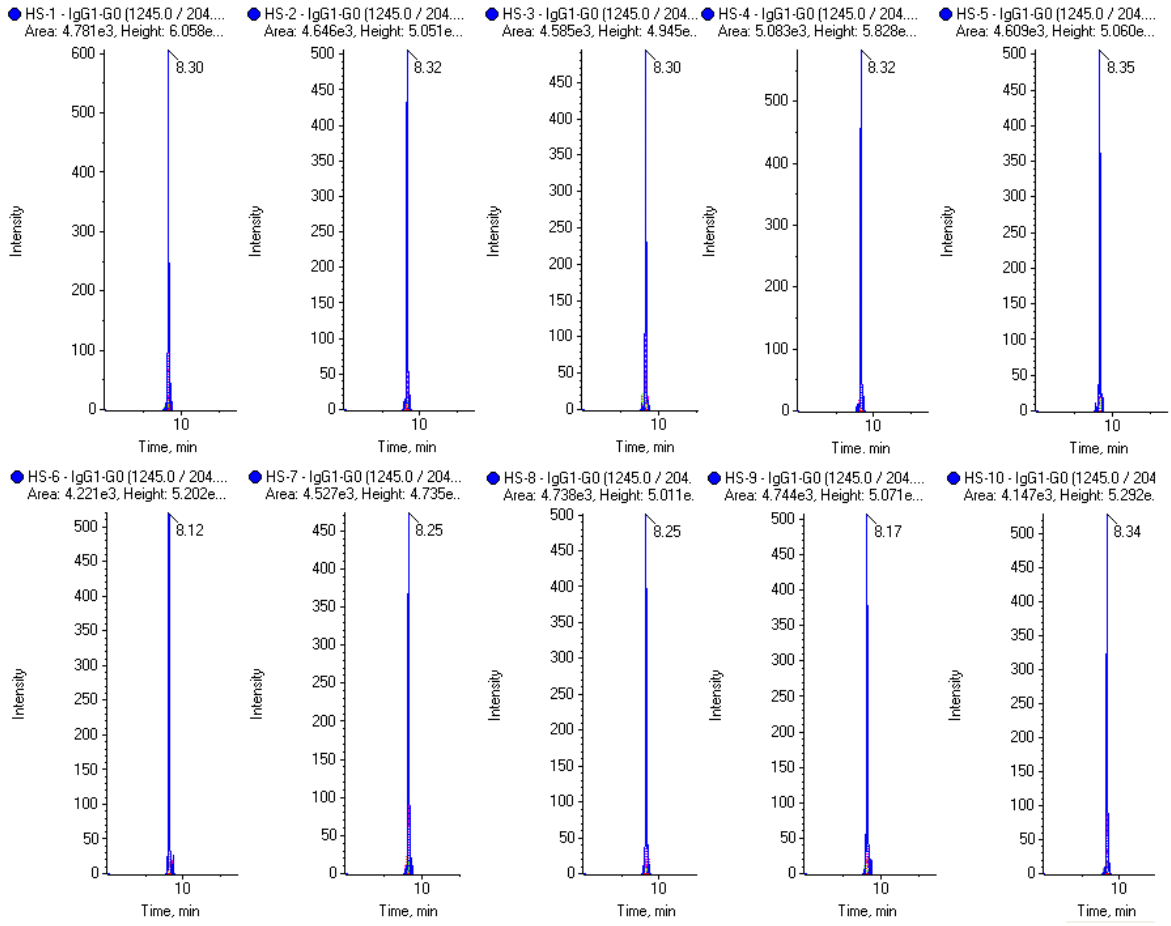


Figure S4-8. IgG1-G2 glycopeptide quantification analyzed with transition 1407.2/204.1 in positive polarity with equation of the line, y=7.49364x-634.28300 (r^2=0.99317) under MRM. Shown below are 10 representative spectra from n=20.

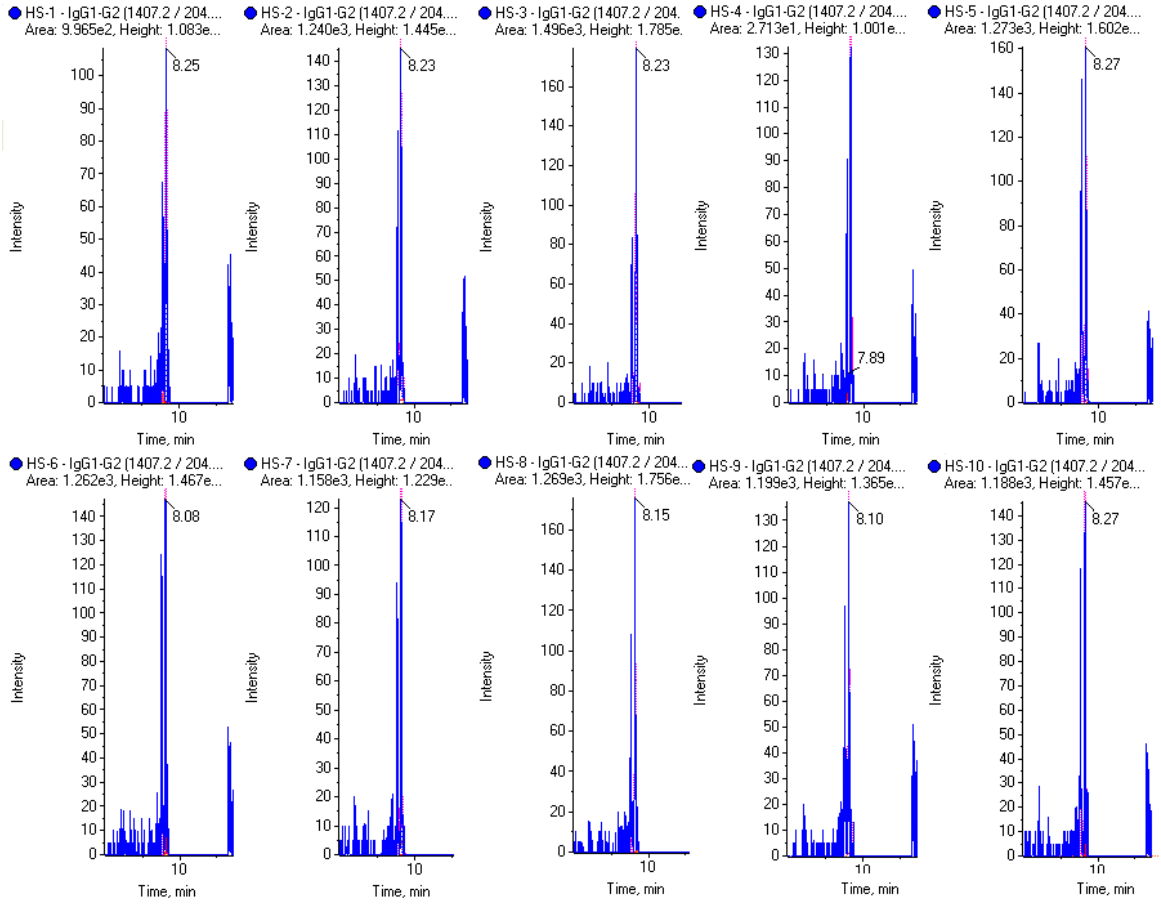


Figure S4-9. IgG1-G2S2 glycopeptide quantification analyzed with transition 1132.2/204.0 in positive polarity with equation of the line, y=1.77922x-88.12468 (r^2=0.99730) under MRM. Shown below are 10 representative spectra from n=20.

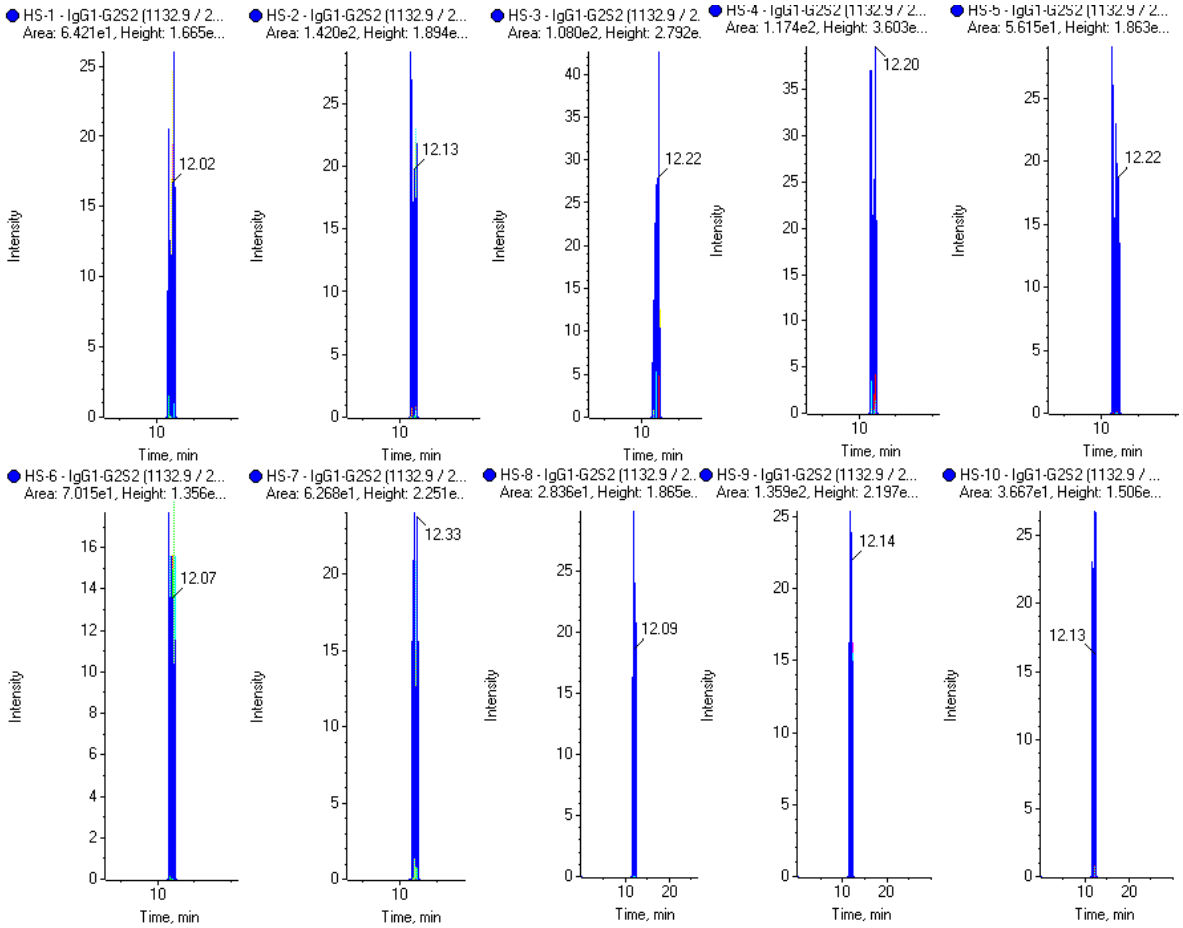


Figure S4-10. IgG2-G0 glycopeptide quantification analyzed with transition 1229.0/204.1 in positive polarity with equation of the line, y=118.47528x-4951.12629 (r^2=0.99928) under MRM. Shown below are 10 representative spectra from n=20.

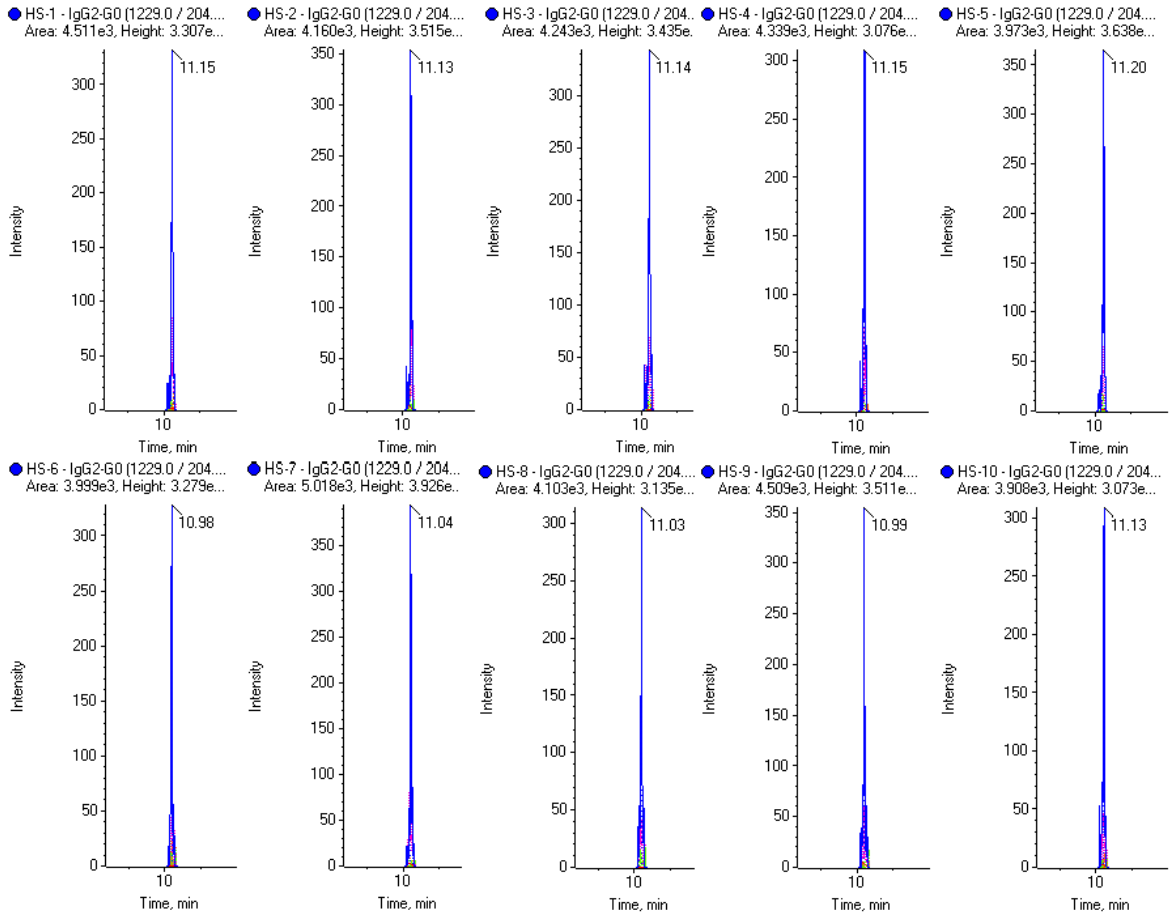


Figure S4-11. IgG2-G2 glycopeptide quantification analyzed with transition 1390.7/204.1 in positive polarity with equation of the line, y=19.15928x-624.74676 (r^2=0.99994) under MRM. Shown below are 10 representative spectra from n=20.

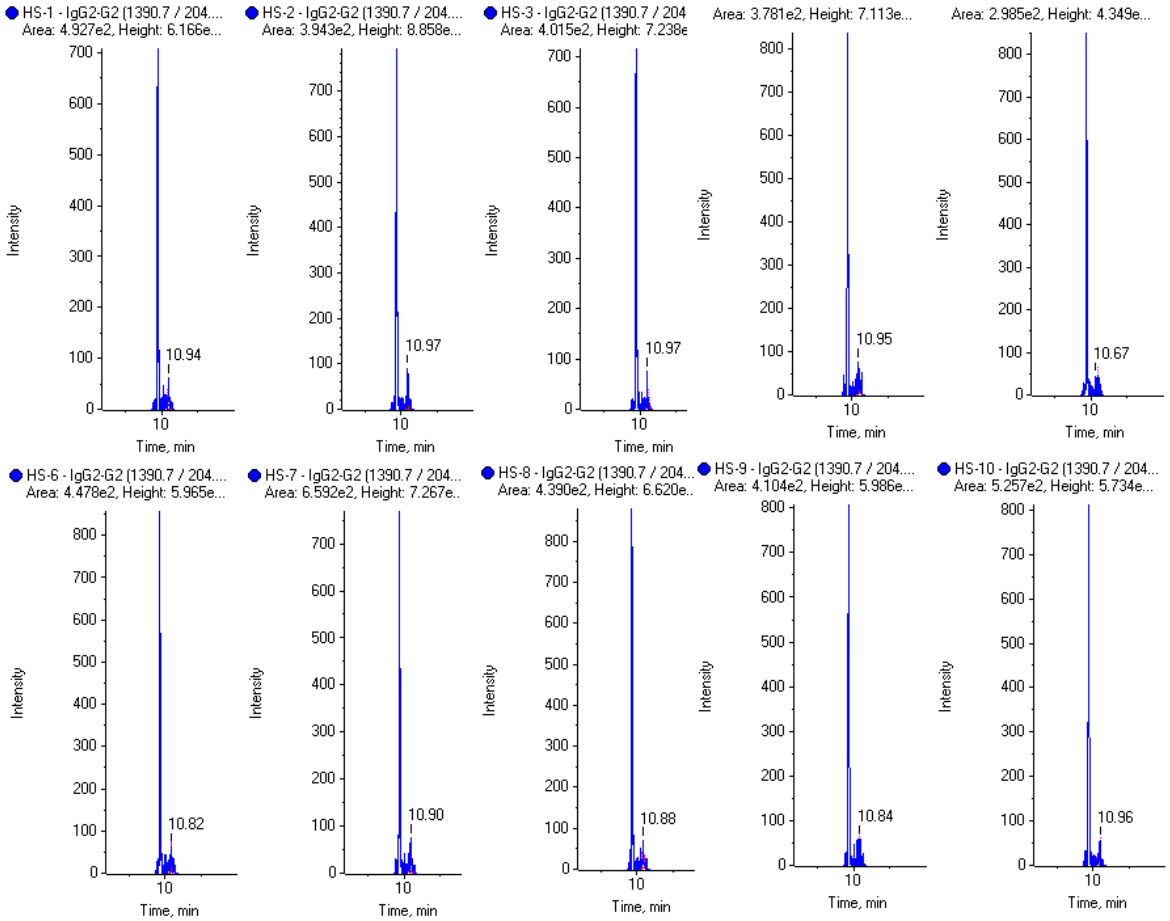


Figure S4-12. IgG2-G2S2 glycopeptide quantification analyzed with transition 1121.9/204.1 in positive polarity with equation of the line, y=157.20091x-4658.29151 (r^2=0.99974) under MRM. Shown below are 10 representative spectra from n=20.

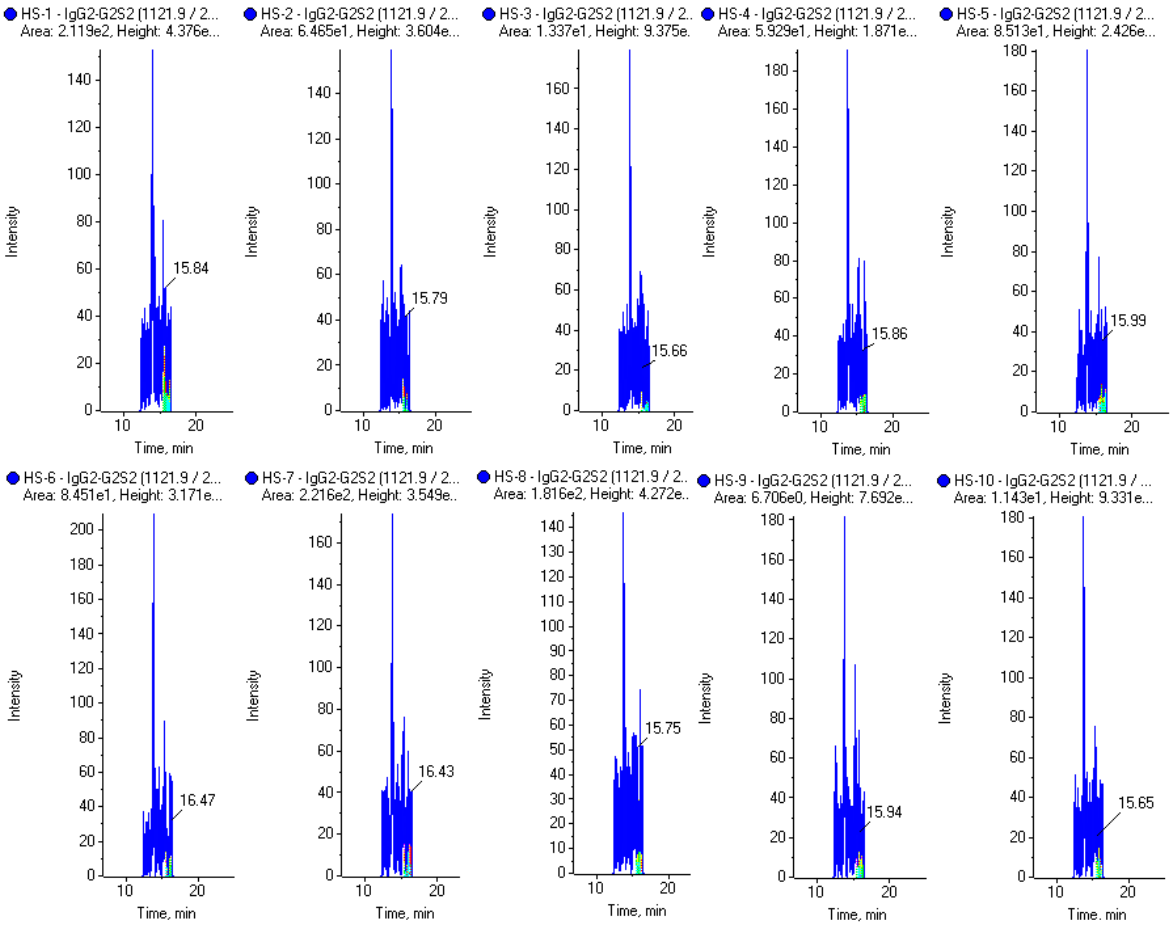
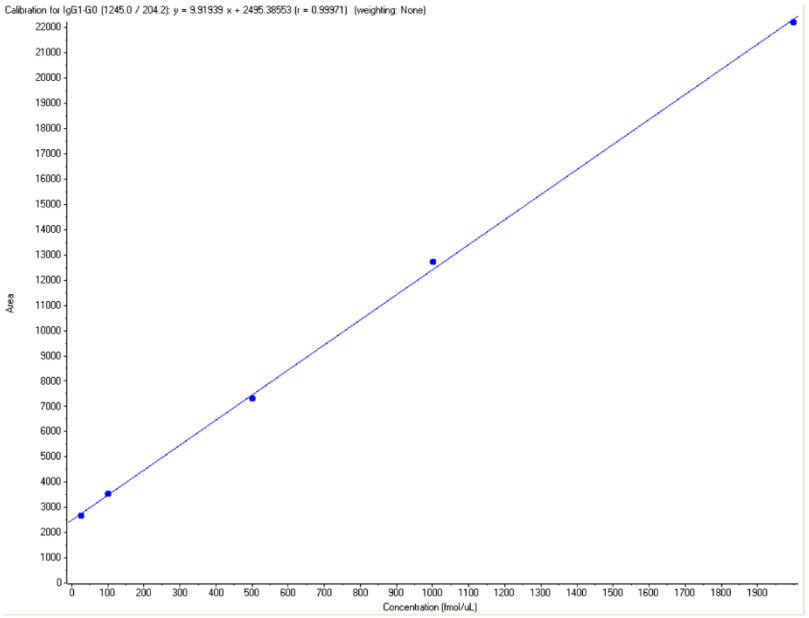


Figure S4-13. Spiking of IgG1-G0 tryptic glycopeptide in human serum analyzed with transition 1245.0/204.2with equation of the line, y=9.91939x+ 2495.38553 (r^2=0.99942) in positive polarity under Scheduled MRM. A. Calibration curve. B. Elution time. The standards added are 25, 100, 500, 1000, and 2000 fmol/µL. The bottom spectra in B are the expanded versions of the top spectra.

A.



B.

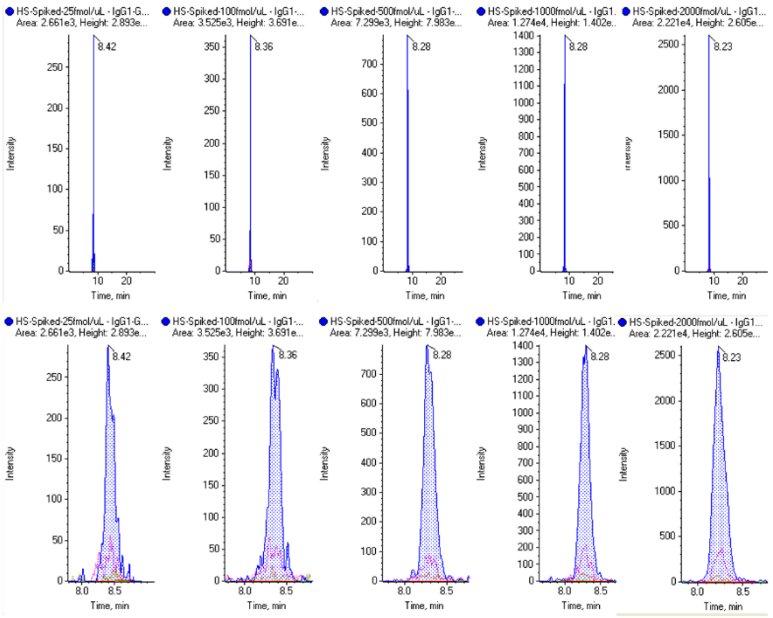
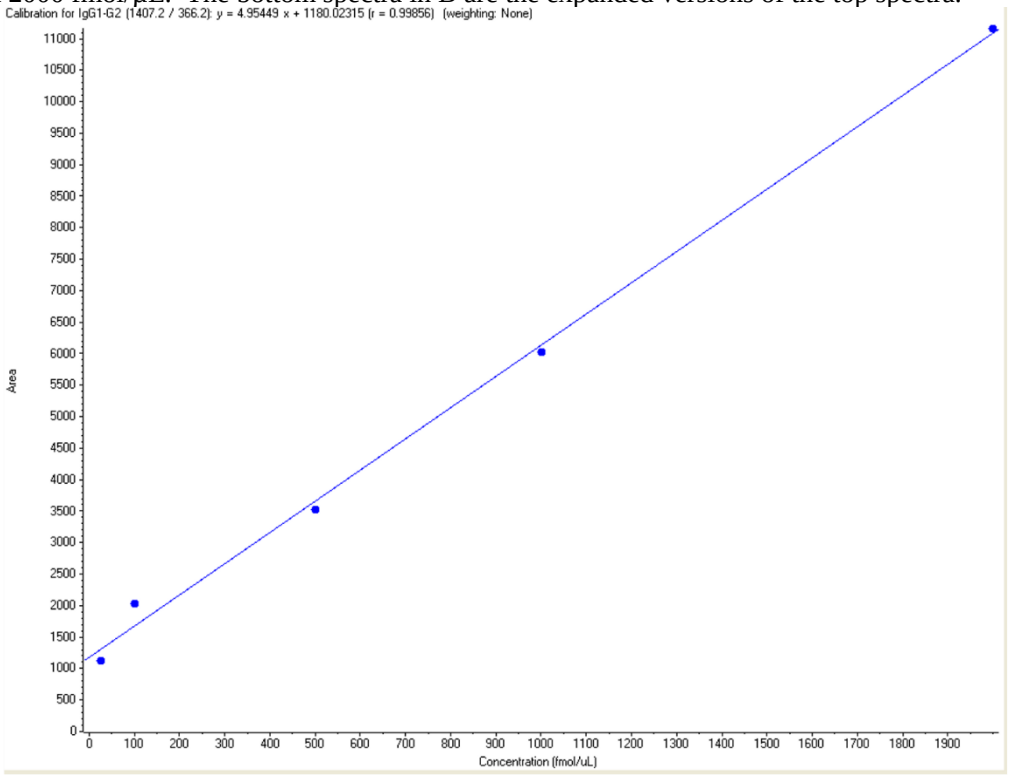


Figure S4-14. Spiking of IgG1-G2 tryptic glycopeptide in human serum analyzed with transition 1407.2/204.1 with equation of the line, y=3.87027x+ 750.42853 (r^2=0.99870) in positive polarity under Scheduled MRM. A. Calibration curve. B. Elution time. The standards added are 25, 100, 500, 1000, and 2000 fmol/µL. The bottom spectra in B are the expanded versions of the top spectra.

A.



B.

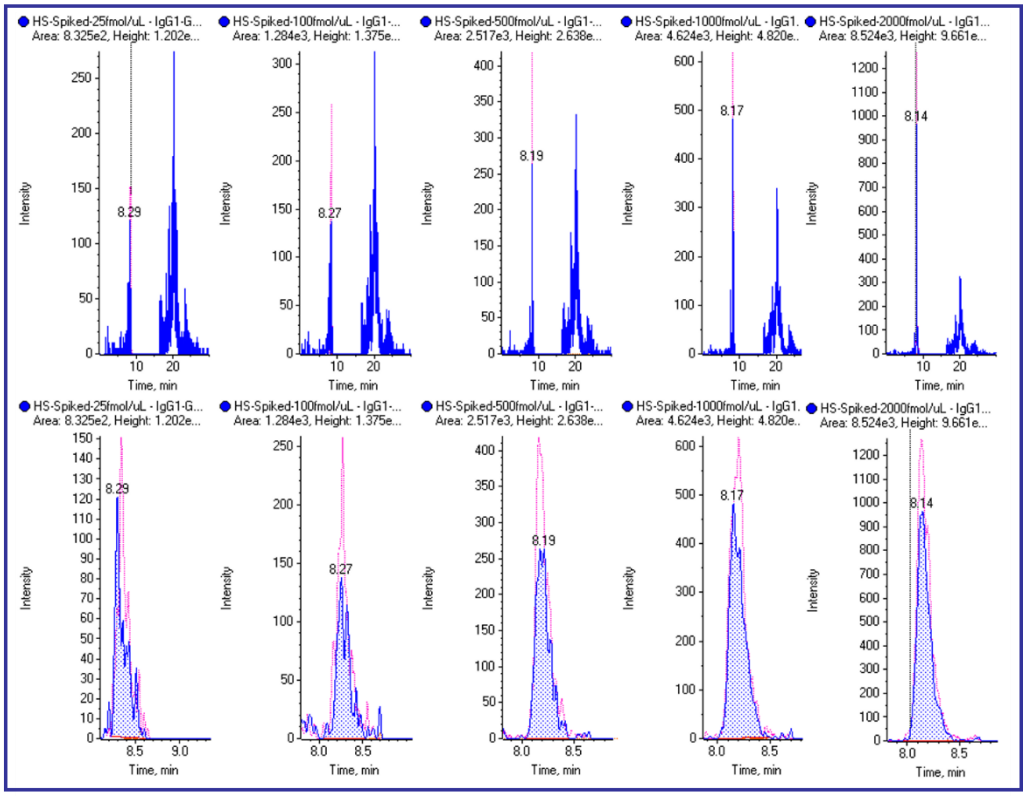
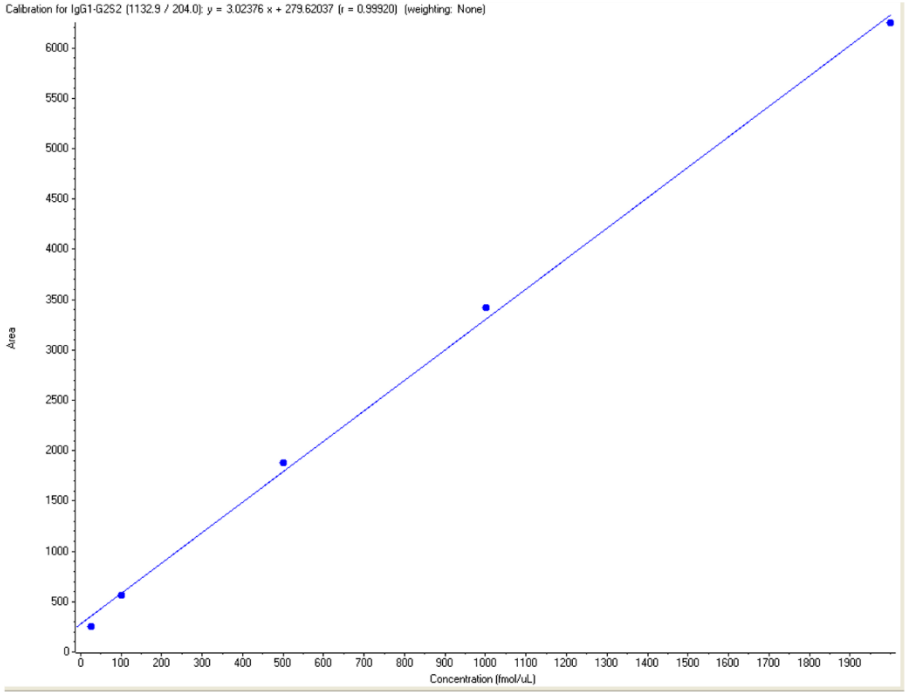


Figure S4-15. Spiking of IgG1-G2S2 tryptic glycopeptide in human serum analyzed with transition 1132.9/204.0 with equation of the line, y=3.02376x+ 279.62037 (r^2=0.99840) in positive polarity under Scheduled MRM. A. Calibration curve. B. Elution time. The standards added are 25, 100, 500, 1000, and 2000 fmol/µL. The bottom spectra in B are the expanded versions of the top spectra.

A.



B.

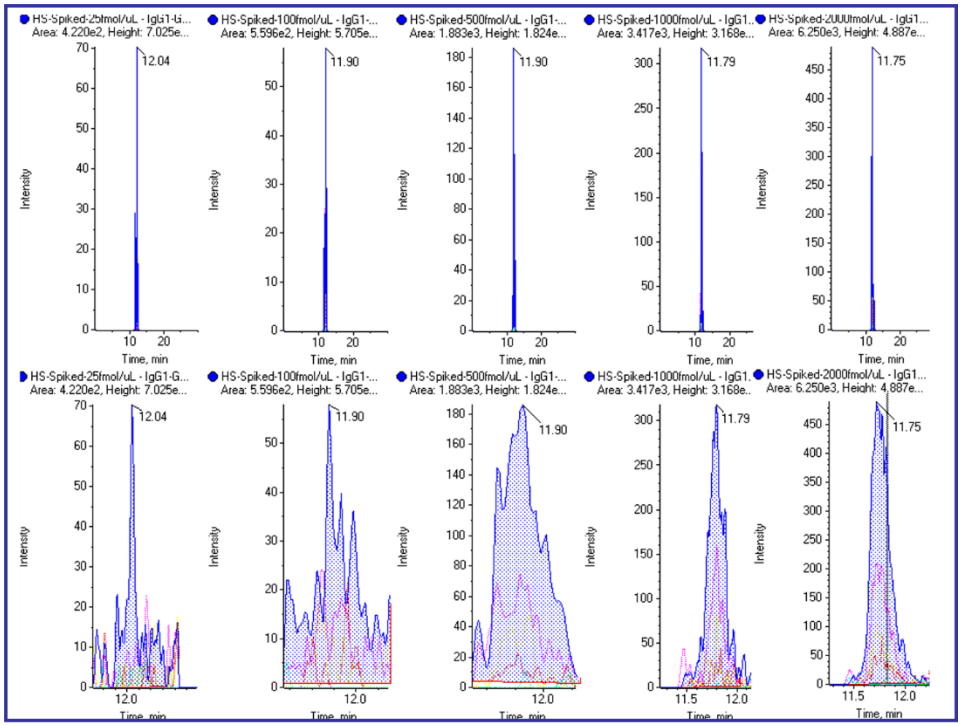
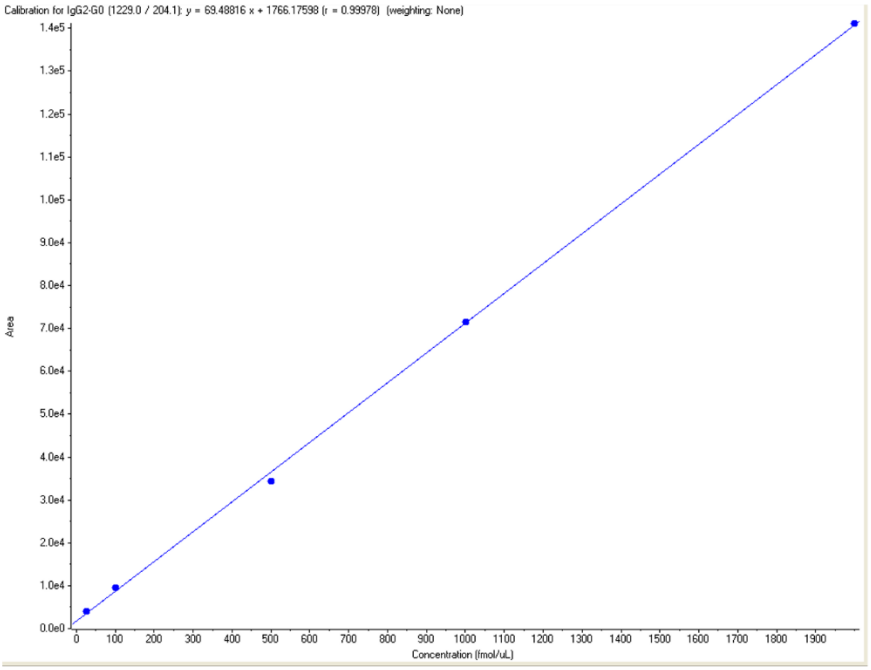


Figure S4-16. Spiking of IgG2-G0 tryptic glycopeptide in human serum analyzed with transition 1229.0/204.1 with equation of the line, y=69.48816x+ 1766.17598 (r^2=0.99956) in positive polarity under Scheduled MRM. A. Calibration curve. B. Elution time. The standards added are 25, 100, 500, 1000, and 2000 fmol/µL. The bottom spectra in B are the expanded versions of the top spectra.

A.



B.

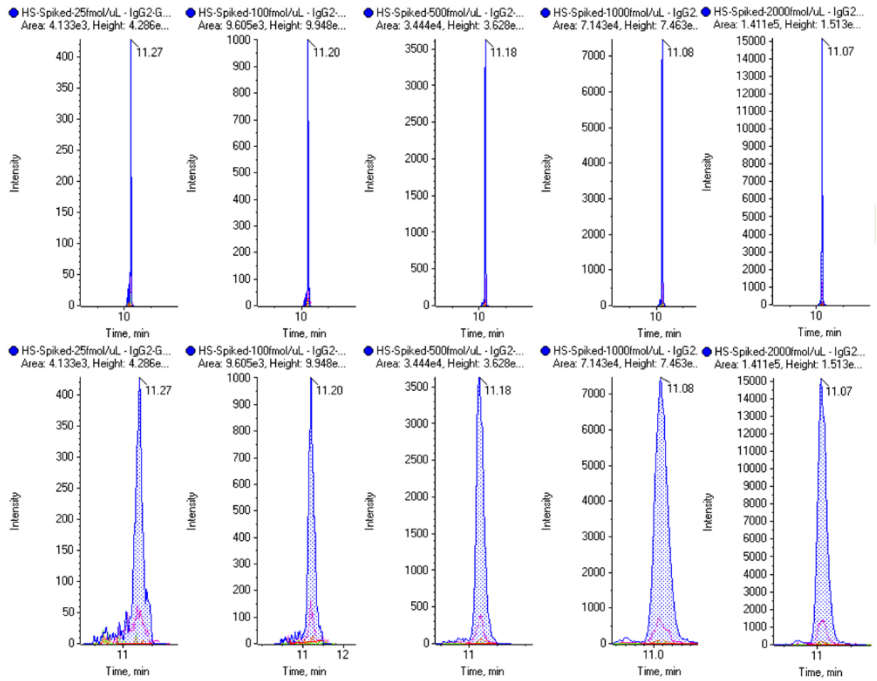
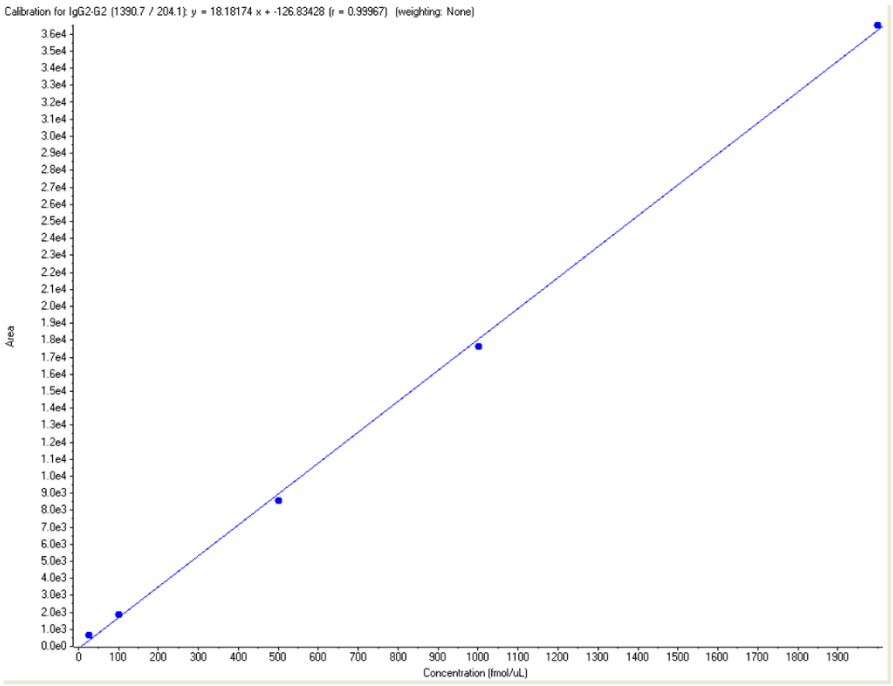


Figure S4-17. Spiking of IgG2-G2 tryptic glycopeptide in human serum analyzed with transition 1390.7/204.1 with equation of the line, y=18.18174x- 126.83428 (r^2=0.99934) in positive polarity under Scheduled MRM. A. Calibration curve. B. Elution time. The standards added are 25, 100, 500, 1000, and 2000 fmol/µL. The bottom spectra in B are the expanded versions of the top spectra.

A.



B.

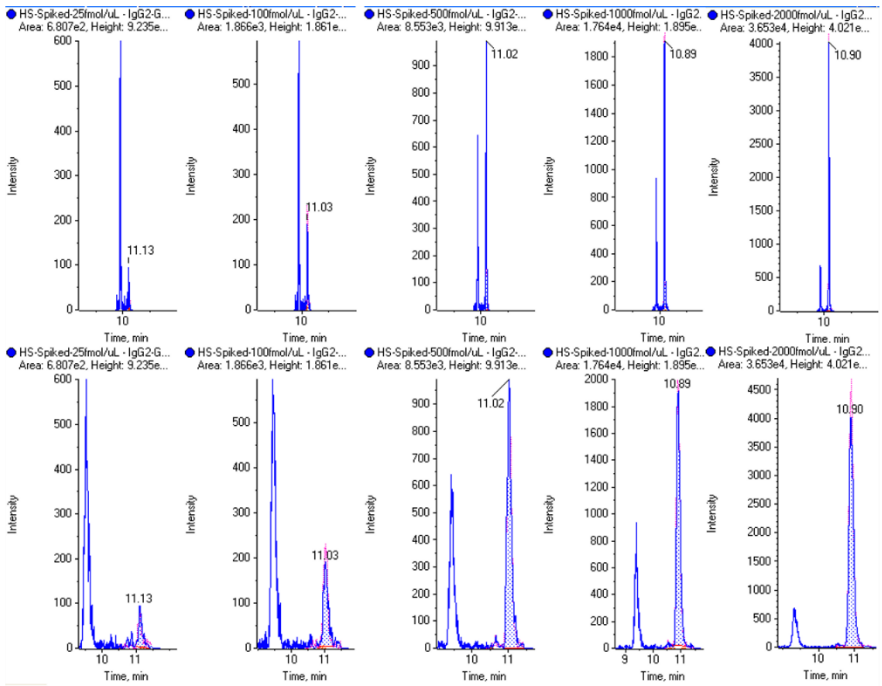


Figure S4-18. Spiking of IgG2-G2S2 tryptic glycopeptide in human serum analyzed with transition 1121.9/204.1 with equation of the line, y=121.94474x+ 461.94449 (r^2=0.99946) in positive polarity under Scheduled MRM. The standards added are 25, 100, 500, 1000, and 2000 fmol/µL.

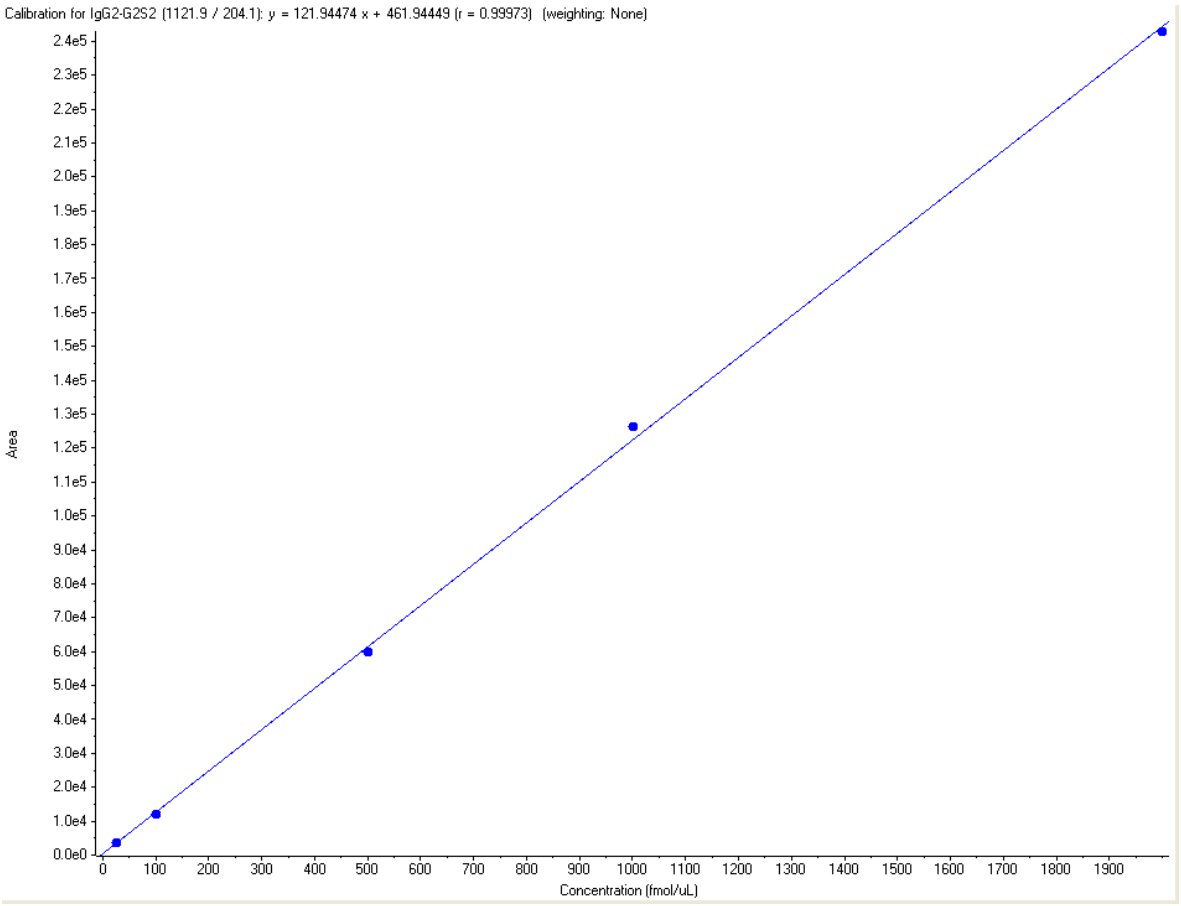


Table S4-4. Calibration standard for sialic acid-containing IgG glycoforms analyzed in negative polarity mode. The calibration standards used were 0, 25, 50, 100, 200, 500, 1000, 2000, and 5000 fmol/ µL.

Glycopeptide	*Transition*	*Equation of the Line*	*r^2*
IgG1-G2S2	1130.6/290.0	y=8.46638x-379.29915	0.99270
IgG2-G2S2	1119.8/290.2	y=37.06224x-2199.87043	0.99654

Figure S4-19. Calibration of IgG1-G2S2 glycopeptide standard analyzed with transition 1130.6/290.0 in negative polarity with equation of the line, y=8.46638x-379.29915 (r^2=0.99270) under MRM. The standards added are 0, 25, 50, 100, 200, 500, 1000, 2000 and 5000 fmol/µL dissolved in water.

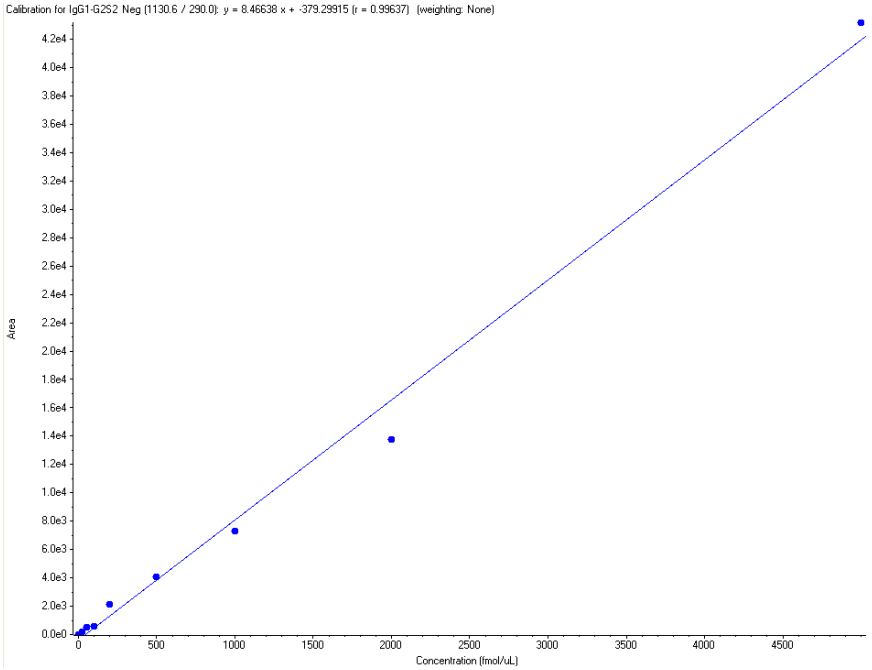


Figure S4-20. Calibration of IgG2-G2S2 glycopeptide standard analyzed with transition 1119.8/290.2 in negative polarity with equation of the line, y=37.06224x-2199.87043 (r^2=0.99654) under MRM. The standards added are 0, 25, 50, 100, 200, 500, 1000, 2000 and 5000 fmol/µL dissolved in water.

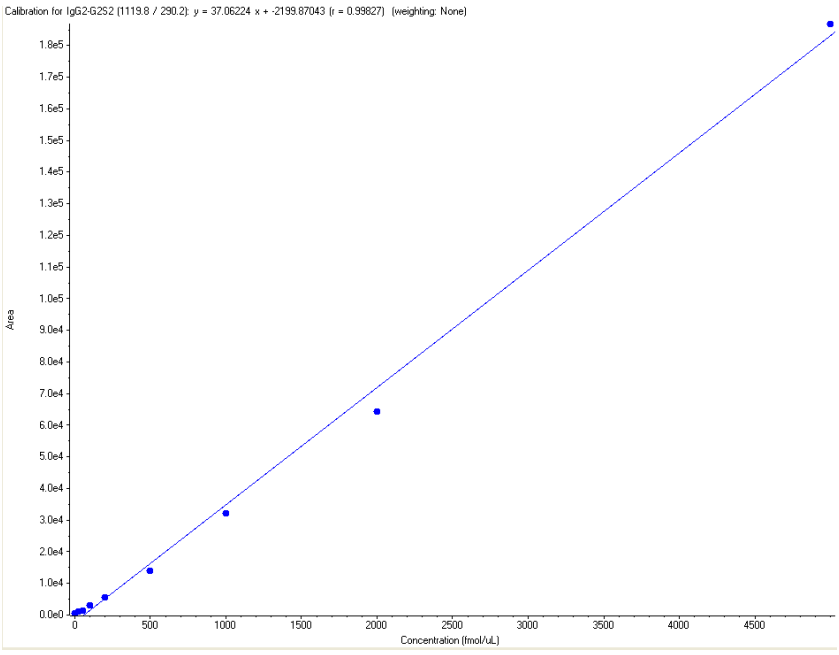


Figure S4-21. Comparison of the presence of IgG Glycopeptides in human serum samples taken from Stage 4a and Stage 4b pancreatic cancer patients. Concentrations are expressed in fmol/µL. The middle line for each group is expressed as mean of the results while the outer lines depicts standard deviation. All Stage 4a and Stage 4b IgG glycopeptides tested does not show any significant difference.

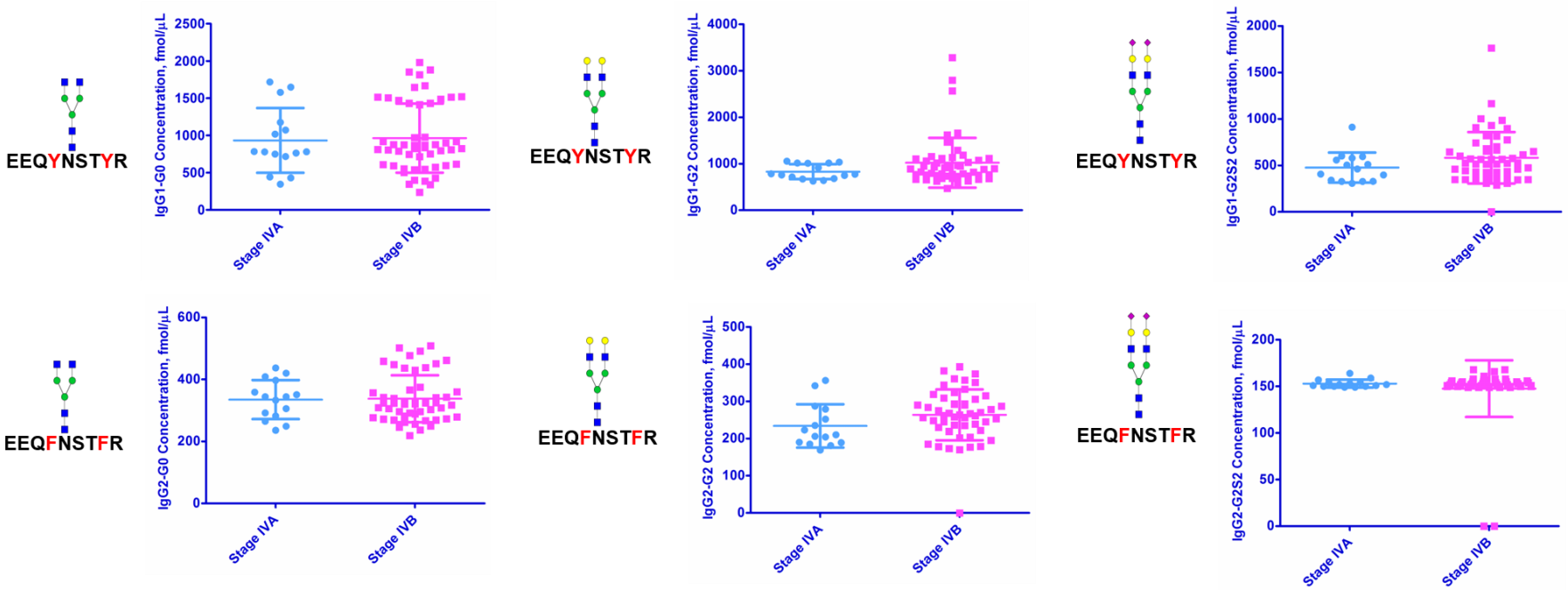
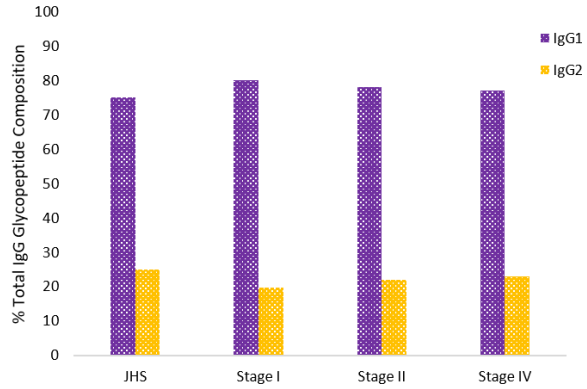


Figure S4-22. Total Glycopeptides Present in Non-cancer and Pancreatic Cancer Japanese Human Serum Samples. A. Tally of glycopeptides for each IgG Class. B. Histogram for total IgG glycopeptides at each state. C. Pie-chart composition IgG glycopeptides at each state. JSH; Japanese non-cancer human serum. Concentrations are expressed in fmol/µL. When total glycopeptides were compared at each state, not a lot of difference is seen. As such, individual glycopeptide assessment better gives a clears difference.

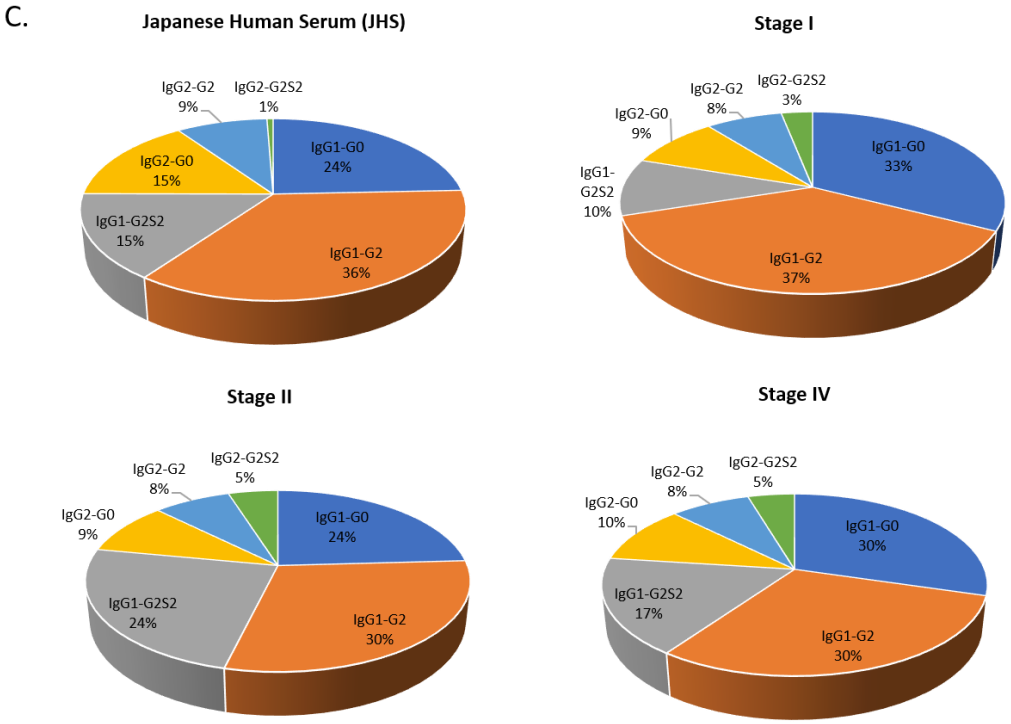
A.

	IgG1	IgG2
JHS	75	25
Stage I	80	20
Stage II	78	22
Stage IV	77	23

B.

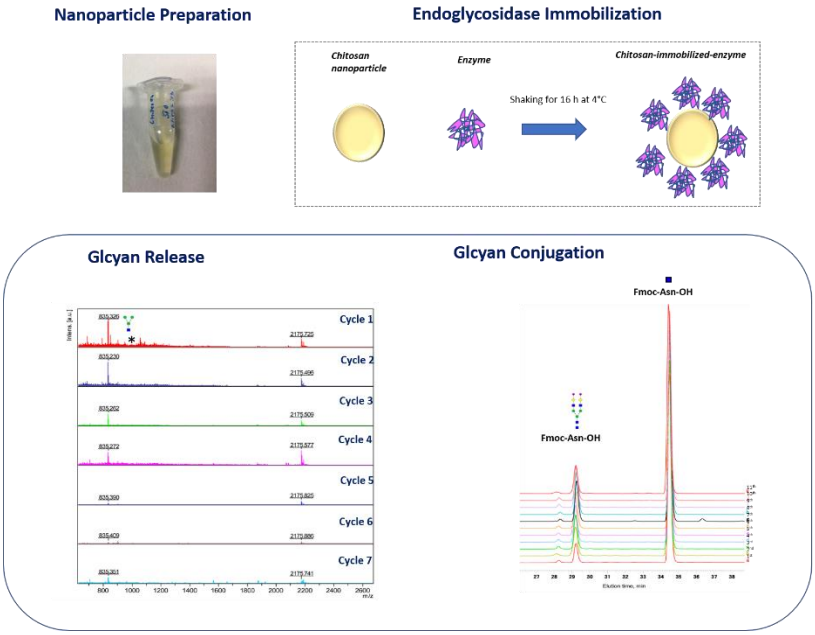


C.



Chapter 5

Endoglycosidase Immobilization: The Use of Nanoparticle-Chitosan Beads for Glycan Release and Transglycosylation Reactions



A version of this work will be submitted for publication

5.1 Chapter Summary

Endoglycosidases (Endos) have been important enzymes for releasing different glycan structures and attaching sugars to different glycoconjugate acceptors. Without protein expression systems, Endos are only available commercially; with which limits usage and are denatured right after use. As such, strategies for reusing these precious enzymes in different applications is needed. Chitosan nanoparticle is known as an effective enzyme immobilization technique; but, was never applied to Endos. In this work, commercial Endos are bound onto a chitosan-nanoparticle immobilization support which shows reusability for glycan release and transglycosylation reactions assessed through glycoblotting methodology and HPLC separation, respectively. The procedure presented herein provides for an economically sound and viable immobilization technique for Endos.

Keywords: *Enzyme immobilization, chitosan, $Man_3GlcNAc$, Endoglycosidases, Endo-D, Endo-M, transglycosylation, glycan release, chemoenzymatic, glycopeptide synthesis*

5.2 Introduction

Endoglycosidases (Endos) are enzymes that cleave in between the chitobiose core structure of *N*-glycans attached to an asparagine (Asn) amino acid.[1-3] Endos show a dual-property of hydrolyzing glycans, and at the same time, is able to catalyze the attachment of truncated glycans to a GlcNAc residue of a glycosyl acceptor.[4-7] The latter termed as a transglycosylation reaction.[7] A wide array of Endos, each specific to a different glycan-type, has been isolated and mutated. With this, a glycan or glycoconjugate structure can be obtained by using an appropriate endoglycosidase. The advent of glycobiology require an abundant supply of glycan structures and glycoproteins for structure-specific activity studies and as quantitative calibration standards. Commercial glycosidases are expensive and obtaining valuable amount of glycans or glycopeptides would mean an abundant supply of Endos. Thus, a supplement to protein-expression systems and the development of enzyme-immobilization procedure for the reusability of Endos is necessary.

Numerous enzyme immobilization strategies have been developed.[8-10] A simple yet effective strategy for immobilization is attaching enzymes onto chitosan nanoparticles.[11] Chitosan-based enzyme immobilization techniques have been conducted previously for other enzymes[11], e.g., trypsin[12,13] and galactosidases[14-17], but not Endos for glycoconjugate applications. In this work, the applicability of chitosan-immobilized Endos on glycan release and transglycosylation reaction is demonstrated. Highlighted also in this work, is the use of the glycoblotting methodology to assess for glycan release, immobilization efficiency, and reusability of chitosan-immobilized Endos. HPLC on the other hand was used to separate transglycosylated products from starting glycosyl acceptors. Looking forward, using chitosan-immobilized Endos can be applied in either lab or large-scale chemo-enzymatic glycoconjugate preparations.

5.3 Materials and Methods

5.3.1 Chitosan Nanoparticle Preparation

Chitosan nanoparticles were prepared by modification of previous protocols[18,19] prepared for a 1-mL scale. A 950 µL of 0.25% Chitosan-500 (Wako Chemicals) dissolved in 2% aq. acetic acid containing 1% Tween-20 was prepared. The solution was sonicated while 50 µL 20% sodium sulfate solution was added drop-by-drop. The mixture was further sonicated for 2h and was centrifuged at 5 RCF for 10 mins at 20 ° C. The supernatant was removed, and the precipitate washed twice with water. Chitosan nanoparticles were resuspended in 1 mL 0.05 M PBS (pH 7.4). Addition of 100 µL 10% sodium phosphate in water followed while the nanoparticle solution was constantly sonicated. Crosslinking with 50 µL 25% glutaraldehyde followed and the mixture was stirred for 30 min. Excess glutaraldehyde was washed with 100 µL water and centrifuged under 5 RCF for 10 mins at 20 °C, then the water removed. Cross-linked particles were resuspended in 1 mL water.

5.3.2 Endogoglycosidase- Immobilization

5.3.2.1 Chitosan-immobilization

300 µL of the nanoparticle suspension from the previous step was taken and centrifuged in 10,000 RPM for 10 mins at 20 °C. The solvent was removed, and the nanoparticles were added with the endoglycosidase of interest. For Endo-D (New England Biolabs, Inc) immobilization, 2, 5, 10, and 20 U of 1U/µL were added for dose-dependent enzyme immobilization. For Endo-M-(Tokyo Chemical Industry Co. Ltd) immobilization, 2 mU (2 µL of 1mU/µL) was added. A total volume of 30 µL containing the enzyme and PBS buffer (0.05 M pH 7.4 for Endo-D, and 60 mM pH 6.4 for Endo-M) was added and vortexed continuously for 16 h at 4 ° C. The supernatant was removed after centrifugation in 10,000 RPM for 10 mins at 20 ° C. An additional 50 µL of water was added to wash unimmobilized enzymes, repeated twice, and the 3 washings lyophilized separately. Immobilization efficiency was analyzed indirectly by checking whether the supernatant contains unimmobilized

Endo-D that can release $Man_3GlcNAc$ glycan. The immobilized endoglycosidase on chitosan was then used for glycan release (Endo-D), or for transglycosylation reaction (Endo-M).

5.3.3 Endoglycosidase-D(Endo-D) Releases N-Glycans from Japanese Quail Egg White

The procedure is similar to Chapter 3. [20] Briefly, 0.5 mg of lyophilized Japanese quail egg white was added with 20 µL 200 mM ammonium bicarbonate, followed by the addition of a 26 µL 100 µM disialyloctasaccharide (Tokyo Chemical Industry Co., LTD.) internal standard and mixed well. A 54 µL mixture of 0.06% 1-propanesulfonic acid (PHM), 2-hydroxyl-3-myristamido with 12 mM dithiothreitol (DTT) in 105 ammonium bicarbonate (ABC) was added and the solution incubated at 60° C for 90 min. Addition of 10 µL 123 mM iodoacetamide (IAA) followed and then the solution incubated in the dark for 1 h at room temperature. Addition of 10 µL of 40 U/µL Trypsin (Sigma-Aldrich) was performed and the solution incubated for 24 h at 37° C. The trypsin was then heat-deactivated for 10 min at 90° C.

Glycan-release was performed under a free or a chitosan-immobilized Endo-D system. For comparison, 2U Endoglycosidase-D (Endo-D) (New England Biolabs) in either free or immobilized form was added and incubated at 37° C for 24 h. For dose-dependent comparison on glycan release in egg white, 2U, 5U, 10U, and 20U-Endo-D immobilized in chitosan nanoparticles were added. Glycan-released samples were lyophilized. For $Man_3GlcNAc$ release from JQEW using immobilized Endo-M, all the reagents and egg white were first mixed in a separate Eppendorf tube and the trypsin-heat deactivated before adding the solution mixture onto the chitosan nanoparticle. To quantitate the release of $Man_3GlcNAc$ in egg whites, the glycoblotting methodology was conducted. Same methodology was used to assess immobilization efficiency and reusability of chitosan-immobilized Endo-D for $Man_3GlcNAc$ cleavage. For the reusability assay, the chitosan nanoparticle was recovered through centrifugation at 5 RCF for 10 mins at 20 ° C separating it from the solution mixture. The nanoparticle was then washed with water and pooled with the solution mixture. The chitosan nanoparticle was re-dispersed before adding 0.5 mg JQEW dissolved in initial reagents following the glycan release methodology.

5.3.3.1 Glycoblotting Methodology for $Man_3GlcNAc$ Quantitation from Endo-D Catalyzed Japanese Quail Egg White

The procedure was similar to previous reports on egg white *N*-glycan analysis. [20,21, 22,] 500 µL BlotGlycoH™ bead suspension (10 mg/mL) were pipetted into 96-well multiScreen Solvinert filter plate and securely attached in a vacuum-manifold. The water was removed. From the lyophilized sample in the previous section, 20 µL MilliQ water and 180 µL of 2% acetic acid/acetonitrile were added to reconstitute the sample. The Solvinert plate was then incubated at 80° C for 45 min or until dry. A series of two-200 µL washings was performed sequentially using 2M guanidine-HCl in 16.6 mM ammonium bicarbonate, MilliQ water, and 1% triethylamine in methanol followed. The unreacted hydrazine beads were then added with 100 µL 10 % acetic anhydride in methanol, incubated at RT for 30 minutes, and aspirated. Consecutive addition of 200 µL of 10 mM HCl, methanol, dioxane was then conducted. Washing with each solvent was performed twice. Addition of 100 µL of 100 mM 3-methyl-1-*p*-tolyltriazene (MTT) in dioxane was then conducted and incubated at 60° C until dry. Washing twice with 200 µL of dioxane, water, methanol, and water was then conducted sequentially. A 20 µL of 50 mM O-benzyloxyamine hydrochloride with 180 µL of 2 % acetic acid/acetonitrile was then added to label the glycans. The Solvinert plate was then placed in an incubator at 80 °C until dry. After incubation, addition of 100 µL of MilliQ water was performed to elute the labeled *N*-glycans and dried in a SpeedVac. Labeled-glycans were reconstituted with 20 µL of MilliQ water and subjected to Mass Spectrometry Analysis.

5.3.4 Endoglycosidase-M Transglycosylation Reaction

5.3.4.1 Fmoc-Asn (GlcNAc)-OH Glycosyl Acceptor Synthesis

Fmoc-Asn(GlcNAc)-OH was prepared similarly to Chapter 3 [20] using an N-α-Fmoc-N-γ-(2-acetamido-2-deoxy-3,4,6-tri-O-acetyl-β-D-glucopyranosyl)-L-asparagine (Fmoc-Asn(Ac_3β-GlcNAc)-OH) (Medicinal Chemistry Pharmaceuticals, Japan). The Fmoc-Asn(GlcNAc)-OH served as the GlcNAc-glycosyl acceptor for the transglycosylation reaction conducted using a free and chitosan-immobilized Endo-M.

5.3.4.2 Transglycosylation Reaction with Endoglycosidase-M (Endo-M)

The procedure involved a mixture containing 5 µL 10 mM IgG1-GlcNAc tryptic peptide, 5 µL 50 mM SGP (Tokyo Chemical Industry Co. Ltd), 18 µL 60 mM potassium phosphate buffer (pH 6.4), and 2 µL of 1 mU/µL free Endo-M or 2 mU chitosan-immobilized Endo-M in an Eppendorf tube and the solution incubated at 37° C with constant shaking. Aliquots of 1 µL were taken at different time intervals and added to a 99 µL water. Aliquot samples were taken after 1h, 2h, 6h, and 24 h for free Endo-M incubation while after 1h, 2h, 4h, and 6h for chitosan-immobilized-Endo M were taken. Analytical HPLC separation was then conducted with solvents, A. water B. acetonitrile, both containing 0.1% TFA under a gradient of 0%B (0 min)→60%B (50min)→95%B (50.1 min)→95%B (60min) in an Inertsil-ODS 3 column (4.5x250 mm) set at a 265 nm wavelength. The percentage transglycosylation yield was based on the HPLC peak areas of the transglycosylated product, Fmoc-Asn($Hex_5HexNAc_4NeuAc_2$)-OH, as compared to the total peak areas coming from both Fmoc-Asn($Hex_5HexNAc_4NeuAc_2$)-OH and Fmoc-Asn(GlcNAc)-OH. The % transglycosylation yield using the free and chitosan-immobilized Endo-M were compared. The eluted fractions(~30min) containing the transglycosylated product were taken and checked in MALDI-TOF/MS for confirmation.

Basing on the results of the transglycosylation yield monitoring, the applicability of reusing the immobilized-Endo-M was performed. Reusability was monitored after an hour of incubation at 37ºC conducted at different days (1^{st}, 3^{rd}, 10^{th}, and 15^{th} day) after enzyme immobilization to monitor changes in transglycosylation reaction activity. The chitosan-immobilized Endo was placed in 4ºC when it is not incubated with the transglycosylation reactants. There is a difference between aliquot acquisition for the free and immobilized enzyme. For the chitosan immobilized Endo-M, centrifugation was necessary to separate the chitosan from the solution mixture, and then 1 µL aliquot was taken. The solution mixture and the chitosan nanoparticle were re-dispersed and incubated and taken out again for the next time interval. For the free-Endo-M, the solution was just vortexed, aliquot was taken, and afterwards, the tube was placed back in the shaking incubator.

5.3.5 Mass Spectrometry Analysis

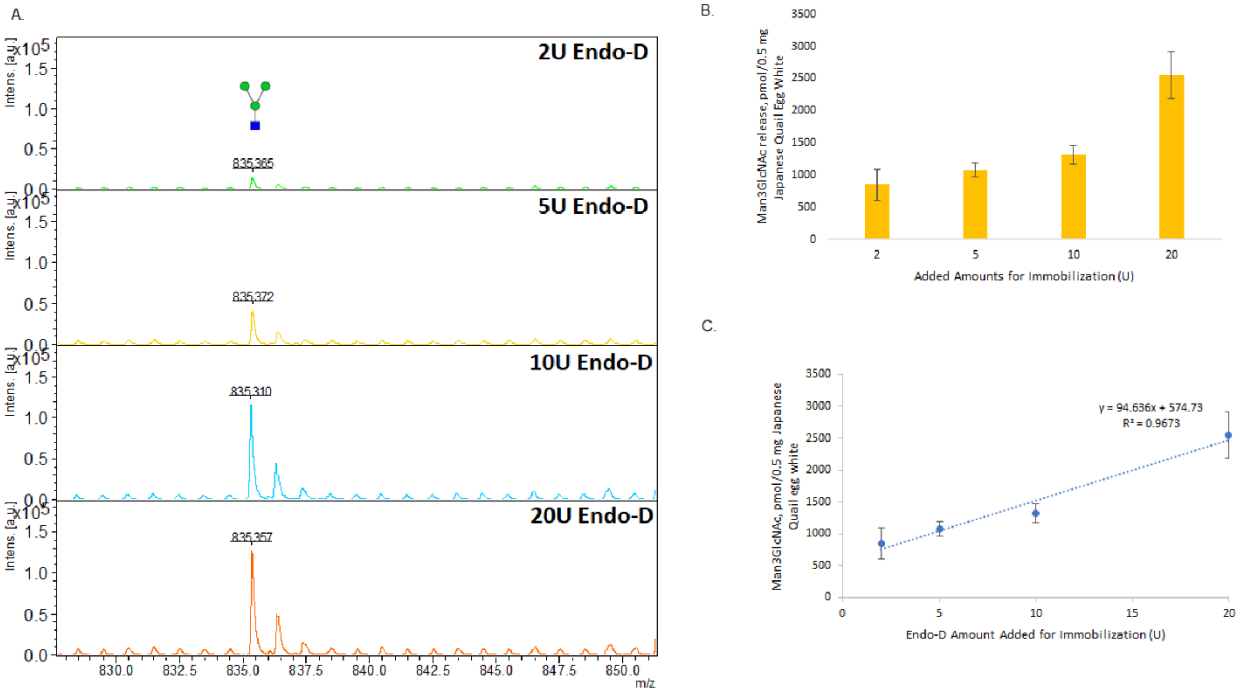
The matrix consists of 20 mg/mL 2,5-dihydroxybenzoic acid and 1 mM NaCl dissolved in 30% acetonitrile containing 0.1% trifluoroacetic acid (TFA) solution. A 1 µL of the prepared matrix was spotted, dried, followed by spotting of the reconstituted sample into an Anchorchip MTP 384 Target Plate (polished steel TF, Bruker), and then dried. *FlexAnalysis* 3.0 software (Bruker Daltonics) was used to obtain spectra of experimental masses. MALDI-TOF and MALDI-TOF/TOF parameters using an Ultraflex III (Bruker Daltonics) are the same as that of Chapter 3. [20] Samples in MS were analyzed in reflector, positive ion mode, typically totaling 2000 shots with settings of acceleration voltage, reflector voltage, and pulsed ion extraction at 25.3 kV, 26.4 kV, and 100 ns, respectively. In MS/MS Mode, chosen parent peaks were initially accelerated to 8 kV and further accelerated to 20.1 kV. *N*-glycan structures were prepared by using GlycoWorkBench.[23]

5.4 Results and Discussion

5.4.1 Enzyme Immobilization for Glycan Cleavage Purposes using Endo-D

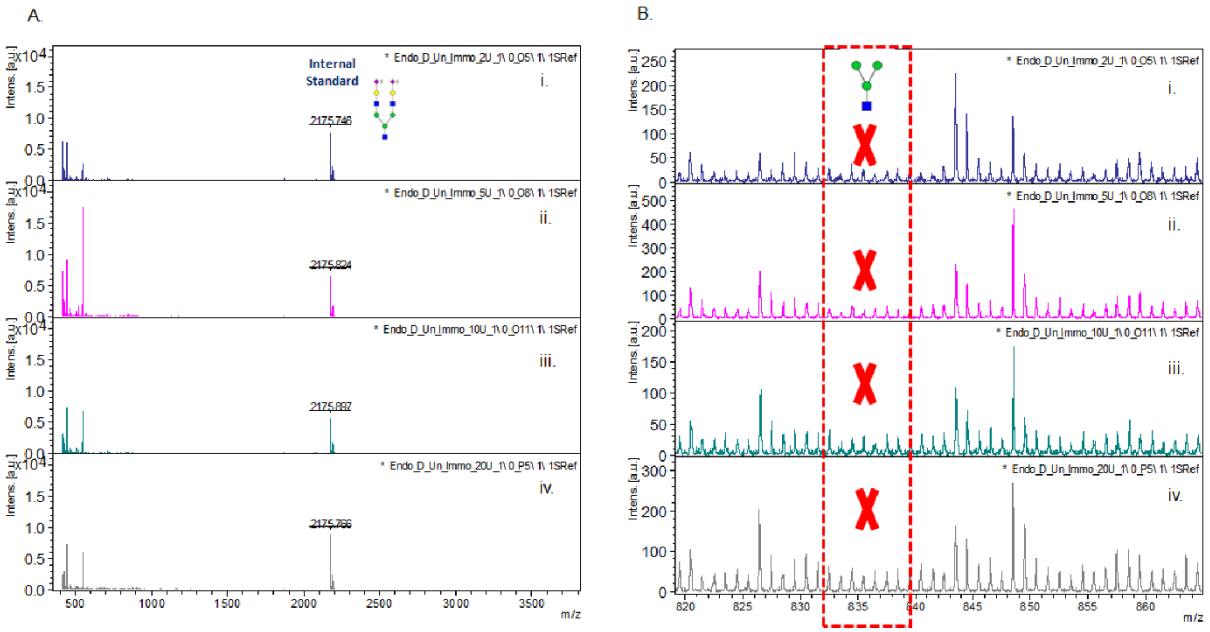
Comparison of free and immobilized-enzyme toward glycan release was performed. The results showed no significant difference on $Man_3GlcNAc$ amount at 99.9% confidence interval with 849±241 pmol and 1095±196 pmol for the immobilized and free enzyme, respectively. Enzyme immobilization at different concentrations were also conducted as shown in Figure 5-1. The varying activity of chitosan-immobilized Endo-D showed a linear plot for $Man_3GlcNAc$ release seen as a Na adduct at 835 *m/z*. The results revealed that there is a dose-dependent response towards endoglycosidase immobilization which would relate to complete enzyme immobilization and that even using up to 2 U of Endo-D is possible.

Figure 5-1. Dose-dependent enzyme immobilization onto chitosan nanoparticles. A.) MALDI-TOF/MS Spectra for $Man_3GlcNAc$ (at ~835 *m/z*) released using chitosan immobilized Endo-D with varying activity (2U, 5U, 10U, and 20 U from top to bottom, respectively.). B.) Total $Man_3GlcNAc$ in pmol/0.5 mg sample. C.) Linear plot for $Man_3GlcNAc$ released using chitosan immobilized Endo-D. The results suggest that the chitosan nanoparticles can immobilize an enzyme even up to 2 U of Endo-D and is dose-dependent.



Efficiency of immobilization was further checked by recovering the supernatant after endoglycosidase conjugation. With SDS-Gel, no bands for solution washing, immobilized-Endo-D, and Endo-D control were detected (See Figure S5-1). It could be that the amount is too small to be seen in SDS. With this, glycoblotting was used as an alternative methodology to check for unimmobilized Endo-D. By using the supernatant after immobilization, lyophilized pooled washings and following the glycan release procedure, indirectly, the release of $Man_3GlcNAc$ would signify presence of Endo-D in the washings. In the result, no $Man_3GlcNAc$ peak was found in the expanded spectra of Figure 5-2 for the supernatant which would mean that no unretained Endo-D exist from the washings in all concentrations tested. The results may support that complete immobilization was successful. As there is no difference on the action of glycan release between the free and immobilized endoglycosidase, the next step was to check for reusability.

Figure 5-2. Immobilization efficiency as checked through glycoblotting. A.) MALDI-TOF/MS Spectra of 0.5 mg JQEW digested with the final washing for chitosan nanoparticles and analyzed following glycoblotting procedure. Top to bottom represents final washings from 2U, 5U, 10 U, and 20 U chitosan nanoparticle immobilized Endo-D B.) Expanded MALDI-TOF/MS region at 820-865 *m/z*. No peak at ~835 *m/z* was seen in the spectra signifying that no $Man_3GlcNAc$ was released from the supernatant solution for the final washings of the chitosan-immobilized Endo-D's at varying activities. Using SDS-Gel, no bands were seen for Endo-D as shown in Figure S5-1 in the supplementary information.



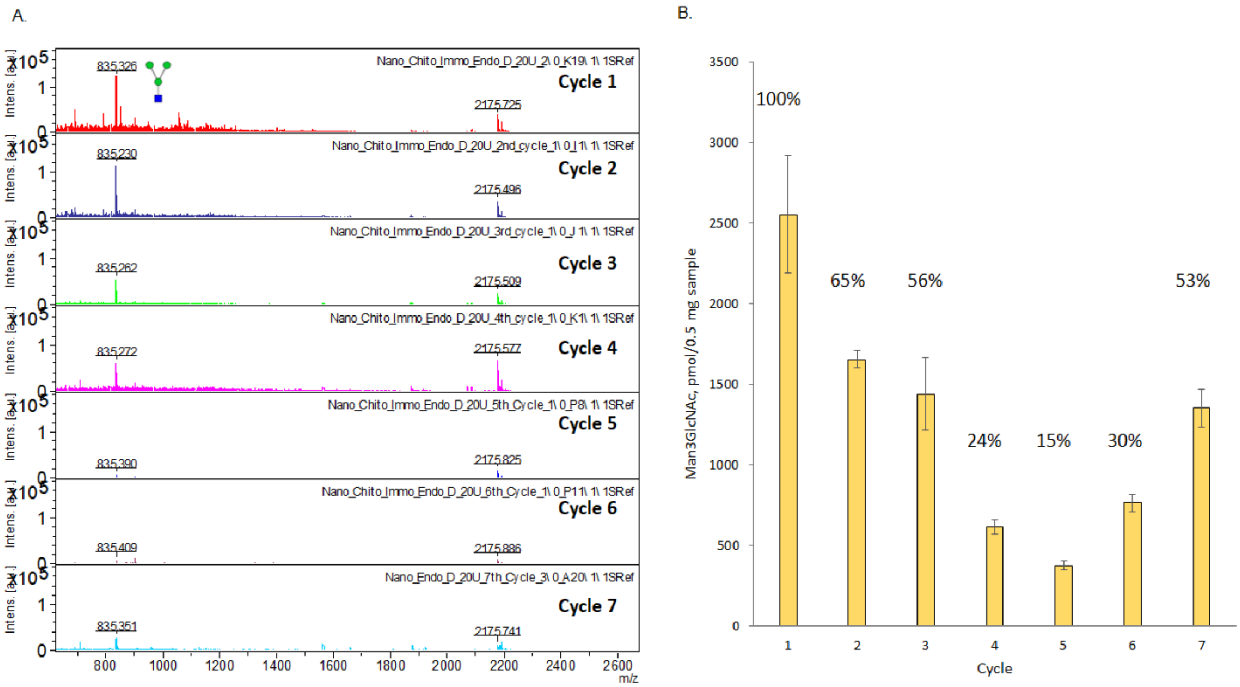
5.4.1.1 Reusability of Endo-D for Glycan Release

A 20U immobilized Endo-D on chitosan nanoparticle was used for the reusability assay. Quantification of $Man_3GlcNAc$ release was based on the glycoblotting methodology conducted after egg white incubation on the chitosan-nanoparticle immobilized Endo-D. From the 2nd up to the 7th cycle shown in Figure 5-3, percentage was normalized first on the internal standard (2600 pmol) and then further normalized based on the glycan amount found in the 1st cycle, assigned as 100%. The reusability of Endo-D showed a 65% and 56% $Man_3GlcNAc$ release in the 2nd and 3rd cycles, respectively. Further usage for 4th to 6th cycles showed 24, 15, and 30% of $Man_3GlcNAc$ released. To circumvent the decrease in glycan amount, it is recommended to increase the incubation time of the catalysis so that more of glycans can be recovered as presented in the 7th cycle wherein 53% $Man_3GlcNAc$ was obtained. This feat was achieved by incubating the sample with the chitosan-immobilized Endo D for 36 hours instead of the normal 24 h conducted for previous cycles. The reusability was only tested up to the 7th cycle as the chitosan nanoparticle ultimately changed in color compared to the initial brownish-yellow color. There was also an obvious decrease in the $Man_3GlcNAc$ as cycles were added.

Even if the reusability is shown to decrease for releasing $Man_3GlcNAc$, this is understandable as the reaction conditions, i.e. reagents (ABC, PHM, DTT, IAA), may also affect the activity of the immobilized Endos or the nanoparticle in general. The effect of various reagents on chitosan-immobilized enzyme activity was showcased in a previous work. [19,24] It is also worth noting that the incubation temperature was set at 37º C for 24 h in each cycle, totalling 7.5 days. Removal of the chitosan nanoparticle from the solution was based on centrifugation. Accumulation of the egg white as well as the reagents may likely happen. This might cause a problem as some precipitates on the egg white sample, proteins included, may attach on the chitosan that would in turn may limit the catalysis for glycan release. The egg white sample itself is a viscous heterogeneous mixture of glycoproteins. An actual real sample was used for glycan release and not as a single purified glycopeptide; with this, the results are expected. The decrease in yield for reusability was also seen in previous chitosan-based enzyme immobilization reports[17,24] even with different conditions. In the next application of the work,

transglycosylation reactions, only reactants were added in the chitosan-immobilized Endo and the buffer.

Figure 5-3. Reusability of chitosan nanoparticle immobilized Endo-D. The chitosan nanoparticle immobilized Endo-D containing 20U initial activity was reused for $Man_3GlcNAc$ (at ~835 *m/z*) release using 0.5 mg JQEW at each run with an internal standard (at 2175 *m/z*). A. MALDI-TOF/MS Spectra for 7 cycles. B. Total amount of $Man_3GlcNAc$ pmol/0.5 mg of JQEW comparison for different cycles. The 7th cycle increased compared to the 6th cycle as incubation of the tube was extended for 48h.



Although there are low abundant glycans released using Endo-D as discussed previously, [20] it was only the amount of $Man_3GlcNAc$ that was accounted in this work as it was highest. Presence of other low abundant *N*-glycans released by Endo-D seen previously [20] had low intensity to no peaks in the reusability cycles which was not quantifiable. For consistency, the glycan release was only based on $Man_3GlcNAc$ normalized with the internal standard. The $Man_3GlcNAc$ glycan is a truncated version of $Man_3GlcNAc_2$ which is a conserved structure found in all *N*-glycans. Acquiring the truncated structure is significant for the attachment of this conserved glycan onto any glycosyl acceptors which can be further extended with the aid of glycosyltransferases as mentioned in our previous works. [20, 22] The use of chitosan-immobilized Endo-D for $Man_3GlcNAc$ release has been

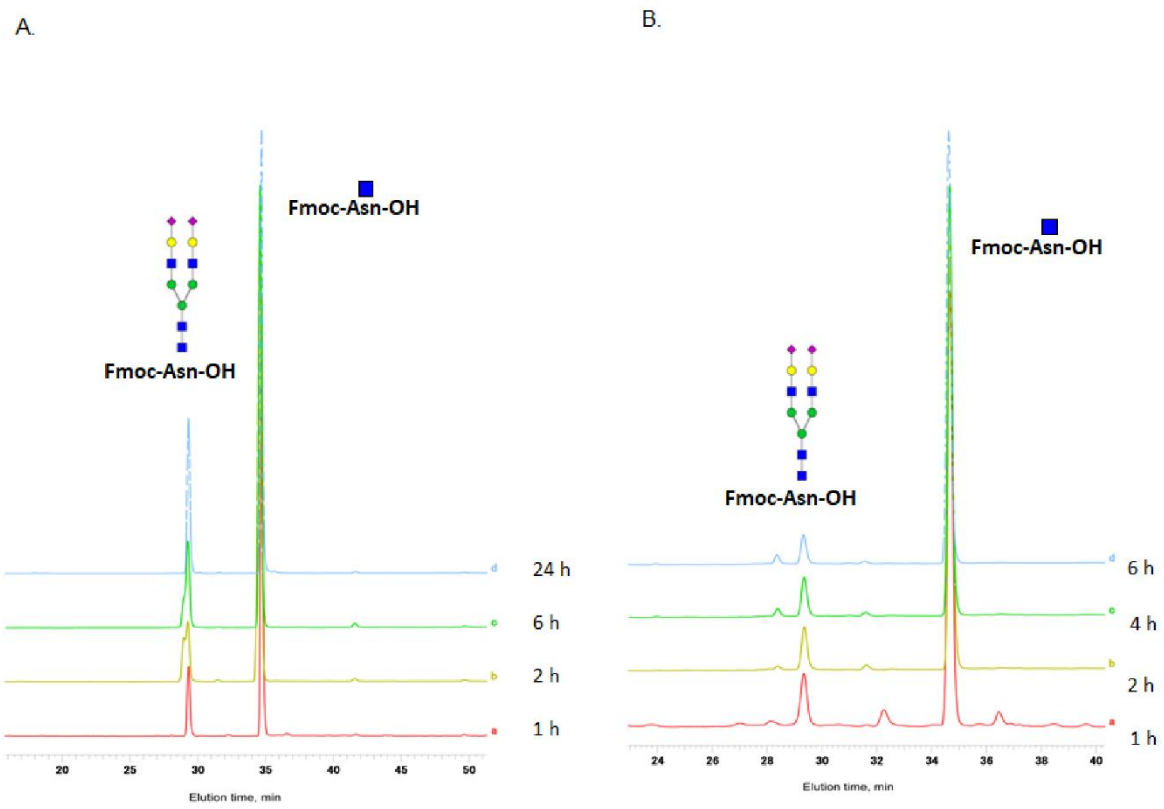
shown to be effective even with decrease enzyme activity for successive re-use. The important thought here is that instead of only using a free Endo-D once, it can now be re-used many times and by just extending the incubation time at later cycles, more of the glycan is released. Furthermore, even with a complex sample like an egg white, the immobilized-Endo on chitosan is shown here to be reusable.

5.4.2 Enzyme Immobilization for Transglycosylation Reaction

Endo-M enzyme was conjugated onto the chitosan-nanoparticle to check whether effective transglycosylation reaction will occur from the sialylglycopeptide donor to the IgG1-GlcNAc tryptic glycopeptide acceptor. Even with only 2 mU of the enzyme used for immobilization, results still show good transglycosylation yield for both free and immobilized Endo-M. Comparison of the yields of the transglycosylation reaction is shown in Figure 4. The percent transglycosylation yield was taken from the transglycosylated product peak divided by the total peak areas of the glycosyl acceptor and transglycosylated product.

Highest transglycosylation yield was seen for longer incubation with the free Endo-M while a decreasing trend was observed or chitosan-immo-Endo-M as time of incubation is increased. Although a somewhat trend can be seen, the time course-reaction was only conducted once and should be taken with a grain of salt. The transglycosylation yield was 15, 12, 22, and 28 % for the free Endo-M during 1h, 2h, 6h, and 24 h, respectively. Prolonging the incubation with the free Endo-M for 24h yielded the highest amount of transglycosylated product, this is expected. The somehow decrease in yield for chitosan immobilized Endo-M, may be due to some of the glycopeptides attaching to the chitosan when incubated at a longer time. As mentioned, both time course reactions were only conducted once and the change in % transglycosylation yield is only subtle, i.e., 12, 10, 9, and 7 % at 1, 2, 4, and 6 h, respectively, for the immobilized enzyme. No 24 h incubation was conducted for the immobilized Endo-M. From the result, a 1-h incubation time for the chitosan-immo Endo-M yield around the same amount as the free enzyme. This is considered the best time for the reusability assay wherein maybe not lots of of glycopeptides stick to chitosan particles. The main goal of this work was to be able to reuse the chitosan-immobilized Endo and this is the highlight.

Figure 5-4. Tranglycosylation yield comparison of 2mU free (A) and chitosan-immobilized Endo-M (B) at different time intervals. Analytical HPLC separation was conducted with solvents, water and acetonitrile (ACN), both containing 0.1% TFA under a gradient of 0%ACN (0 min)→60%ACN (50min)→95%b (50.1 min)→95%ACN (60min) in an Inertsil-ODS 3 column (4.5x250 mm) set at a 265 nm wavelength. The transglycosylated product was observed at ~30 minutes. The glycosyl acceptor was observed at ~35 min.



The identity of the transglycosylated product, Fmoc-Asn($Hex_5HexNAc_4NeuAc_2$)-OH, was confirmed in both MS(shown in Figure 5-6) and MS/MS spectra (shown in Figure 5-7). As expected, the reflector positive mode gave out small peaks for the compound due to the easily ionized sialic acids. But, when switched to a linear negative mode, the highest peak observed was for the Fmoc-Asn($Hex_5HexNAc_4NeuAc_2$)-OH transglycosylated product at 2560 *m/z*. The parent ion peak was chosen for MS/MS fragmentation and the fragment ions annotated. An additional confirmation was conducted using an LC-MS (see Figure S5-2). With the efficiency of chitosan-nanoparticle immobilization established, effectivity towards transglycosylation reaction repeated usage was conducted next.

5.4.2.1 Reusability of Endo-M for Transglycosylation Reaction

A new chitosan-immobilized Endo-M was prepared for the reusability assay since the previous immobilized enzyme was used for the time-course reaction. As for the cycles mentioned herein, new reactants were used for each cycle placed in the same chitosan-immobilized-Endo-M nanoparticle particularly for the reusability assay. The 2mU immobilized Endo-M on chitosan nanoparticle showed high reusability up to 13 cycles at (1^{st}, 3^{rd}, 10^{th}, and 15^{th} day after enzyme immobilization) shown in Figure 5-5. The percent transglycosylation yields are 16±3(n=6), 18±4(n=5), 10, and 9, for the 1^{st}, 3^{rd}, 10^{th}, and 15^{th} day of usage, respectively. Overall, the average percentage transglycosylation yield is 16±4 for all 13 cycles.

Stability of chitosan-immobilized enzyme was observed previously for 3 weeks when stored at 4ºC with a 15% reduction on enzymatic activity after 8 days[19]; the conditions and the enzyme used in this work is different. Higher amounts of transglycosylation yield were seen in the middle cycles as compared to the free enzyme's which can probably be the results of the enzyme's attachment on a surface similar to its natural environment in cells. [10] More cycles can be conducted further but was discontinued after finishing the 13^{th} cycle for this work to limit the use of expensive glycosyl donors and acceptor. Although, it was mentioned that below pH 6.5 causes the solubilization of chitosan [10], the buffer used in the transglycosylation in this work was pH 6.4, which is the Endo-M manufacturer's

recommended pH. The pH 6.4 buffer volume used was just 18µL while the chitosan nanoparticle used was ~22.8 mg. A more acidic pH would probably solubilized chitosan, e.g. 2% aq. acetic acid used in chitosan nanoparticle preparation. A more comprehensive analysis on the effect of acidic solutions on chitosan solubilization was conducted previously using acetic acid [25] and hydrochloric acid [26] wherein an intrinsic p*K* value of pK_0=6.0 at α=0.5 for both was observed. The soluble-insoluble transition of chitosan occurs between pH 6 and 6.5. [27] In this work, no chitosan solubilization was observed in pH 6.4.

Figure 5-5. Reusability of chitosan-immobilized Endo-M. A. HPLC Spectra of repeated cycles. The transglycosylated product was observed at ~30 minutes. The glycosyl acceptor was observed at ~35 min. B. Histograms of repeated usage(cycle). Cycles 1 to 6 were conducted for 1st day (yellow) after Endo-M immobilization, cycles 7 to 11 were after the 3rd day(pink), cycle 12 for the 10th day (green), and cycle 13 for the 15th day(violet). The same chitosan nanoparticle was used all throughout with new reagents incubated at each cycle. Analytical HPLC separation was conducted with solvents, water and acetonitrile (ACN), both containing 0.1% TFA under a gradient of 0%ACN (0 min)→60%ACN (50min)→95%b (50.1 min)→95%ACN (60min) in an Inertsil-ODS 3 column (4.5x250 mm) set at a 265 nm wavelength.

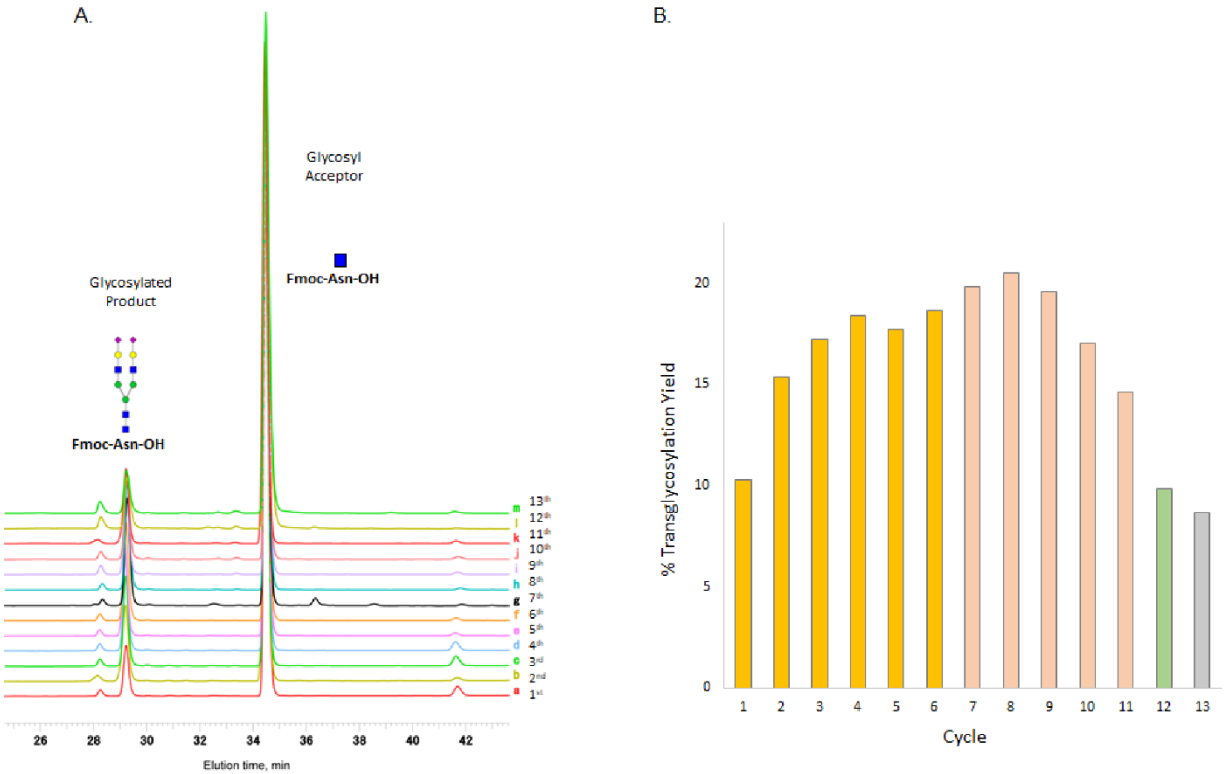


Figure 5-6. MS confirmation of transglycosylated product, Fmoc-Asn ($Hex_5HexNAc_4NeuAc_2$)-OH in both reflector positive (A.) and linear negative (B.) ion modes in MALDI-TOF/MS. C. Expanded spectra of (A.) showing positive adducts. D. Expanded spectra of (B.) showing negative adducts. It can be clearly seen that the peak at 2560 m/z for the negative H adduct has the highest intensity. There is not a big difference in the mass of the adduct in either positive or negative mode. The $[M+H]^+$ adduct theoretically at 2558 *m/z* was integrated in FlexAnalysis software as 2560 *m/z* for the positive mode. The $[M+Na]^+$ adduct in C was used for MS/MS confirmation as shown in Figure 7. Sample was taken from the analytical HPLC fraction. The purified sample detected in MALDI-TOF was for the reaction incubated with chitosan nanoparticle used after 15 days from Endo-M immobilization which is considered the 13th cycle for reusability assay.

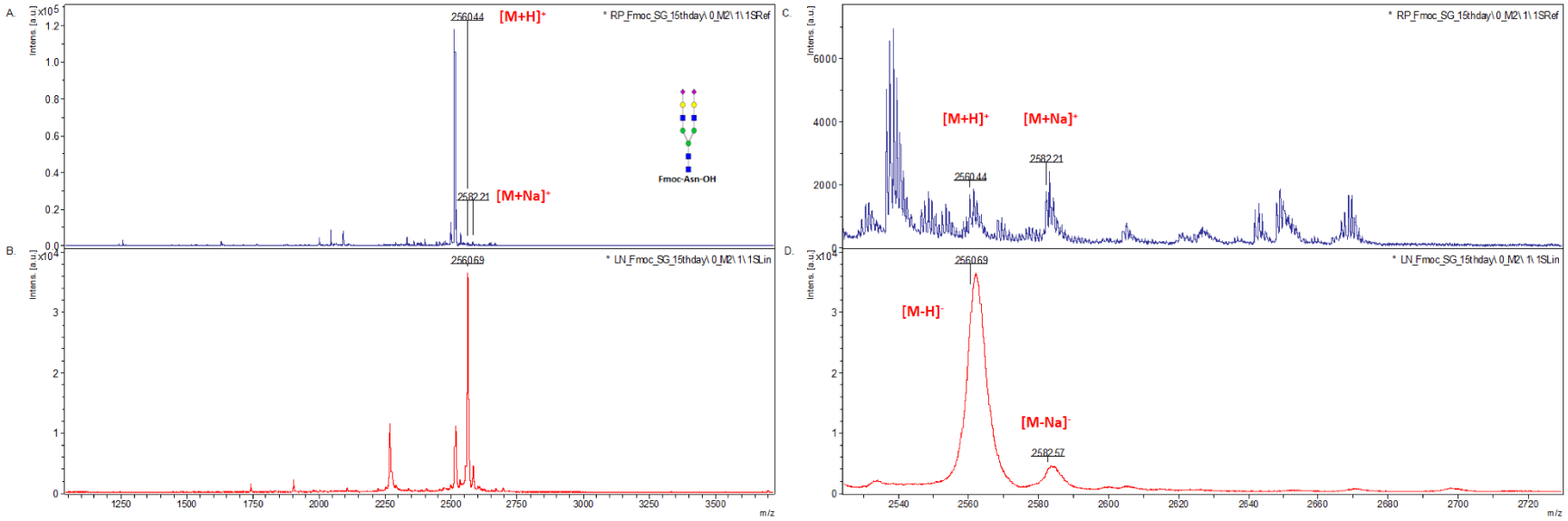
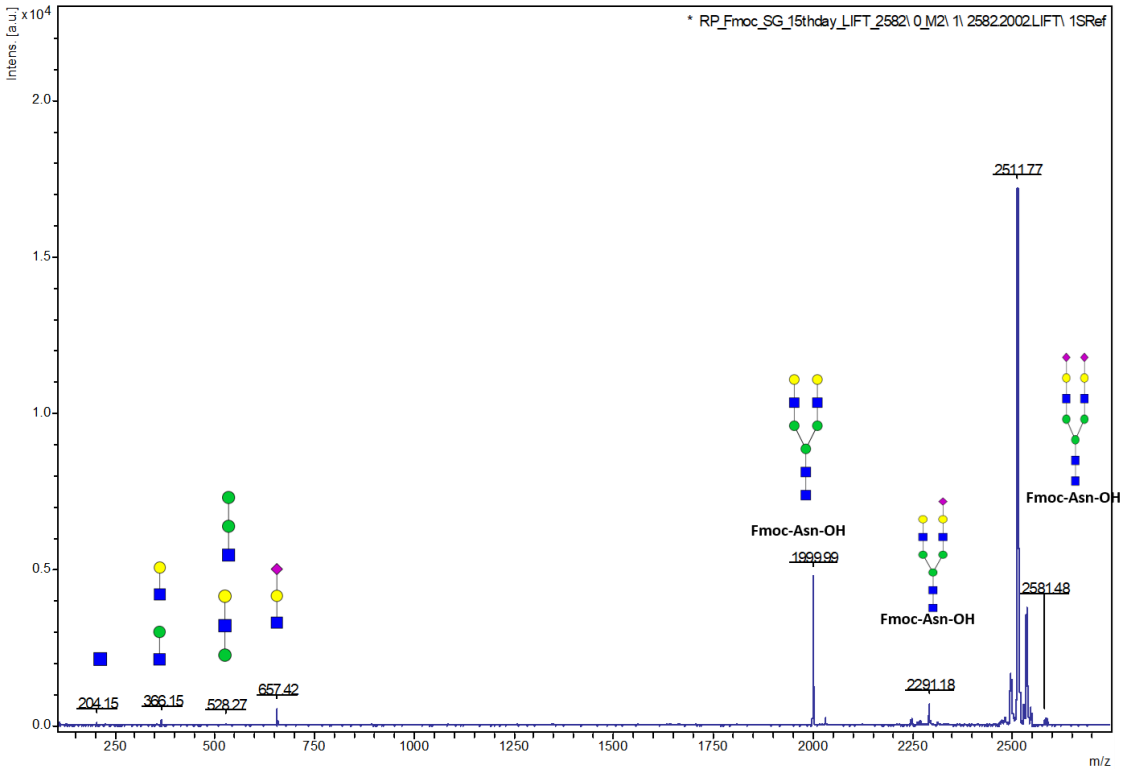


Figure 5-7. Confirmation of Fmoc-Asn($Hex_5HexNAc_4NeuAc_2$)-OH in MS/MS Spectra. Sample was taken from the analytical HPLC fraction. The purified sample detected in MALDI-TOF was for the reaction incubated with chitosan nanoparticle used after 15 days from Endo-M immobilization which is considered the 13th cycle for reusability assay. The *m/z* 204, 366, 528, and 657 are $[M+H]^+$ adducts, while, 1999, 2291, and 2281 are all $[M+Na]^+$ ions.



5.4.3 Using Chitosan Nanoparticle for Endo-immobilization

Chitosan has been used in the immobilization of different enzymes. [28-29] Chitosan, a derivative of chitin, is highly positive comprising of polymers of *N*-acetyl-D-glucosamine and D-glucosamine residues [10] which enables formation of ionic bridges to anionic residues of enzymes [17]. Wild-type Endo-D and Endo-M is highly negative at pH 7 and 6.4, respectively, (please see Tables S5-2 to S5-7) and would have a greater tendency to attach to the highly positive chitosan. Although the Endo used in this work is a recombinant Endo-D (New England Biolabs), it is assumed only a minor change in the overall charge at different pH will be likely. Chitosan has reactive amino and hydroxyl groups susceptible to chemical modification including reactions with crosslinking agents. [30] The addition of glutaraldehyde crosslinks [31] the nanoparticle and the enzyme which provides for a greater attachment. The properties mentioned may somehow explain that no unimmobilized Endo was seen in the results specified.

In lab scale, the only way of recovering the chitosan nanoparticle is through centrifugation to separate the supernatant containing the reactants and product of the reaction. Separation of the chitosan through gravity filtration as in a filter membrane, i.e., in a column, would make some of the chitosan nanoparticle non-recoverable as it would clog membranes (see Figure S5-3).

A decreasing trend for $Man_3GlcNAc$ released was seen for immobilized Endo-D. Even with a decrease in activity due to the sample used, reusing it outweighs activity reduction as more amount of $Man_3GlcNAc$ was acquired as compared to just using the enzyme once. The decreasing activity maybe negligible in large-scale synthesis. Improvement on methodologies for substrate recovery during reusability assays is important. There is not much change in the % transglycosylation yield in the chitosan-immobilized Endo-M. The subtle increase of transglycosylation yield seen in the chitosan-immobilized Endo-M maybe related to structural stability of the enzyme being attached to chitosan wherein the Endos' natural environment in membranes of cell is somehow mimicked [10]. Since chitosan has a significant affinity for proteins because of the presence of hydroxyl and amino groups [30], non-specific binding of free sites would likely happen.

Other endoglycosidase immobilization techniques were performed, an Endo-M-N175Q microbioreactor for sialic-acid oxazoline attachment to an Fmoc-Asn(GlcNAc)-OH[32] and a site-specific covalent Endo-S2 and Endo-S2-D184M immobilization for glycan remodelling of anitbodies [33]. The microbioreactor technique was based on the functionality of *N*-hydroxysuccinimide (NHS)-activated resin to couple primary amines of Endo-M-N175Q for attachment. [32] For the Endo-S and mutant, the agarose NHS-resin was further modified with diethylene glycol bis(3-aminopropyl) ether for covalent immobilization of the Endos. [33] The microbioreactor used 40 mU of the Endo-M-N175Q for immobilization while the Endo-S2 and Endo-S2-D184M used 0.5 mg/mL. In this work, 2 mU (1mU/µL) was used for Endo-M immobilization, while, various units for Endo-D was used, i.e., 2, 5, 10, and 20 U (1U/µL). These values are very minute quantities as compared to previously immobilized enzymes on chitosan. The chitosan endoglycosidase immobilization technique is easily performed without the need for difficult synthesis procedure but could be further improved. Methodologies for large scale-isolation of glycans are lacking. [11] Coupled with this problem is the efficient, stable, and economically sound enzymatic *N*-glycan release need appropriate strategies for immobilization. [11] Diversity of tools and improvement of the technology for Endo enzyme immobilization is relevant in this age of glycobiology and glycoconjugate preparations.

5.5 Chapter Conclusion

The use of chitosan-nanobased particles have been shown in this work as an Endo-immobilization support for glycan release and transglycosylation reactions requiring only small quantities of Endos. Owing to high amino groups present in chitosan that provides positive charges and with the addition of glutaraldehyde, an excellent cross-linker, it is assured that enzyme-attachment is successful. The glycoblotting methodology found new uses in quantifying truncated-released glycans released by immobilized-Endos which includes reusability and immobilization efficiency monitoring. The tools provided here are important strategies for chemoenzymatic synthesis of glycopeptides.

5.6 References

1 Fairbanks, A. J. (2017). The ENGases: versatile biocatalysts for the production of homogeneous N-linked glycopeptides and glycoproteins. *Chem. Soc. Rev.*, *46*(16), 5128-5146.

2 Karamanos, Y., Bourgerie, S., Barreaud, J. P., & Julien, R. (1995). Are there biological functions for bacterial endo-N-acetyl-β-D-glucosaminidases?. *Res. Microbiol.*, *146*(6), 437-443.

3 Maley, F., Trimble, R. B., Tarentino, A. L., & Plummer Jr, T. H. (1989). Characterization of glycoproteins and their associated oligosaccharides through the use of endoglycosidases. *Anal. Biochem.*, *180*(2), 195-204.

4 Yamamoto, K., Fujimori, K., Haneda, K., Mizuno, M., Inazu, T., & Kumagai, H. (1997). Chemoenzymatic synthesis of a novel glycopeptide using a microbial endoglycosidase. *Carbohydr. Res.*, *305*(3-4), 415-422.

5 Takegawa, K., Yamaguchi, S., Kondo, A., Iwamoto, H., Nakoshi, M., Kato, I., & Iwahara, S. (1991). Transglycosylation activity of endo-beta-N-acetylglucosaminidase from Arthrobacter protophormiae. *Biochemistry international*, *24*(5), 849-855.

6 Takegawa, K., Yamaguchi, S., Kondo, A., Kato, I., & Iwahara, S. (1991). Enzymatic synthesis of novel oligosaccharides by use of transglycosylation activity of endo-beta-N-acetylglucosaminidase from Arthrobacter protophormiae. *Biochem. Int.*, *25*(5), 829-835.

7 Cote, G. L., & Tao, B. Y. (1990). Oligosaccharide synthesis by enzymatic transglycosylation. *Glycoconj. J.*, *7*(2), 145-162.

8 Dwevedi, A. (2016). Basics of enzyme immobilization. In *Enzyme Immobilization* (pp. 21-44). Springer, Cham.

9 Homaei, A. A., Sariri, R., Vianello, F., & Stevanato, R. (2013). Enzyme immobilization: an update. *J. Chem. Biol.*, *6*(4), 185-205.

10 Krajewska, B. (2004). Application of chitin-and chitosan-based materials for enzyme immobilizations: a review. *Enzyme Microb. Technol.*, *35*(2-3), 126-139.

11 Karav, S., Cohen, J. L., Barile, D., & de Moura Bell, J. M. L. N. (2017). Recent advances in immobilization strategies for glycosidases. *Biotechnol. Prog.*, *33*(1), 104-112.

12 An, X., & Su, Z. (2001). Characterization and application of high magnetic property chitosan particles. *J. Appl. Polym. Sci.*, *81*(5), 1175-1181.

13 Sun, J., Xu, B., Shi, Y., Yang, L., & Ma, H. L. (2017). Activity and stability of trypsin immobilized onto chitosan magnetic nanoparticles. *Adv. Mater. Sci. Eng.*, *2017*.

14 Sheu, D. C., Li, S. Y., Duan, K. J., & Chen, C. W. (1998). Production of galactooligosaccharides by β-galactosidase immobilized on glutaraldehyde-treated chitosan beads. *Biotechnol. Tech.*, *12*(4), 273-276.

15 Carrara, C. R., & Rubiolo, A. C. (1994). Immobilization of beta-galactosidase on chitosan. *Biotechnol. Prog. 10*(2), 220-224.

16 Taqieddin, E., & Amiji, M. (2004). Enzyme immobilization in novel alginate–chitosan core-shell microcapsules. *Biomaterials*, *25*(10), 1937-1945.

17 Wahba, M. I. (2017). Porous chitosan beads of superior mechanical properties for the covalent immobilization of enzymes. *Int. J. Biol. Macromol.*, *105*, 894-904.

18 Berthold, A., Cremer, K., & Kreuter, J. S. T. P. (1996). Preparation and characterization of chitosan microspheres as drug carrier for prednisolone sodium phosphate as model for anti-inflammatory drugs. *J. Control. Release*, *39*(1), 17-25.

19 Biró, E., Németh, Á. S., Sisak, C., Feczkó, T., & Gyenis, J. (2008). Preparation of chitosan particles suitable for enzyme immobilization. *J. Biochem. Biophys. Methods*, *70*(6), 1240-1246.

20 Sanes, J.T., Hinou, H., Nishimura, S-I. (2019). Sourcing N-Glycans from a Heterogeneous Mixture of Glycoproteins in Japanese Quail Egg White: A Methodology to Obtain Man3GlcNAc Utilized Directly for Oxazoline and Transglycosylation Reactions. *Unsubmitted.*

21 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*(21), 4757-4774.

22 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*(1), 531-540.

23 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*(4), 1650-1659.

24 Cakmak, U., & Ertunga, N. S. (2016). Gene cloning, expression, immobilization and characterization of endo-xylanase from Geobacillus sp. TF16 and investigation of its industrial applications. *J. Mol. Catal. B: Enzym.*, *133*, S288-S298.

25 Rinaudo, M., Pavlov, G., & Desbrieres, J. (1999). Influence of acetic acid concentration on the solubilization of chitosan. *Polymer*, *40*(25), 7029-7032.

26 Rinaudo, M., Pavlov, G., & Desbrieres, J. (1999). Solubilization of chitosan in strong acid medium. *Int. J. Polym. Anal. Charact.*, *5*(3), 267-276.

27 Yi, H., Wu, L. Q., Bentley, W. E., Ghodssi, R., Rubloff, G. W., Culver, J. N., & Payne, G. F. (2005). Biofabrication with chitosan. *Biomacromolecules*, *6*(6), 2881-2894.

28 Muzzarelli, R. A. (1980). Immobilization of enzymes on chitin and chitosan. *Enzyme Microb. Technol.*, *2*(3), 177-184.

29 Juang, R. S., Wu, F. C., & Tseng, R. L. (2002). Use of chemically modified chitosan beads for sorption and enzyme immobilization. *Adv. Environ. Res.*, *6*(2), 171-177.

30 Peter, M. G. (1995). Applications and environmental aspects of chitin and chitosan. *J. Macromol. Sci. A.*, *32*(4), 629-640.

31 Chen, H., Zhang, Q., Dang, Y., & Shu, G. (2013). The effect of glutaraldehyde cross-linking on the enzyme activity of immobilized β-galactosidase on chitosan bead. *Adv. J. Food Sci. Technol. 5*(7), 932-935.

32 Haneda, K., Oishi, T., Kimura, H., & Inazu, T. (2018). Development of a microbioreactor for glycoconjugate synthesis. *Bioorg. Med. Chem.*, *26*(8), 2092-2098.

33 Li, T., Li, C., Quan, D. N., Bentley, W. E., & Wang, L. X. (2018). Site-specific immobilization of endoglycosidases for streamlined chemoenzymatic glycan remodeling of antibodies. *Carbohydr. Res. 458*, 77-84.

5.7 Supplementary Information

Supplementary Methodologies

5.7.1 SDS-Page

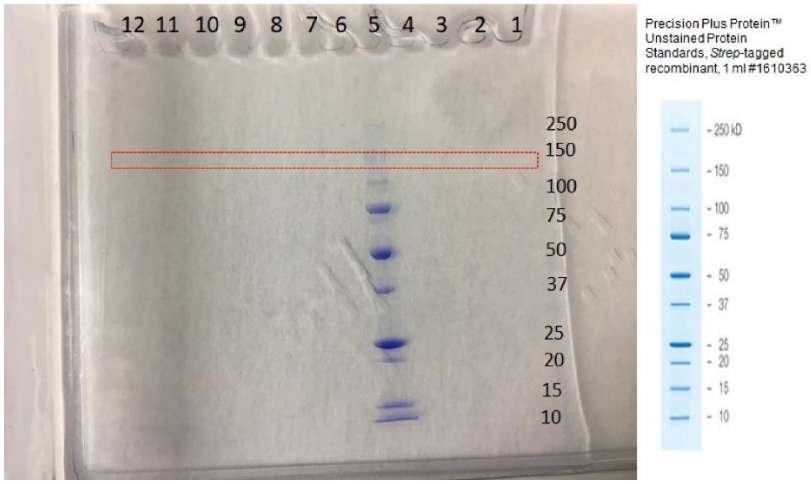
The samples were analyzed in Compact PAGE ATTO AE-7300 with sample loading buffer (0.125 M Tris-HCl pH 6.8, 20% glycerol, 4 % SDS, 2% Beta-mercaptoethanol, 0.02% Bromphenol blue) and running buffer (25 mM Tris, 192 mM glycine, and 0.1% SDS) for ~30 minutes. Lyophilized samples were reconstituted in 7.5 uL MilliQ water with 2.5 uL sample loading buffer added. As for the chitosan nanoparticle with immobilized Endo-D, 7.5 µL was taken. A total of 10 uL is needed for sample placement in a precast gel using CPAGEL 520-L Precast Get (Atto Corporation, Japan). The coomassie blue staining solution is composed of 0.1% Coomassie R-250 in 40% ethanol and 10% acetic acid.

5.7.2 Recovery of chitosan nanoparticles using a filter membrane.

After chitosan-immobilized Endo-D digestion for 24h at 37ºC, the solution mixture together with the chitosan nanoparticle was carefully transferred to a YM-10 centrifugal filter membrane (Merck, Millipore) and centrifuged at 5 RCF for 10 mins at 20 ° C to separate the nanoparticle from the reaction mixture. The reaction mixture was analyzed through the glycoblotting methodology described in the main text while the chitosan nanoparticle was reused for another round of $Man_3GlcNAc$ release. Only 2 cycles were conducted as incomplete chitosan recovery was observed due to some nanoparticles stuck on the membrane.

Supplementary Figures and Tables.

Figure S5-1. SDS-Gel checking for immobilization and unretained Endo-D in the nanoparticles. The corresponding lanes are labeled in the table below. Only the protein ladder was seen and maybe the sample quantities are too small to be observed, even for the control. Endo-D apparently can be seen at around 140 kDa (dashed red rectangle), but, was not observed. As such, an indirect checking was conducted through glycoblotting and MS as shown in Figure 2.



Lane	Sample Identity
1	Lyophilized washing from 10 U
2	Lyophilized washing from 20 U
3	Endo-D supernatant for 10 U
4	Endo-D supernatant for 20 U
5	Protein Marker, BioRad
6	Endo-D Control, 2U
7	Endo-D 20 U Supernatant from 1st Washing
8	Endo-D 20 U Supernatant from 2nd Washing
9	Endo-D 20 U Supernatant from 3rd Washing
10	2U Endo-D in Nanoparticle
11	5U Endo-D in Nanoparticle
12	10U Endo-D in Nanoparticle

Table S5-1. Lane sample identity for enzyme immobilization checking.

pH	Charge
4.0	160.7
4.5	91.2
5.0	29.1
5.5	-5.0
6.0	-22.0
6.5	-34.1
7.0	-43.5
7.5	-49.1
8.0	-53.0
8.5	-58.8
9.0	-72.1
9.5	-103.4
10.0	-157.4

Table S5-2. Estimated charge over pH range for Wild-type Endo D, Uniprot Q93HW0. Calculation was conducted using http://protcalc.sourceforge.net/

Residue			Number Found
A	Ala	Alanine	149
R	Arg	Arginine	53
N	Asn	Asparagine	78
D	Asp	Aspartate	108
Q	Gln	Glutamine	67
E	Glu	Glutamate	146
G	Gly	Glycine	119
H	His	Histidine	30
I	Ile	Isoleucine	50
L	Leu	Leucine	128
K	Lys	Lysine	150
M	Met	Methionine	17
F	Phe	Phenyalanine	58
P	Pro	Proline	69
S	Ser	Serine	107
T	Thr	Threonine	110
Y	Tyr	Tyrosine	56
V	Val	Valine	124
W	Trp	Tryptophan	25
C	Cys	Cysteine	2

Table S5-3. Wild-type Endo-D (Uniprot Q93HW0 identifier) amino acid content. Calculation was conducted using http://protcalc.sourceforge.net/

pH	Charge
4.0	68.6
4.5	41.1
5.0	16.5
5.5	2.5
6.0	-5.4
6.5	-12.4
7.0	-18.4
7.5	-22.5
8.0	-26.1
8.5	-31.2
9.0	-39.4
9.5	-54.0

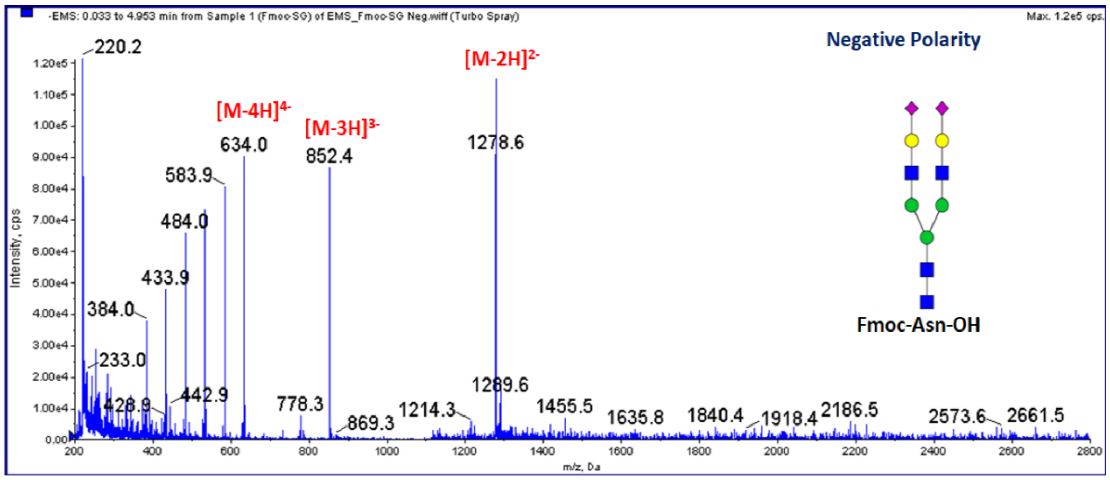
Table S5-4. Estimated charge over pH range for wild-type Endo M, Uniprot Q9C1S6. Calculation was conducted using http://protcalc.sourceforge.net/

Residue			Number Found
A	Ala	Alanine	34
R	Arg	Arginine	27
N	Asn	Asparagine	45
D	Asp	Aspartate	48
Q	Gln	Glutamine	22
E	Glu	Glutamate	52
G	Gly	Glycine	49
H	His	Histidine	20
I	Ile	Isoleucine	44
L	Leu	Leucine	63
K	Lys	Lysine	50
M	Met	Methionine	14
F	Phe	Phenyalanine	39
P	Pro	Proline	35
S	Ser	Serine	53
T	Thr	Threonine	43
Y	Tyr	Tyrosine	37
V	Val	Valine	39
W	Trp	Tryptophan	20
C	Cys	Cysteine	10

Table S5-5. Wild-type Endo-M (Uniprot Q9C1S6 identifier) amino acid content. Calculation was conducted using http://protcalc.sourceforge.net/

Figure S5-2. LC-MS confirmation of transglycosylated product, Fmoc-Asn ($Hex_5HexNAc_4NeuAc_2$)-OH using 4000QTRAP (AB-Sciex). A. Enhanced Mass Spectrometry Mode wherein 2^-, 3^-, 4^- charges were observed. B. Enhanced Resolution Mode for the $[M-2H]^{2-}$ adduct. Both modes were detected in negative polarity. Sample was taken from the eluted fraction at ~30 minutes during analytical HPLC separation with solvents, water and acetonitrile (ACN), both containing 0.1% TFA under a gradient of 0%ACN (0 min)→60%ACN (50min)→95%b (50.1 min)→95%ACN (60min) in an Inertsil-ODS 3 column (4.5x250 mm) set at a 265 nm wavelength.

A.



B.

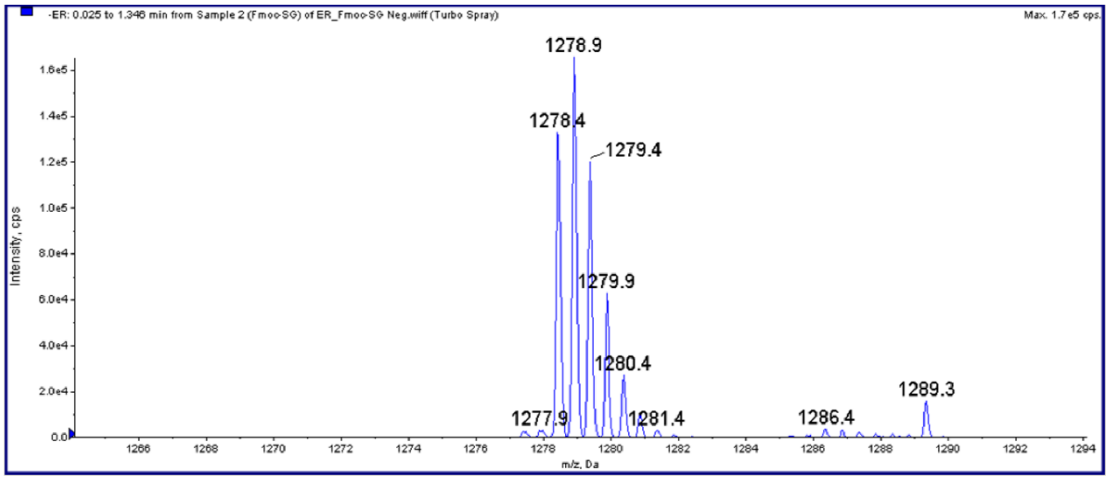
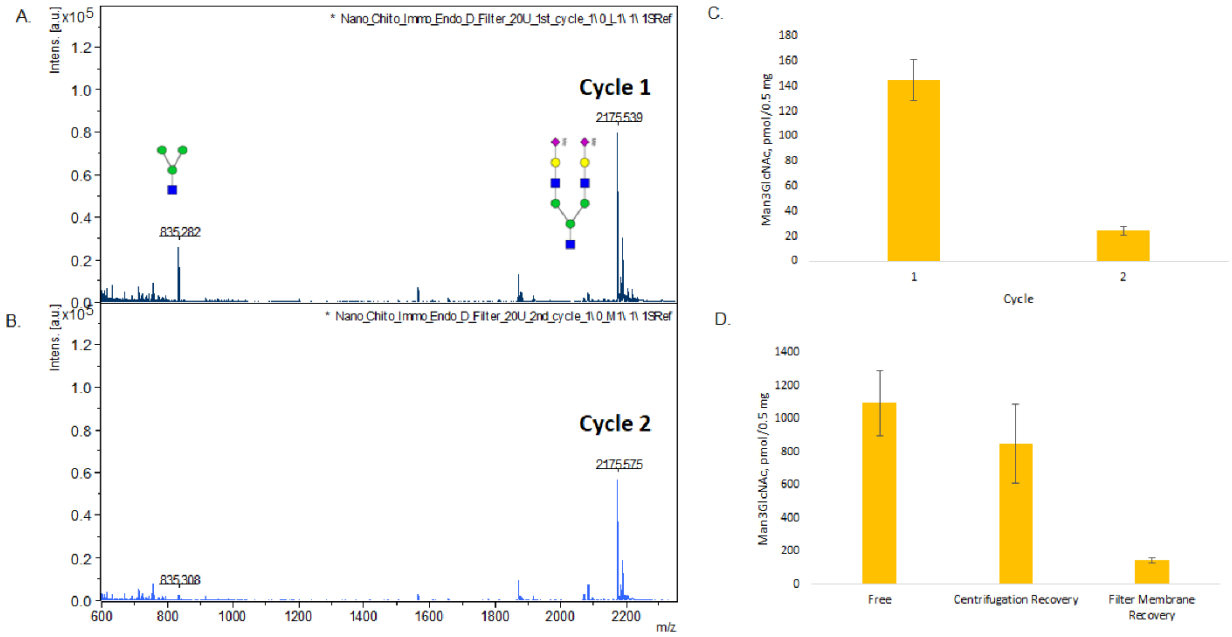


Figure S5-3. $Man_3GlcNAc$ quantitation for chitosan-immobilized Endo-D recovered using a filter membrane. MALDI-TOF/MS Spectra of BOA-labeled $Man_3GlcNAc$ at 835 *m/z* with methylated internal standard at 2175 *m/z* of different cycles, Cycle 1 (A) and Cycle 2 (B). Analysis of the *m/z* was conducted in reflector positive mode. C. Histogram comparison of Cycle 1 and Cycle 2. D. Histogram comparison of the $Man_3GlcNAc$ release among free Endo-D, centrifuged-recovered chitosan-immobilized-Endo-D, and filter membrane-recovered chitosan-immobilized Endo-D. Undoubtedly, recovering the chitosan-immobilized Endo-D was better with centrifugation in an Eppendorf tube as opposed to using a filter membrane due to incomplete recovery of the chitosan particle.



Chapter 6

Concluding Remarks

The glycobiology field is flourishing, however, the biological roles and functions of glycans can only be assessed by having specific structures first-hand. With the community of glycoenthusiast developing tools, i.e., glycan-oxazolines, screening and identification of endoglycosidases roles, chemical and enzymatic strategies, etc., the glycobiomolecules can be prepared. Having such glycoconjugate structures will help rationalize the function of glycans in biological processes. Achieving this feat need collaboration of scientist in different fields.

In my work, the assessment of the glycan components was made possible through the glycoblotting methodology that made it easier to analyze and quantify glycans in the egg white samples. The glycome diversity in organisms is one area that need to be tackled on. An investigation among quail egg white samples as presented in Chapter 2 gave an insight into glycan diversity in the Galliformes order of birds. Glycan screening in different organisms such as in egg white and plants can provide natural glycans necessary for glycoconjugate procedures. The glycoblotting methodology hastens the screening of natural glycans.

With my work, screening natural samples for glycan components and using it directly for glycocojugate preparations will provide for new methodologies in supplying the glycofield with these valuable structures. The glycan isolation and endoglycosidase immobilization strategies presented in Chapters 3 and 5, respectively, can be upscaled and utilized in the production of homogeneous glycoconjugates with a correctly defined glycan component that can be used in different assays such as in biomarker quantitation.

The connection of glycans in different diseased state is gaining traction. Screening of clinical therapeutic antibodies is important and quality control checks are necessary to elicit the same healing effects on individuals with debilitating diseases. In so doing, having glycoform calibration standards to quantitate biomarkers is the next area that need to be addressed. Chapter 4 of this dissertation provides a multiplexing procedure to quantitate possible IgG tryptic antibody biomarkers using chemoenzymatically synthesized glycopeptides as standards. Undoubtedly, the use of different mass spectrometry techniques, MALDI-TOF/TOF and LC-MS/MS have remarkably made structural confirmations and quantifications easier. The dissertation presented gives insight into glycan diversity

and addresses relevant strategies needed for chemoenzymatic preparation of concretely defined *N*-glycopeptides utilizing mass spectrometry techniques to achieve the objectives in this work.