DEVELOPMENT OF GLYCOANALYTICAL METHODS FOR

CHARACTERIZATION OF BIOSIMILAR MONOCLONAL ANTIBODIES

by

Katharina Yandrofski

B. S. Biology (University of Maryland) 2014

THESIS

Submitted in partial satisfaction of the requirements

for the degree of

MASTER OF SCIENCE

in

BIOMEDICAL SCIENCE

in the

GRADUATE SCHOOL

of

HOOD COLLEGE

December 2019

Accepted:

_____	_____
Ann L. Boyd, Ph.D. Committee Member	Ann L Boyd , Ph.D. Director, Biomedical Science Program
_____	
Craig Laufer, Ph.D. Committee Member	
_____	_____
John E. Schiel, Ph.D. Thesis Adviser	April M. Boulton, Ph.D. Dean of the Graduate School

STATEMENT OF USE AND COPYRIGHT WAIVER

I authorize Hood College to lend this thesis, or reproductions of it, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

NIST Disclaimer

Values reported herein were collected during NISTmAb qualification and/or certification and current at the time of publication. Users should always refer to the Report of Investigation for their specific material lot for the most up to date values and uncertainty ranges. Certain commercial equipment, instruments, or materials are identified to adequately specify the experimental procedure. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.

Portions of this thesis contain work that was previously published in: Kashi L, Yandrofski K, Preston RJ, Arbogast LW, Giddens JP, Marino JP, Schiel JE, Kelman Z. 2018. Heterologous recombinant expression of non-originator nistmab. mAbs. 10(6):922-933. https://www.tandfonline.com/doi/full/10.1080/19420862.2018.1486355

ACKNOWLEDGEMENTS

I would like to thank my advisor Dr. John Schiel for his guidance throughout the course of this project and for providing me with foundational knowledge in mass spectrometry. Dr. Zvi Kelman and Lila Kashi for their assistance in production and processing of the non-originator cell lines. I would like to thank Dr. Ann Boyd and Dr. Craig Laufer as my committee members for their guidance throughout my graduate studies. I am grateful for the instrument and application specific knowledge provided by Agilent Technologies.

TABLE OF CONTENTS

ABSTRACT

Monoclonal antibodies are the fastest growing class of therapeutics with more than $100 billion in yearly revenue. Many of the world's top selling biotherapeutic mAbs will lose their patent protection by the end of 2019 and a large number of biosimilars are emerging. Detailed characterization of physiochemical and biophysical attributes, such as glycosylation and aggregation, are of critical importance to their safety and efficacy are critical to establishing analytical similarity. Herein we describe development of three analytical assays; size exclusion chromatography (SEC), intact mass spectrometry, and released glycan analysis using the NISTmAb Reference Material – an open innovation tool for the development of innovative analytical approaches. These analytical methods were then applied to a series of non-originator cell lines expressing the NISTmAb that were developed as a pre-competitive biosimilarity test case. The analytical methods and dataset provide the foundational infrastructure for continued development of this resource for the biopharmaceutical community.

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

FDA	Food and Drug Administration
IgG	Immunoglobulin G
CDC	Complement-dependent cytotoxicity
ADCC	Antigen-dependent cellular cytotoxicity
Fab	Antigen binding fragment
Fc	Crystallizable fragment
CDR	Complementary determining region
PTM	Post-translational modification
Dol-P	Dolichol phosphate
OST	Oligosaccharyltransferase
CQA	Critical quality attribute
HPLC	High-pressure liquid chromtography
MS	Mass spectrometry
HRMS	High-resolution intact mass spectrometry
ESI	Electrospray ionization
TOF	Time-of-flight
Q-TOF	Quadrupole time-of-flight
CID	Collision associated disassociation
NIST	National Institute of Standards and Technology
RM 8671	NISTmAb Reference Material 8671
PQA	Product quality attributes
RP-HPLC	Reversed-phase high-pressure liquid chromatograph

mAb	Monoclonal antibody
UV	Ultraviolet
LC/MS	Liquid chromatography/mass spectrometry
SEC	Size exclusion chromatography
BEH	Ethylene bridged hybrid
PS 8670	Primary Standard 8670
FLD	Fluorescence detection
PNGase F	N-glycosidase F
2-AB	2-Aminobenzamide
HILIC	Hydrophilic interaction liquid chromatography
ACN	Acetonitrile
DMF	Dimethylformamide
ISD	In-source decay
BEx	Buffer exchange
m/z	mass-to-charge ratio
NeuGc	N-glyconeuraminic acid
GalNAc	N-acetylglucosamine

Introduction

The United States biotherapeutic monoclonal antibody (mAb) market generates over $100 billion in yearly revenue with this value expected to surpass $150 billion in the next 5 years (Walsh 2014; Newswire). In addition, they are the fastest-growing class of therapeutics with more than 30% of drug products approved over the next 10 years expected to be monoclonal antibodies (Newswire ; Jefferis 2009; Nelson et al. 2010; Schiel et al. 2014). With an increase in the impact of these biomolecules on the biopharmaceutical industry, the need for detailed characterization for physiochemical and biophysical attributes is of critical importance.

Monoclonal Antibodies

There are 5 classes of antibodies (IgA, IgD, IgE, IgG, and IgM) with each having a specific function in the immune response. IgG is currently the only class that has US Food and Drug Administration (FDA) approved therapeutics on the market. Extensive research has been performed on the mechanism of action of IgG-based therapeutics which includes activation of complement-dependent cytotoxicity (CDC) and antigen-dependent cellular cytotoxicity (ADCC) (Schiel 2012; Mimura et al. 2018; Liu et al. 2008; van de Bovenkamp et al. 2016). The common structure of an IgG molecule has been well documented and characterized (Nelson et al. 2010; Beck et al. 2013; Cymer et al. 2018; Jefferis 2014). The structure is a heterodimer (~150kDa) consisting of 4 chains, 2 light (~25kDa) and 2 heavy (~50kDa). The Y-shaped structure is divided into 2 specific areas via the hinge region, the antigen binding fragment (Fab) region and the crystallizable fragment (Fc) (Figure 1). The Fab region contains both light and heavy chains which are connected via a single disulfide

bond and noncovalent interactions. Each chain consists of a variable (V_L and V_H) and constant (C_L and C_H1) domains. The variable regions of the both the light and heavy chain contain areas of unique sequences that vary between individual antibodies which are known as complementary determining regions (CDRs) (Owen et al. 2013). These areas determine antibody specificity in terms of what antigens it will recognize and bind. The light chain contains a single constant domain on each fragment while the heavy chain contains a total of 3 constant domains, 1 in the Fab region (C_H1) and 2 in the Fc region (C_H2, C_H3) (Owen et al. 2013; Wang et al. 2007). The 2 heavy chains are connected via 2 inter-disulfide bonds with an additional 12 intra-molecular disulfide bonds spread out through the antibody in each of the regions described above. The two regions of the antibody each serve a specific purpose with the Fab area binding to antigen to form antigen-antibody complexes while the Fc region can bind to Fc receptors on phagocytic cells or other immune effector molecules. (Owen et al. 2013; Formolo et al. 2015).

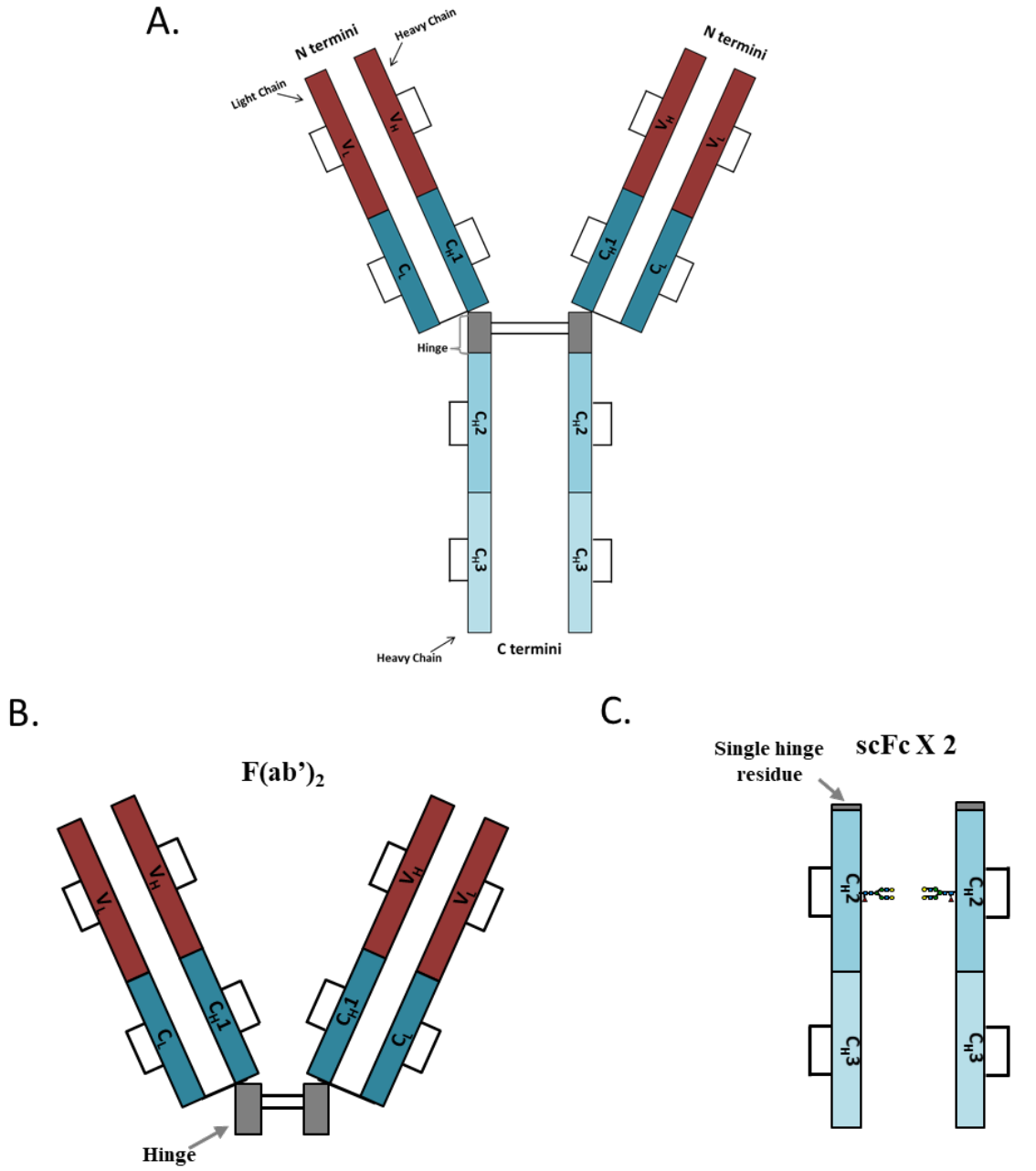


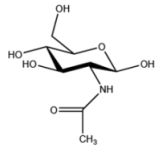
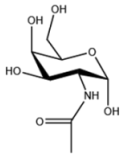
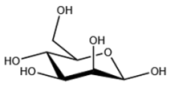
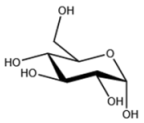
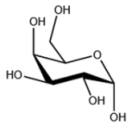
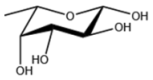
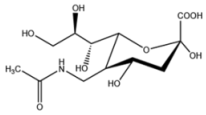
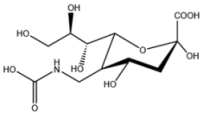
Figure 1. Structure of an IgG molecule. A. Intact antibody, B. Fab region including hinge, C. Fc region including attached glycan structures. Disulfide bonds are indicated by inter- or intra-chain black lines.

Post-Translational Modifications

Monoclonal antibodies can undergo an array of post-translational modifications (PTMs) during production and storage that can affect their immunogenicity, efficacy, and safety (Mimura et al. 2018; Jefferis 2014; Zhang et al. 2016). There are several common mAb PTMs which contribute to increased heterogeneity including glycosylation, disulfide bond formation, deamidation, oxidation, phosphorylation, N-terminal pyroglutamate formation, and C-terminal lysine clipping (Liu et al. 2008; Walsh 2010; Zhang and Ge 2011; Barton et al. 2014). N-linked Glycosylation is the most abundant PTM and main source of heterogeneity in monoclonal antibodies. This modification involves the covalent attachment of oligosaccharides to asparagine residues of proteins at the consensus sequence Asn-Xaa-Ser/Thr, where Xaa can be any amino acid except proline. This addition of carbohydrate residues has a profound effect on specific immune functions including antibody dependent cell-mediated cytotoxicity (ADCC) and complement mediated cytotoxicity (CDC). IgG1 antibodies contain a single N-glycosylation site located in the Fc region on each heavy chain CH2, and less frequently may contain N-glycosylation in the Fab (Schiel 2012; Zhou and Qiu 2018; Stanley et al. 2009).

The biosynthesis of N-glycans is a complicated process involving 3 main steps that takes place in the endoplasmic reticulum and the golgi apparatus. This process is essentially a complex series of additions and deletions of monosaccharide building blocks to form the larger, more complex glycan structure. The structure for each of the monosaccharide constituents that make up N-glycans are detailed in Table 1 below depicting both the representative structure and the symbolic representation for each. N-glycan structures will be presented using the symbolic representation throughout this paper for simplicity.

Table 1. Structure and symbolic representation of N-glycans

Glycan Species	Abbreviation	Structure	Mass	Symbol
N-Acetylglucosamine	GalNAc		203.0794	
N-Acetylgalactosamine	GlcNAc		203.0794	
Mannose	Man		162.0528	
Glucose	Glc		162.0528	
Galactose	Gal		162.0528	
Fucose	Fuc		146.0579	
N-Acetylneuraminic Acid	NeuAc		291.0954	
N-Glyconeuraminic Acid	NeuGc		307.0903	

Synthesis begins on the cytoplasmic face of the ER where the membrane-bound form of Dolichol phosphate (Dol-P) as shown in Figure 2 has an N-acetylglucosamine residue (GlcNAc-P) added to form Dol-P-P-GlcNAc. (Stanley et al. 2009)

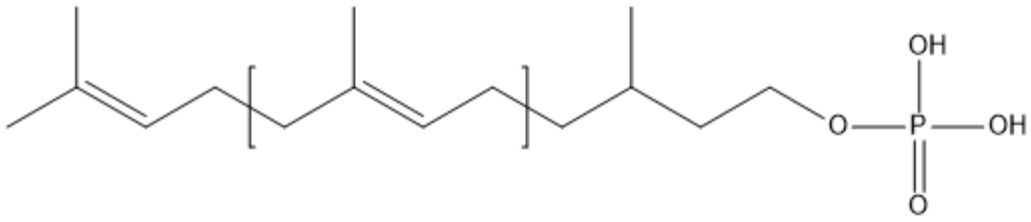


Figure 2. Structure of dolichol phosphate (Dol-P).

The sequential addition of 14 monosaccharide residues form the N-glycan precursor $Glc_3Man_9GlcNAc_2$, which is comprised of simple and complex carbohydrate molecules (Zhou and Qiu 2018; Stanley et al. 2009). The entire structure is then transferred to the Asn-X-Ser/Thr sequence in the protein by the oligosaccharyltransferase (OST) (Stanley et al. 2009). Once the precursor is bound to the protein it goes through a number syntheses using glycosidases and glycotransferases to remove and add monosaccharide residues, creating one of the 3 types of N-glycan structures (oligomannose, complex, and hybrid (Zhou and Qiu 2018; Stanley et al. 2009) (Figure 3). All 3 types contain the same trimannosyl core structure of Man3GlcNAc2 which is depicted in Figure 3.

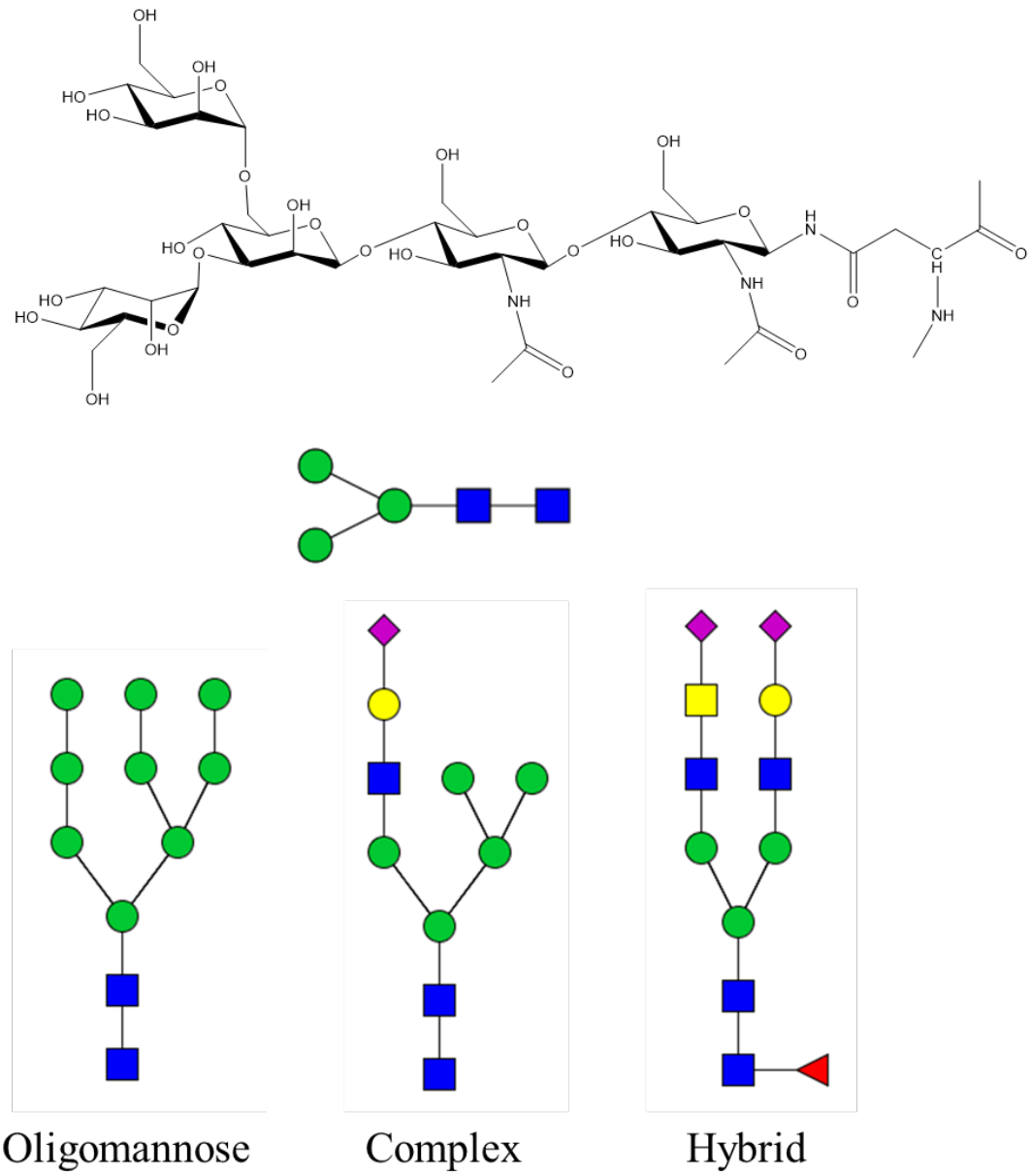


Figure 3. Representative core structure and symbolic representative structures for oligomannose, complex, and hybrid N-glycans

The biosynthesis of oligosaccharides is not a template driven process, instead it involves a series of enzymatic reactions which is affected by factors including the expression system and upstream/downstream processing. Due to the increased heterogeneity that N-glycans add to mAbs, glycosylation is considered a critical quality

attribute (CQA). CQAs are physical, chemical, biological, or microbiological characteristics that must stay within a specific, predetermined range in order to ensure product quality as they are essential to the safety and efficacy of the drug product (Formolo et al. 2015; Reusch and Tejada 2015). The overall glycan profile should be maintained during a product's lifecycle. In addition, specific glycan structures are often critically monitored as they have been shown to elucidate certain immune responses and/or affect pharmacokinetics. Afucosylated N-glycan species, for example, have been shown to induce ADCC by enhancing binding affinity between IgG and the Fcγ receptor on NK cells (Mimura et al. 2018; Zhang et al. 2016; Aich et al. 2016). An increase in terminal galactose has similar effects to afucosylation by leading to the activation of CDC by binding to C1q initiating the complement cascade. Tailoring a glycoprofile to suite the intended mechanism of action (ADCC vs. CDC) has been a strategy used to increase potency.

Although N-glycans do provide a number of positive benefits to monoclonal antibodies, there are certain factors that can have negative impacts including increased clearance and immunogenicity. It has been well documented that the addition of a Galα1-3Gal species effects recombinant antibody half-life due to its high concentration in human serum and results in recognition by antibodies present in the human body (Jefferis 2009; Liu et al. 2008). This specific glycan is produced by murine myeloid cell lines such as NS0 and SP2/0 so monitoring for this glycan in these cell expression systems is of critical importance. In addition, a sialic acid species (N-glycylneuraminic acid, NeuGc) expressed by NS0 cells is not observed in humans and can lead to immunogenic response (Cymer et al. 2018; Zhou and Qiu 2018; Reusch and Tejada 2015). Lastly, high-mannose has been connected to a faster clearance of IgGs which can be linked to the mannose receptor that is

present on macrophages and dendritic cells (Mimura et al. 2018). Table 2 describes a general overview of known glycan species effect on immunogenicity, efficacy, or clearance.

Table 2. Impact of glycan structure modifications on immunogenicity, efficacy, and clearance

Glycan Species	Property	Impact
Galα1-3Gal	Immunogenicity	Negative
Afucosylation	Efficacy	Positive
Terminal Galactose	Efficacy	Positive
High Mannose	Clearance	Negative
N-glycylneuraminic acid	Immunogenicity	Negative

Due to the broad range of effects brought about by these structures it is of critical importance to characterize and monitor these species during production, processing, and storage of monoclonal antibodies. A multitude of analytical methods are required to do complete glycoprotein characterization with high-performance liquid chromatography (HPLC) and mass spectrometry (MS) among the most common techniques used. High resolution intact mass spectrometry (HRMS) is used for preliminary analysis to determine overall mass and dominant glycoforms present in monoclonal antibodies. In order to do a full characterization and identify structures confidently, the N-glycan structures require release from the antibody followed by labeling and analysis using mass spectrometry and/or fluorescence to detect the glycan species.

Glycan Release and Labeling

Detailed glycan analysis is a multi-step process involving release, labeling and detection of released glycans. Sample preparation begins with the addition of a denaturant solution followed by heat to unfold the protein. Denaturation exposes the glycan structures attached to the Asn residues in the Fc region of the antibody. Next, the glycan species are removed from the protein using an enzyme; the most common enzyme used for N-glycan removal is peptide N-glycosidase F (PNGase F). PNGase F cleaves between the asparagine residue and the first GlcNAc residue on the core glycan structure (Man3GlcNAc2) of high mannose, complex, and hybrid structures (Anthony and Plummer 1987; Prien et al. 2015; Ruhaak et al. 2010). This reaction results in the conversion of the asparagine residue to an aspartic acid and an intermediate glycosylamine oligosaccharide (Figure 4) (Anthony and Plummer 1987; Prien et al. 2015; Ruhaak et al. 2010).

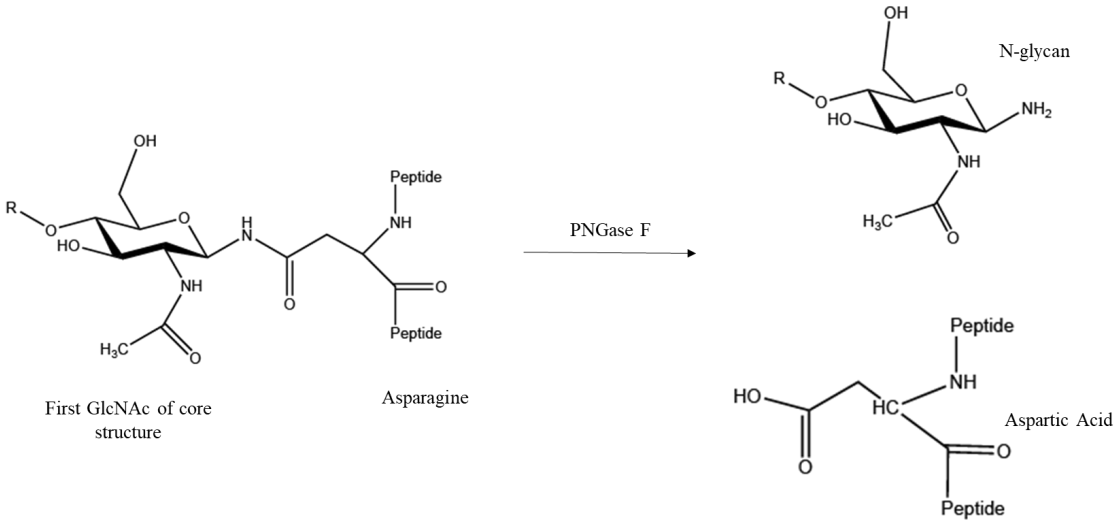
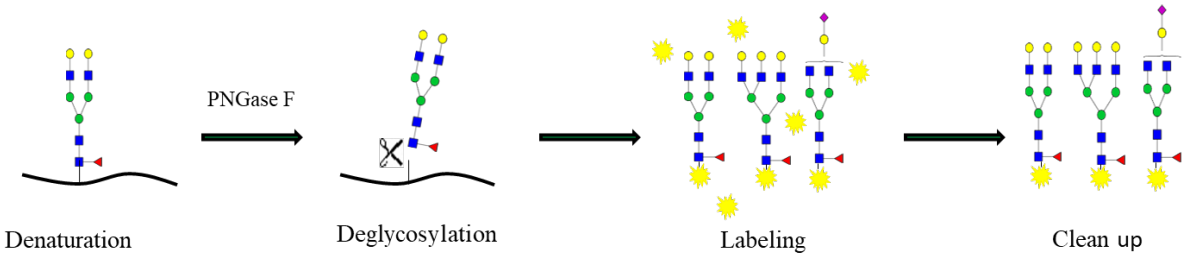


Figure 4. Removal of glycans from proteins using PNGase F

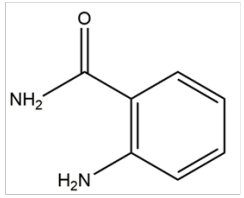
Released glycans are then labeled with a flurophore followed by the removal of any additional free flurophore that has not bound to glycan with a clean-up step. Figure 5A shows a general overview of the glycan release and labeling process. It is worth noting that the newest generation of labels (e.g. InstantPC in Figure 5B) rapidly react with the glycosylamine intermediate. Legacy labels (e.g. 2AB) required extended sample preparation times wherein the glycosylamine is allowed to exchange with water to form free reducing terminus, followed by a Shiff's base reaction to attach the dye. The increased efficiency of sample preparation afforded by the newer generation of dyes make them an attractive option in part due to resultant increased overall analytical throughput.

A.



B.

2-Aminobenzamide (2-AB)



InstantPC

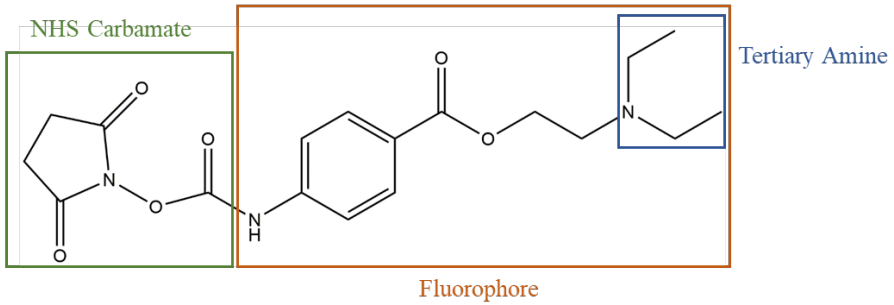


Figure 5. A. Overview of released glycan labeling and sample prep; B. Labels used for glycan analysis; 2-Aminobenzamide (2-AB) and InstantPC.

These new labeling techniques are preferred to methods such as the 2-AB label due to the higher ionization efficiency. The RapiFluor and InstantPC labels consist of a NHS-carbamate reactive group and a fluorophore containing a tertiary amine (Figure 6) (Lauber et al. 2015; Vainauskas et al. 2018). The NHS-carbamate group undergoes a self-quenching step resulting in amine byproducts which allows for labeling of N-glycans in very short period of time (minutes). This reaction results in a stable urea linkage between the N-glycan and fluorophore in under 5 minutes (Jones et al. 2016; Yan et al. 2019) (Figure 6).

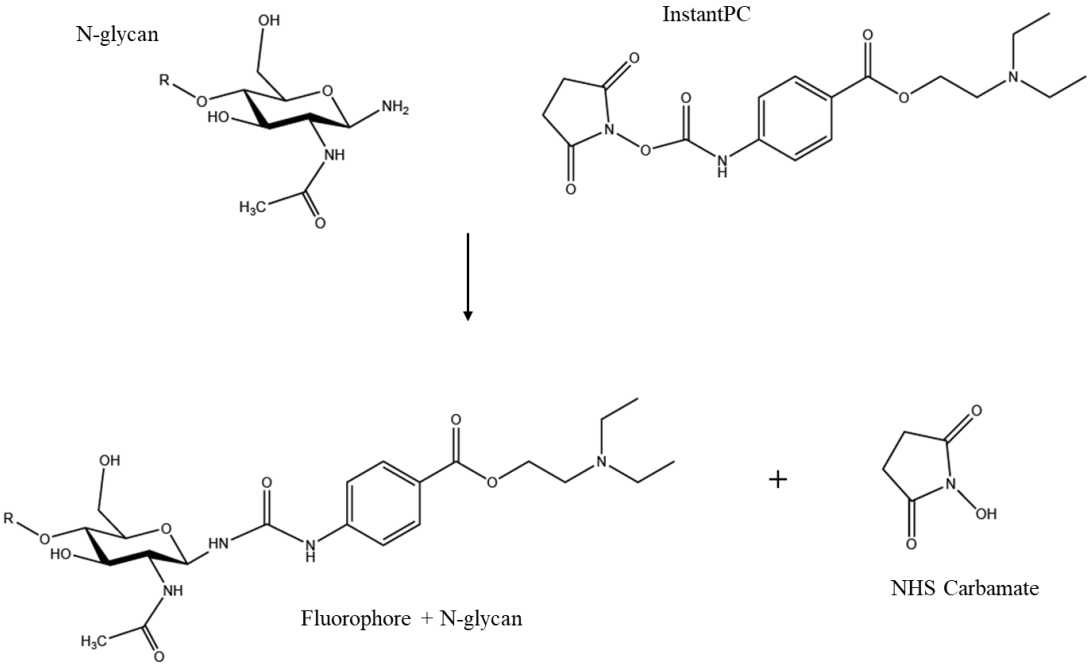


Figure 6. Labeling of N-glycan structure using InstantPC.

The addition of the tertiary amine structure on the flurophore which is easily protonated increases the sensitivity of the MS due to its high ionization efficiency. The increased sensitivity and rapid labeling rate of this compound makes it a predictable choice in

evaluating released glycans. Increased MS sensitivity allows for detection of low abundant glycans that are unable to be identified using other labels as well as with fluorescence.

Labeled glycans are separated using HPLC and then detected using fluorescence detection (FLD), mass spectrometry, or a combination of both. Hydrophilic interaction liquid chromatography (HILIC) is used to separate the glycan species before detection in an effort to yield quantitative information as each glycan has a single fluorophore attached to the reducing end of the glycan structure (Higel et al. 2013). HILIC is the preferred chromatographic method over RPLC for analysis of released glycans due to their hydrophilic composition. In RPLC glycans are not sufficiently retained on the column (Ruhaak et al. 2010; Guile et al. 1996). HILIC uses a high organic mobile phase (i.e. acetonitrile) with a polar stationary phase; analytes are eluted from the column in order of polarity with increasing aqueous concentrations (~50% in glycan analysis) (Guile et al. 1996; Han and Costello 2013). The eluted analytes are then detected using a high-pressure liquid chromatography-fluorescence detection (LC-FLD) or high-pressure liquid chromatography-mass spectrometry (LC/MS).

Mass Spectrometry

Mass spectrometry (MS) is an analytical technique used to characterize biological molecules using the mass-to-charge (m/z) ratio. Analysis involves ionization of a compound into the gas phase and measurement of the ion's mass-to-charge ratio and relative intensity to produces a mass spectrum (Dass 2007). Mass spectrometers are most commonly coupled to a high-pressure liquid chromatography (HPLC) system for analysis of proteins. HPLC uses a high pressure pump to flow a sample onto a tightly packed column containing a stationary phase that is used to separate the biomolecule of interest. The flow from the column enters the source of the mass spectrometer as a liquid which is then ionized for detection using the mass analyzer. One of the most popular ionization methods for proteins is electrospray ionization (ESI). ESI is considered a soft ionization technique that produces highly charged droplets that through the combination of heat/flow of nitrogen drying gas are shrunk to individual ions which are then detected by the mass analyzer (Dass 2007; Ho et al. 2003; Hoffman and Stroobant 2007; Gaskell 1997) (Figure 7 below).

Figure 7. Principle of mass spectrometry

There are many different types of mass analyzers with Time-of-Flight (TOF) being one of the most popular systems currently available. This system works by measuring the mass of the ions utilizing what is known as a flight tube. A high voltage ion pulser sends the

ions into the flight tube, reflected off an ion mirror that sits at the top of the tube, and then into the detector (Dass 2007; Hoffman and Stroobant 2007). The amount of time that it takes for the ions to reach the detector is used to determine the mass of the ion. The mass of the ion is directly proportional to the time at which it arrives to the detector; ions of smaller mass will arrive at the detector earlier than larger ions.

An additional set of mass analyzer components can be coupled to the system before the time-of-flight resulting in a tandem mass spectrometer known as a quadrupole time-of-flight (Q-TOF). A quad can either be operated as an all ion transfer mode or as a single filter-like system that allows only a single parent ion with a specific mass-to-charge ratio (m/z) to pass through it (Dass 2007; Hoffman and Stroobant 2007; Aebersold and Mann 2003). Often an individual parent ion is then be passed through a collision cell which causes collision associated disassociation (CID) by forcing the ions to collide with an inert gas such as nitrogen (Dass 2007; Aebersold and Mann 2003). The resulting fragments are then directed into the TOF and a spectrum of the ion fragments can be produced. A fragmentation spectrum can be used to confirm the original composition of the parent ion when parent ion m/z is not sufficient for confident identification. Mass spectrometry can be used to confirm the primary structure of a monoclonal antibody using peptide mapping, identify the mass and main glycoforms using intact mass analysis, and characterize N-linked glycosylation using a released glycan assay.

Intact Mass Spectrometry

High resolution intact mass spectrometry (HRMS) represents one of multiple orthogonal measurements of protein primary structure (Formolo et al. 2015). It provides a rapid method for determining the average mass of monoclonal antibodies in addition to identifying moderate to high abundance PTMs including C-terminal lysine clipping, N-terminal pyroglutamination and glycosylation. Intact mass analysis has the advantage over other middle down and peptide mapping approaches in that no sample preparation is required, thus allowing proteoform heterogeneity to be assessed at the intact level. Analysis includes a chromatographic separation of the protein using reversed-phase-high-performance liquid chromatography (RP-HPLC) followed by detection using a mass spectrometer. RP-HPLC uses a hydrophobic stationary phase that allows for the separations of species that have minor variations in structural heterogeneity (Hage and Carr 2011; Chen and Horváth 1995). Analysis begins with a high aqueous concentration which allows the biomolecules that are flowing through the column to adsorb to the stationary phase. A gradient elution with an increasing organic solvent (e.g. Acetonitrile) will decrease the polarity of the mobile phase which allows for elution of the biomolecules from the column (Hage and Carr 2011; Chen and Horváth 1995; Le and Bondarenko 2005; Sousa et al. 2017). A relatively high column temperature (~80°C) is recommended which decreases backpressure allowing for a higher flow rate resulting in enhanced elution and producing a desired peak shape (Hage and Carr 2011; Le and Bondarenko 2005; Sousa et al. 2017).

Intact mass analysis provides preliminary conformation on the primary amino acid sequence of a protein in addition to the major glycoforms that are common to all mAbs.

Identification and quantification of these glycoforms is performed using electrospray ionization (ESI) and data analysis is completed using deconvoluted spectra. ESI is used for large proteins to charge them with multiple ions which yields the m/z value is in a mass range that is detectable by current state of the art Q-TOFs. An m/z value is the ratio of relative mass to z (the number of charges) and is calculated using the equation below.

$$\frac{m}{z} = \frac{MW + zH}{z}$$

Mass-to-charge ratio. Where MW is molecular weight, z is the charge, and H is the mass of a proton (a constant of 1.007825).

Proteins have multiple ionizable groups (i.e. multiple lysines) which are ionized over multiple charge states requiring protein deconvolution for data analysis. This is completed using software with a mathematical algorithm to determine the real molecular weight from the measured m/z values. Species of the same mass that are ionized with different m/z values are combined to predict the zero order charge of the protein and produce a deconvoluted mass spectrum. The mass error (ppm) for the theoretically calculated masses of the proteoforms vs the mass provided from the deconvoluted spectra is used to confirm sequence identity as shown in the equation below.

$$\text{mass error (ppm)} = \frac{m_{theoretical} - m_{measured}}{m_{theoretical}} \text{ x } 10\text{^}6$$

Mass accuracy. Where $m_{theoretical}$ is the calculated mass based off of the known structure of mAbs and $m_{measured}$ is the mass calculated using the deconvluted spectrum.

Theoretical mass values are compared to experimental data for proteoform identification and are calculated based off of the common mAb structure which includes 2 light chains, 2 heavy chains, and 16 disulfide bonds. In addition, a pyroglutamic acid at the N-terminus end of the heavy chain and the absence of C-terminal lysine is also included in the calculation as these are common modifications observed in mAbs brought upon by processing during protein production (Formolo et al. 2015). An observed ppm value of ≤ 50 ppm from the theoretically calculated value is used as an indicator of confirmed sequence identity. A detailed analysis of analytical methods including the development of a high resolution intact mass spectrometry method for analysis of NISTmAb RM 8671 and several non-originator cell lines expressing the NISTmAb sequence will be described in the following sections.

NISTmAb Reference Material

The National Institute of Standards and Technology (NIST) mAb (NISTmAb) is a publicly available IgG1κ antibody performance standard useful for evaluation/development of state-of-the-art and emerging analytical measurement technologies. The material has been used extensively to evaluate current best practices and develop innovative analytical technologies. Preliminary characterization was performed on a single production lot, NISTmAb Primary Sample 8670 (PS 8670), currently reserved as the NIST in-house primary standard. Initial characterization identified the antibody as a ~150kDa homodimer that is composed of 2 heavy and 2 light chains that are linked through 16 total inter- and intra-chain disulfide bonds (Figure 8). The protein has a number of identified PTMs including C-terminal lysine clipping, N-terminal pyroglutamination, glycosylation of the heavy chains at Asn297, and deamidation as described in the ACS Symposium Book Series (Schiel et al. 2014; Schiel et al. 2015a; 2015b). Additional lots of the molecule were pooled and vialed at 10 mg/mL by NIST as the homogenized NISTmAb Reference Material 8671 (RM 8671) (Schiel and Turner 2018; Kashi et al. 2018). The physicochemical reference values for RM 8671 were assigned using qualified test methods representative of industry best practices and demonstrated to be homogeneous and stable (Kashi et al. 2018; Schiel et al. 2018). NISTmAb was used for all sample prep and analytical method development as outlined in the following 2 chapters.

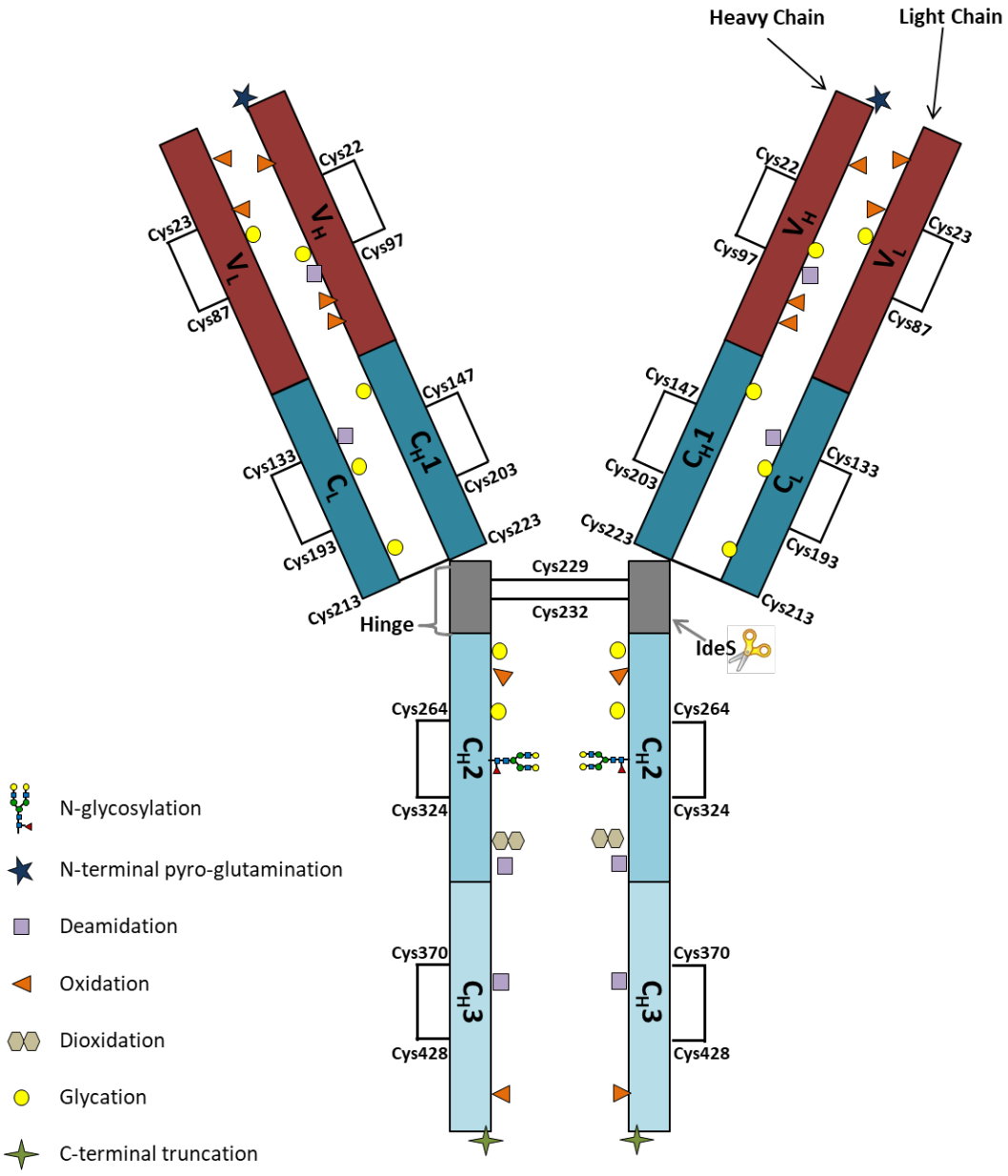


Figure 8. Structure of the NISTmAb reference material.

Biosimilars

There are more than 3,000 candidate biotherapeutics in some stage of clinical development in the United States today (Walsh 2014). Development of biosimilar alternatives is accelerating as many of the top selling biotherapeutic mAbs will lose their patent protection by the end of 2019, which accounts for more than $50 billion in the biologics market (Walsh 2014). In addition, the growth for biosimilars is aided by the pressure to reduce healthcare costs and the prevalence of chronic diseases in the aging population. The US Food and Drug Administration (FDA) has approved seven mAb biosimilar alternatives as of December 2017 and as these products become more prevalent in the biopharmaceutical industry the need for a pre-competitive, industrially relevant expression system increases (Biosimilar product information 2019). This is mainly due to the fact that mAbs are produced in mammalian cells with differences in growth conditions such as media composition (i.e. nutrient concentration), media conditions (i.e. pH), growth time, etc., affecting the yield, structure, PTMs, and function of the purified mAb. In addition, even if expressed in the same cell type, clonal variation can affect protein characteristics.

While biosimilars are being developed and approved, there is currently no publicly available material or dataset that can be used to define "how similar is similar". The development of a pre-competitive expression system would accelerate development of originator and follow on products as a collaborative test case to develop innovative process technology as continuous manufacturing, process intensification strategies, and process analytical technology. Advancements demonstrated with such a reference cell could then be adopted by biopharmaceutical manufacturers and may result in improvements in our

ability to prepare and define product quality attributes (PQAs), predict and tune PQAs through process control, and quantitatively correlate structural elements responsible for biological activity, among others (Kashi et al. 2018).

Non-originator NISTmAb

A series of non-originator NISTmAb cell lines that encode for the same primary amino acid sequence, but are not the cell line used for production of the NISTmAb, were developed to fulfill the gap of a pre-competitive biosimilarity test case. Three non-originator NS0 expression cell lines (referred to as NS0-59, NS0-60 and NS0-66) were developed and further evaluated for protein expression. The DNA construct encoded in the cell lines was constructed to directly encode the same primary amino acid sequence, although the specific codons used for the NISTmAb were unknown and were optimized for mouse expression using a third-party subcontractor. Expression using the non-originator cell lines was conducted in a shaker flask mode, which affords little process control capability other than the starting media, growth conditions, and cell density. This form of expression represents the earliest stage of bench-scale expression and under these conditions, each of the NS0 clones were capable of expressing >100 mg of antibody per L of culture.

In the following chapters, a series of analytical assays (high resolution mass spectrometry, size exclusion chromatography, glycan release analysis) were developed and employed to characterize the non-originator materials and directly compare them to the NISTmAb. These analyses are completed to indicate if these non-originator cell lines have the ability to express antibody that resembles the NISTmAb and if they can be considered suitable for further process optimization. The following two chapters will detail analytical assay development and analysis of both NISTmAb and a direct comparison to the three non-originator cell lines (NS0-59, NS0-60, and NS0-66) (Kashi et al. 2018).

Materials and Methods:

NISTmAb Reference Samples

Analytical analysis was performed on NISTmAb Reference Material (RM 8671 lot 14HB-D-002) concentration 10 mg/mL. Detailed information can be found in multiple recently published sources (Schiel and Turner 2018; Schiel et al. 2018; Turner et al. 2018; Turner and Schiel 2018).

Intact Mass Analysis

Non-Originator Samples

Expression and purification methods of non-originator samples from three different cell lines (NS0-59, NS0-60, NS0-66) can be found in the recently published article "Heterologous recombinant expression of non-originator NISTmAb"(Kashi et al. 2018). Protein concentration was based on UV measurements at 280 nm (using an extinction coefficient of 1.42 ml-1 cm-1/mg) and determined to be 3.75, 2.4, and 4.2 mg/m for cell lines NS0-59, NS0-60, and NS0-66 respectively.

Intact Mass Analysis

All analyses were performed on an Agilent Technologies 1290 Infinity II series liquid chromatography system coupled to an Agilent 6545 Accurate-Mass Q-TOF LC/MS. Separation was performed on a Polymeric Reversed Phase (PLRP-S) Column (5 µm particle size, 1000 Å pore size, 50 x 2.1 mm length, Agilent Technologies PN PL1912-1502). 0.1% Formic Acid in Acetonitrile (PN LS120) and 0.1% Formic Acid in water (PN LS118) were from Fisher Scientific. Vials (PN 5188-6591) and vial caps (PN 5182-0717) were from Agilent Technologies. ESI-L tune mix (PN G1969-85000) and ES-TOF

biopolymer reference mass kit (PN G1969-85003) were from Agilent Technologies. Samples were diluted to a 1mg/mL solution using 0.1% Formic acid in water.

For the optimized analysis of both NISTmAb and non-originator samples, the 1mg/mL dilution was made and a 1 µg aliquot was injected at flow rate of 0.4 mL/min (mobile phase A = 0.1% formic acid in water and mobile phase B = 0.1% formic acid in acetonitrile). The initial conditions of 5% B were held for 2 minutes, followed by a linear gradient of 5% to 80% B over 2 minutes. The column was held at 80% B for an additional 3 minutes, adjusted to 5% B over 2 minutes, and allowed to equilibrate for 4 minutes prior to the next run. Mass analysis of the protein was performed in high mass using a mass range of 500-5,000 *m/z* and the following instrument parameters: gas temperature 350°C, dry gas 8 L/min, nebulizer 45 psig, capillary voltage 5000 V, fragmentor 400V, skimmer voltage 65V, and sheath gas temperature 275°C. Instrument calibration was performed with Agilent ESI-L tune mix followed by reference ion calibration at *m/z* 1221.9906. Table 3 details the original and optimized method parameters below.

Table 3. Chromatographic and source parameters for intact mass analysis

Chromatographic Method		
Parameter	*Original Method*	*Optimized Method*
Flow Rate	0.400 ml/min	0.400 ml/min
Gradient	20-80% B over 18 minutes	20-80% B over 7 minutes
Column Temperature	80°C	80°C
Injection Quantity	1ug	1ug
Source Conditions		
Gas Temp/Flow Rate	350°C – 8L/min	350°C – 8L/min
Sheath Gas Temp/Flow Rate	275°C – 11L/min	275°C – 11L/min
VCap	5000V	5000V
Nozzle Voltage	1000V	2000V
Fragmentor	380V	380V
Skimmer	65V	65V
Mass Range	500-5,000 m/z	500-5,000 m/z

Detailed discussion of method optimization leading to the chromatographic and mass spectrometry instrumentation parameters will be described in the discussion section below. Theoretical mass values were calculated using the NIST Mass and Fragment Calculator (Software available at https://www.nist.gov/services-resources /software/nist-mass-and-fragment-calculator-software). Deconvolution of mass spectra was performed using the protein deconvolution feature of MassHunter Bioconfirm 7 and Qualitative Analysis 7 software and the maximum entropy algorithm. Deconvolution parameters are outline in Table 4 below.

Table 4. Deconvolution parameters for intact mass analysis

Average Scans	10% of peak height
Exclude TOF spectra if above	40% of saturation
Mass Range	140,000 – 150,000 Daltons
Baseline Factor	7.00
Peak Signal-to-Noise	30.0

SEC Analysis.

SEC was performed according to the previously qualified method (Turner et al. 2018). Briefly, all samples were analyzed on a Thermo Fisher Scientific/Dionex U3000 high-pressure liquid chromatography system using isocratic elution (100 mmol/L sodium phosphate supplanted with 250 mmol/L sodium chloride, pH 6.8) at 0.30 mL/min and monitored at 280 nm. Each sample was injected neat and equivalent to 60 µg of protein, with no dilution or buffer exchange, onto an Acquity UPLC Protein BEH SEC Column from Waters (1.7 µm particle size, 200 Å pore size, 4.6 x 150mm length). NISTmAb RM 8671 lot 14HB-D-002 concentration was previously determined to be 10.0 mg/mL, therefore 6 µL was injected. Injection volumes for each of the following cell lines were based on the UV measurements described above and are as follows; NS0-59 (16µL), NS0-60 (25µL), and NS0-66 (14.3µL).

Glycan Release Assay

Non-Originator Samples

Expression and purification methods of non-originator samples from three different cell lines (NS0-59, NS0-60, NS0-66) can be found in the recently published article "Heterologous recombinant expression of non-originator NISTmAb" (Kashi et al. 2018). Protein concentration was based on UV measurements at 280 nm (using an extinction coefficient of 1.42 ml-1 cm-1/mg) and determined to be 3.75, 2.4, and 4.2 mg/mL for cell lines NS0-59, NS0-60, and NS0-66 respectively. Samples were further concentrated prior to analysis using Amicon Ultra centrifugal-10k molecular weight filters (PN: UFC 501096) and the final concentrations are outlined in Table 5 below. 40µg of each sample was then used for glycan release analysis.

Table 5. Concentration of non-originator cell lines NS0-59, NS0-60, and NS0-66 post buffer exchange.

Sample	Sample Concentration 1 (mg/ml)	Sample Concentration 2 (mg/ml)	Average (mg/ml)
Clone 59	9.064	9.017	9.041
Clone 60	9.843	9.872	9.858
Clone 66	8.525	8.622	8.574

Glycan Release Analysis

Gly-X Glycan release kits (PN GX24-201PC), G0F/FA2 N-glycan standard (PN GKPC-302), and Human IgG N-linked glycan library (PN GKPC-005) were purchased from Prozyme (Agilent Technologies). All analyses were performed on Agilent Technologies 1290 Infinity II series liquid chromatography system or an Agilent Technologies 1290 Infinity II series liquid chromatography system coupled to an Agilent 6545XT Accurate-Mass Q-TOF. Separation was performed on an AdvanceBio Glycan

Map Rapid Resolution HD column (1.8 µm particle size, 2.1 x 1500 mm length, Agilent Technologies PN 859700-913). LC/MS grade 0.1% Formic Acid in Acetonitrile (PN 34668) and 0.1% Formic Acid in Water (PN 60-044-849) were from Honeywell. 50mM ammonium formate buffer was prepared by dissolving 3.1527g of ammonium formate in 900mL of LC/MS grade water pH was adjusted to 4.4 using formic acid and diluted to 1000mL with LC/MS grade water.

For the optimized method of both NISTmAb and non-originator samples, the samples were diluted 1 part sample with 3 parts of 1:1 acetonitrile:dimethylformamide (ACN:DMF) and a 4 µl aliquot was injected at a flow rate of 0.400 mL/min (mobile phase A = 50mM ammonium formate pH 4.4 and mobile phase B = 0.1% formic acid in acetonitrile). The initial conditions of 82% B were held for 2 minutes, followed by a linear gradient of 82 to 40% B over 49 minutes. The column was held at 40% B for an additional 1.5 minutes, adjusted to 82% B over 0.5 minutes, and allowed to equilibrate for 18 minutes prior to the next run. Fluorescence detection wavelengths were excitation-285 nm and emission-345 nm. Mass analysis was performed in standard mass using a mass range of 300-3000 m/z and the following instrument parameters: gas temperature 150°C, dry gas 9 L/min, nebulizer 35 psig, capillary voltage 2800 V, fragmentor 100V, skimmer voltage 65V, and sheath gas temperature 350°C. Instrument calibration was performed with Agilent ESI-L tune mix in fragile ion mode followed by reference ion calibration at m/z 1221.9906 and 322.0481. Detailed discussion of method optimization leading to the chromatographic and mass spectrometry instrumentation parameters will be described in the discussion section below. Vials (PN 5188-6591) and vial caps (PN 5182-0717) were

from Agilent Technologies. ESI-L tune mix (PN G1969-85000) and ES-TOF biopolymer reference mass kit (PN G1969-85003) were from Agilent Technologies.

Buffers

Buffers with a concentration of 25mM or 50mM were prepared according to the formulations descried below.

25mM Histidine pH 6.0

Formulation buffer (12.5 mmol/L L-Histidine/12.5 mmol/L L-Histidine HCl, pH 6.0) was prepared by dissolving 1.3129g Histidine monohydrochloride monohydrate and 0.9704 g L-Histidine and dilute with ~450 mL LC/MS water. pH was adjusted to 6.0 using hydrochloric acid and diluted to 500mL with LC/MS grade water. Buffer was filtered using a 0.22 µm cellulose acetate membrane into a sterile plastic bottle

25mM HEPES pH 7.9

Buffer was prepared by dissolving 1.4890g of HEPES in 200mL of LC/MS grade water. pH was adjusted to 7.9 using sodium hydroxide and diluted to 250mL with LC/MS grade water.

50mM HEPES pH 7.9

Buffer was prepared by dissolving 2.9780g of HEPES in 200mL of LC/MS grade water. pH was adjusted to 7.9 using sodium hydroxide and diluted to 250mL with LC/MS grade water.

Samples requiring buffer exchange prior to analysis used the Zeba 7k MWCO spin desalting columns (PN 98883) and the concentration was measured using UV measurements at 280 nm (using an extinction coefficient of 1.42 ml-1 cm-1/mg). Data

analysis of released glycans was performed using MassHunter BioConfirm 10.0 software and PCDL Manager software (Agilent Technologies). Theoretical mass values were calculated using NIST Glyco Mass Calculator https://www.nist.gov/static/glyco-mass-calc/). Analysis parameters are outline in Table 6 below.

Table 6. MassHunter BioConfirm Method Analysis parameters

Match Tolerance	
Masses	+/- 10 ppm
Retention Times	+/- 0.350 minutes
Expansion of Values for Chromatogram Extraction	+/- 20.0
Allowed Glycan Ion Species	+H, +Na, +K
Scoring – Contribution to Overall Score	
Mass Score	100.00
Isotope Abundance Score	60.00
Isotope Spacing Score	50.00
Retention Time Score	100.00
Scoring – Expected Data Variation	
MS mass	2.0 mDa + 5.6 ppm
MS isotope abundance	7.5%
Retention Time	0.250 minutes

Results

Size Exclusion Chromatography

A SEC method was previously qualified during RM 8671 value assignment which yielded a typical mAb profile consisting of a dominant monomer peak, low abundance LMW and HMW species, and a void volume peak arising from the L-histidine buffer (elution time ~6.25 minutes) (Turner et al. 2018). Chromatograms for RM 8671 and each of the non-originator materials are shown in Figure 9. Each non-originator material demonstrated an increase in the multiplicity and relative abundance of high molecular weight species (ranging from ≈4% to ≈9%, Figure 9) as well as a decrease in the LMW species originally identified in RM 8671 (peak at ≈ 5 min was decreased from ≈0.2% in RM 8671 to <0.1% in the non-originator materials). The monomer peak demonstrated increased tailing as evident from the increased peak asymmetry (from ≈1.3 to 1.5, Table 7). The increased peak tailing is likely due to the presence of relatively large fragments and/or disulfide variants that are unresolved by SEC. Lastly, a new species was identified migrating at approximately 5.9 minutes. Preliminary evaluation of cell-free media subjected to downstream processing indicate the peak arises only in the presence of active cell expression. Additional fraction collection and MS analysis will be pursued in the future to conclusively identify this impurity; it was not included in relative abundance calculations in Table 7 (Kashi et al. 2018).

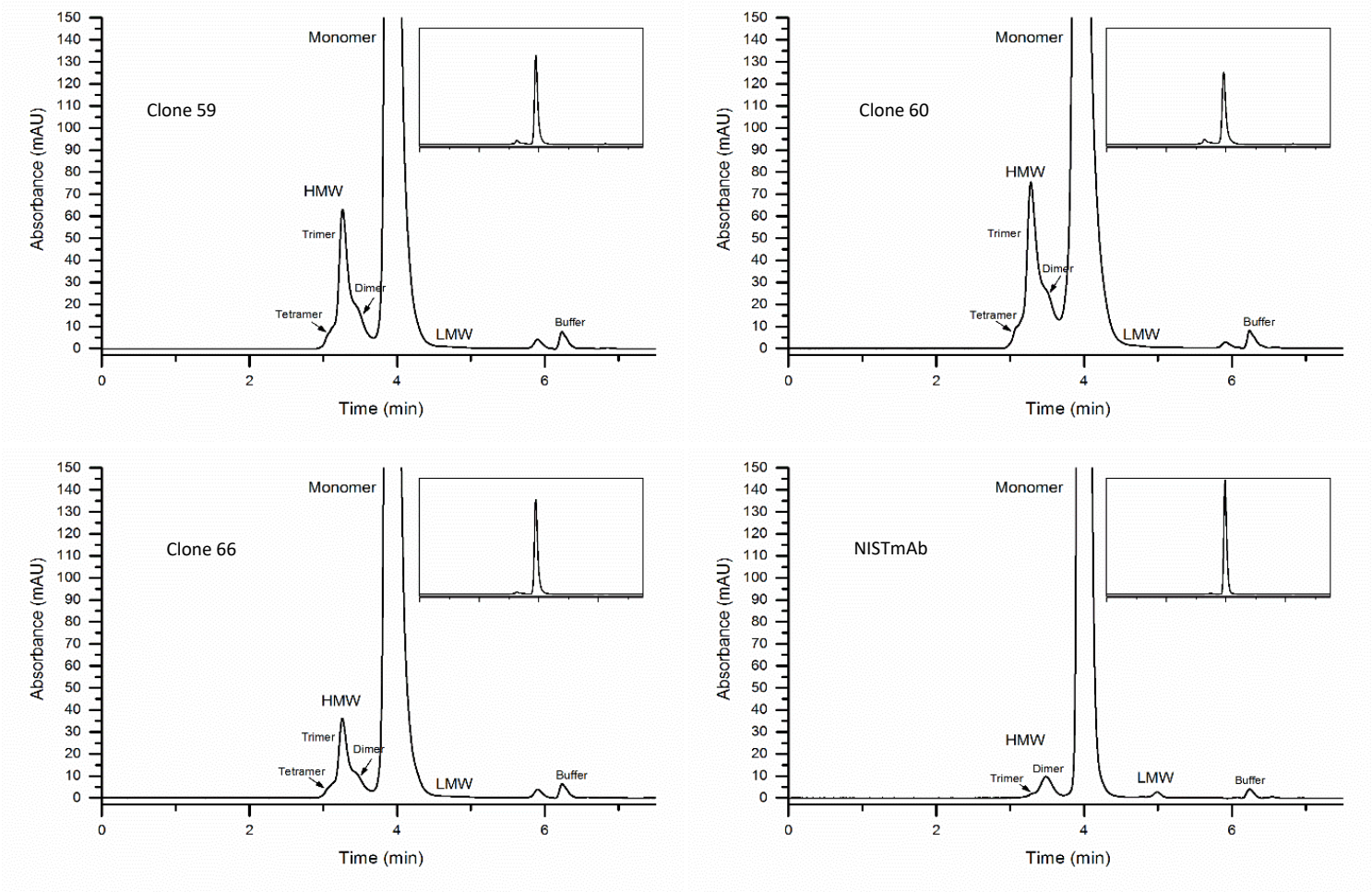


Figure 9. Size-exclusion chromatograms of non-originator cell lines and NISTmAb.
A. NS0-59, B. NS0-60, C. NS0-66, D. NISTmAb RM 8671. HMW, high molecular weight; LMW, low molecular weight. (Kashi et al. 2018)

Table 7. SEC results for non-originator cell lines and NISTmAb (n = 3).(Kashi et al. 2018)

	NISTmAb (8671)	NS0-59	NS0-60	NS0-66
% RA[a] High Molecular Weight (HMW)	3.14 ± 0.04 (Conforms to expectation)	5.24 ± 0.00	9.42 ± 0.04	3.83 ± 0.00
% RA Monomer	96.65 ± 0.04 (Conforms to expectation)	92.92 ± 0.04	90.34 ± 0.04	95.85 ± 0.00
% RA Low Molecular Weight (LMW)	0.20 ± 0.00 (Conforms to expectation)	0.08 ± 0.00	0.09 ± 0.00	0.11 ± 0.00
Total Peak Area (HMW + LMW + Monomer)	193.7± 0.4	201.8 ± 0.1	195.9 ± 0.5	206.9 ± 0.25
Monomer Peak Asymmetry	1.27 ± 0.01	1.5 ± 0.00	1.5 ± 0.00	1.46 ± 0.00

Intact Mass Analysis

Intact mass analysis method optimization

Method optimization for intact analysis focused on two areas, chromatographic separation and source conditions of the mass spectrometer. Intact mass analysis methods have been studied in detail over the last few years as monoclonal antibodies are one of the top selling therapeutic products on the market, due to this there are multiple standard method conditions that were used as starting conditions for this method (Formolo et al. 2015; Wong 2017). The starting method conditions are outline in Table 3 in the materials methods section above. Initial run conditions produced one broad main peak with an intensity of 2.0×10^8 and a small peak co-eluting at the end of the main peak believed to be a hydrophobic variant as demonstrated in Figure 10 (top) A deconvoluted spectra (figure 10, bottom) is evaluated for the main glycoforms known to be present in mAbs (FA2, FA2G1, FA2G2) in addition to high abundance PTMs (C-terminal lysine) (Schiel 2012; Formolo et al. 2015; Prien et al. 2015; Hilliard et al. 2017).

Figure 10 (bottom) demonstrates rather sharp peaks overlaid on a circular arch, which we will refer to as the "signal hump". The hump can be a result of several factors including incomplete desolvation or adduction with salts or other system contaminants. The peak-to-hump ratio can also be small when there is significant heterogeneity in the protein as this splits the protein concentration across multiple proteoform signals. A high peak-to-hump ratio or low signal hump is desired. The peak-to-hump ratio in the original run conditions is higher than desired and in some cases can be minimized with ion source/optics optimization. Individual source parameters were changed and built upon the previous method to observe the effect on the desolvation of the sample.

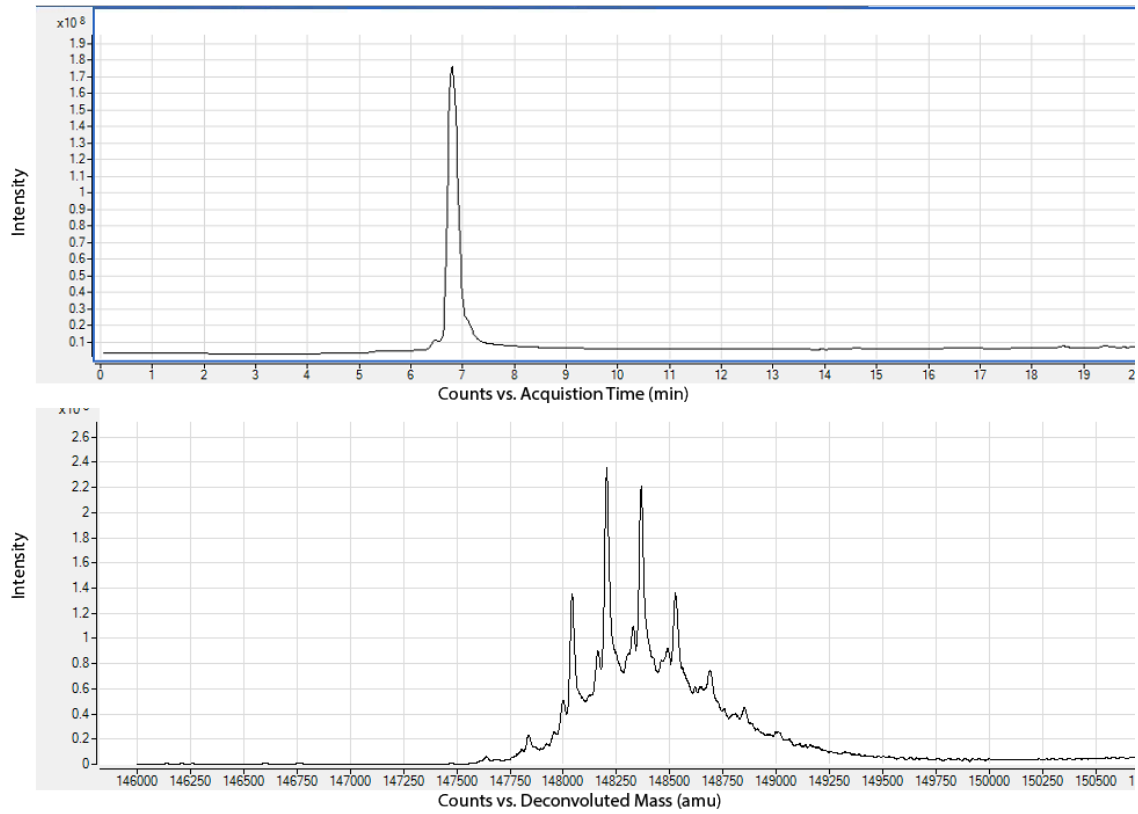


Figure 10. Total Ion Chromtaogram (TIC, top) and deconvoluted spectra for NISTmAb (bottom).

The rate of desolvation is dependent on components of the ion source including the nozzle voltage, nebulizer gas pressure, and gas temperatures/flow rates (Dass 2007; Hoffman and Stroobant 2007; Gaskell 1997; Banerjee and Mazumdar 2012). The "nozzle" is located at the very tip of the electrospray needle, this voltage is the amount of charge that the liquid exiting from the LC is exposed to is as it enters the electrospray needle (Dass 2007; Hoffman and Stroobant 2007; Gaskell 1997; Banerjee and Mazumdar 2012). Monoclonal antibodies are large molecules (~150 kDa) and higher voltage aids in the charging of electrospray droplets which, in turn, increases the number of protons attached to the analyte and thus the charge to allow detection within the mass range of commercial

benchtop mass spectrometers. Sheath gas is high temperature nitrogen gas that surrounds the nebulizer needle and assists in drying the charged droplets that enter the source. Higher temperature can also increase the rate of desolvation but too high of a temperature can cause fragmentation, so optimization is required to prevent this from occurring (Dass 2007; Hoffman and Stroobant 2007; Gaskell 1997; Banerjee and Mazumdar 2012). In terms of chromatographic parameters, the elution gradient is typically the most essential component of a separation in an attempt to provide resolution of various components. In the current case, however, the desire is to have a rapid gradient allowing a narrow peak wherein all glycoforms co-elute in a single band. Differentiation of glycoforms is then performed in the mass spectrometer rather than via chromatographic resolution

Several parameters were changed across 3 methods beginning with increasing the nozzle voltage, skimmer, and sheath gas flow rates/temperature. Nozzle voltage was selected as this will have the greatest effect on the desolvation, followed by skimmer voltage, and sheath gas flow rate/temperature. Changes to these parameters are outlined in Table 8 below.

Table 8. **Source parameters optimized for intact mass analysis analytical method.**

Method 1	Nozzle voltage: 2000V Skimmer: 65V Sheath Gas Temperature: 275°C
Method 2	Nozzle voltage: 2000V Skimmer: 140V Sheath Gas Temperature: 275°C
Method 3	Nozzle voltage: 2000V Skimmer: 140V Sheath gas temperature: 350°C

A comparison of these 4 resulting spectra can be seen in Figure 11 below. Analysis of the deconvoluted spectra identified little variability between the 4 methods with respect to an increase in the peak-to-hump ratio. A slight decrease in signal intensity (~2.5 $x10^8$ vs 2.0 $x10^8$) seen with an increase in the skimmer voltage (method 2) and sheath gas temperature (method 3). Increasing the nozzle voltage resulted in a minimal but non-quantitative increase in the peak-to-hump ratio. Of the 3 source parameters adjusted, nozzle voltage was the only change that was kept as an optimized parameter as there was a minimal change in the peak-to-hump ratio. This was a previously well-developed method prior to optimization so little adjustment on the source parameters was required to obtain optimal results (Wong 2017).

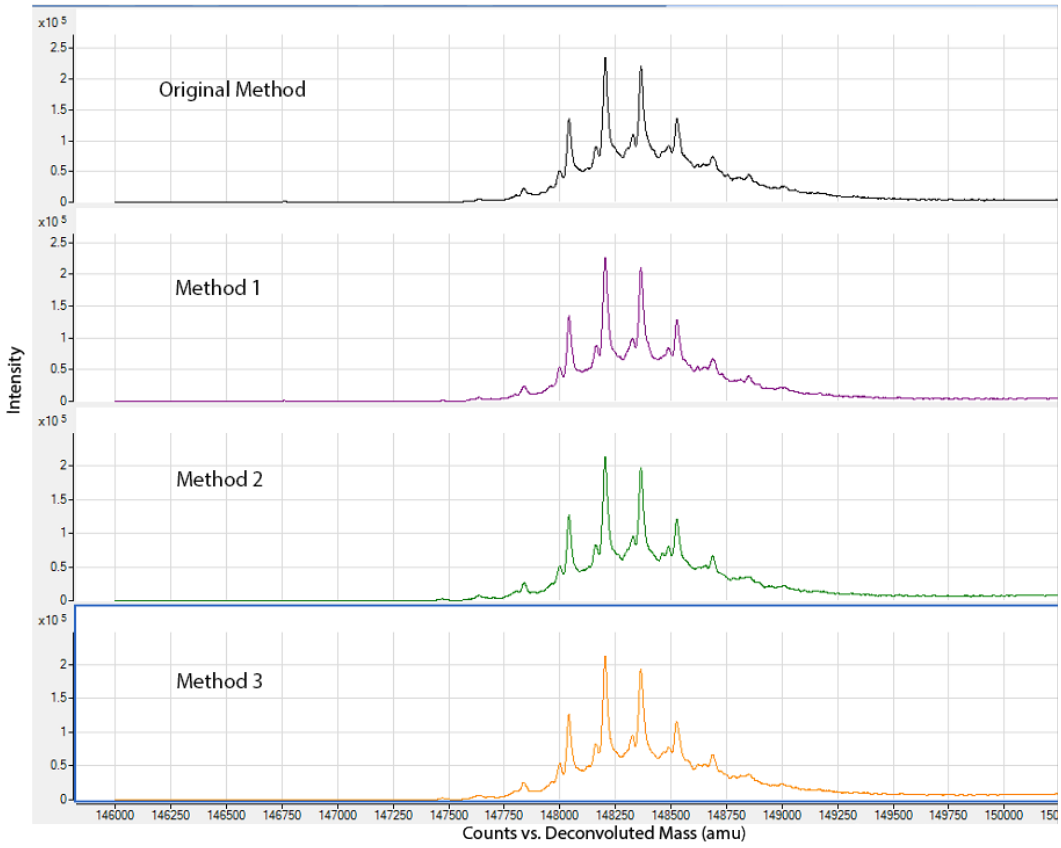


Figure 11. Comparison of deconvoluted spectra of NISTmAb using different source/ion optimization parameters

In addition to achieving an efficient separation, the gradient was adjusted as a high-throughput method was desired without losing spectral quality and to decrease the amount of bleed coming from the column. The column used for separation (polymeric reversed-phase) was identified to have column bleed when using high organic solvents which is required for elution of the intact mAb. The stationary phase of this column was proprietary at the time of analysis, but believed to contain a polymer species which was evident in co-elution with the protein at high organic concentration. The method was changed from a 30-minute total run time to a 13 minute run time which adjusted the gradient). This decreased the length of time the column was exposed to high organic solvent as the protein

was being eluted resulting in a sharper, more resolved peak. Analysis of the shorter gradient shows initially more bleed from the column (peak tailing shown between 3.5-5 minutes, Figure 12). This is likely due to the higher organic concentration earlier in the gradient which allows elution of both the protein and the polymer species more quickly. Although this seems to produce a more undesired result this column bleed is much smaller than that observed in the original method, which occurred between 7-20 minutes and led to an extensive tail on the main peak (Figure 12). The overall run time was also decreased by 17 minutes without effecting the quality or intensity of the spectra producing a method that allows for rapid analysis that can be used for development of non-originator cell lines as discussed in the sections below. In addition, a column pre-wash step of ~1 hour at 95% organic is effective in decreasing the amount of residual polymer present in each injection during the run sequence. This wash was performed before any subsequent analysis.

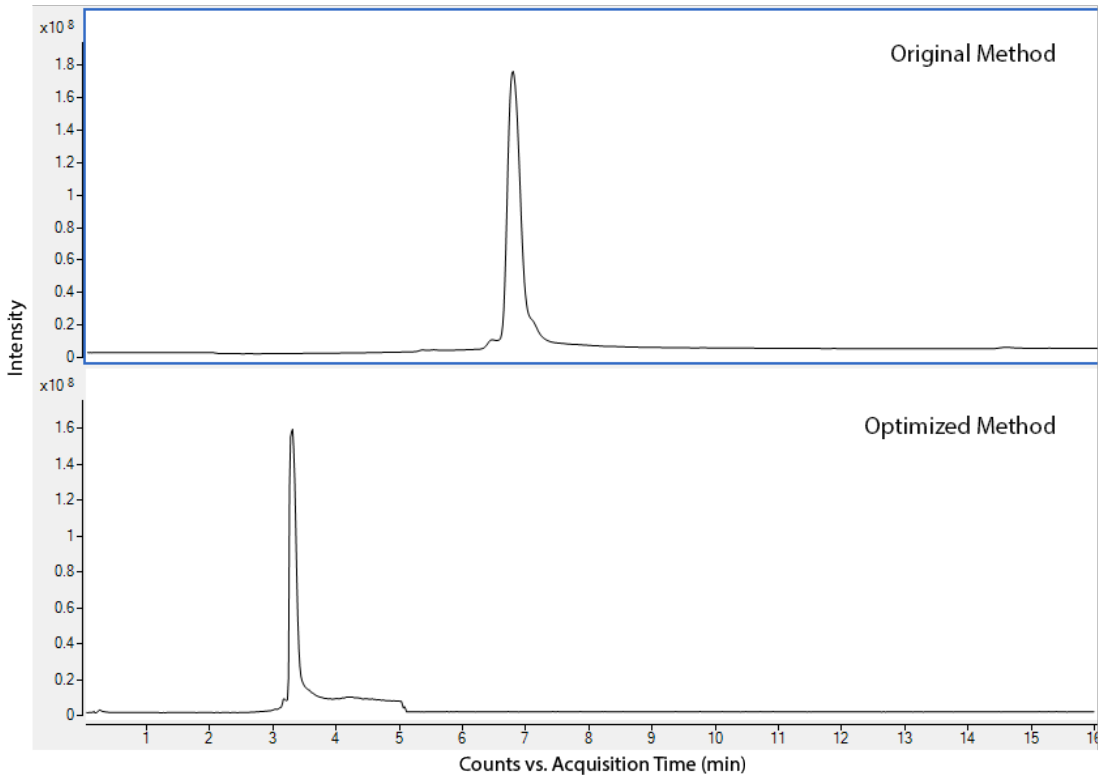


Figure 12. Elution of NISTmAb using extended vs rapid gradient.

Table 9. Gradient parameters for original vs. optimized method of intact mass analysis

Original Method		**Optimized Method**	
Time (min)	*%B (Acetonitrile w 0.1% Formic acid)*	*Time (min)*	*%B (Acetonitrile w 0.1% Formic acid)*
0.0	20	0.0	5
2.0	20	2.0	5
20.0	80	4.0	80
21.0	95	7.0	80
25.0	95	9.0	5
26.0	20	13.0	5
30.0	20		

Intact NISTmAb mass analysis using final method.

Previous studies of PS 8670 identified 3 major glycoforms as G0F/G0F, G0F/G1F, and G1F/G1F each of which were consistent with 2 light chains, 2 heavy chains, the expected 16 disulfide bonds, pyroglutamic acid at both heavy chains N-termini, and no C-terminal lysine on either heavy chains. A series of low to moderate abundance proteoforms were also identified consistent with the addition of C-terminal lysine and/or hexose residues (Formolo et al. 2015; Kashi et al. 2018). Preliminary assignments for RM 8671 using the method described above (Figure 13) are consistent with the proteoforms previously observed for PS 8670 and were measured within 50 ppm of the theoretical values as (Table 10). Peak identification were made by comparing the measured zero-charge state mass to theoretical values based on purported amino acid/glycan compositions and calculated using the NIST Mass and Fragment calculator (Kashi et al. 2018; Kilpatrick et al. 2012)

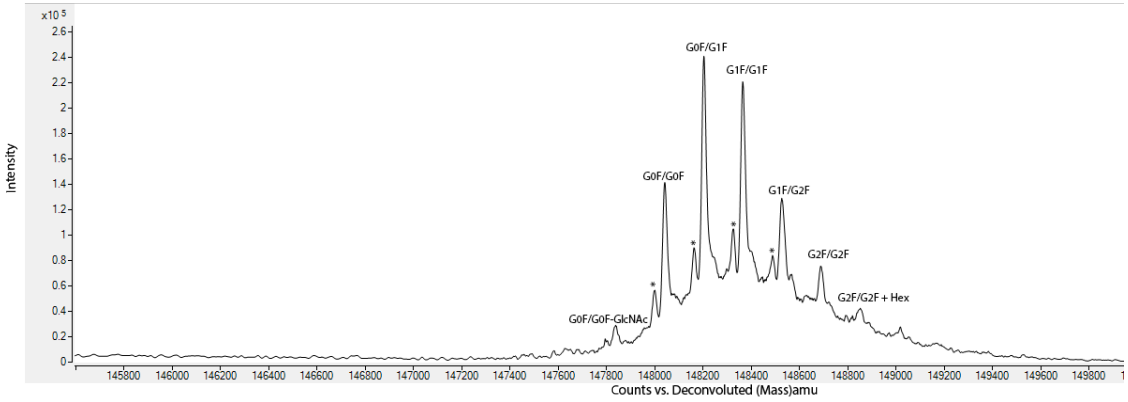


Figure 13. Deconvoluted spectra of NISTmAb using optimized intact mass analysis parameters. Peaks marked with a * indicated the presence of a C-terminal lysine on the previous glycoform.

Table 10. Measured molecular mass values for NISTmAb

Proteoform	[a.]Total Theoretical Mass (Da)	Observed Mass (Da)	PPM
G0F/G0F - GlcNAc	147834.7	147838.0	22.2
G0F/G1F - GlcNAc	147996.1	148000.4	28.8
G0F/G0F	148037.2	148044.1	47.1
G0F/G0F + K*	148165.1	148164.0	-7.4
G0F/G1F	148199.3	148205.4	41.3
G0F/G1F + K*	148327.5	148323.5	-26.9
G1F/G1F	148361.4	148368.3	46.5
G1F/G1F + K*	148489.6	148485.6	-27.1
G1F/G2F	148523.6	148529.0	36.3
G1F/G2F + K*	148651.5	148634.4	-115.3
G2F/G2F	148685.7	148689.3	24.3
G2F/G2F + Hex	148847.7	148847.4	-1.7
G2F/G2F + 2Hex	149010.0	149009.9	-0.7

[a.]Theoretical values include 16 disulfide bonds, two N-terminal pyroglutamic acids, and no C-terminal Lys (K), unless otherwise noted. Theoretical values were calculated using the NIST Mass and Fragment calculator. Peaks labeled with a * indicate presence of a C-terminal lysine on the previous glycoform. Glycan assignments were made based on putative compositions expected from the glycan biosynthetic pathway.

Non-Originator Analysis

Non-originator materials were run using the optimized method parameters described in the materials and methods sections above; deconvoluted intact HRMS spectra for RM 8671 and each of the non-originator materials are shown in Figure 14. The three main glycoforms (G0F/G0F, G0F/G1F, and G1F/G1F) are also present in each of the NS0-59, NS0-60, and NS0-66 non-originator cell lines. This commonality between the protein lots indicates the non-originator cell lines are producing a substantial quantity of proteoform consistent with that of NISTmAb including 2 heavy chains, 2 light chains, 16 disulfide bonds, and pyroglutamic acid at the N-terminus of both heavy chains. There is, however, a significant shift toward higher galactose content for each of the non-originator cell lines; the most dominant glycoform being G1F/G1F or G1F/G2F compared to the G0F/G1F in RM 8671. A series of proteoforms consistent with the addition of one C-terminal lysine are also observed in NISTmAb and all non-originator materials.

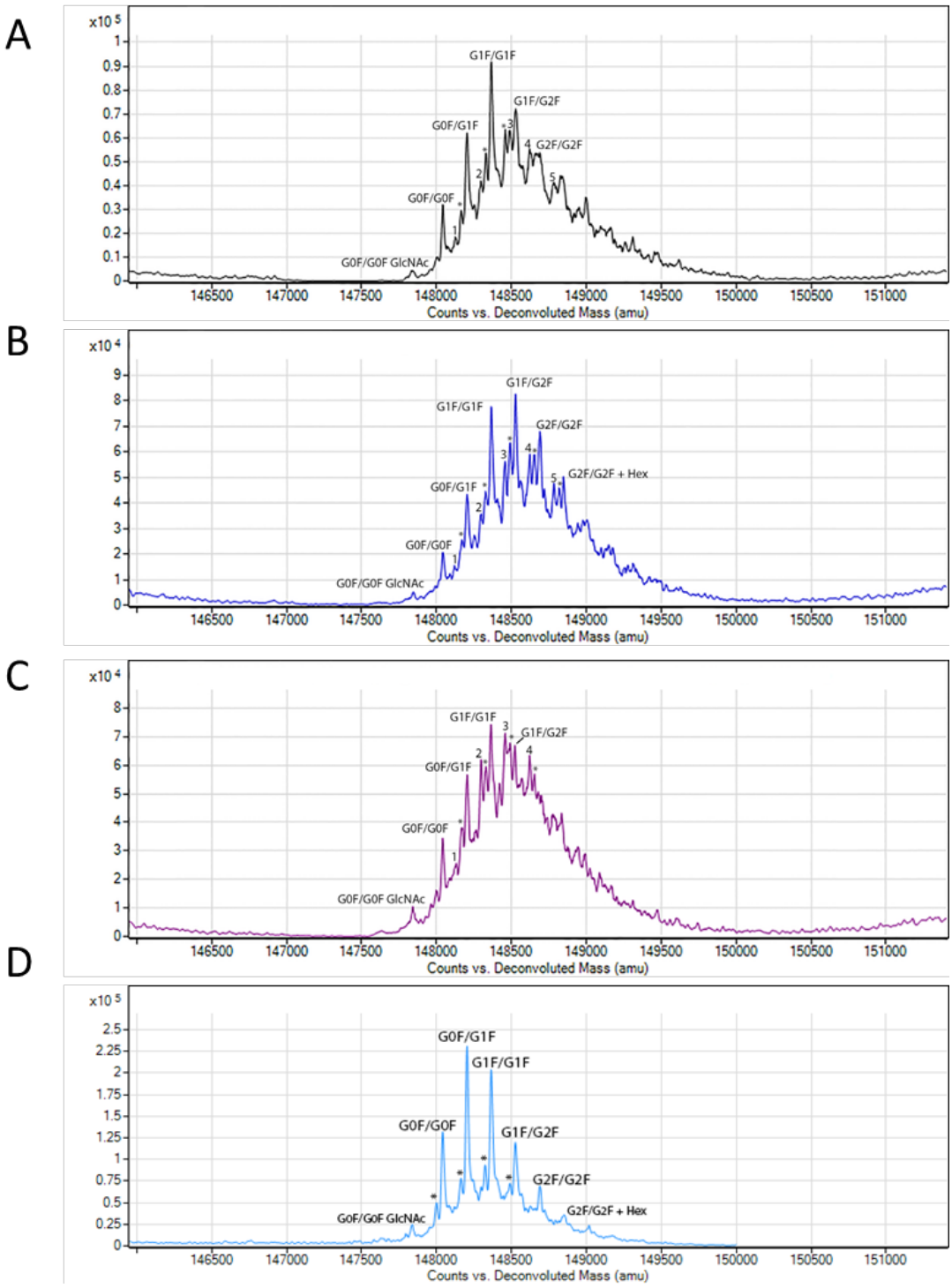


Figure 14. Deconvoluted intact mass spectra of NISTmAb and non-originator cell lines. Peaks marked with a * indicated the presence of a C-terminal lysine on the previous glycoform. Peaks with a number of 1-5 indicate the presence of a partially uncleaved Ser of the signal peptide (Fig. 1). A, NS0-59; B, NS0-60; C, NS0-66; D, NISTmAb RM 8671.(Kashi et al. 2018)

Lastly, a series of new peaks not present in NISTmAb was also observed. These peaks likely correspond to an alternative cleavage site of the light chain signal peptide which resulted in the addition of a serine residue on all of the non-originator lots. Processing of a mAb expressed in murine suspension culture involves the clipping of an N-terminal signal peptide upstream of these penultimate residues. In development of the non-originator cell lines, the signal peptide used for expression of NISTmAb was unknown; therefore murine signal peptides were selected by a third-party contractor (Kashi et al. 2018). Under complete processing of the signal peptide; the entire sequence depicted in Figure 15 will be cleaved, resulting in expression of non-originator products with the same terminal amino acid sequence as NISTmAb (Figure 15 red arrows). However, in a small percentage of product an additional Ser residue was identified which may be a result of promiscuous cleavage at the three Ser stretch of the light chain signal peptide (Figure 15 blue arrow) (Petersen et al. 2011)

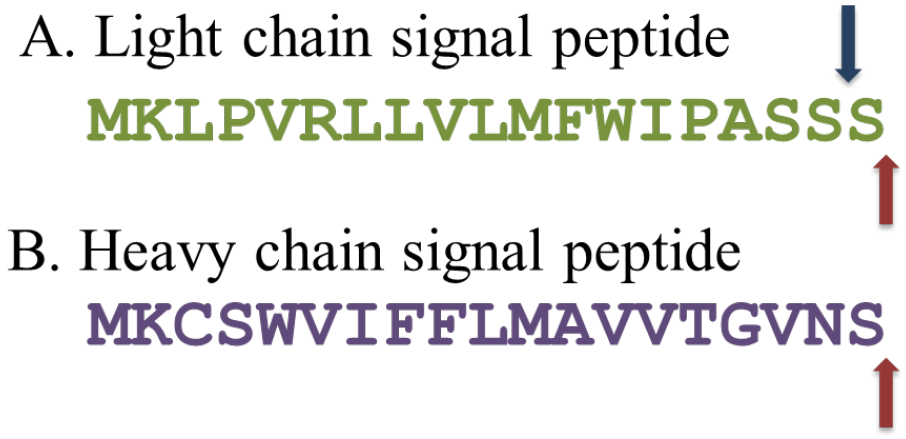


Figure 15. Amino acid sequences of the signal peptides encoded in the non-originator cell lines for the light and heavy chains.
A. The signal peptide of the light chain is shown. The red arrow shows the expected cleavage site and the blue arrow shows the alternate cleavage site.
B. The signal peptide of the heavy chain is shown. The red arrow shows the expected cleavage site.

The addition of this residue can be seen on all major glycoforms of lots NS0-59, NS0-60, and NS0-66 as indicated by the numeric value of 1-5 represented in Figure 14 and Table 11. A decrease in relative peak intensity for all main "target" glycoforms was observed in each of the non-originator lots due to the altered glycoprofile, increase in %C-terminal lysine, and new alternative signal peptide cleavage peaks. In NS0-59 and NS0-60 the target proteoforms still have a high relative peak intensity, yet in NS0-66 the altered primary sequence variants begin to dominate. In this mass range, it is also possible that additional unresolved PTMs differ between the samples including oxidation, deamidation, unpaired cysteines, and/or various combinations of glycoforms. This may have contributed to (in addition to the inherently lower signal to noise) the lower mass accuracy observed for some of the non-originator peaks. A series of changes in PTMs and minor sequence modifications as well as increase in heterogeneity between NS0-50, NS0-60, NS0-66, and NISTmAb was observed based on HRMS. This is not unexpected considering the use of non-originator expression systems and a unique downstream purification process (Kashi et al. 2018). Importantly, however, the major proteoforms expressed in the NISTmAb are also seen in all corresponding non-originator lots and do resemble to some degree the NISTmAb (Kashi et al. 2018). Intact mass analysis provides information with respect to major glycoforms and moderate to high abundance primary sequence modifications, however, additional analysis is required to evaluate several characteristics of the non-originator species including the decrease in relative peak intensity for the main glycoforms, increased heterogeneity, and decrease in signal-to-noise ratio.

Table 11. Measured molecular mass values for NISTmAb and non-originator cell lines NS0-59, NS0-60, and NS0-66.

Proteoform	[a.]Theoretical Mass (Da)	Observed Mass (Da)			
		NIST mAb 8671	*NS0-59*	*NS0-60*	*NS0-66*
G0F/G0F -	147833.9	147838.0	147841.0	ND	147841.8
G0F/G1F -	147996.1	148000.4	ND	ND	ND
G0F/G0F	148037.2	148044.1	148043.9	148044.9 [b.]	148042.9
G0F/G0F + S^1	148124.2	ND	148128.2	148124.5	148127.7
G0F/G0F + K*	148165.3	148164.0	148166.0	148173.1 [b.]	148170.3
G0F/G1F	148199.3	148205.4	148207.7[b]	148205.7	148205.0
G0F/G1F + S^2	148286.4	ND	148298.0	148298.8 [b.]	148299.8 [b.]
G0F/G1F + K*	148327.5	148324.2	148329.9	148329.6	148330.4
G1F/G1F	148361.4	148368.3	148370.0	148367.6	148367.5
G1F/G1F + S^3	148448.5	ND	148461.1	148456.8 [b.]	148460.3 [b.]
G1F/G1F + K*	148489.6	148485.6	148491.6	148492.2	148488.3
G1F/G2F	148523.6	148528.9	148528.5	148527.8	148521.3
G1F/G2F + S^4	148610.7	ND	148626.9	148620.9 [b.]	148622.0 [b.]
G1F/G2F + K*	148651.8	148657.1	ND	148652.0	148651.2
G2F/G2F	148685.7	148689.3	148688.6	148688.9	ND
G2F/G2F + S^5	148772.8	ND	148788.0	148784.9 [b.]	148788.0
G2F/G2F +	148847.7	148847.4	ND	148848.2	ND

[a.]Theoretical values include 16 disulfide bonds, two N-terminal pyroglutamic acids, and no C-terminal Lys (K), unless otherwise noted. Theoretical values were calculated using the NIST Mass and Fragment calculator. Peaks labeled with a * indicate presence of a C-terminal lysine on the previous glycoform. Peaks labeled with a number of S^{1-5} indicate the presence of a partially uncleaved serine of the leader peptide. Glycan assignments were made based on putative compositions expected from the glycan biosynthetic pathway.

Glycan Release Assay

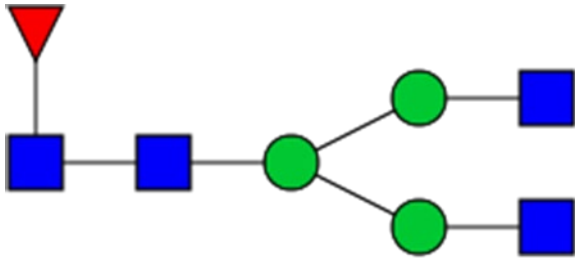
MS Condition Optimization

Before optimizing the sample conditions minimal mass spectrometry parameter adjustment was completed. The Gly-X protocol along with several other publications have outlined specific source and chromatographic conditions for N-glycan analysis (Yan et al. 2019; Hilliard et al. 2017; Prozyme 2017). Preliminary analysis of source conditions used a combination of method parameters from the publications listed above to compare with a method originally used for peptide analysis in house which utilizes high gas flow rates/temperatures; parameters are outline in Table 12 below.

Table 12. Source parameters for original and in-house peptide method MS analysis

Conditions	Original Method	In-House Peptide Method
Gas Temperature	150°C	325°C
Drying Gas	9 L/min	13 L/min
Nebulizer	35 psi	35 psi
Sheath Gas Temperature	350°C	275°C
Sheath Gas Flow	10 L/min	12 L/min
VCap	2800V	4000V
Nozzle Voltage	1500V	2000V
Fragmentor	100V	175V
Skimmer	65V	65V

An IgG N-linked library standard which contains 17 common glycan species present in mAbs was used for method evaluation (Appendix II). During optimization and for simplicity-sake the FA2 peak area was the only area evaluated in both FLD and MS data. FA2 is the most abundant glycan in mAbs consisting of 2 GlcNAc residues one on each arm in addition to the core structure (Figure 16).



Mass: 1723.6921

Figure 16. Structure of FA2

Preliminary analysis of the library standard showed the peptide method produced a 10-fold increase in signal intensity compared to the original method. Although increased signal intensity is desired, this information only indicates the total number of ions that are reaching the detector but additional analysis of the mass spectrum is required to determine the ions that are being detected. Analysis identified a significant amount of in-source fragmentation as seen in the mass spectrum snapshots in Figure 16.

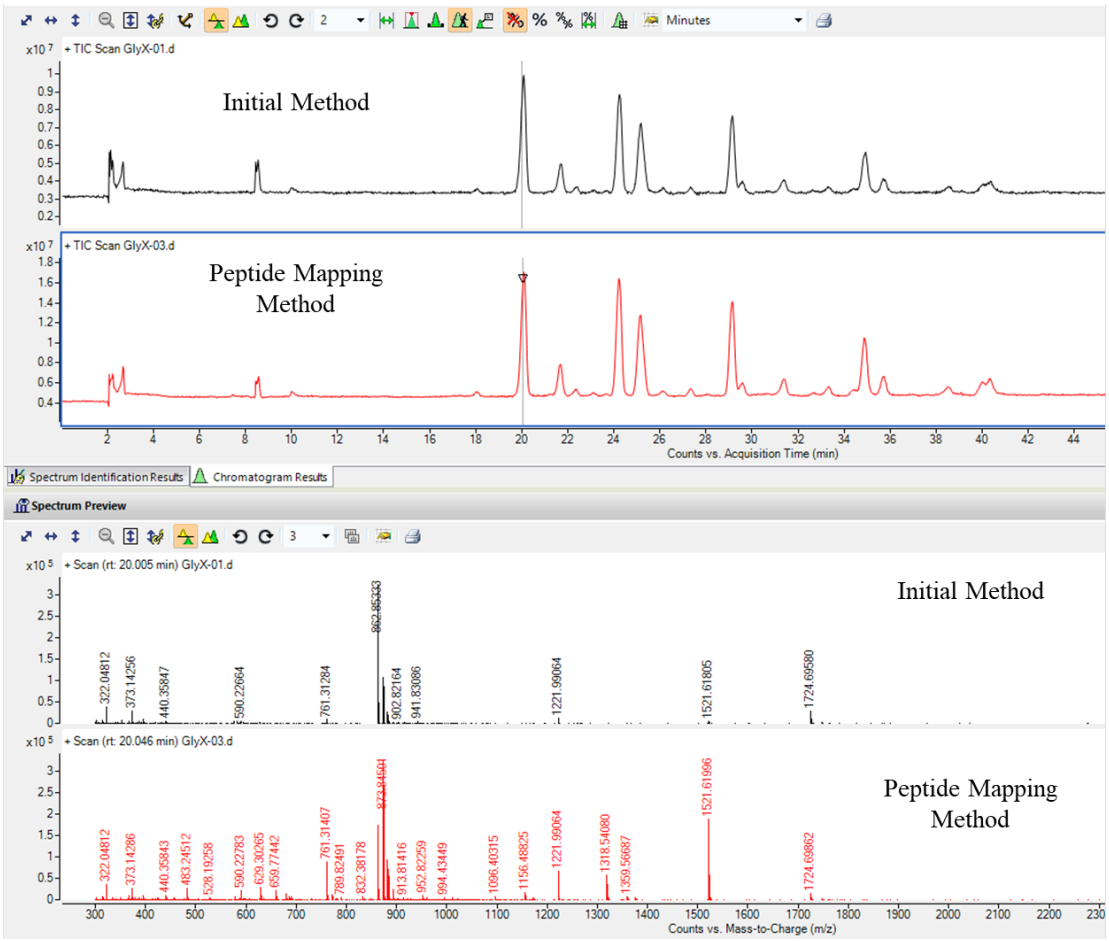


Figure 17. Total ion chromatograms (top) and mass spectrums (bottom) for initial method vs peptide mapping source conditions

The high voltage of the VCap, nozzle voltage, and fragmentor along with the high gas temperature conditions are too severe for the linkages connecting the monosaccharide structures to the initial core structure. The two most common forms of in-source fragmentation (ISD) observed in oligosaccharides is the loss of a GlcNAc residue resulting in a mass difference of ~203 and the loss of a GlcNAc-gal residue resulting in a mass difference of ~365 as seen in Table 13. Decreasing these conditions as seen using the original method produces little to no in-source fragmentation and this will be used for the optimization of sample prep conditions described below. It is important to track the in-

source decay during initial stages of method development/optimization as it could lead to false identification of glycan species and requires ion/source parameter optimization. There is some level of in-source decay in all glycan analysis as they are fragile ions and fragment easily, however it is normally the most abundant glycan species (FA2, FA2G1, and FA2G2) that have detectable levels of fragmentation. In-source fragmentation is easily differentiated from a true glycan as the elution time of the ISD will be identical to that of the glycan from which it occurred (i.e FA2-N will be the identical elution time of FA2). In-source fragmentation of the FA2 species was selected for monitoring throughout the remaining optimization steps as it is the most abundant glycan species in mAbs.

Table 13. Common structures of in-source fragmentation

Name	Mass	Structure
GlcNAc	203.0794	
GlcNAc-Gal	365.1322	

General Labeling Optimization Strategy.

The Gly-X protocol is divided into four main sections: deglycosylation, labeling, cleanup, and elution; this is followed by detection using a fluorescence detector or LC/MS. A general overview of the Gly-X protocol is presented in Figure 18. Within this protocol, several areas were selected for further optimization including initial sample buffer composition, quantity of denaturant, and quantity of enzyme used for deglyocsylation. In addition, the storage conditions for prepared samples was validated to ensure stability of the labeled glycan samples for extended periods of storage. It should be noted that labeling and cleanup conditions were not altered as these steps rely solely on the released glycan, not the starting material (mAb) properties. These steps were optimized by the manufacturer and should be broadly applicable to glycans released from any mAb.

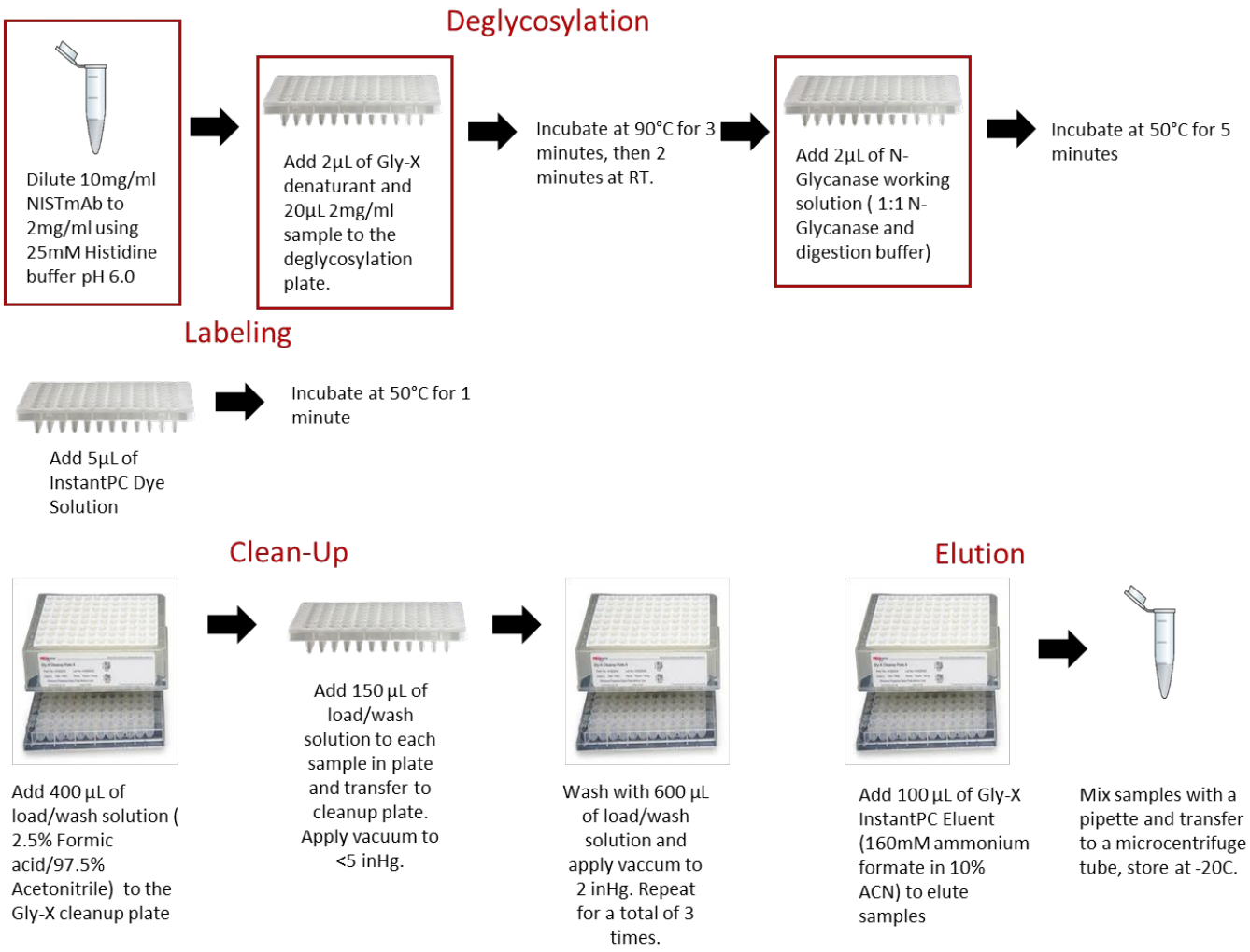
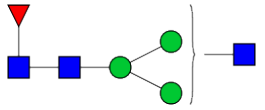
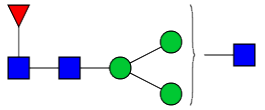
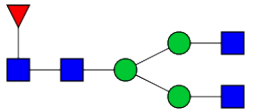
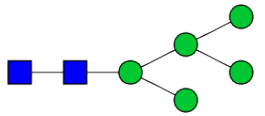
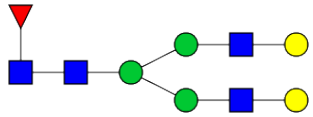


Figure 18. Overview of Gly-X protocol

Sample Stability

The Gly-X protocol recommends storage of the final labeled glycan product at -20°C for up to 6 months or 5 days at 4°C. Samples were intending to be stored at -20°C for no longer than 30 days, however the stability of individual glycan species was important to monitor at the intended storage conditions to ensure artifacts were not induced during the sample optimization process. Initially, storage of samples at -20°C and -80°C was evaluated with a total of 5 glycan structures evaluated for relative quantitation at time of sample prep (day 0) and 30 days post sample prep (day 30). These 5 glycan structures included FA2, FA2-N, Man5, FA2G2, and $[FA2\text{-}N]^{ISD}$, which were chosen as they are a representative group for all the glycan structures that are present in mAb structures (Table 14). Bianntennary glycans are the predominant structures seen in monoclonal antibodies, with FA2 being the most abundant glycan making it an optimum choice for traceability. In addition, high-mannose (man5) and species with galactosylation (FA2G2) play a critical role in pharmacokinetics and efficacy. Lastly, FA2-N (FA2 minus a GlcNAc residue) and $[FA2\text{-}N]^{ISD}$ have identical mass and composition, however $[FA2\text{-}N]^{ISD}$ is due to in-source fragmentation and is not a naturally occurring glycan species.

Table 14. Glycan structures used for optimization of glycan release assay

Glycan Composition	Name	Chemical Formula	Theoretical Mass	Theoretical Mass + InstantPC	Structure
H3N3F1	FA2-N	C48H81N3O35	1259.46506	1520.61279	
H3N3F1	$[FA2\text{-}N]^{ISD}$	C48H81N3O35	1259.46506	1520.61279	
H3N4F1	FA2	C56H94N4O40	1462.54443	1723.69216	
H5N2	Man5	C46H78N2O36	1234.43343	1495.58116	
H5N4F1	FA2G2	C68H114N4O50	1786.65008	2047.79781	

Comparison of these glycans at time of sample preparation with after 30 days of storage showed a slight increase (~1%) in the relative quantitation of FA2 for both -20°C and -80°C (Figure 19). Although this increase was small it was present in both storage temperatures after 30 days and this change can likely be attributed to day-to-day variability of the mass spectrometer. To eliminate the factor of variability a second analysis was completed and all samples were prepared, stored for a specific period of time and then run on the same day. In addition to just evaluating the samples at day 0 and day 30 several additional time points (5,10,15,20,25 days) were added and tested. All samples were placed at -20°C and each sample was moved to -80°C after 5, 10, 15, 20, 25, and 30 days as seen in Table 15 below.

Table 15. Number of vials stored at -20°C and -80°C for stability study

	Day						
Temperature	0	5	10	15	20	25	30
-20	6	5	4	3	2	1	0
-80	0	1	2	3	4	5	6

After 30 days at the completion of the study all samples were run using MS analysis on the same day to eliminate inter-day variability. In addition, the order of the samples was randomized using a random order generator did confirm that if degradation did occur during the study it would not be attributed to the amount of time that the samples were kept at 4°C during the run analysis which had a total run time of 19 hours. Table 16 shows randomized run order with corresponding FA2 relative quantitation values. Little to no variability (~0.25% quantitative area) was observed between the samples over a 30-day period, indicating the stability of the InstantPC label and sample at -20°C (Figure 20).

Storage conditions for all samples will be -20°C as no temperature showed superiority over the other and this was recommended by the manufacturer for samples and some Gly-X components/standards.

Table 16. Length of storage at -20C and resulting % relative quantitation for FA2

Storage at -20°C	FA2 %Quant Area
25 days	80.66
20 days	80.64
15 days	80.66
5 days	80.61
30 days	80.81
10 days	80.54

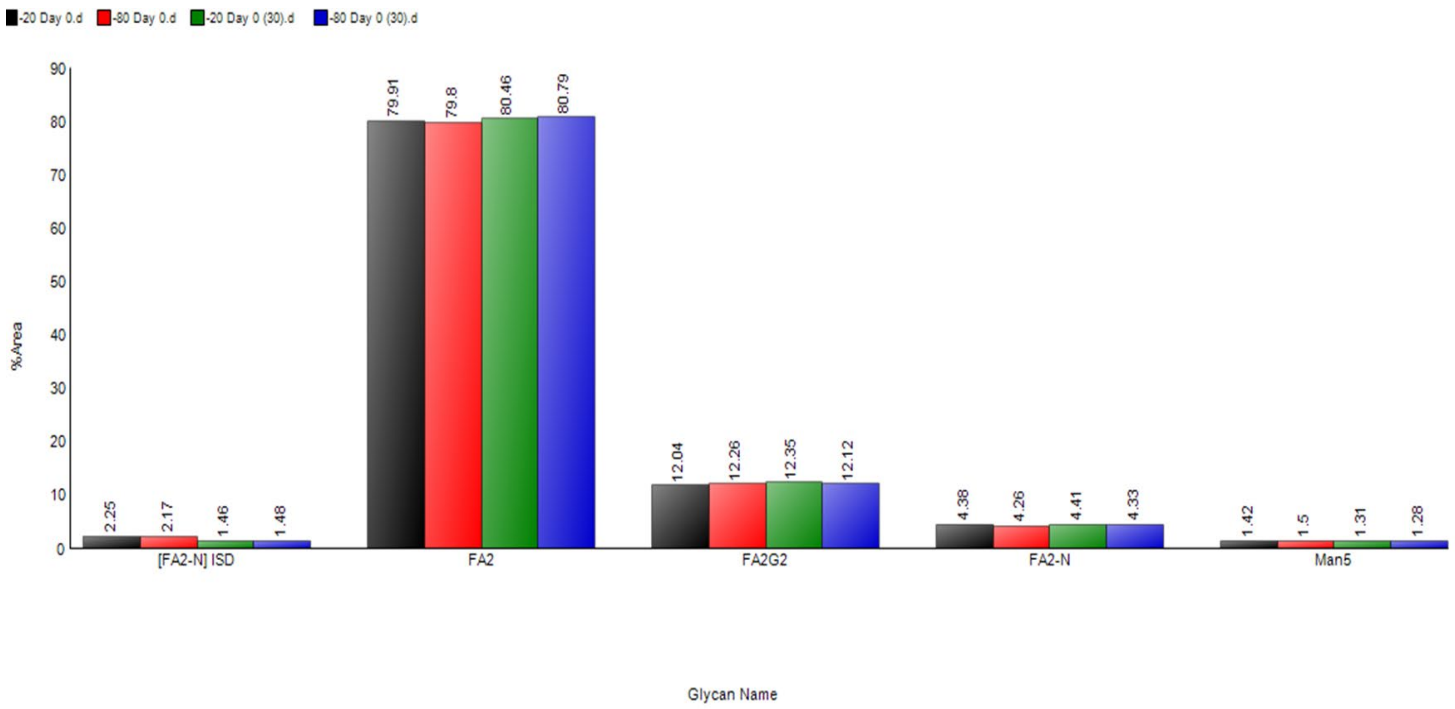


Figure 19. Storage conditions for NISTmAb released glycans at -20°C and -80°C.

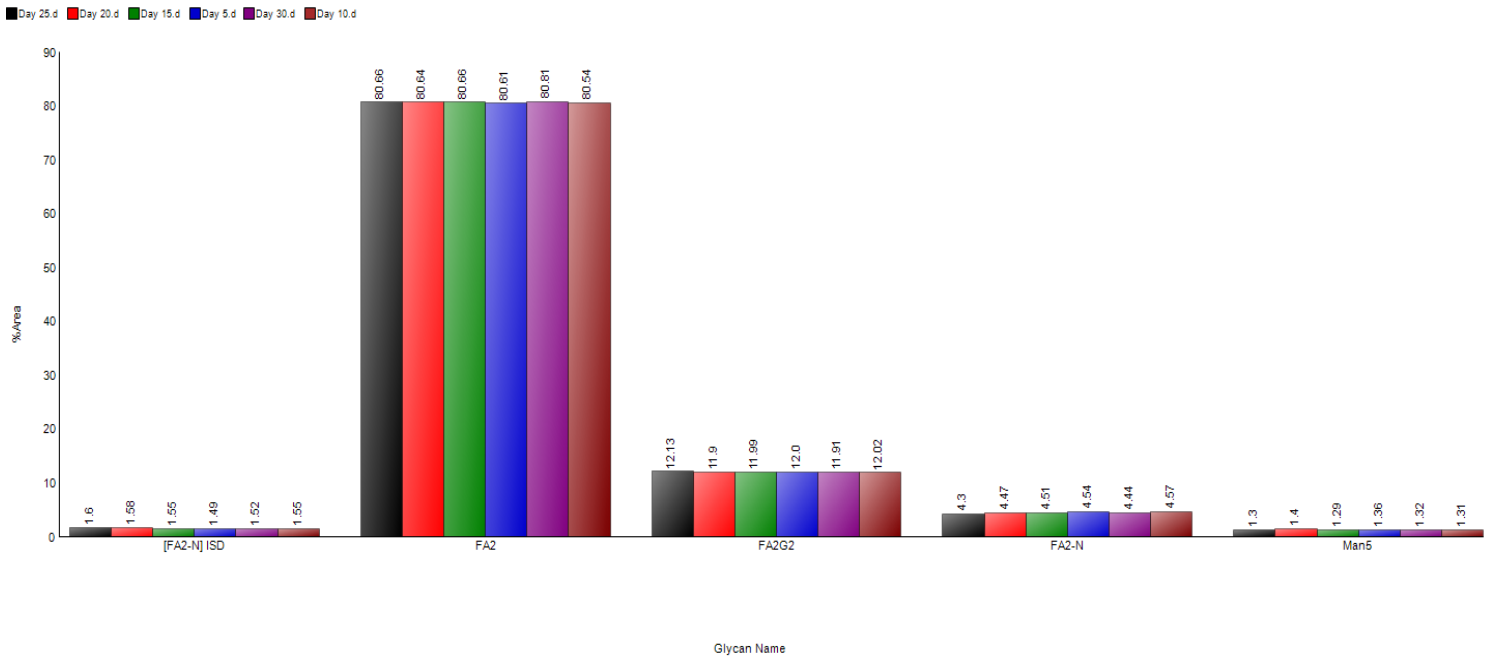


Figure 20. Stability of NISTmAb released glycan samples stored 5-30 days at -20°C and -80°C.

Mass spectrometry was used as the sole detection method to evaluate storage conditions as the stability of the individual glycans was of main concern. Glycans are positively identified by mass spectrometry using the m/z value, providing accurate identification but adding a level of complexity to sample analysis. This is due to several factors such as in-source fragmentation, additional data analysis requirements and instrument constraints. Fluorescence detection is the preferred method for determining labeling efficiency as it only detects the presence of a fluorophore attached to a glycan structure. This simplifies the sample optimization process providing a clear distinction between increased or decreased signal based on small changes in sample preparation such as sample buffer composition and solution volumes. It should be noted that preferred analysis of the samples is fluorescence detection in line with mass spectrometry, however instrument constraints prevent this from occurring. The electrospray needle in the source of the mass spectrometer is short with a small internal diameter creating backpressure on the fluorescence detector that is connected to the MS. This backpressure is too great and can damage the flow cell of the detector which has a pressure limit of ~ 10 bar. Possible solutions to this include the addition of a solvent splitter which will directly only have the amount of flow coming from the LC to the MS, thus decreasing the back pressure. This can lead to a significant amount of sample loss, as a result the samples were run simultaneously on a stand-alone LC system for fluorescence detection and a LC/MS system. Each sample above was split into 2 post labeling and run on each system, this allowed for collection of both sets of data.

Sample Buffer Composition Optimization

The composition of the histidine buffer solution that NISTmAb is currently stored in is 12.5 mmol/L L-Histidine and 12.5 mmol/L L-Histidine hydrochloride for a total concentration of 25mM and a pH of 6.0. The InstantPC label contains an NHS-ester that may react with amine components such as histidine, arginine, glycine, and tris. These buffers will prevent complete labeling as they will compete with the glycan for the label. Buffer incompatibility was confirmed with the formation of a precipitate following addition of the denaturant solution to the sample. Manufacturer recommendations provided with the kit indicated non-amine, salt-containing buffers up to 150mM would be compatible for all solutions included in the kit.

HEPES buffer was selected due to being a low salt buffer and recommendations by the kit manufacturer suggest this type of neutral buffer. The pH of the buffer was also increased to 7.9 which will induce the nucleophilic reaction of the amine compound on the InstantPC label during the labeling process. 25mM HEPES pH 7.9 was prepared and a direct comparison between buffer exchange vs dilution of samples was also evaluated. Diluting the samples instead of buffer exchanging will leave a small concentration of histidine remaining in the sample (~5mM), however there will be no sample loss as seen with the samples that have been buffer exchanged. Buffer exchanged samples concentration was measured using UV prior to analysis as seen in Table 17 below.

Table 17. Concentrations of 25mM HEPES and water post-buffer exchange

Buffer	Sample Concentration 1 (mg/ml)	Sample Concentration 2 (mg/ml)	Average	Desired Concentration (mg/ml)
25mM HEPES pH 7.9	8.745	8.677	8.711	2
Water	8.182	8.134	8.156	2

Analysis of FA2 area for FLD and MS showed nearly identical results for all samples with 25mM HEPES having approximately 2 times the signal intensity than histidine (Figure 21). A slight increase in intensity was observed for the buffer exchanged vs non-buffer exchanged HEPES sample (fluorescence: 322 vs. 279) which can be attributed to the small amount of histidine buffer that remains in the sample after dilution (Figure 21).

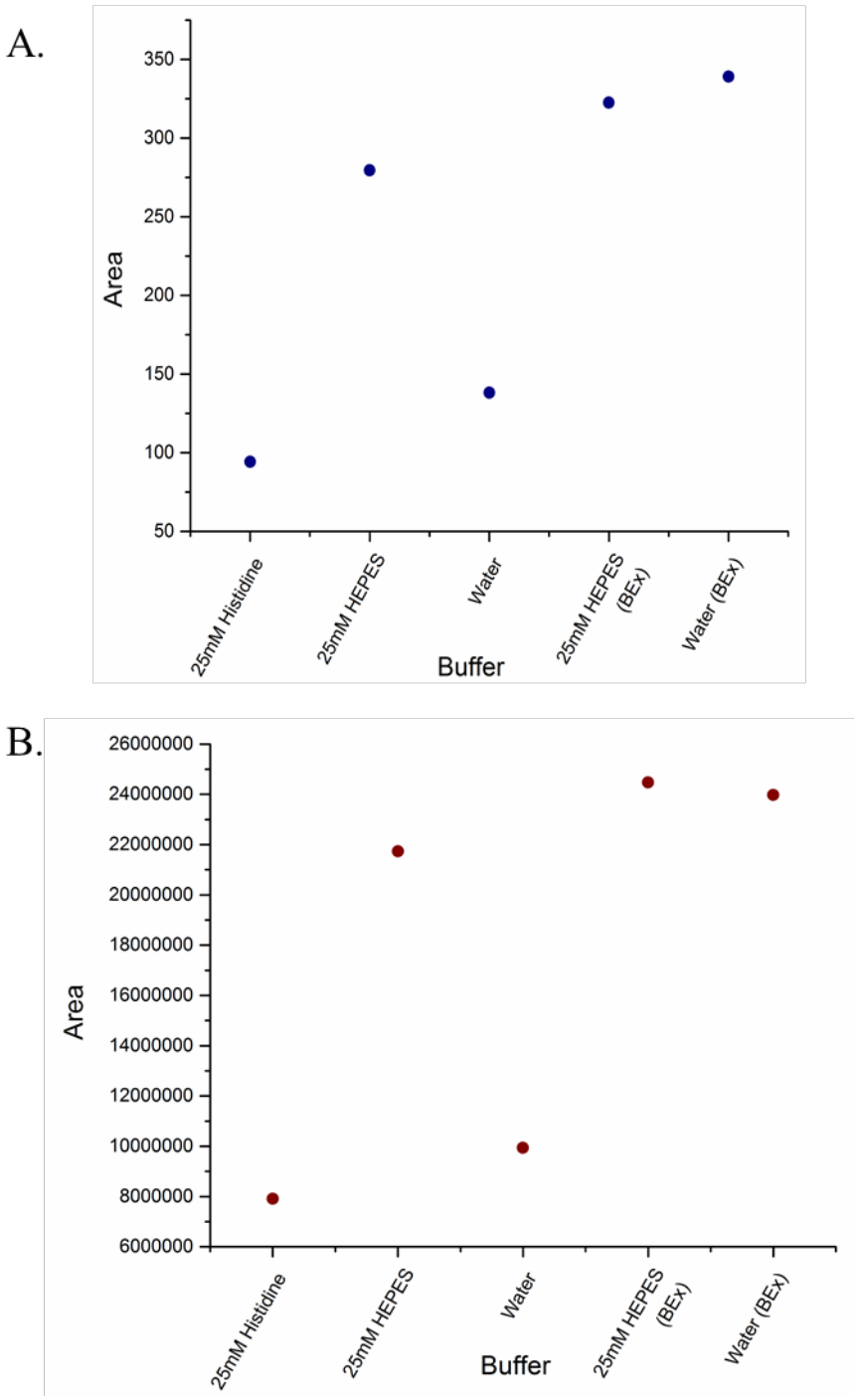


Figure 21. Comparison of FLD and MS spectra for histidine, 25mM HEPES and water. BEx = buffer exchanged samples

Samples diluted with water had little improvement from the histidine sample, however buffer exchange doubled signal intensity from 138 to 339. Both MS and FLD data was collected for the first round of optimization, however both sets of data showed identical trends in terms of signal intensity. Fluorescence provides the best measure of signal intensity and as a result it will be used for data collection for the remaining optimization studies. A second analysis was done using 25mM HEPES, water, and 50mM HEPES, to evaluate the reproducibility of buffer exchange. Although 25mM HEPES provided promising results, 50mM HEPES was evaluated as this was referenced in the manual as a possible option for sample composition analysis. This increase in buffer concentration should not have a significant effect on signal intensity or total area of the peak, it should be evaluated as it is mentioned in the protocol. As done in the previous run, one set of samples was buffer exchanged while the other was simply diluted with the desired buffer. Concentrations of samples post-buffer exchange are shown in Table 18 below. 50mM HEPES sample was run in duplicate to have 2 replicates for each preparation as 25mM HEPES and water were previously run under identical conditions.

Table 18. Concentrations of 25, 50mM HEPES, and water post buffer exchange

Buffer	Sample Concentration 1 (mg/ml)	Sample Concentration 2 (mg/ml)	Average	Desired Concentration (mg/ml)
25mM HEPES pH 7.9	8.662	8.689	8.675	2
50mM HEPES pH 7.9	8.825	8.724	8.775	2
Water	7.728	7.827	7.775	2

Fluorescence data showed higher levels of signal intensity for FA2 peak area in buffer exchanged samples for all three solutions (25mM HEPES, 50mM HEPEs, water) in comparison to the samples that were diluted (Figure 22). Although the water samples that were buffer exchanged showed little variability between the 2 samples it will not be used as the sample buffer due to the variation in sample concentration between the first and second analysis. Water samples that were simply diluted had too much inter-day variability and the lowest signal of any of the samples. A comparison between the 25mM and 50mM HEPES showed minimal variability between the buffer exchanged samples (340 vs. 327) while the 25mM diluted HEPES sample had higher signal intensity (283 vs 239) in both instances.

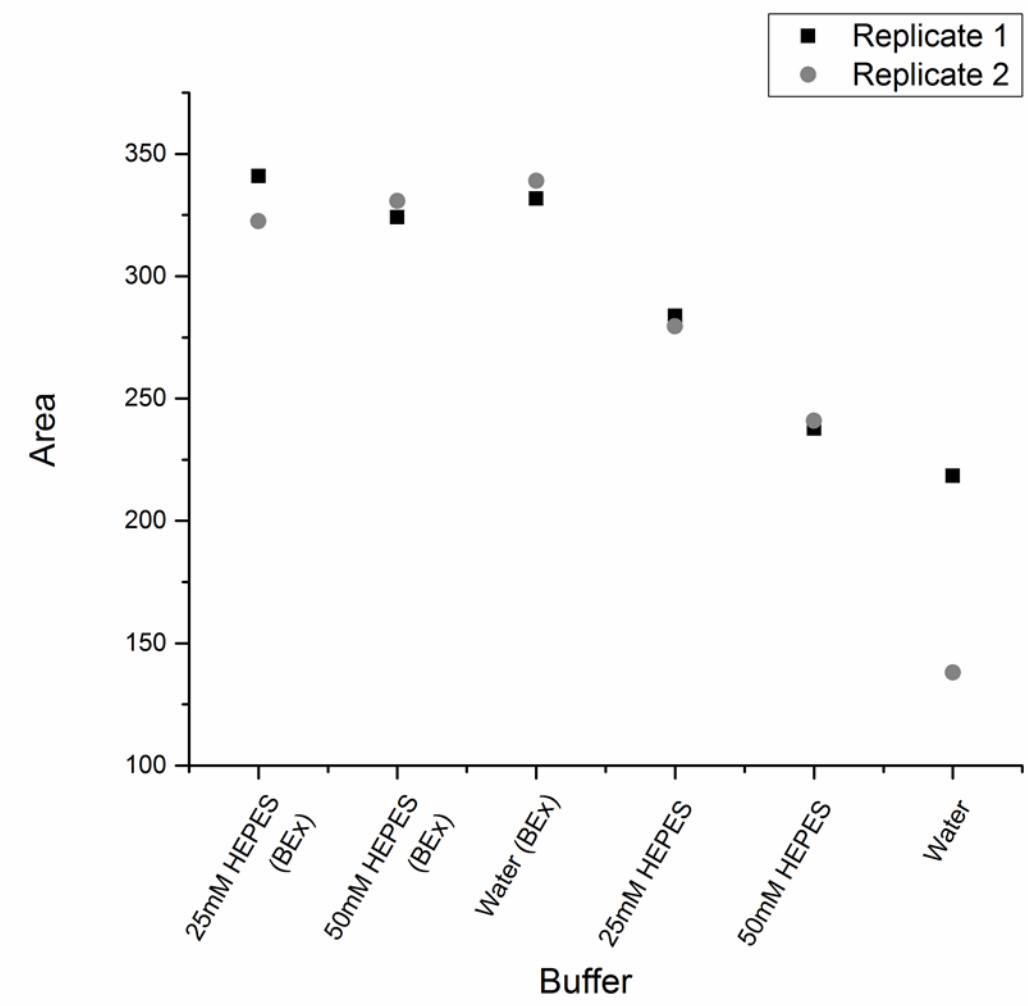


Figure 22. Signal intensity for varying HEPES buffer concentrations

While the buffer exchanged samples do provide a higher signal intensity there are multiple factors that should be considered in determining which sample preparation method will provide the most efficient data. Samples that are diluted will provide more robust data as they start with the same concentration each time (10 mg/ml), while samples that are buffer exchanged will have a different concentration each time requiring measurement with UV. While the preparation of the diluted samples will not completely remove the histidine, it will significantly decrease the concentration to ~5mM which is quite small and does not impact the signal intensity and reproducibility of the method. The reproducibility is of

critical importance for this method and as a result samples will be diluted with 25mM HEPES buffer instead of using a buffer exchanged method.

Denaturant

Complete protein denaturation is essential in glycan release and labeling assays in order for the enzyme to access the asparagine residue and cleave the glycans from the protein. Incomplete denaturation can cause partial deglycosylation leading to incomplete labeling, therefore the total quantity of denaturant was evaluated. Figure 23 shows the fluorescence intensity for the FA2 glycan where increments of 2,4, and 6 µl of denaturant were used in the sample preparation protocol. Surprisingly, the FA2 area decreased with increasing volume of denaturant (Figure 23). The larger volume of denaturant does dilute the total volume of sample that is cleaved by PNGase F, which may have adversely affected the enzyme kinetics. More likely, however, is that the denaturant itself may reduce enzyme activity. Regardless of the root cause, 2 uL of denaturant was selected for all further analysis as this provided the most complete glycan release.

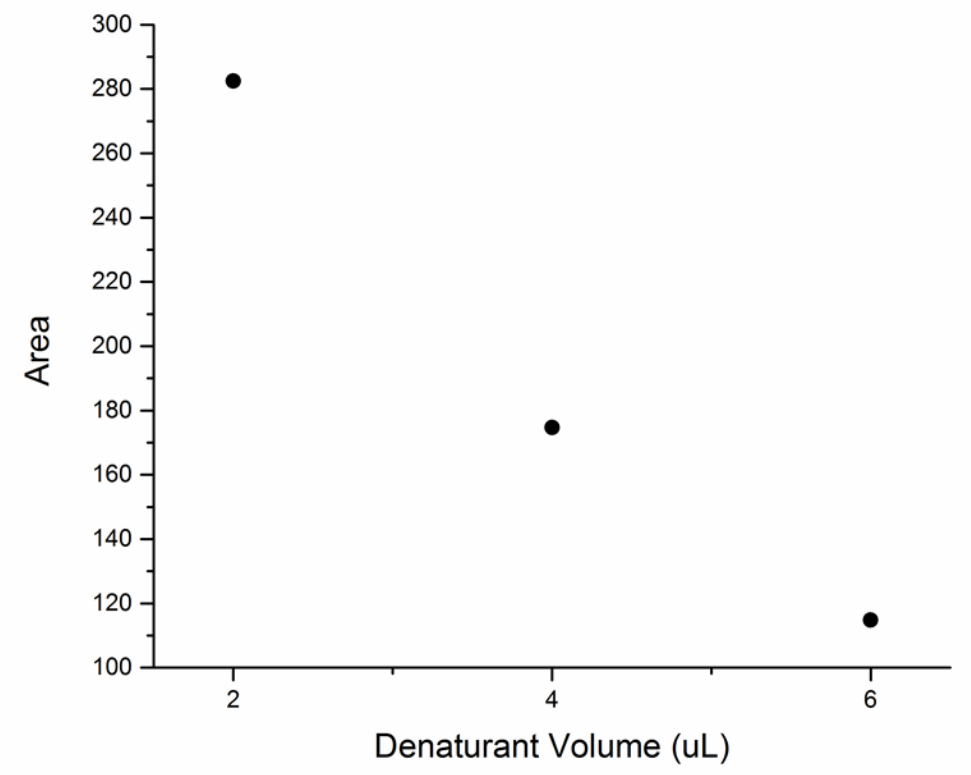


Figure 24. Effect of denaturant volume on FA2 fluorescence signal intensity

Working Solution

The working solution is comprised of a 1:1 mixture of PNGase F and digestion buffer, where PNGase F is responsible for cleavage of the glycans from the protein at the asparagine residue Varying concentrations of working solution were evaluated in increments of 2, 4, and 6 µl. Although the effectiveness of enzyme cleavage was of focus here, instead of just increasing the enzyme concentration the entire working solution volume was increased to ensure the ratio remained 1:1 and all volumes were consistent. Analysis of FA2 area showed little variability in terms of different volumes, with all signal intensity being above 300 demonstrating that cleavage was occurring at each condition (Figure 24). Although the 4 µL volume of working solution did show the highest signal intensity there was only a difference of ~20 between the volume and the 2 µL volume. As

a result, the 2µL working solution volume will be used in the optimized method to conserve reagents as it is still providing good signal intensity. The kit manufacturer released the optimized method in terms of labeling with the InstantPC fluorophore. However, efficient labeling cannot occur if the glycan structures are not effectively released from the protein; optimization of these deglycosylation steps ensured optimal cleavage was occurring.

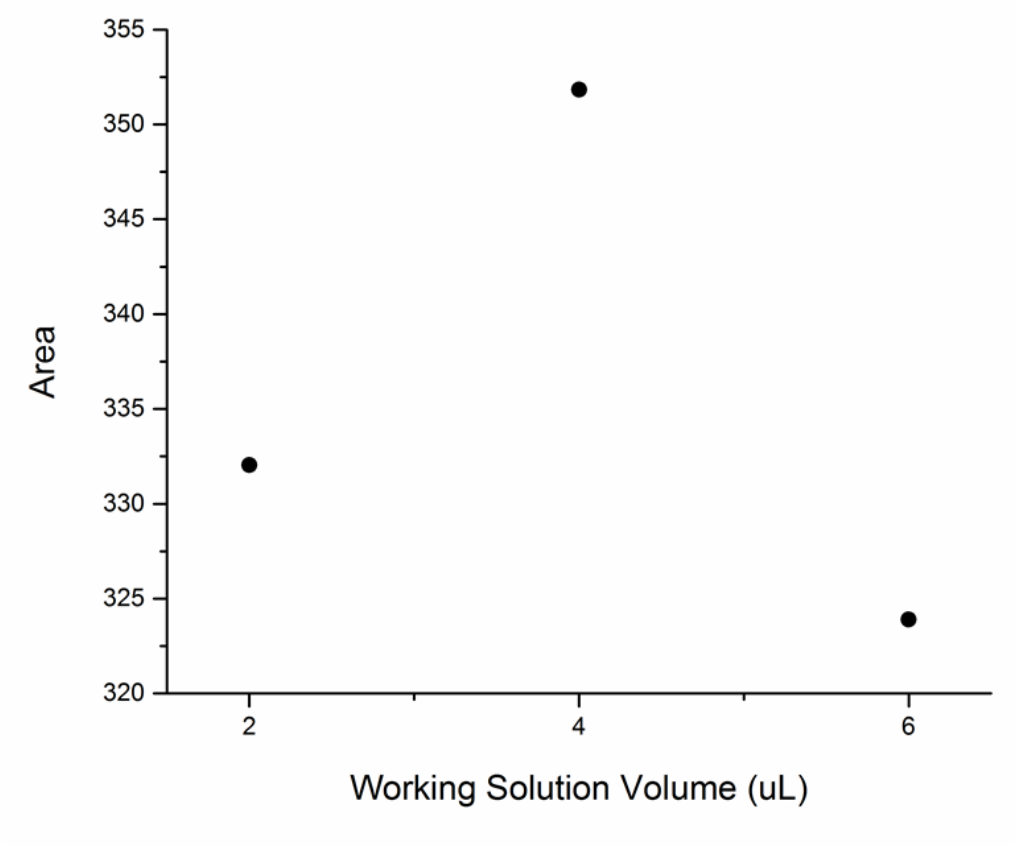


Figure 24. Effect of working solution on fluorescence signal intensity

NISTmAb Glycan Analysis

The optimized sample preparation and analysis method was used to characterize the identity and relative abundance of glycans present in the NISTmAb. Resultant data was analyzed using Agilent's MassHunter BioConfirm software in combination with the Agilent PCDL Manager. PCDL Manager uses a theoretical library that contains all of the N-glycan structures that the MassHunter software will look for in each sample. Initial analysis of NISTmAb used an N-glycan library that contained all possible glycan structures reported to date in monoclonal antibodies. This comprehensive list of more than 100 glycans was comprised from multiple sources beginning with a library provided by Agilent Technologies. This library was cross-referenced with literature on NISTmAb glycans previously reported using the traditional labeling mechanisms (2-AB) (Prien et al. 2015; Hilliard et al. 2017). Any glycans not included in Agilent's library were added to form a comprehensive search space. The NISTmAb glycan data collected herein was searched against the comprehensive glycan library to identify glycans present in the NISTmAb, each of which must then be manually verified.

The overall data analysis workflow is described in Figure 25 and includes software search parameters followed by manual validation. The workflow begins with the software using the PCDL Manager glycan library to search for m/z in the LC-MS data that have the same mass as those in the library. The software will only report a putative glycan structure if the data matches the theoretical m/z AND appropriate theoretical isotopic profile. The BioConfirm software will populate a list of potential biomolecule matches and generate an extracted ion chromatogram that incorporates the signal intensities of all related ions (e.g. parent m/z value, additional charge states, associated isotopes, and Na^+ and K^+ adducts).

The software uses these three components to build a score for similarity to the theoretical values for each glycan. Biomolecules yielding a value of ≥75 were reported by the software for further manual verification. While this matching is correct more often than not, characterization data must be manually verified. Each biomolecule match was reviewed for determination of any false positive peaks associated with in-source decay events and mis-assignment of the monoisotopic m/z.

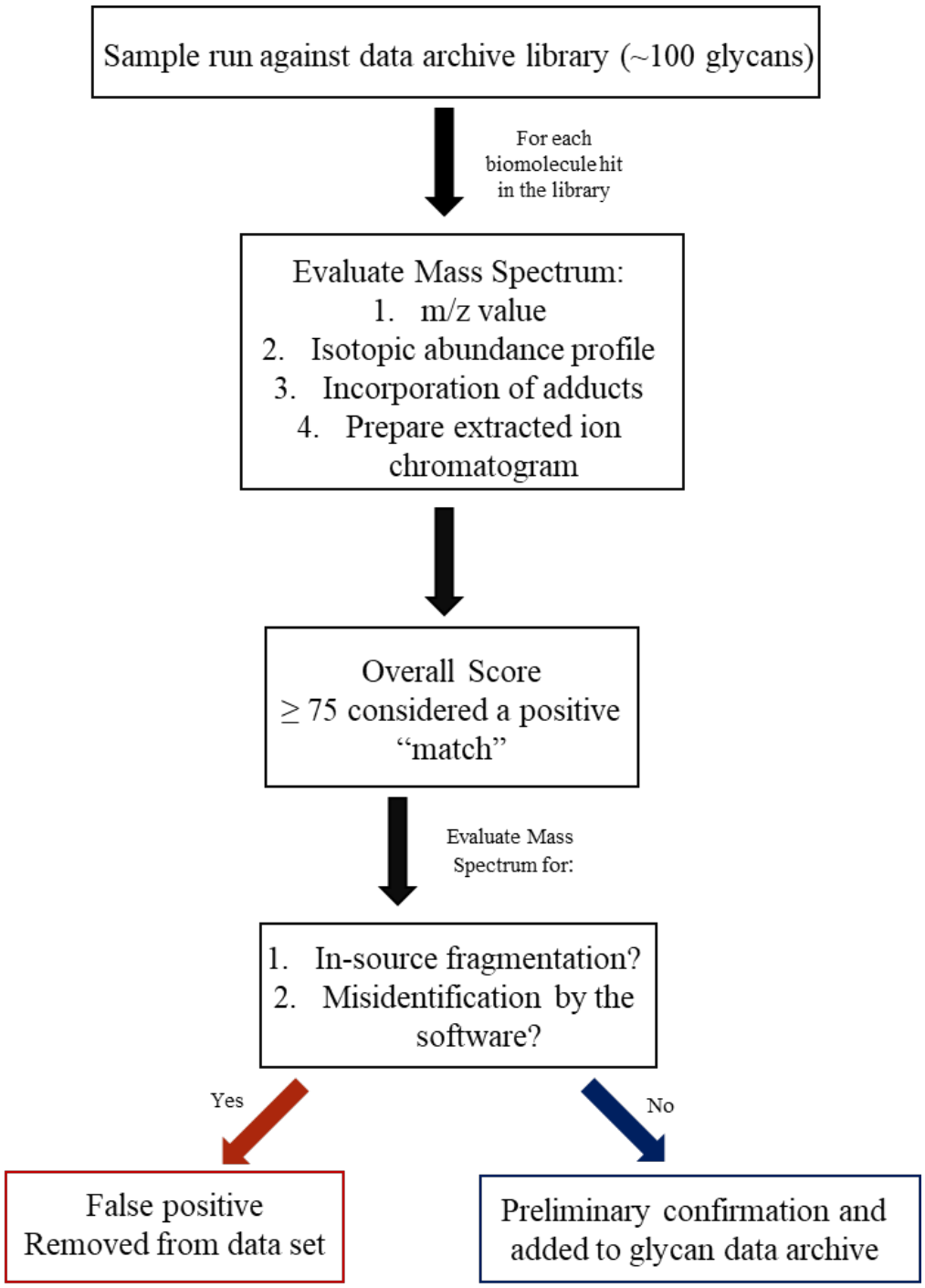


Figure 25. BioConfirm workflow for analysis and identification of released glycans

An example of a common false positive identification is that due to in-source fragmentation is shown in Figure 26. ISD can be identified through a combination of mass error and retention time. In-source fragmentation will have the identical retention time and peak shape to that of the parent glycan species which is demonstrated with the glycan FA2G1 and its resulting ISD (FA2G1-N) as shown in Figure 26 below. FA2G1 has a theoretical mass of 1885.7449 while FA2G1-N (an in-source decay product) and the isomeric glycan FA1G1 have a theoretical mass of 1682.6656. The initial data archive search performed by MassHunter does not incorporate retention time, and therefore initially identified the blue peak as FA1G1. Manual review of the data demonstrates the coincident retention with FA2G1 at 24.34 minutes, indicating this peak is due to ISD. The true FA1G1 glycan with a mass of 1682.6656 elutes earlier in the gradient at 22.69 minutes and was also identified before the one with a larger mass due to the concept of HILIC described above. All potential identifications were reviewed for in-source fragmentation mis-identifcation in a similar manner by overlaying the 2 extracted ion chromatograms.

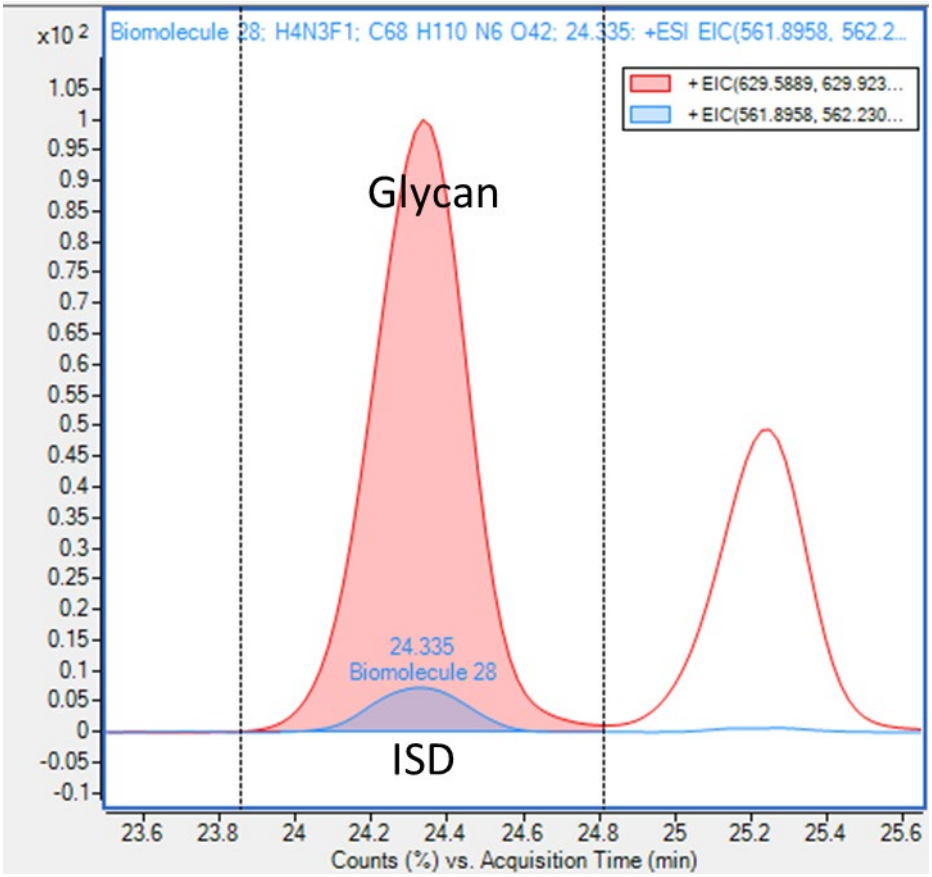
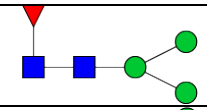
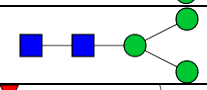
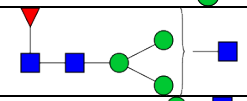
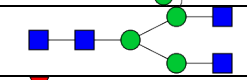
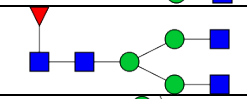
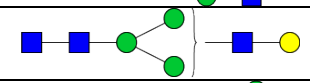
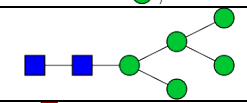
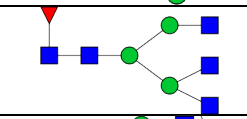
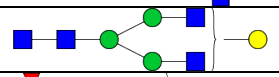
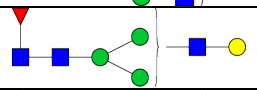
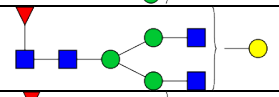
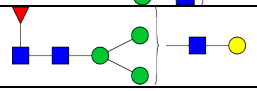


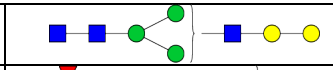
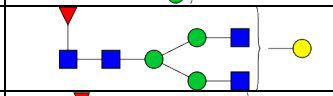
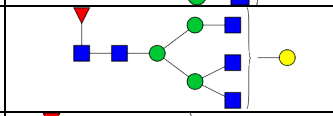
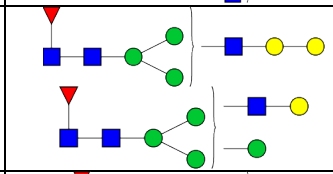
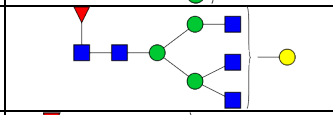
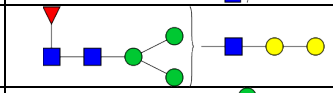
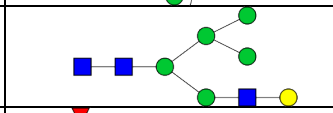
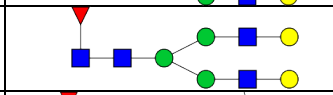
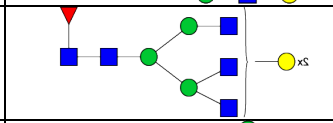
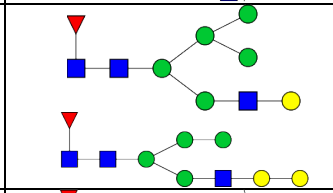
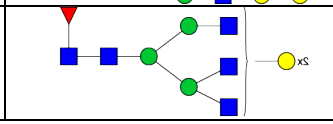
Figure 26. Extracted ion chromatogram showing a glycan (red) and its in-source decay fragmentation (blue, ISD).

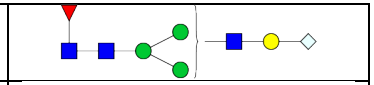
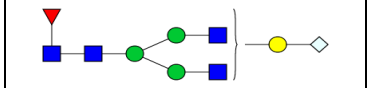
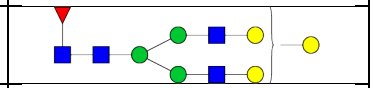
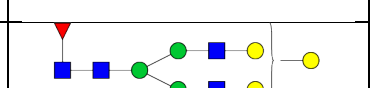
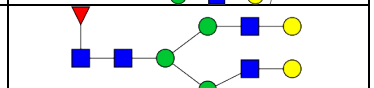
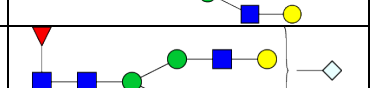
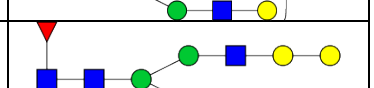
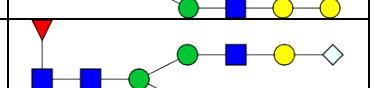
Table 19 shows a comprehensive list of all glycans identified in the NISTmAb. Bianntennary (2 branches linked to the GlcNAc core) species were the predominant structure identified in NISTmAb with FA2, FA2G1, and FA2G2 having the highest quantitation with 42, 24 (FA2G1[a]), 11 (FA2G1[b]), and 9 % percent's respectively. Several low abundance and sialylated species were also identified with the majority of these structures having the N-glyconeuraminic acid (NeuGc) terminal decorations due to expression in NS0 cell lines. Both fucoysylated and afucosylated glycans were identified with the afucosylated species detected in very low abundance; A1G1 (~0.07) and A2G1 (~0.04). In some cases, a single composition is representative of multiple isomeric glycans

as observed in several species in Table 19 below (ie H5N3F1). An example is the proposed composition of H5N3F1; this structure could be identified as FM5A1G1 or FM4A1G1Ga1. These putative structural assignments are based off of historical data on these glycan species, however, all glycan species identified will evaluated using MS/MS in future analysis to confirm structural identity.

Table 19. Composition of glycan classification groups in NISTmAb.

Name	Proposed Composition	Theoretical Mass	Observed Mass	Mass Error (ppm)	% Quant	Retention Time (min)	Structure
FM3	H3N2F1	1317.5334	1317.5344	0.7248	0.13	14.1375	
A1	H3N3	1374.5548	1374.5553	0.2692	0.3425	15.1507	
FA1	H3N3F1	1520.6127	1520.6133	0.3518	2.5075	17.4115	
A2	H3N4	1577.6342	1577.6341	-0.1172	0.0875	18.1602	
FA2	H3N4F1	1723.6921	1723.6934	0.7194	42.0425	20.1765	
A1G1	H4N3	1536.6077	1536.6031	-3.0326	0.0725	20.3687	
M5	H5N2	1495.5811	1495.5819	0.5115	0.765	21.3510	
FA3	H3N5F1	1926.7715	1926.7719	0.1739	0.405	22.2815	
A2G1	H4N4	1739.6870	1739.6805	-3.7966	0.04	22.4530	
FA1G1	H4N3F1	1682.6656	1682.6658	0.1278	2.9375	22.6920	
FA2G1[a]	H4N4F1[a]	1885.7449	1885.7446	-0.2201	24.6	24.3395	
[FA2G1]-N	H4N3F1-ISD	1682.6656	1682.6673	0.9746	0.2275	24.3395	

A1G1Ga1	H5N3	1698.6605	1698.6578	-1.6366	0.05	24.4487	
FA2G1[b]	H4N4F1[b]	1885.7449	1885.7448	-0.1273	11.76	25.2490	
FA3G1[a]	H4N5F1[a]	2088.8243	2088.8227	-0.7947	0.3225	26.2310	
FA1G1Ga1 or FM4A1G1	H5N3F1[a]	1844.7184	1844.7193	0.4526	0.125	26.5480	
FA3G1[b]	H4N5F1[b]	2088.8227	2088.8232	0.2513	0.3625	27.0210	
FA1G1Ga1	H5N3F1[b]	1844.7184	1844.7185	0.0189	1.21	27.6707	
M5A1G1	H6N3	1860.7133	1860.7133	-0.0134	0.1425	29.0270	
FA2G2	H5N4F1	2047.7978	2047.7966	-0.5787	6.995	29.2507	
FA3G2[a]	H5N5F1[a]	2250.8771	2250.8748	-1.0796	0.26	30.7475	
FM5A1G1 or FM4A1G1Ga1	H6N3F1	2006.7712	2006.7703	-0.4784	0.09	30.8150	
FA3G2[b]	H5N5F1[b]	2250.8771	2250.8750	-0.9574	0.085	31.4487	

FA1G1Sg1	H4N3F1Sg1	1989.7559	1989.7546	-0.6659	1.2375	31.4697	
FA2G1Sg1[a]	H4N4F1Sg1	2192.8353	2192.8331	-1.0352	0.0775	32.1870	
FA2G2Ga1[a]	H6N4F1[a]	2209.8506	2209.8475	-1.4164	1.1725	33.2157	
FA2G1Sg1[b]	H4N4F1Sg1	2192.8353	2192.8544	8.7124	0.3225	33.3045	
FA2G2Ga1[b]	H6N4F1[b]	2209.8506	2209.8478	-1.2806	0.28	33.6525	
FA3G3	H6N5F1	2412.9300	2412.9274	-1.0713	0.1375	34.9362	
FA2G2Sg1	H5N4F1Sg1	2354.8881	2354.8845	-1.5563	0.28	36.7967	
FA2G2Ga2	H7N4F1	2371.9034	2371.9015	-0.8369	0.6225	37.2957	
FA2G2Ga1Sg1	H6N4F1Sg1	2515.9205	2515.9323	4.6464	0.195	40.1850	

[a,b]. isomer for the proposed glycan/composition was identified. Glycans labeled with an [a] eluted first, followed by the second isomer labeled with an [b.] Glycan assignments were made based on putative compositions expected from the glycan biosynthetic pathway.

Non-Originator Analysis

Following completion of the NISTmAb reference data archive the non-originator NISTmAb expressed from various cell lines (NS0-59, NS0-60, and NS0-66) were evaluated using the optimized released glycan assay. Data analysis was performed using the NISTmAb-specific data archive (Table 19 above) for evaluation of the similarity of non-originator expression products to the NISTmAb. LC-MS data collected for the non-originator NISTmAb samples were searched for the 32 glycans identified above using Bioconfirm. An overview of the % relative quantitation values for each material's respective glycans are outlined in Table 20 and Figure 27 below. Evaluation of non-originator glycan species showed significant changes in glycosylation in terms of terminal decorations (Table 20). Terminal β-(1-4) galactosylation increased from ~47.6% in NISTmAb to 51.3% and 57.4% in NS0-59 and NS0-60, respectively; while a decrease in galactosylation was seen in NS0-66 (~41.4%). Terminal α-(1-3)galactosylation increased from ~3.7% in NISTmAb to 4.9% and 8.0% in NS0-59 and NS0-60 with NS0-66 showing a similar trend from the terminal-β-(1-4) with a slight decrease to ~3.4%. However, all three cell lines showed substantial increases in the amount of sialyated species with NS0-66 having the highest amount with ~20% followed by NS0-59 and NS0-60 with 14% and 8% respectively. The biosynthetic pathway required\s terminal galactosylation prior to sialylation, indicating extended processing of the non-originator glycans to include terminal decorations. An increase in these species is of critical importance to document when developing a biosimilar or non-originator biomolecule as the immunogenicity could be greatly influenced by these structural changes. Galactosylation and sialylation have both been reported to be affected by bioreactor temperature and will likely be the subject

of a future growth temperature study (Agarabi et al. 2015). Additional analysis identified a considerable increase in the amount of afucosylated species between the NISTmAb and non-originator glycan species. Afucosylated species identified in the NISTmAb were ~1.7% while non-originator product expressed in NS0-59, NS0-60, and NS0-66 presented values of 6.1%, 6.3%, and 5.5%. Afucosylation is modulated by the biosynthetic pathway and shown to greatly increase the activity of ADCC. Fucosylation has been reported to be affected by decreases in bioreactor temperature and regulating the osmolality (Agarabi et al. 2015; Konno et al. 2012). Lastly, the concentration of high mannose containing glycans was evaluated which showed a minimal increase in 2 of the non-originator expressing cell lines. NS0-60 and NS0-66 showed increases of 2.3% and 1.3% while NS0-59 showed nearly identical high mannose concentration to that of the NISTmAb (0.94% vs 0.91%). While these values are not a substantial increase in high mannose containing structures this will affect the pharmacokinetics of these products. Regulation of bioreactor conditions including temperature and osmolality have been shown to affect the biosynthetic pathway resulting in an increase or decrease in high-mannose specific species (Pacis et al. 2011). The analysis reported above serves as a preliminary analysis of a T-flask expression system and additional studies monitoring bioreactor conditions will provide a better understanding of the "tunability" of each cell line to produce a more highly similar glycoprofile.

Table 20. Comparison of glycan classification groups between NISTmAb and non-originator cell lines.

Glycan Classification Groups	Afucosylated	High Mannose	Terminal β-(1-4) galactosylated	Terminal α-(1-3) galactosylated	NeuGc-Containing
NISTmAb	1.78	0.91	47.65	3.74	2.04
NS0-59	6.10	0.94	51.39	4.98	14.70
NS0-60	6.33	2.34	57.43	8.01	8.05
NS0-66	5.51	1.33	41.43	3.41	20.05

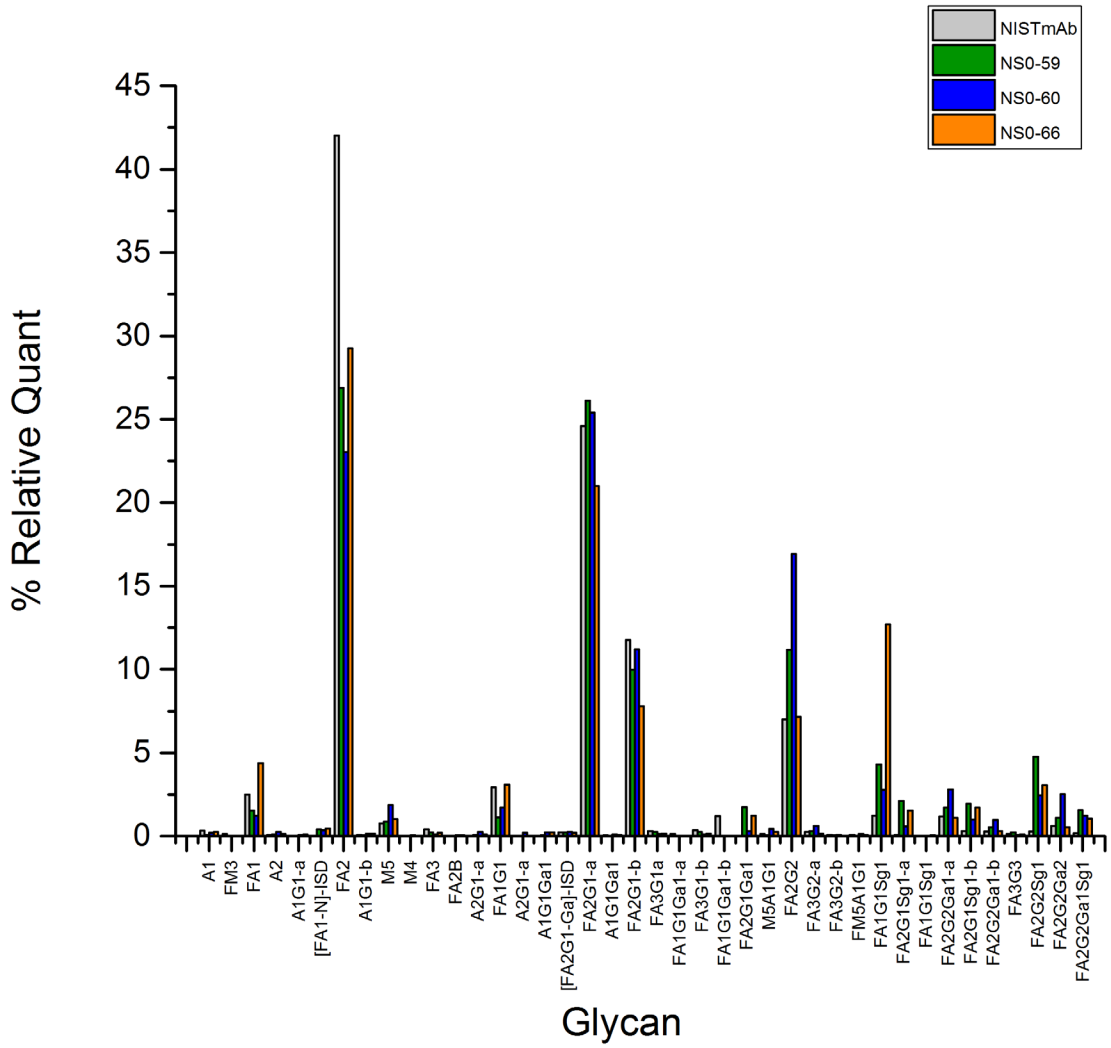


Figure 27. Relative quantitation values for glycan classification groups in NISTmAb and non-originator cell lines.

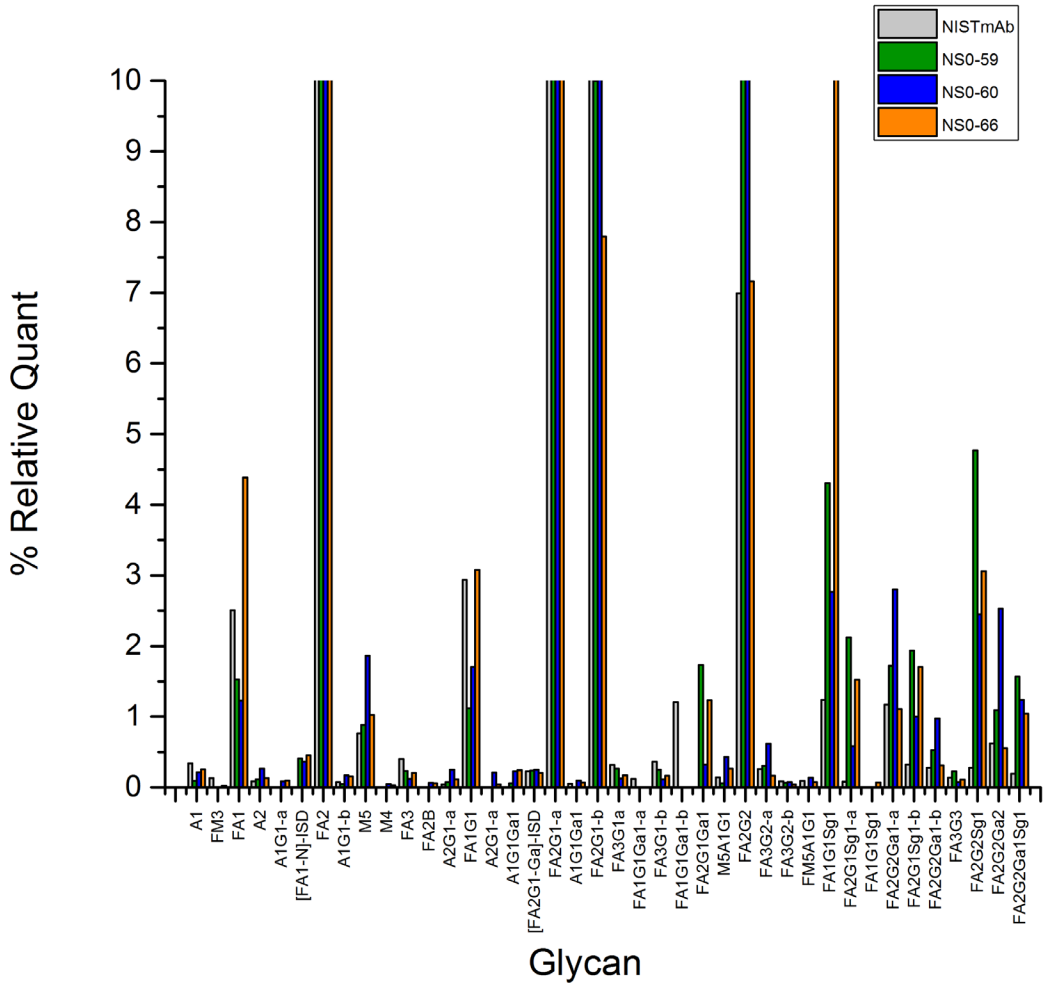


Figure 28. Relative quantitation values for glycan classification groups in NISTmAb and non-originator cell lines (zoomed).

Discussion

This thesis detailed the development of analytical methods for the characterization of monoclonal antibodies including non-originating expressing cell lines. Three cell lines NS0-59, NS0-60, and NS0-60 were made to nominally express the NISTmAb sequence and evaluated using SEC, intact mass analysis, and a glycan release assay to assess analytical similarity.

Aggregation of biopharmaceuticals is a CQA that is typically monitored because it can affect immunogenicity. Size exclusion chromatography (SEC) is considered the 'gold standard' for evaluation of size heterogeneity. A previously optimized method was used here to screen the non-originator materials for gross increases in aggregation that might indicate significant upstream/downstream issues prior to embarking on more advanced characterization techniques. SEC is an efficient method for monitoring low molecular weight fragments (LMW) and higher order aggregates such as dimers, trimers, and tetramers (HMW). The amount of aggregation as determined by SEC varied between cell lines (4-9%) and from the initial NISTmAb value of ~3%. The total amount of aggregation did not correlate with the decrease in relative peak intensity as observed in the intact mass analysis data. Additional downstream processing to include additional chromatography purification using methods such as cation exchange (CEX) or size exclusion (SEC) should be employed to remove these possible variants or impurities.

Intact mass analysis is a rapid method used to inform on primary structure and moderate to high abundant PTMs. The developed method was optimized for an efficient and high-throughput separation for NISTmAb and the non-originator expressing cell lines. Each of the cell lines demonstrated a reasonable degree of similarity, but did show differences in glycoslylation, N- and C-terminal processing, among other proteoforms

differences. The resulting heterogeneity caused a rather low signal to noise ratio in the intact mass spectrometry results. The N-terminal sequence heterogeneity due to incomplete processing of the signal peptide results in an additional serine residue on the light chain signal. This difference in primary structure is a small, albeit important factor in considering the potential analytical similarity of these species. Additional differences were notes in the glcyoprofile, although intact mass spectrometry allows only identification of relatively high abundance glycoforms. The second phase of this study involved additional glycoanalytical characterization to determine the degree of glycan heterogeneity.

A glycan release assay is an efficient tool in determining complete glycan heterogeneity including the detection of low abundant glycans. The creation of a data archive encompassing identified glycan structures in NISTmAb was used for preliminary analysis completed on the glycan structures in the 3 cell lines expressing the non-originator NISTmAb. Evaluation of the non-originators demonstrated an increase in terminal galactosylation, sialylation, high-mannose, and afucosylated species compared to the NISTmAb. Each of these characteristic glycan species play an important role in efficacy, immunogenicity, and pharmacokinetics of monoclonal antibodies and must be monitored during drug product production and development. The glycan release assay did show some level of similarity between the NISTmAb and these non-originator cell lines; it ultimately showed that these cell lines are more similar to one another then to the NISTmAb. The next stage of this project will require bioreactor screening of all 3 cell lines to observe optimum growth conditions and evaluate analytical similarity. Specific glycan characteristics such as % fucosylation, % sialyation and % galactosylation can be altered

using bioreactor conditions which will be instrumental as the clonal selection and process development optimization progresses (Agarabi et al. 2015).

The non-originator cell lines are primarily intended as an open innovation test case for upstream and downstream technology innovation, however the obvious implications toward biosimilar-driven research warrant discussion, and selection of the ideal clone would have highly similar glycosylation. The FDA guidance on scientific considerations for demonstrating biosimilarity recommends a stepwise approach in which each milestone indicates sufficient similarity to proceed, with any residual uncertainty being resolved through a hierarchy of measurements (FDA 2015a). In general, these steps can include demonstration of similarity with regard to structure and binding activity (analytical similarity), animal toxicity, human pharmacokinetics, and clinical safety and efficacy (FDA 2015a). Ultimately an intended biosimilar is granted biosimilar status upon demonstration that the "biological product is highly similar to the reference product notwithstanding minor differences in clinically inactive components" considering the totality of evidence (FDA 2015b).

The NISTmAb and non-originator materials described herein are not intended for clinical use, therefore they are most relevant to the analytical similarity stage of biosimilar development. Currently, the "goalpost" for analytical similarity is defined by the historical performance of the originator's process and process changes/improvements via characterization of multiple commercial drug product lots to form an analytical target profile (ATP) (FDA 2015a; Bandyopadhyay et al. 2015; Liu et al. 2016; Chow et al. 2017)] Of particular interest to these cell lines is that the RM 8671 ATP is that of a single homogenized batch. Therefore, the target profile of the requisite product quality attributes

will likely be small in comparison to a product whose lifecycle includes inevitable process changes. Such a system provides an analytical similarity challenge as the target is narrowed to the capability of current-state-of-the-art analytical characterization methods.

The ATP may include analytical characterization (primary structure, HOS, size variants, charge variants, glycosylation and other PTMs, etc.), biophysical characterization (HOS, colloidal stability, conformational stability), as well as biological characterization (in vitro binding studies, potency assays, etc.) (Bandyopadhyay et al. 2015; Liu et al. 2016). A subset of analytical measurements was used to characterize the non-originator materials described here demonstrating moderate differences. It is suspected that with further process optimizing (feed media, supplementation strategy, downstream purification, etc.) in one or more of these cell lines, a product with a higher degree of analytical similarity to the NISTmAb can be achieved. Extended characterization to include additional orthogonal analytical measurements combined with well documented process history would provide a unique pre-competitive analytical similarity exercise.

The non-originator NISTmAb characterization described throughout this thesis was broken down into a 3 phase study beginning with the determination of "high-level" similarity using intact analysis, evaluation of monomeric purity via SEC, and followed by a released glycan assay performed here. Each of these analytical assays are being used to guide clonal selection and for process feedback optimization and control. The intact mass analysis demonstrated a reasonable degree of sameness with regard to primary structure, however an N- and C-terminal sequence heterogeneity was noted as a potential development risk. The SEC studies indicated only moderate differences in aggregation profile that should be easily removable using an additional (e.g. cation exchange

chromatography) purification step. The more extensive released glycan analysis also demonstrated a reasonable degree of sameness between the non-originator materials and NISTmAb to warrant further bioreactor screening (e.g. temperature, feeding strategy, etc.). The analytical data described here provides a snapshot of product quality attributes for the first batches produced. The analytical methods will become critical components of the non-originator cell line lifecycle and continue to provide data as continued bioreactor optimization samples are produced.

References

Walsh G. 2014. Biopharmaceutical benchmarks 2014. Nature biotechnology. 32(10):992.

Newswire P. Global monoclonal antibodies market hit $100 billion in 2017: Report.

Jefferis R. 2009. Glycosylation as a strategy to improve antibody-based therapeutics. Nature reviews Drug discovery. 8(3):226.

Nelson AL, Dhimolea E, Reichert JM. 2010. Development trends for human monoclonal antibody therapeutics. Nature reviews drug discovery. 9(10):767.

Schiel J, Davis D, Borisov O. 2014. State-of-the-art and emerging technologies for therapeutic monoclonal antibody characterization volume 1. Monoclonal antibody therapeutics: Structure, function, and regulatory space preface. American Chemical Society, ACS Symposium Series Washington, DC. p. Ix-Xi.

Schiel JE. 2012. Glycoprotein analysis using mass spectrometry: Unraveling the layers of complexity. Analytical and bioanalytical chemistry. 404(4):1141-1149.

Mimura Y, Katoh T, Saldova R, O'Flaherty R, Izumi T, Mimura-Kimura Y, Utsunomiya T, Mizukami Y, Yamamoto K, Matsumoto T. 2018. Glycosylation engineering of therapeutic igg antibodies: Challenges for the safety, functionality and efficacy. Protein & cell. 9(1):47-62.

Liu H, Gaza-Bulseco G, Faldu D, Chumsae C, Sun J. 2008. Heterogeneity of monoclonal antibodies. Journal of pharmaceutical sciences. 97(7):2426-2447.

van de Bovenkamp FS, Hafkenscheid L, Rispens T, Rombouts Y. 2016. The emerging importance of igg fab glycosylation in immunity. The Journal of Immunology. 196(4):1435-1441.

Beck A, Wagner-Rousset E, Ayoub D, Van Dorsselaer A, Sanglier-Cianferani S. 2013. Characterization of therapeutic antibodies and related products. Anal Chem. 85(2):715-736.

Cymer F, Beck H, Rohde A, Reusch D. 2018. Therapeutic monoclonal antibody n-glycosylation–structure, function and therapeutic potential. Biologicals. 52:1-11.

Jefferis R. 2014. Monoclonal antibodies: Mechanisms of action. State-of-the-art and emerging technologies for therapeutic monoclonal antibody characterization volume 1

monoclonal antibody therapeutics: Structure, function, and regulatory space. American Chemical Society. p. 35-68.

Owen JA, Punt J, Stranford SA. 2013. Kuby immunology. WH Freeman New York.

Wang W, Singh S, Zeng DL, King K, Nema S. 2007. Antibody structure, instability, and formulation. Journal of pharmaceutical sciences. 96(1):1-26.

Formolo T, Ly M, Levy M, Kilpatrick L, Lute S, Phinney K, Marzilli L, Brorson K, Boyne M, Davis D. 2015. Determination of the nistmab primary structure. State-of-the-art and emerging technologies for therapeutic monoclonal antibody characterization volume 2 biopharmaceutical characterization: The nistmab case study. ACS Publications. p. 1-62.

Zhang P, Woen S, Wang T, Liau B, Zhao S, Chen C, Yang Y, Song Z, Wormald MR, Yu C. 2016. Challenges of glycosylation analysis and control: An integrated approach to producing optimal and consistent therapeutic drugs. Drug discovery today. 21(5):740-765.

Walsh G. 2010. Post-translational modifications of protein biopharmaceuticals. Drug discovery today. 15(17-18):773-780.

Zhang H, Ge Y. 2011. Comprehensive analysis of protein modifications by top-down mass spectrometry. Circulation: Cardiovascular Genetics. 4(6):711-711.

Barton C, Spencer D, Levitskaya S, Feng J, Harris R, Schenerman MA. 2014. Heterogeneity of iggs: Role of production, processing, and storage on structure and function. State-of-the-art and emerging technologies for therapeutic monoclonal antibody characterization volume 1 monoclonal antibody therapeutics: Structure, function, and regulatory space. American Chemical Society. p. 69-98.

Zhou Q, Qiu H. 2018. The mechanistic impact of n-glycosylation on stability, pharmacokinetics and immunogenicity of therapeutic proteins. Journal of pharmaceutical sciences.

Stanley P, Schachter H, Taniguchi N. 2009. N-glycans. In: nd, Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, Etzler ME, editors. Essentials of glycobiology. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press

The Consortium of Glycobiology Editors, La Jolla, California.

Reusch D, Tejada ML. 2015. Fc glycans of therapeutic antibodies as critical quality attributes. Glycobiology. 25(12):1325-1334.

Aich U, Lakbub J, Liu A. 2016. State-of-the-art technologies for rapid and high-throughput sample preparation and analysis of n-glycans from antibodies. Electrophoresis. 37(11):1468-1488.

Anthony L, Plummer TH. 1987. [63] peptide-n4-(n-acetyl-β-glucosaminyl) asparagine amidase and endo-β-n-acetylglucosaminidase from flavobacterium meningosepticum. Methods in enzymology. Academic Press. p. 770-778.

Prien JM, Stöckmann H, Albrecht S, Martin SM, Varatta M, Furtado M, Hosselet S, Wang M, Formolo T, Rudd PM et al. 2015. Orthogonal technologies for nistmab n-glycan structure elucidation and quantitation. State-of-the-art and emerging technologies for therapeutic monoclonal antibody characterization volume 2 biopharmaceutical characterization: The nistmab case study. American Chemical Society. p. 185-235.

Ruhaak L, Zauner G, Huhn C, Bruggink C, Deelder A, Wuhrer M. 2010. Glycan labeling strategies and their use in identification and quantification. Analytical and bioanalytical chemistry. 397(8):3457-3481.

Lauber MA, Yu Y-Q, Brousmiche DW, Hua Z, Koza SM, Magnelli P, Guthrie E, Taron CH, Fountain KJ. 2015. Rapid preparation of released n-glycans for hilic analysis using a labeling reagent that facilitates sensitive fluorescence and esi-ms detection. Analytical chemistry. 87(10):5401-5409.

Vainauskas S, Kirk CH, Petralia L, Guthrie EP, McLeod E, Bielik A, Luebbers A, Foster JM, Hokke CH, Rudd PM. 2018. A novel broad specificity fucosidase capable of core α1-6 fucose release from n-glycans labeled with urea-linked fluorescent dyes. Scientific reports. 8(1):9504.

Jones A, Kimzey M, Yan J, Sharma V, Guerrero A, Gyenes A, Hyche J, Dale E, Haxo T, Vlasenko S. 2016. Development of a 5-minute deglycosylation method and instant labeling dye for high-throughput n-glycan analysis by mass spectrometry. No. 0959-6658.

Yan J, Guerrero A, Galermo A, Haxo T, Vlasenko S, Hyche J, Rice T, Jones A. 2019. Streamlined workflows for n-glycan analysis of biotherapeutics using instantpc and 2-ab with lc-fld-ms. ProZyme, Agilent Technologies

Higel F, Demelbauer U, Seidl A, Friess W, Sörgel F. 2013. Reversed-phase liquid-chromatographic mass spectrometric n-glycan analysis of biopharmaceuticals. Analytical and bioanalytical chemistry. 405(8):2481-2493.

Guile GR, Rudd PM, Wing DR, Prime SB, Dwek RA. 1996. A rapid high-resolution high-performance liquid chromatographic method for separating glycan mixtures and analyzing oligosaccharide profiles. Analytical biochemistry. 240(2):210-226.

Han L, Costello CE. 2013. Mass spectrometry of glycans. Biochemistry (Moscow). 78(7):710-720.

Dass C. 2007. Fundamentals of contemporary mass spectrometry. John Wiley & Sons.

Ho CS, Lam C, Chan M, Cheung R, Law L, Lit L, Ng K, Suen M, Tai H. 2003. Electrospray ionisation mass spectrometry: Principles and clinical applications. The Clinical Biochemist Reviews. 24(1):3.

Hoffman ED, Stroobant V. 2007. Mass spectrometry: Principles and applications. West Sussex: John Wiley & Sons, Bruxellas, Bélgica. 1(2):85.

Gaskell SJ. 1997. Electrospray: Principles and practice. Journal of mass spectrometry. 32(7):677-688.

Aebersold R, Mann M. 2003. Mass spectrometry-based proteomics. Nature. 422(6928):198.

Hage DS, Carr JD. 2011. Analytical chemistry and quantitative analysis. Prentice Hall Boston.

Chen H, Horváth C. 1995. High-speed high-performance liquid chromatography of peptides and proteins. Journal of Chromatography A. 705(1):3-20.

Le JC, Bondarenko PV. 2005. Trap for mabs: Characterization of intact monoclonal antibodies using reversed-phase hplc on-line with ion-trap mass spectrometry. Journal of the American Society for Mass Spectrometry. 16(3):307-311.

Sousa F, Gonçalves VM, Sarmento B. 2017. Development and validation of a rapid reversed-phase hplc method for the quantification of monoclonal antibody bevacizumab from polyester-based nanoparticles. Journal of pharmaceutical and biomedical analysis. 142:171-177.

Schiel JE, Davis DL, Borisov OV. 2015a. State-of-the-art and emerging technologies for therapeutic monoclonal antibody characterization volume 2. Biopharmaceutical characterization: The nistmab case study. ACS Publications.

Schiel JE, Davis DL, Borisov OV. 2015b. State-of-the-art and emerging technologies for therapeutic monoclonal antibody characterization volume 3. Defining the next generation of analytical and biophysical techniques. ACS Publications.

Schiel JE, Turner A. 2018. The nistmab reference material 8671 lifecycle management and quality plan. Analytical and bioanalytical chemistry. 410(8):2067-2078.

Kashi L, Yandrofski K, Preston RJ, Arbogast LW, Giddens JP, Marino JP, Schiel JE, Kelman Z. 2018. Heterologous recombinant expression of non-originator nistmab. mAbs. 10(6):922-933.

Schiel JE, Turner A, Mouchahoir T, Yandrofski K, Telikepalli S, King J, DeRose P, Ripple DC, Phinney KW. 2018. The nistmab reference material 8671 value assignment, homogeneity, and stability. Analytical and bioanalytical chemistry.

Biosimilar product information. 2019. [accessed]. https://www.fda.gov/drugs/biosimilars/biosimilar-product-information.

Turner A, Yandrofski K, Telikepalli S, King J, Heckert A, Filliben J, Ripple D, Schiel JE. 2018. Development of orthogonal nistmab size heterogeneity control methods. Analytical and bioanalytical chemistry. 410(8):2095-2110.

Turner A, Schiel JE. 2018. Qualification of nistmab charge heterogeneity control assays. Analytical and bioanalytical chemistry. 410(8):2079-2093.

Wong DL. 2017. Precise characterization of intact monoclonal antibodies by the agilent 6545xt advancebio lc/q-tof. Agilent Technologies Application Note.

Hilliard M, Alley WR, Jr., McManus CA, Yu YQ, Hallinan S, Gebler J, Rudd PM. 2017. Glycan characterization of the nist rm monoclonal antibody using a total analytical solution: From sample preparation to data analysis. mAbs. 9(8):1349-1359.

Banerjee S, Mazumdar S. 2012. Electrospray ionization mass spectrometry: A technique to access the information beyond the molecular weight of the analyte. International journal of analytical chemistry. 2012.

Kilpatrick EL, Liao WL, Camara JE, Turko IV, Bunk DM. 2012. Expression and characterization of 15n-labeled human c-reactive protein in escherichia coli and pichia pastoris for use in isotope-dilution mass spectrometry. Protein expression and purification. 85(1):94-99.

Petersen TN, Brunak S, Von Heijne G, Nielsen H. 2011. Signalp 4.0: Discriminating signal peptides from transmembrane regions. Nature methods. 8(10):785.

Prozyme. 2017. Gly-x n-glycan rapid release and labeling with instantpc kit (user manual).

Agarabi CD, Schiel JE, Lute SC, Chavez BK, Boyne II MT, Brorson KA, Khan M, Read EK. 2015. Bioreactor process parameter screening utilizing a plackett-burman design for a model monoclonal antibody. Journal of pharmaceutical sciences. 104(6):1919-1928.

Konno Y, Kobayashi Y, Takahashi K, Takahashi E, Sakae S, Wakitani M, Yamano K, Suzawa T, Yano K, Ohta T. 2012. Fucose content of monoclonal antibodies can be controlled by culture medium osmolality for high antibody-dependent cellular cytotoxicity. Cytotechnology. 64(3):249-265.

Pacis E, Yu M, Autsen J, Bayer R, Li F. 2011. Effects of cell culture conditions on antibody n-linked glycosylation—what affects high mannose 5 glycoform. Biotechnology and bioengineering. 108(10):2348-2358.

FDA US. 2015a. Guidance for industry: Scientific considerations in demonstrating biosimilarity to a reference product. Maryland: FDA.

FDA U. 2015b. Quality considerations in demonstrating biosimilarity of a therapeutic protein product to a reference product. USA: FDA.

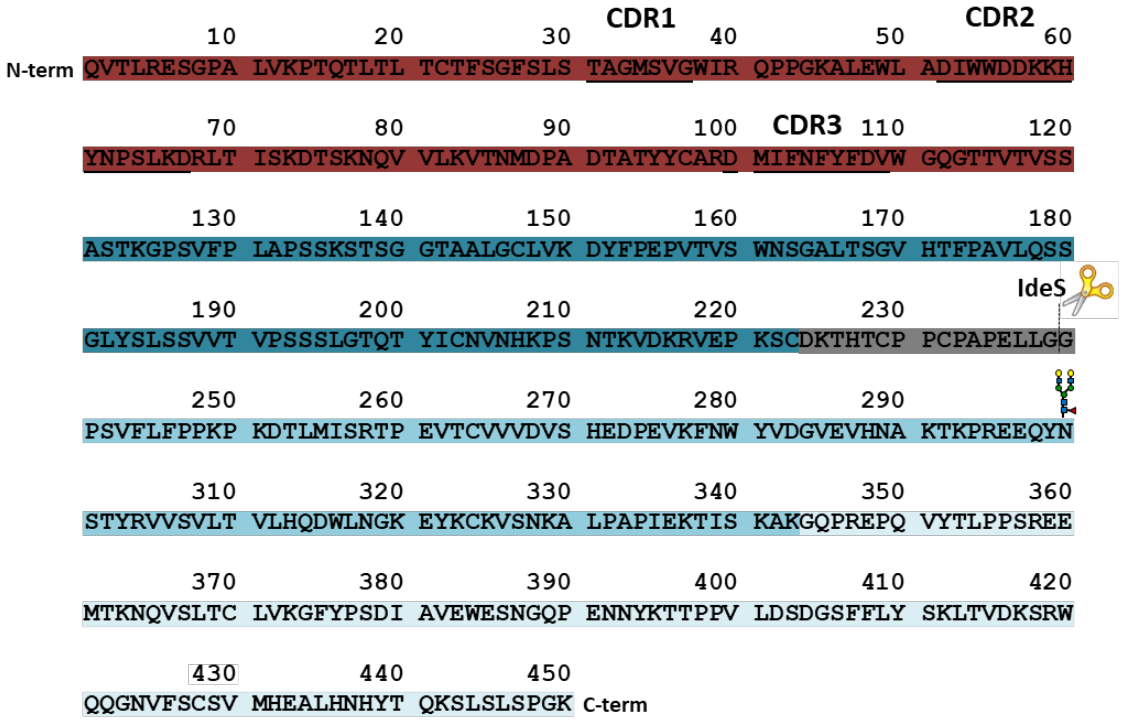
Bandyopadhyay S, Mahajan M, Mehta T, Singh AK, Parikh A, Gupta AK, Kalita P, Patel M, Mendiratta SK. 2015. Physicochemical and functional characterization of a biosimilar adalimumab zrc-3197. Biosimilars. 5:1.

Liu J, Eris T, Li C, Cao S, Kuhns S. 2016. Assessing analytical similarity of proposed amgen biosimilar abp 501 to adalimumab. BioDrugs. 30(4):321-338.

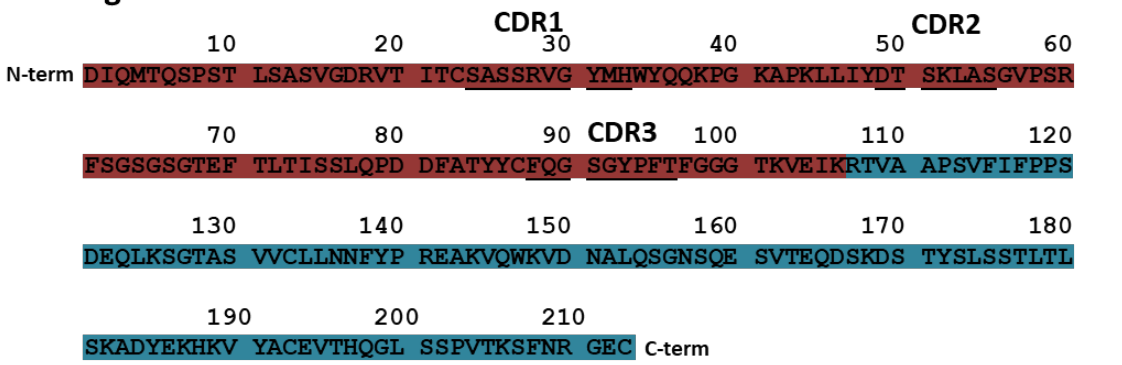
Chow SC, Song F, Bai H. 2017. Sample size requirement in analytical studies for similarity assessment. Journal of biopharmaceutical statistics. 27(2):233-238.

Appendix I. NISTmAb Primary Sequence.

Heavy Chain



Light Chain



Appendix II. InstantPC Human IgG N-Glycan Library

Glycan	Mass
A2	1577.6343
FA2	1723.6922
FA2B	1926.7716
A2G1	1739.6871
FA2G1	1885.7450
FA2BG1	2088.8244
A2G2	1901.7399
FA2G2	2047.7978
FA2BG2	2250.8772
FA2G1S1	2176.8404
A2GS1	2192.8353
FA2G2S1	2338.8932
FA2BG2S1	2541.9726
A2G2S2	2483.9307
FA2G2S2	2629.9886
FA2BG2S2	2833.0650

Appendix III. Composition of glycan classification groups between NISTmAb and non-originator cell lines

Name	Proposed Composition	NISTmAb	NS0-59	NS0-60	NS0-66
A1	H3N3	0.343	0.0900	0.2167	0.2567
FM3	H3N2F1	0.130	ND	ND	0.0267
FA1	H3N3F1	2.508	1.5267	1.2267	4.3833
A2	H3N4	0.088	0.1167	0.2700	0.1300
A1G1[a]	H4N3	ND	ND	0.0867	0.1000
[FA1-N][ISD]	[H3N3F1-N][ISD]	ND	0.4100	0.3633	0.4533
FA2	H3N4F1	42.030	26.8867	23.0333	29.2467
A1G1[b]	H4N3	0.073	0.0467	0.1700	0.1567
M5	H5N2	0.765	0.8833	1.8600	1.0267
M4	H4N2	ND	ND	0.0467	0.0300
FA3	H3N5F1	0.405	0.2333	0.1200	0.2067
FA2B	H3N5F1	ND	ND	0.0633	0.0600
A2G1[a]	H4N4	0.040	0.0733	0.2500	0.1133
FA1G1	H4N3F1	2.938	1.1200	1.7067	3.0800
A2G1[b]	H4N4	ND	ND	0.2133	0.0433
A1G1Ga1	H5N3	ND	0.0600	0.2300	0.2467
[FA2G1]-Ga[ISD]	H4N3F1	0.228	0.2367	0.2500	0.2050
FA2G1[a]	H4N4F1	24.593	26.1333	25.4000	21.0133
A1G1Ga1	H5N3	0.050	ND	0.0967	0.0700
FA2G1[b]	H4N4F1	11.753	9.9933	11.2167	7.7967
FA3G1[a]	H4N5F1	0.320	0.2667	0.1233	0.1700
FA1G1Ga1[a] or FM4A1G1	H5N3F1-A	0.123	ND	ND	ND
FA3G1[b]	H4N5F1	0.363	0.2500	0.1133	0.1667
FA1G1Ga1[b]	H5N3F1-B	1.205	ND	ND	ND
FA2G1Ga1	H5N4F1	ND	1.7333	0.3233	1.2367
M5A1G1	H6N3	0.143	0.0600	0.4333	0.2700
FA2G2	H5N4F1	6.995	11.1767	16.9267	7.1600
FA3G2[a]	H5N5F1	0.260	0.3067	0.6167	0.1633
FA3G2[b]	H5N5F1	0.085	0.0633	0.0767	0.0400
FM5A1G1 or FM4A1G1Ga1	H6N3F1	0.090	ND	0.1367	0.0767
FA1G1Sg1[a]	H4N3F1Sg1	1.238	4.3033	2.7700	12.7133
FA2G1Sg1[a]	H4N4F1Sg1	0.078	2.1233	0.5833	1.5233

FA1G1Sg1 [b]	H4N3F1Sg1	ND	ND	ND	0.0667
FA2G2Ga1[a]	H6N4F1	1.173	1.7233	2.8033	1.1100
FA2G1Sg1[b]	HN4F1Sg1	0.323	1.9333	1.0033	1.7033
FA2G2Ga1[b]	H6N4F1	0.278	0.5300	0.9767	0.3133
FA3G3	H6N5F1	0.138	0.2300	0.0733	0.1100
FA2G2Sg1	H5N4F1Sg1	0.280	4.7667	2.4533	3.0633
FA2G2Ga2	H7N4F1	0.623	1.0933	2.5300	0.5533
FA2G2Ga1Sg1	H6N4F1Sg1	0.195	1.5700	1.2400	1.0433

[a,b]. isomer for the proposed glycan/composition was identified. Glycans labeled with an [a] eluted first, followed by the second isomer labeled with an [b]. Glycan assignments were made based on putative compositions expected from the glycan biosynthetic pathway.