

Method development and its application for *N*-glycan analysis of influenza A virus antigens

Dissertation

Zur Erlangung des akademischen Grades

Doktoringenieur

(Dr.-Ing.)

von: Dipl.-Ing. Alexander Pralow

geb. am: 18. April 1990 in Magdeburg

genehmigt durch die Fakultät für Verfahrens- und Systemtechnik
der Otto-von-Guericke-Universität Magdeburg

Promotionskommission: Prof. Dr.-Ing. Andreas Seidel-Morgenstern

Prof. Dr.-Ing. Udo Reichl

Prof. Dr. Bernd Lepenies

Dr. Erdmann Rapp

eingereicht am: 29.09.2022

Promotionskolloquium am: 13.04.2023

KURZFASSUNG

Die Glykosylierung gilt als kritisches Qualitätsmerkmal für die Herstellung von rekombinanten Biopharmazeutika wie monoklonalen Antikörpern – jedoch nicht für virale Impfstoffe. Die Antigene des Influenza-A-Virus (IAV), Hämagglutinin (HA) und Neuraminidase (NA), haben mehrere potentielle *N*-Glykosylierungsstellen. In der Vergangenheit wurden starke Auswirkungen auf Virulenz und Immunogenität in Abhängigkeit der Veränderung der Mikro- und Makroheterogenität der *N*-Glykosylierung nachgewiesen.

Früheren analytischen Verfahren fehlt jedoch die Fähigkeit, sowohl die fein-strukturelle Glykan Information (spezifische Struktur der Oligosaccharose), als auch die Position eines Glykans auf dem Glykoprotein (Ortsspezifität) parallel zu identifizieren. Um diese Informationen zu erhalten, wurden in dieser Arbeit zwei Flüssigchromatographie-gekoppelte Massenspektrometrie-basierte Methoden kombiniert. Zusätzlich wurde zur Identifizierung aller potenziellen *N*-Glykosylierungsstellen der IAV-Glykoproteine, eine neu entdeckte „Glyko“-Spezifität des Enzyms Flavastacin validiert, die zu einer spezifischen Spaltung am C-Terminus von *N*-glykosyliertem Asparagin führt. Somit konnte die Anzahl identifizierter Glykosylierungsstellen optimiert werden. Darüber hinaus wurde die Flüssigchromatographie mittels porösem graphitisierten Carbon mit einem nach der Säule geschaltetem zusätzlichen Fluss mit 100% Acetonitril bestückt, da aufgrund der hohen Oberflächenspannung des wässrigen Puffers die Bildung von Tröpfchen erfolgte. Dies führte zu einem stabilen Spray und somit zu einer verbesserten Signaldetektion von freien *N*- und *O*-glykanen, insbesondere bei geringerer Retentionszeit. Dadurch wurden die Anzahl und die Qualität von identifizierten fein-strukturellen Glykanen optimiert.

Die Kombination beider Erfindungen ermöglichte die umfassende *N*-Glykosylierungsanalyse beider IAV-Antigene und damit eine kombinierte Charakterisierung der spezifischen *N*-Glykan-Strukturen (Verlinkung, Verzweigung, Art des Epitopes) der viralen Antigene sowie deren *N*-Glykan-Mikro- und Makroheterogenitäten. Dieser Arbeitsablauf wurde mit dem IAV Stamm A/PR/8/34 H1N1 getestet, welcher in zwei verschiedenen "Madin-Darby Canine Kidney" (MDCK)-Zelllinien vermehrt wurde, nämlich einer adhärennten und einer suspensions MDCK-Zelle-Linie.

Die Analyse der viralen Antigen-Glykosylierung beider Wirtszelllinien führte zu ähnlichen *N*-Glykanen, jedoch mit unterschiedlichen relativen Häufigkeiten der einzelnen *N*-Glykan-Strukturen. Eine detaillierte Analyse der HA-*N*-Glykan-Mikroheterogenität zeigte im Vergleich zur HA Stammregion eine Zunahme der *N*-Glykan-Variabilität und -Komplexität für *N*-Glykosylierungsstellen, die näher an der Kopfregion des Moleküls liegen. Im Gegensatz dazu wurde festgestellt, dass NA ausschließlich an der Stelle N73 *N*-glykosyliert ist. Weiter wurde gezeigt, dass fast alle *N*-Glykan-Strukturen am Kern fucosyliert waren (meistens in Kombination mit einer Antennen-Fucosylierung. *N*-Glykan Strukturen am HA und NA-Antigen waren ausschließlich hybrid- und komplexartige Strukturen, welche teilweise mit alpha-verknüpften Galactosen (Galili-Epitop) sowie Epitopen der Blutgruppe H Typ 2 und der Blutgruppe A terminiert sind. Mit Galili-Epitopen dekorierte Strukturen sind in menschlichen Zellen nicht zu finden und könnten daher im Zusammenhang mit einem Influenza Impfstoff zu einer erhöhten Immunogenität führen und als eine Art Adjuvans fungieren. Eine bessere Charakterisierung der *N*-glykosylierung von viralen Antigenen durch diesen Arbeitsablauf generiert umfassende Datensätze, welche in Kombination mit Tierversuchen bezüglich der Immunogenität, eine Evaluierung der Bedeutung der Glykosylierung von viralen Proteinen als Qualitätsattribut in der Impfstoffherstellung ermöglichen.

ABSTRACT

Glycosylation is considered as a critical quality attribute for the production of recombinant biopharmaceuticals such as monoclonal antibodies – but not for viral vaccines. The influenza A virus (IAV) antigens hemagglutinin (HA) and neuraminidase (NA) have multiple potential *N*-glycosylation sites. In the past, strong effects on virulence and immunogenicity upon alteration of the *N*-glycan micro- and macroheterogeneity were demonstrated.

Previous analytical methods lack the ability of providing both, the fine structural glycan information (specific structure of the oligosaccharose) and the location of a glycan on the glycoprotein (site-specificity). Here, two liquid chromatography coupled to mass spectrometry (LC-MS) based methods were combined to obtain all this information from the glycoproteins. Additionally, to identify all potential *N*-glycosylation sites of the IAV glycoproteins, a newly identified “glyco”-specificity of the enzyme flavastacin was found resulting in a specific cleavage of the C-terminus of *N*-glycosylated asparagine. This optimized the number of detected glycosylation sites. Furthermore, porous grafitized carbon (PGC)-LC was equipped with a post-column make-up flow of 100% acetonitril, because the high surface tension of the aquatic buffer system resulted in the formation of droplets. This resulted in a more stable spray and therefore in an increased MS signal detection of released *N*- and *O*-glycans, especially at lower retention time. Thereby the number and quality of identified fine-structural glycans was optimized.

Combining both approaches enabled the comprehensive *N*-glycosylation analysis of both IAV antigens and provided a detailed characterization of the viral antigen *N*-glycan structures (linkages, branching, type of epitopes) and of the *N*-glycan micro- and macroheterogeneities. This workflow was tested with

ABSTRACT

the IAV strain A/PR/8/34 H1N1 propagated in two different Madin-Darby Canine Kidney (MDCK) cell lines, namely an adherent MDCK cell line and its corresponding suspension cell line.

Viral antigen glycosylation of both host cell lines resulted in similar *N*-glycan patterns, but with different relative abundances of the individual *N*-glycan structures. Detailed analysis of the HA *N*-glycan microheterogeneity showed an increase of the *N*-glycan variability and complexity for *N*-glycosylation sites located closer to the head region of the molecule. In contrast, NA was found to be exclusively *N*-glycosylated at site N73. In general, almost all *N*-glycan structures were found to be core-fucosylated (mostly in combination with an antenna-fucosylation). Furthermore, HA and NA *N*-glycan structures were found to be exclusively hybrid- and complex-type structures, to some extent terminated with alpha-linked galactoses (Galili-epitope) as well as blood group H type 2 and blood group A epitopes. Galili-epitope decorated structures cannot be found in human cells and could therefore lead to increased immunogenicity, functioning as a sort of adjuvant in combination with an influenza vaccine. A better characterization of viral antigen *N*-glycosylation by using this workflow resulting in very comprehensive data sets, that, together with animal studies on immunogenicity, will allow evaluating the importance of viral protein glycosylation as a quality attribute in vaccine manufacturing.

ERKLÄRUNG

Ich erkläre hiermit, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht.

Insbesondere habe ich nicht die Hilfe einer kommerziellen Promotionsberatung in Anspruch genommen. Dritte haben von mir weder unmittelbar noch mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form als Dissertation eingereicht und ist als Ganzes auch noch nicht veröffentlicht.

Magdeburg, am

Dipl.-Ing. Alexander Pralow

CONTENTS

KURZFASSUNG	III
ABSTRACT.....	V
ERKLÄRUNG	VII
CONTENTS.....	VIII
LIST OF TABLES.....	XI
LIST OF FIGURES.....	XII
LIST OF ABBREVIATIONS	XIII
1. MOTIVATION	1
2. BACKGROUND	4
2.1. Glycoproteins.....	4
2.1.1. Glycobiology: from sugar to biological function	4
2.1.2. Glycan structures	8
2.1.3. Protein glycosylation: <i>N</i> - and <i>O</i> -glycosylation	10
2.2. Analytics in glycobotechnology	11
2.2.1. Glycomics	11
2.2.2. Glycoproteomics	14
2.2.2.1. Digestion of glycoproteins	14
2.2.2.2. LC-MS measurement of glycopeptides	16
2.2.3. Enrichment strategies for glyco(-conjugate) analysis.....	20
2.3. Influenza A virus	21
2.3.1. Structure	21
2.3.2. Surface antigens – hemagglutinin and neuraminidase.....	22
2.3.3. Replication cycle	24
2.4. Influenza vaccines.....	26
2.5. The importance of the glycosylation of influenza A virus proteins on vaccine preparations	28
3. MATERIALS AND METHODS.....	32
3.1. Cell lines, cell cultivation and virus infection	33
3.2. Virus harvesting, purification and inactivation	34

CONTENTS

3.3.	Protein concentration assay.....	34
3.4.	<i>N</i> -glycan analysis: Sample preparation, measurement and data analysis.....	34
3.4.1.	Chemicals, solvents and buffers.....	34
3.4.2.	<i>N</i> - and <i>O</i> -glycan release.....	35
3.4.3.	Desalting and enrichment of glycans.....	36
3.4.4.	<i>N</i> - and <i>O</i> -glycan measurement by nano-PGC-LC-MS(/MS).....	37
3.4.5.	Analysis of PGC-LC-MS(-MS) data from releases glycans.....	39
3.5.	Glycopeptide analysis: Sample preparation, measurement and data analysis.....	39
3.5.1.	Chemicals, solvents and buffers.....	39
3.5.2.	Proteolytic digestion via filter-aided sample preparation.....	40
3.5.3.	Enrichment of glycopeptides.....	41
3.5.4.	<i>N</i> -glycopeptide measurement by RP-LC-MS(/MS).....	41
3.5.1.	Analysis of RP-LC-MS(/MS) data from <i>N</i> -glycopeptides.....	42
3.5.2.	Relative quantification of <i>N</i> -glycopeptides.....	44
3.6.	Graphical illustration.....	45
4.	RESULTS AND DISCUSSION.....	46
4.1.	Improvement of the glycoproteomic toolbox with the discovery of a unique C-terminal cleavage specificity of flavastacin for <i>N</i> -glycosylated asparagine.....	46
4.1.1.	Introduction.....	46
4.1.2.	Results and discussion.....	48
4.2.	Improvement of electrospray stability in negative ion mode for nano-PGC-LC-MS glycoanalysis via post-column make-up flow.....	56
4.2.1.	Introduction.....	56
4.2.2.	Results.....	56
4.2.3.	Discussion.....	63
4.3.	Comprehensive <i>N</i> -glycosylation analysis of the IAV proteins HA and NA from adherent and suspension MDCK cells.....	65
4.3.1.	Introduction.....	65
4.3.2.	Results.....	66
4.3.2.1.	Characterization of influenza <i>N</i> -glycans fine-structure.....	66
4.3.2.2.	Characterization of influenza glycopeptides.....	73

CONTENTS

4.3.2.3.	Assembling of structural and site-specific <i>N</i> -glyco analysis results.....	81
4.3.3.	Discussion.....	81
4.3.3.1.	Novel methods used for influenza <i>N</i> -glycan analysis	82
4.3.3.2.	Influenza antigen glycosylation.....	86
5.	CONCLUSIONS	91
6.	OUTLOOK.....	94
	REFERENCES.....	96
	APPENDIX.....	110
A.	Supplementary information	110
A.1.	Improvement of the glycoproteomic toolbox with the discovery of a unique C-terminal cleavage specificity of flavastacin for <i>N</i> -glycosylated asparagine	110
A.2.	Comprehensive <i>N</i> -glycosylation analysis of the IAV proteins HA and NA from adherent and suspension MDCK cells.....	113
B.	List of materials.....	137
C.	Publications.....	142
D.	Attended conferences	144
D.1.	Talks.....	144
D.2.	Posters	144
E.	Supervised projects.....	145
E.1.	Students	145
E.2.	Other projects	145

LIST OF TABLES

Table 2-1: <i>List of monosaccharides used in this work.</i>	5
Table 2-2: <i>Overview of proteases used for mass spectrometry-based protein analysis.</i>	15
Table 2-3: <i>Monoisotopic masses of oxonium ions.</i>	19
Table 4-1: <i>Manually annotated N-glycopeptide sequences of human lactotransferrin (hLTF) from MS(/MS) spectra of nano-RP-LC-ESI-OT-MS²(HCD) measurements after sequential digestion of hLTF with trypsin and flavastacin.</i>	48
Table 4-2: <i>Proteome Discoverer results of a nano-RP-LC-ESI-OT-MS²(HCD) measurement after sequential digestion of BSA with trypsin and flavastacin.</i>	53
Table 4-3: <i>N- and O-glycan compositions released from bovine fetuin detected via nano-PGC-LC-ESI-MS(/MS) with and without the use of a PCMF.</i>	62
Table 4-4: <i>Identified (glyco)-peptide sequences of IAV surface glycoproteins hemagglutinin and neuraminidase.</i>	75

LIST OF FIGURES

Figure 2-1: <i>Illustration of a β 1-4 glycosidic bond between two galactoses.</i>	6
Figure 2-2: <i>Importance of glycobiology for the human organism.</i>	7
Figure 2-3: <i>Blood group antigens.</i>	8
Figure 2-4: <i>Exemplary structures of N- and O-glycans.</i>	9
Figure 2-5: <i>Nomenclature of fragment ions derived from a peptide sequence.</i>	18
Figure 2-6: <i>Specific B-fragment ions of an N-glycopeptide with the corresponding monoisotopic masses.</i>	19
Figure 2-7: <i>Schematic illustration of IAV.</i>	22
Figure 2-8: <i>Three dimensional structure of IAV glycoproteins HA (A) and NA (B).</i>	23
Figure 3-1: <i>Workflow.</i>	32
Figure 3-2: <i>Schematic of the PCMF setup.</i>	38
Figure 4-1: <i>Base peak ion-chromatogram (BPC) and oxonium ion related extracted ion-chromatograms (EIC) of MS(/MS) spectra of nanoRP-LC-ESI-OT-MS²(HCD) measured hLTF after sequential digestion with trypsin and flavastacin.</i>	49
Figure 4-2: <i>Fragment ion spectra of nanoRP-LC-ESI-OT-MS²(HCD) measured hLTF N-glycopeptides after sequential digestion with trypsin and flavastacin.</i>	51
Figure 4-3: <i>Specificity of flavastacin for the C-terminus of N-glycosylated asparagine.</i>	54
Figure 4-4: <i>Comparison of the spray stability during elution when using nano-PGC-LC-ESI-MS(/MS) in negative ion mode.</i>	57
Figure 4-5: <i>Comparison of the BPC of eluted N- and O-glycans released from bovine fetuin with (A) and without (B) PCMF supplementation.</i>	58
Figure 4-6: <i>Comparison of the EIC of selected masses corresponding to N- and O-glycans released from bovine fetuin before (A) and after (B) PCMF supplementation.</i>	60
Figure 4-7: <i>Separation of N-glycan isomers using nano-PGC-LC-MS(/MS).</i>	68
Figure 4-8: <i>Negative ion mode HCD fragmentation of different N-glycan isomers.</i>	70
Figure 4-9: <i>Comparison of the most abundant N-glycan structures.</i>	72
Figure 4-10: <i>Site-specific glycopeptide analysis using nano-RP-LC-MS(/MS).</i>	74
Figure 4-11: <i>Graphical illustration of the trimeric IAV antigen HA.</i>	78
Figure 4-12: <i>Graphical illustration of the tetrameric IAV antigen NA.</i>	80

LIST OF ABBREVIATIONS

ABC	ammonium bicarbonate
ACN	acetonitrile
AIDS	acquired immune deficiency syndrome
AA	amino acid
BPC	base peak chromatogram
BSA	bovine serum albumin
hLTF	lactotransferrin from human milk
C18	octadecyl carbon chain
CDG	congenital disorders of glycosylation
CE	capillary electrophoresis
CID	collision-induced dissociation
CQA	critical quality attribute
CVV	candidate vaccine virus
CHMP	Committee for Medicinal Products for Human Use
dHex	deoxyhexose
DTT	DL-dithiothreitol
ECACC	European collection of authenticated cell cultures
EIC	extracted ion chromatogram
EMA	European Medicines Administration
ER	endoplasmic reticulum
ESI	electrospray ionization
ETD	electron-transfer dissociation
FASP	filter-aided sample preparation
FCS	fetal calf serum
FDA	Food and Drug Administration
FLD	fluorescence detector
Fuc	fucose
α -Gal	galactose- α -(1,3)-galactose
GalNAc	<i>N</i> -acetylgalactosamine
Gal	galactose
GlcNAc	<i>N</i> -acetylglucosamine
GMEM	Glasgow minimum essential medium
HA	hemagglutinin
HCD	higher-energy collisional dissociation
HEK293	human embryonic kidney 293 cells
Hex	hexose
HexNAc	<i>N</i> -acetylhexosamine
HILIC	hydrophilic interaction liquid chromatography
HPLC	high-performance liquid chromatography

LIST OF ABBREVIATIONS

IAA	iodoacetamide
IAV	influenza A virus
ID	identifier
InD	inner diameter
IM	ion-mobility
IT	ion-trap
KOH	potassium hydroxide
LacNAc	<i>N</i> -acetylglucosamine
LC	liquid chromatography
LC-MS	liquid chromatography coupled to mass spectrometry
M1	matrix protein 1
M2	matrix protein 2
mAb	monoclonal antibody
MALDI-TOF-MS	matrix assisted laser desorption ionization — time of flight mass spectrometry
Man	mannose
MDBK	Madin-Darby bovine kidney cells
MDCK	Madin-Darby canine kidney cells
MDCK.ADH	adherent MDCK cells
MDCK.SUS2	suspension MDCK cells
Mp	mass of the peptide moiety
MWCO	molecular weight cut-off
MS	mass spectrometry
MS/MS	tandem mass spectrometry
NA	neuraminidase
NaBH ₄	sodium borohydride
NeuAc	<i>N</i> -acetylneuraminic acid
NCE	normalized collision energy
NP	non-structural protein 1
OD	outer diameter
OT	orbitrap
PCMF	post-column make-up flow
PDB	protein data bank
PGC	porous graphitized carbon
PNGase F	peptide <i>N</i> -glycosidase F
PA	polymerase acidic protein
PB1	polymerase basic protein 1
PB2	polymerase basic protein 2
RBS	receptor binding site
rEPO	recombinant erythropoietin
rHA	recombinant HA
RNA	ribonucleic acid

LIST OF ABBREVIATIONS

RP	reversed-phase
RT	room temperature
SP-D	surfactant protein-D
SPE	solid phase extraction
TFA	trifluoroacetic acid
tR	retention time
UPLC	ultra performance liquid chromatography
Vero	African green monkey kidney cells
WHO	World Health Organization

1. MOTIVATION

Infectious diseases are one of the major causes of death, with the most cases leading to respiratory diseases, diarrhea, Acquired Immune Deficiency Syndrome (AIDS), tuberculosis and malaria [1]. Viral infections can be treated in some cases with antiviral drugs (class of drugs that can inhibit a specific type or a wide range of viruses) or prevented via vaccination – the presentation of viral antigens to the human immune system to build up resistance against the virus.

In this thesis the influenza A virus (IAV) is used, a virus known for leading to several pandemics in the past with the most severe in 1918 known as the Spanish Flu. Influenza viruses have the ability to change the structure of the major antigens hemagglutinin (HA) and neuraminidase (NA) via antigenic drift and shift. Therefore, adaptive immunity via vaccination can be ineffective against evolving influenza virus strains. To overcome this, annually adjustment of the vaccine (comprising the major circulating influenza A and influenza B strains) and vaccination of the recommended population must be achieved.

Most influenza vaccines are produced in embryonated chicken eggs, but also animal cell-based and recombinant vaccines propagated in insect cell lines are licensed and applied to humans. Defined animal cell-based systems are more efficient compared to traditional egg-based systems, because of a more flexible and faster production of vaccines. One major limiting factor in animal cell-based vaccine production is the surface of the cultivation system when using adherent cells. To enhance the capacity of the cultivation system by increasing the amount of cells producing viruses, suspension cell-lines are of increasing interest for the biopharmaceutical industry. In addition they are considered for the transition from batch/fed-batch cultivation to more efficient semi-continuous production systems. Furthermore, upscaling processes of suspension cell-lines is more capable compared to adherent cell lines and they can be operated in high cell density cultivation processes.

Influenza A virus belongs to the class of enveloped viruses, which are surrounded by a lipid membrane. This membrane incorporates the two major antigens and glycoproteins HA and NA. The glycosylation (oligosaccharide linked to the protein backbone) of HA and NA was demonstrated to influence the virulence and immunogenicity of influenza viruses [2-4]. However, until now, analytical studies only provided glycan fingerprints showing differences between e.g. different cell lines. To unveil the importance of the glycosylation of viral antigens, new methods and workflows need to be established suitable to identify the structure of the attached glycans (including detailed information about linkages, branching and anomericity (type of geometric variation found at certain atoms) of a glycan). In addition, the position of the glycan linkage to the protein backbone needs to be examined. Currently, it is not possible to answer both questions with a single method. Site-specific glycoproteomics elucidates the locus of a glycan composition to the glycoprotein, but lacks most of the fine structural information. Glycomic analysis of released glycans enables the possibility to gain structural information of the target glycome, but the link to a specific locus of the glycoprotein is lost.

The target of this work was to combine both pieces of information by establishing a comprehensive workflow combining two state-of-the-art LC-MS based methods to enable fine structure and site-specific glycoanalysis of viral antigens. First, enzymes had to be identified to generate glycopeptides with appropriate amino acid lengths and single *N*-glycosylation sites for site-specific analysis. Furthermore, the LC-MS setup for fine-structural glycomic analysis using nano-instrumentation had to be optimized, to increase the robustness of the method. Finally, the method established was proved to detect not only smaller glycans at lower retention times (tR) e.g. *O*-glycans, but enabled also a better separation of glycans at higher tR, which resulted in a higher identification rate of different glycan structures (isomers) from glycans with the same molecular mass. As application, the HA and NA glycosylation of the IAV strain A/Puerto Rico/ 8/34 H1N1 propagated in two Madin-Darby Canine Kidney (MDCK) cells (adherent

MDCK.ADH and suspension MDCK.SUS2) was analyzed to characterize variations in glycosylation of HA and NA due to the different producing cell lines.

2. BACKGROUND

This chapter is about the main forms of protein glycosylation and their biological meaning. Furthermore, an excursus of state-of-the-art analytical methods in the field of glycomics and glycoproteomics is given, explaining the advantages and drawbacks of the different methods and demonstrating in consequence the approach of the here developed new workflow. Afterwards, the biological background of influenza A virus, including the structure, diversity and replication cycle is described. Subsequently, different kinds of influenza vaccines and production systems are introduced. Finally, the importance of influenza virus antigen glycosylation is addressed.







2.1. Glycoproteins

This section comprises the meaning of glycobiology, including glycan structures as well as protein glycosylation, followed by a short overview of state-of-the-art analytical methods to elaborate the structure and the location of a glycan on a protein.

2.1.1. Glycobiology: from sugar to biological function

Glycobiology comprises the structure of mono- to oligosaccharides, up to their function and importance in biological systems. Monosaccharides are carbon-based ring-shaped organic compounds, resulting from oxidation of polyhydric alcohols. They are forming the basis for oligosaccharides, also called glycans (taken from the Greek word “glykòs” for “sweet”) [5]. In Table 2-1 all monosaccharides utilized in this work are depicted. These monosaccharides are the major building blocks of the human glycome [5]. For the symbolic representation of monosaccharides and glycans in this work *GlycoWorkbench* was used following the Symbol Nomenclature for Graphical Representations of Glycans (SNFG) [6].

Table 2-1: **List of monosaccharides used in this work.** Illustration according to Symbol Nomenclature for Graphical Representations of Glycans [6].

Monosaccharide	Abbreviation	Symbol
Galactose [*]	Gal	
Fucose [#]	Fuc	
Mannose [*]	Man	
<i>N</i> -acetylgalactosamine [§]	GalNAc	
<i>N</i> -acetylglucosamine [§]	GlcNAc	
<i>N</i> -acetylneuraminic acid	NeuAc	

*- hexose (Hex), #- deoxyhexose (dHex), §- *N*-acetylhexosamine (HexNAc),

The reaction of two monosaccharides takes place between the reduced end and the hydroxyl group of both monosaccharides under loss of a water molecule (glycosidic bond) [7] (Figure 2-1). This energetically unfavorable reaction is performed with the help of a respective glycosyltransferase and a nucleotide sugar together with the dephosphorylation of two highly energetic adenotriphosphates. Dependent of the individual position of the anomeric C1-atoms, at equal stereochemistry we are talking about an α - and at unequal stereochemistry about a β -configuration [7]. Therefore, the nomenclature for the structural description of glycans comprises the involved monosaccharides, their conformation, as well as the position of the individual C-atom of the glycosidic bond (e.g. β 1-4 in Figure 2-1).

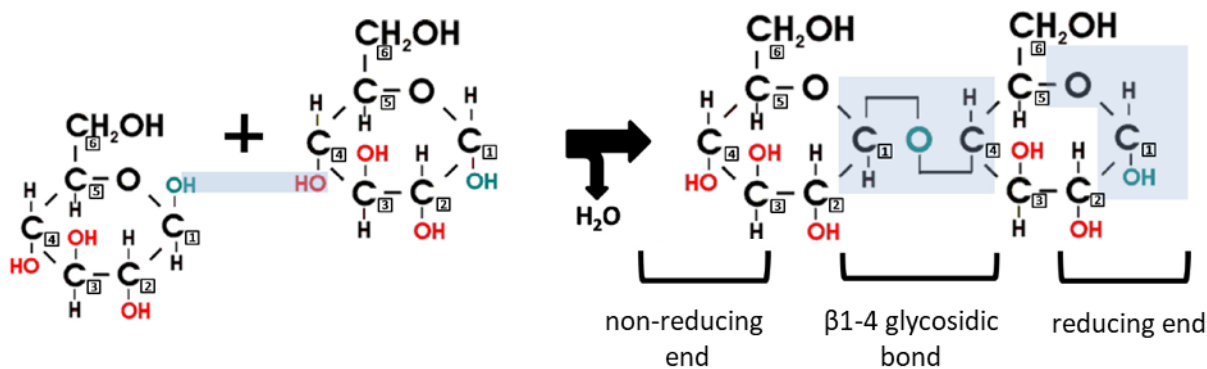


Figure 2-1: *Illustration of a β 1-4 glycosidic bond between two galactoses.* Adapted from Hennig et al. 2009 [8].

In biological systems, glycans exist as free sugars and as conjugates with different macromolecules (glycoconjugates). Glycoconjugates are, for example, glycoproteins and glycolipids. The modification of lipids, or (co-)post-translational modification of proteins with sugars is called glycosylation [5]. All glycans and glycoconjugates of an organism can be called glycome, which is highly diverse between different organisms depending on the host-specific repertoire of endoglycosidases and endoglycosyltransferases (group of enzymatic proteins responsible for the specific cleavage and bond of different glycans and linkages).

Analysis of glycans is very important for production of biopharmaceuticals (critical quality attribute (CQA) for recombinant glycoproteins) as well as for medical applications (drugs) and diagnostics (biomarker). The importance of glycosylation for different biological functions within the human body, as well as associations to diseases, is illustrated in Figure 2-2.

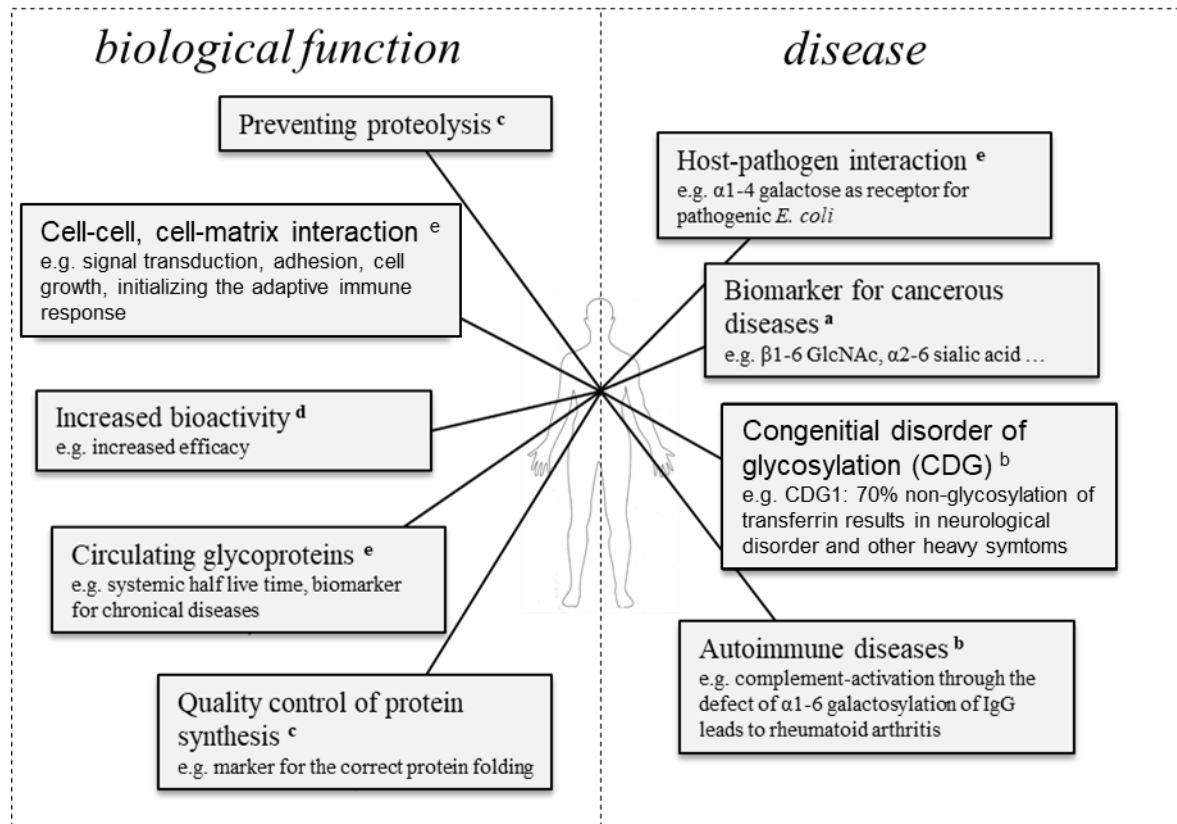


Figure 2-2: **Importance of glycobiochemistry for the human organism (a)[9], (b)[10], (c)[11], (d)[12], (e)[13].**

The efficacy and safety of recombinant glycoproteins (e.g. monoclonal antibodies (mAb) and hormones) depends tremendously on the glycosylation. One example is the recombinant erythropoietin (rEPO). The degree of sialylation (terminal binding of NeuAc) impacts the half-life time of rEPO [14] (higher degree of sialylation results in an increased half-life time). Therefore, additional glycosylation sites decorated with highly sialylated glycans were added to rEPO. To be able to intervene in this glycosylation process to improve quality and function of biopharmaceutical products is the basis for the upcoming field of glycoengineering, which covers the manipulation of glycosylation of a biopharmaceutical.

In medicine, for example, the glycosylation of surface proteins of cancer cells is analyzed for diagnostic and therapeutic purpose. Different glycan structures as surface antigens can be used as highly specific

biomarkers [9]. Furthermore, these structures are suitable targets for antibody-based therapeutic strategies [15].

Glycobiology helps to understand different biochemical processes in the human organism. Especially for immunological reactions the human glycome (glycans and glycoconjugates of an organism) has a great importance [13]. For example, several glycoproteins on the surface of endothelial cells assist the targeted immigration of circulating cells to the location of an inflammation through the interaction of glycans to specific receptors. Another example is the binding of pathogens via surface glycoproteins and the resulting contribution to the activation of the innate immune response via cell-cell interactions mediated via glycan structures [13]. Another example of the importance of glycoconjugates is the ABO blood group system. This system relies on specific blood group antigens (terminal structures of glycans, Figure 2-3) that can be found predominantly on red blood cells, but also, for example, on epithelial cells (like MDCK cells) [16]. Blood group antigens are described to have an impact in infectious diseases, multifactorial diseases as well as cancer via glycan structure mediated interactions [17].

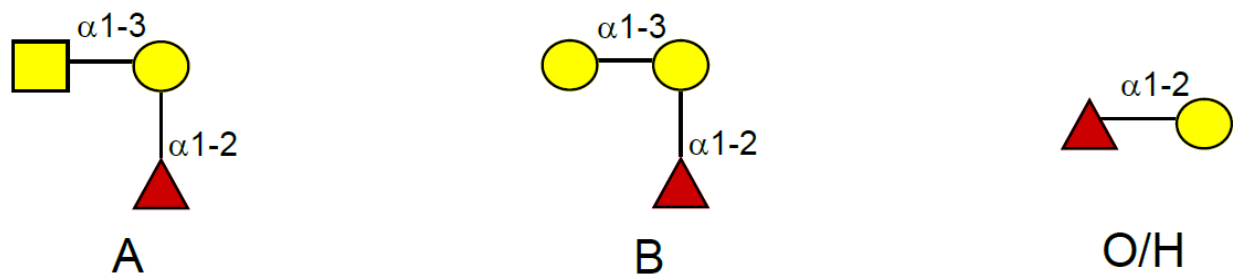


Figure 2-3: **Blood group antigens.** Illustration of the three blood group antigens A (left), B (middle) and O/H (right) that are mainly known as specific epitopes on red blood cells.

2.1.2. Glycan structures

Glycan structures bear a higher variability compared to peptide sequences. They build up branched complex structures from different core structures (Figure 2-4) with a variety of terminal modifications. Besides the monosaccharide composition of a glycan, the structural linkages are essential for the binding

abilities and biological functions [5]. Terminal monosaccharides of a glycan structure are called non-reducing end. The other end of a glycan (e.g. binding site to a protein) is called reducing end [5]. In Figure 2-4 some basic structures of *N*- and *O*-glycans are illustrated. *N*-glycans consist of a core structure of two GlcNAc (chitobiose core) and three Man (trimanosyl core) [7]. Based on the modification of the trimanosyl core, we differentiate among three different types of *N*-glycans: Oligomannose type, complex type and hybrid type (Figure 2-4). Oligomannose type structures have branched linkages of up to nine Man on both alpha linked Man of the trimanosyl core. Complex type *N*-glycans have a much higher variability of possible monosaccharide compositions compared to oligomannose type.

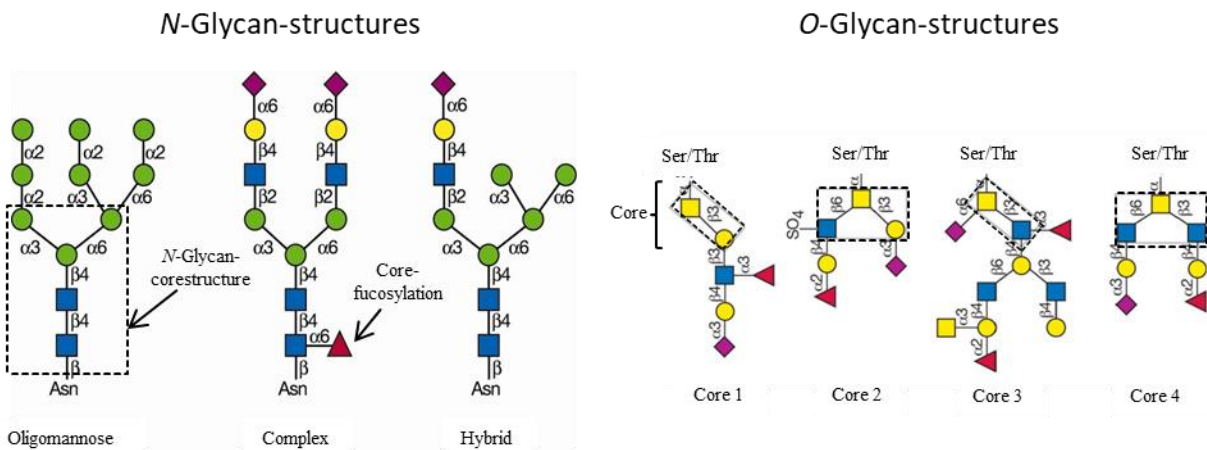


Figure 2-4: **Exemplary structures of *N*- and *O*-glycans.** For *N*-glycans, the three different types are shown. For *O*-glycans, the four main structures of the eight mucine-type core-structures are illustrated. Adapted from Varki et al. (2017) [5]. According to the SNFG nomenclature [6].

When a GlcNAc (complex, hybrid) or a Man (oligomannose, hybrid) binds to one of the α -linked Man this results in a so called antenna. These antennas can be further terminal modified with different monosaccharides (depends on the host specific repertoire of endoglycosidases and endoglycosyltransferases). Besides di-antennary complex type structures (occupation of both terminal mannoses of the trimanosyl core) the core structure can be decorated with up to four antenna. Furthermore, the first GlcNAc of the chitobiose core can be fucosylated (core fucosylation, α 1-6 linkage

in mammals). Antennaries can be fucosylated as well. Hybrid type *N*-glycan structures have the characteristics of both, oligomannose and complex type *N*-glycans [5].

Mucin type *O*-glycans can be grouped in eight different core-structures [5]. The only similarity is the GalNAc bound to the serine or threonine AA. The core structure is defined based on the different bindings of Gal, GlcNAc and GalNAc [5]. These core structures can be further terminal modified.

2.1.3. Protein glycosylation: *N*- and *O*-glycosylation

Proteins are glycosylated via a covalent bond of a glycan structure to a protein sequence. Besides numerous other types of protein glycosylation, the focus in the following one will only be on the two most prominent forms in eukaryotes, the so called *N*- and *O*-glycosylation [7].

N-glycosylation of proteins starts via pyrophosphate binding of dolichol to the inner surface of the endoplasmic reticulum. The initial structure $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$ is build up via different chitobiosyldiphosphodolichol beta-mannosyltransferases. This structure is post-translationally linked to the amino group of the asparagine with the consensus sequence (NXS/T(C/V); $X \neq P$) of the protein via the oligosaccharyltransferase [7]. At this stage the correct folding of the protein is controlled [11]. Further modifications of the *N*-glycan happen in the Golgi apparatus. Thereby, the initial structure $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$ is trimmed down to the core structure $\text{Man}_3\text{GlcNAc}_2$ (Figure 2-4) via exoglycosidases and further build up co-post-translational via glycosyltransferases [7].

Mucin-type *O*-glycans are build up stepwise to a protein sequence via sequential enzymatic transfer of individual monosaccharides in the Golgi apparatus. Therefore, GalNAc binds to the hydroxyl group of predominantly serine or threonine (preferential in the vicinity of proline or aliphatic amino acids (AAs)). Exemplarily for the core 1 structure (Figure 2-4) the *O*-glycan is further build with a β 1-3 glycosidic bond of Gal to GalNAc in the Golgi apparatus[5]. Overall, there are eight different core structures.

Glycoproteins can have one single (like most mAbs) but also multiple glycosylation sites (like hormones or viral proteins). The difference of glycan structures at a single glycosylation site at the same glycoprotein is called microheterogeneity. Macroheterogeneity describes the occupancy (yes (grade) / no) of the different glycosylation sites of a glycoprotein [18].

The functional importance of site-specific glycan structures is rarely described. One example are congenital disorders of glycosylation (CDGs) that can be differentiated in two types. CDGs of type 1 are genetic malfunctions resulting in the partly or complete non-glycosylation of a specific glycosylation site of a specific glycoprotein. CDGs of type 2 result in an incomplete terminal modification of glycans and therefore in truncated glycan structures [10].

2.2. Analytics in glycobotechnology

The analysis of glycoproteins can be divided into the fields of glycan- (glycomic) and glycoprotein- (glycoproteomic) analysis. Glycomics is the conjugate independent analysis of the glycome (entirety of glycans) in a global systemic approach or of an individual protein [18]. State-of-the-art analytical techniques allow the fine-structural glycomic analysis in high throughput with high resolution and software-assisted data evaluation [19, 20]. But glycomics of glycoproteins lacks the information of micro- and macroheterogeneity [18], which can only be obtained by MS-based glycoproteomic approaches.

2.2.1. Glycomics

This paragraph is partly adapted from Pralow et al. (2020) [20].

The vital role glycans have in biotechnology and biology makes it ever more important to characterize and understand glycan structures in the context of such complex bio(techno)logical processes [5, 21, 22]. Glycosylation is a non-genomic template-driven co- and posttranslational modification of glyco-

conjugates. To analyse non-linear oligosaccharide structures confers a level of complexity that emphasizes the need to use high-resolution analytical techniques rather than referring to transcriptomes or genomes to effectively profile such molecules [23, 24]. By definition, a glycomic investigation is aimed at characterizing the entire glycan population in a heterogeneous biological sample. This is not a trivial exercise as a glycomic profile can contain structures differing in linkages, branching and anomericity [5, 25]. Furthermore, glycans can exist in either neutral or charged forms, for instance due to carrying a combination of sialic acids as well as further modifications such as sulfation or phosphorylation [5, 26-29].

Usually, glycan release is performed via enzymatic digestion using peptide *N*-glycosidase F (PNGase F). PNGase F is an amidase able to cleave the binding between GlcNAc and the asparagine residue of an *N*-glycosylation. During this cleavage asparagine is deamidated to aspartic acid and its molecular mass is shifted by 1 Da [30].

Central to high-performance glycoanalysis are the advances made in ultra/high-performance liquid chromatography (U/HPLC) [31-33], capillary electrophoresis (CE) [34-38] and mass spectrometry (MS) [39-43], which now enable the extensive structural analysis of released glycans. Analytical techniques primarily employed include hydrophilic interaction liquid chromatography with fluorescence detection (HILIC-FLD) [31, 32], liquid chromatography coupled with electrospray ionization (tandem) mass spectrometry (LC-ESI-MS(/MS)) [40, 41, 43, 44], matrix assisted laser desorption/ionization with time-of-flight MS(/MS) (MALDI-TOF(/TOF)-MS(/MS)) [42, 45, 46] and more frequently also multiplexed capillary gel electrophoresis with laser-induced fluorescence detection (xCGE-LIF) [2, 35-37, 47-51]. To increase sensitivity or facilitate quantitation, glycans can undergo chemical modification such as permethylation [46, 52, 53], esterification [42] or numerous fluorescent labelling like 2-aminobenzamide [54-56], 2-anthranilic acid [57], procainamide [58], RapiFluor-MS [59] and 8-aminopyrene-1,3,6-trisulfonic acid

trisodium salt [36-38]. Each of these modification techniques affects certain properties of glycans such as their hydrophilicity (HILIC), electrokinetic mobility (CE), composition (MS), as well as their fragmentation behavior (MS/MS). Complete characterization of co-eluting glycan structures or glycan isomerism (same molecular mass of different glycan structures) requires additional orthogonal methods such as sequencing via exoglycosidase digestion [48-50, 60, 61].

Beside these more established analytical methods for glycan analysis, emerging technologies like porous graphitized carbon liquid chromatography coupled to MS(/MS) (PGC-LC-MS) and ion-mobility mass spectrometry (IM-MS) are utilized more frequently. Glycan analysis using PGC-LC coupled with negative ion mode MS detection and MS/MS fragmentation is gaining traction as a method for comprehensive MS-based glycoanalysis [40, 41, 62-70]. Very little modification of the released native glycan is required (compared to the previously mentioned more established analytical methods) other than reduction of the anomeric ring of the glycan's reducing end. The analysis produces information-rich data from chromatographic separation of even isobaric *N*- and *O*-glycans on the PGC column (even further compared to HILIC columns [71]). Additional MS/MS fragmentation and detection in negative ion mode can yield diagnostic fragment ions of rich information owing to cross-ring fragments, which can be used to confirm a glycan's topology, branching and monosaccharide linkages [41, 72-74]. Both neutral and negatively charged glycans can be detected using negative ion polarity MS [73-75]. The mechanism behind glycan separation using PGC columns remains unclear. Therefore, the prediction of glycan structure-specific tR using *in silico* models is not possible [76]. Compared to the more established glycan analysis methods, automation and therefore high throughput sample analysis using PGC-LC-MS is not possible, so far, due to missing normalization standards and validated analysis software. However, decent glycan retention libraries and elution order rules exist to assist the manual data analysis [77-80]. Because PGC-LC has the ability to resolve isomeric glycan structures in combination with gaining fine-

structural information using negative mode MS/MS, PGC-LC-MS(/MS) was the method of choice to obtain an optimized workflow for the comprehensive analysis of glycans from IAV glycoproteins.

2.2.2. Glycoproteomics

In the past decades, the field of site-specific glycoprotein analysis methods has grown tremendously and emphasized the importance of protein glycosylation analyses with respect to micro- and macroheterogeneity. So far, the method of choice for site-specific glycoproteomics is the analysis of specifically or unspecifically digested glycoproteins via LC-MS. In addition, a glycopeptide enrichment can be performed prior to analysis, i.e. by HILIC-solid phase extraction (SPE) [18] that will be portrayed later on.

2.2.2.1. Digestion of glycoproteins

This paragraph is partly adapted from Pralow et al. (2017) [81].

Crucial for identifying glycopeptides is a suitable protein digest. The most used enzyme is the serine-protease trypsin, which cleaves specific N-terminal at the alkaline AAs lysine and arginine [82]. Because of the high specificity of trypsin, the variability of originated peptides is low, facilitating automated (glyco-) protein searches and quantification of glycopeptides [83]. However, tryptic glycopeptides can be too large (> 20 AA) to be analyzed by MS or carry multiple glycosylation sites.

In addition to tryptic digestion of glycoproteins, sequential treatment with other proteases is used to overcome low charge density and sequence constraints of glycopeptides with a large peptide moiety (Table 2-2).

Table 2-2: *Overview of proteases used for mass spectrometry-based protein analysis.*

Endopeptidase	Type	Specificity
AspN (e.g. Flavastacin)	Metalloendoprotease	Aspartic Acid (N-terminal)
Chymotrypsin	Serine protease	Tyrosine, Tryptophane, Phenylalanine
Glu C	Serine protease	Aspartic Acid, Glutamic Acid
Lys C	Serine protease	Lysine
Proteinase K	Serine protease	hydrophilic AA
Trypsin	Serine protease	Arginine, Lysine

The endoproteinase AspN is a zinc metalloendoproteinase produced in *Pseudomonas fragi* (Boehringer Ingelheim, Uniprot: Q9R4J4), which selectively cleaves peptide bonds N-terminal to aspartic acid [84]. AspN is also known for N-terminal cleavage at cysteine and glutamic acid [85, 86]. Its primary use in glycoproteomic experiments so far, has involved the cleavage of deamidated asparagine after *N*-glycan release by PNGase F to assess *N*-glycan presence and location [87]. Flavastacin (New England Biolabs, Uniprot: Q47899), which is produced in *Flavobacterium meningosepticum*, has been described to behave similar to the AspN from *Pseudomonas fragi* [88, 89]. Therefore, it is also called AspN, despite its quite different AA sequence and thus, protein identity. However, both proteins belong to the family of metalloendoproteases.

Unspecific proteases like the serine-protease proteinase K are able to resolve glycosylation sites that cannot be uncovered by trypsin because of its non-selective cleavage specificity for aromatic and aliphatic AAs [90]. However, because of unspecific cleavage behavior, the identification of glycopeptides is more difficult and a single glycosylation site can have multiple redundant peptide moieties [91]. Analysis time using search engines is increased and quantification is more difficult.

A protein digest can be performed in-gel or in-solution. In-gel digest has a lower risk of sample contamination but the workflow is time and sample consuming [92, 93]. In-solution digest can be automatized and has lower hands-on time. But the digest can be affected by contaminants (e.g. salts, detergents) [92]. An alternative to both methods is the filter-aided sample preparation (FAST) approach, which was used in this thesis. With a molecular weight cut-off (MWCO) of 10 kDa, the filter allows to pass contaminants (like salt) and molecules with a low molecular weight (like peptides). Intact proteins retain on the filter unit. Therefore, the desalting and removal of contaminants can be achieved. Furthermore, after proteolytic digest, peptides can pass the filter and intact proteins or enzymes are retained on the filter unit [92].

2.2.2.2. LC-MS measurement of glycopeptides

The technique of choice for site-specific glycopeptide analysis is (LC)-MS(/MS). Analysis of biomolecules using MS can be broken down to two ionization techniques, MALDI and ESI. Both can comprise different pre-separation techniques (e.g. LC, CE), ion fragmentation units, e.g. collision induced dissociation (CID), higher energy collisional induced dissociation (HCD), electron-transfer dissociation (ETD) and mass analyzers/detectors, e.g. ion-trap (IT), orbitrap (OT), TOF. The classic MALDI-TOF-MS has a high mass accuracy and allows the analysis within a wide mass range. Because of its robustness, workflows for the analysis of glycans and glycopeptides using MALDI-TOF-MS are often used in the biopharmaceutical industry [20, 94].

In this work, LC-ESI-MS was used exclusively that will be more explained in detail. Compared to MALDI-TOF-MS, LC-ESI-MS bears a higher sensitivity and has a more soft ionization mechanism, avoiding possible post-ionization decay of e.g. sialic acids [91, 95]. Because of the on-line separation of the analyte using LC prior to MS, more complex samples can be analyzed compared to solely MS instrumentation [96]. One major drawback in the analysis of glycopeptides using MS is the missing

information of the glycan structure. Only partial structural information of a glycopeptides glycan moiety can be obtained using the ratio of specific fragment ions [97] or separating fragment ions via IM [98].

Separation of (glyco-) peptides prior to MS analysis is performed predominantly on a reversed-phase (RP) column based on the different hydrophobicity of the molecules [99]. The stationary phase of RP columns consists mostly of silica gel modified with octadecyl-carbon chain (C18) [18] or others, e.g. C8 or phenyl. The sample is loaded on the stationary phase via an aqueous mobile phase. Peptides bind via ion pairing to the stationary phase. Using a binary continuous elution gradient (mixture of aqueous and organic mobile phase), with an increasing organic proportion, the hydrophobicity is increased and peptides from lower to high hydrophobicity are released from the stationary phase [99]. RP is not able to separate glycopeptide isomers (same peptide sequence, same glycan composition, but different glycan structure, e.g. branching, linkage). Other phases like HILIC or PGC can separate isomeric structures, but are normally used for the separation of released glycans [20]. Prior to peptide separation, most setups (especially when using nano instrumentation) use trap columns (same phase as the separation column) to desalt and concentrate the sample [100].

Glycopeptide analysis in this work was performed using an Orbitrap Elite MS. Here, the eluting peptides are directed to a nanospray needle. The nanospray ionization is performed with a low flowrate and needs less sample compared to conventional ESI [101]. After ionization, the samples are directed to the mass analyzer (IT and OT). Precursor ions can be selected for fragmentation using a collision gas (helium). Usually, CID has limitations when fragmenting glycopeptides. When analyzing peptides without post-translational modifications, CID fragmentation results in b- (N-terminus) and y-ions (C-terminus) of a peptide sequence (Figure 2-5) [102]. Glycopeptides are predominantly fragmented at the glycosidic bonds into B- (non-reducing end) and Y-ions (reducing end) providing information about the monosaccharide-composition of a glycan and the peptide mass, but often lacks b- and y-peptide moiety

fragment ion series of the glycopeptides for sequence verification [103-105]. An alternative fragmentation method is ETD. Peptides are fragmented into c- (N-terminus) and z-fragment ions (C-terminus) (Figure 2-5), while the glycosylation of a glycopeptide stays intact. Therefore, this method provides information about the peptide sequence and the glycosylation site [18, 106]. In this work, HCD fragmentation was performed. Because of the adjustable normalized collision energy (NCE), HCD generates peptide-specific b- and y-ions and provides information about the glycan composition and the peptide sequence [97]. However, due to the neutral loss of the glycan moiety, B- and Y-ion series are underrepresented [107, 108]. Finally, specific fragment ions allow the identification of *N*- and *O*-glycopeptides and provide information rich data with high confidence in identification when using automated glycopeptide analysis software [83, 97, 109].

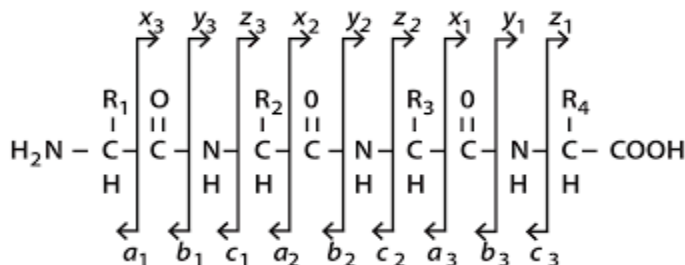


Figure 2-5: *Nomenclature of fragment ions derived from a peptide sequence [110].*

Fragment ion spectra of glycopeptides can be identified based on specific oxonium ions (B-fragment ions, Figure 2-6). Oxonium ions used for glycopeptide identification are listed in Table 2-3. To identify the glycan composition, specific mass differences among glycopeptide fragment ions (Y-fragment ions) are used [91, 111].

BACKGROUND

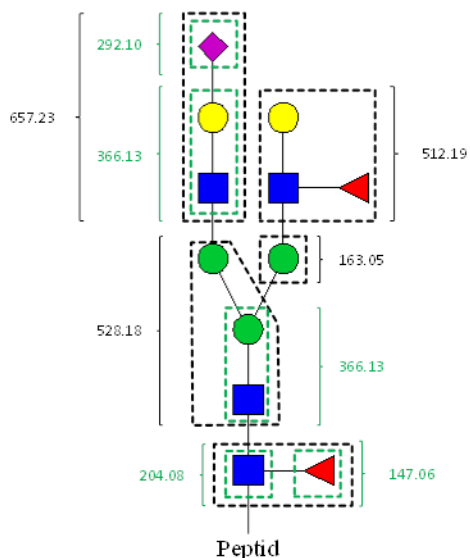


Figure 2-6: **Specific B-fragment ions of an N-glycopeptide with the corresponding monoisotopic masses.** Illustration according to the symbol nomenclature for graphical representations of glycans [6].

Table 2-3: **Monoisotopic masses of oxonium ions.** Illustration according to the symbol nomenclature for graphical representations of glycans [6].

Monosaccharide/ Polysaccharide	Monoisotopic mass [M+H] ⁺	Symbol
Gal	163.05	●
Man	163.05	●
GalNAc	204.08	■
GlcNAc	204.08	■
Fuc	147.06	▲
Neu5Ac	292.10	◆
Gal-GlcNAc	366.13	●-■
Gal-GlcNAc-Fuc	512.19	●-■-▲
Neu5Ac-Gal	454.15	◆-●
Neu5Ac-Gal-GlcNAc	657.23	◆-●-■
Neu5Ac-Gal-GlcNAc-Fuc	803.29	◆-●-■-▲
Man-Man-GlcNAc	528.18	●-■-●

Different software tools exist to assist the manual annotation of glycopeptide fragment ion spectra via (semi-)automated pipelines for glycopeptide LC-MS/MS data processing. The open source software glyXtoolMS [83] (detailed description in the Methods and Materials section) and one of the leading commercial software for MS-based glycoprotein analysis, called Byonic from ProteinMetrics, were used in this thesis [112, 113].

2.2.3. Enrichment strategies for glyco(-conjugate) analysis

Enrichment of glycopeptides and released glycans with depletion of non-glycopeptides, impurities and salts is crucial. Because of the glycan microheterogeneity, glycopeptides can have a low relative abundance. Additionally, because of the hydrophilic glycan moiety, glycopeptides ionize less efficient compared to non-glycosylated peptides. Therefore, it is necessary to perform a glycopeptide enrichment step prior to MS analysis [18, 114].

Most common strategies for glycopeptide and glycan enrichment are lectin and HILIC based [18, 114]. For glycans, PGC-SPE was used in this work.

Lectins are a group of proteins specifically binding to glycans. Therefore, lectins can be used for identification or fractioning via selectively binding specific glycan structures [115, 116]. However, because of the heterogenic binding abilities, lectins are not suited to enrich equally complex mixtures of very different glycopeptide or glycan fractions [18]. Hydrophilic molecules bind to HILIC material. Therefore, HILIC is well suited for the enrichment of glycans and glycoconjugates. There are many different commercial phases on the market providing highly selective enrichment. However, because of the low hydrophilicity, small glycans attached to peptide moieties with many hydrophobic AAs are often lost during the enrichment step [117].

2.3. Influenza A virus

Influenza viruses belong to the family of orthomyxoviridae and comprise four different genera, divided into influenza A, B, C [118] and recently described D virus [119]. Only influenza A and B viruses are known to cause seasonal epidemics, whereby only IAV can occasionally cause pandemic outbreaks and is responsible for the main hospitalizations each year. Further classification of IAVs into subtypes relies on the antigenic properties of their two surface glycoproteins HA (H1 to H16) and NA (N1 to N9). Four influenza pandemics occurred in the last century with the most severe in 1918 (H1N1, Spanish influenza) causing over 50 million deaths [120].

2.3.1. Structure

Influenza virions have a spherical to polymorphic shape with a diameter of 80-120 nm surrounded by a lipid membrane [121, 122]. A schematic illustration of IAV is given in Figure 2-7. Inserted into the membrane and exposed to the outside are two glycoproteins HA (major antigen) and NA. Also embedded in the lipid membrane is the matrix protein 2 (M2). Under the membrane a shell is built by the matrix protein 1 (M1), providing strength and rigidity to the lipid membrane. Influenza A and B virions have eight negative-sense single-strand ribonucleic acid (RNA) segments, encoding for the viral proteome. These RNA segments are packed as a viral ribonucleoprotein (vRNP) in complex with non-structural protein 1 (NP), polymerase acidic protein (PA), polymerase basic protein 1 (PB1) and polymerase basic protein 2 (PB2). As this thesis focuses on the analysis of the surface proteins HA and NA, detailed information about these proteins is given in the next paragraph.

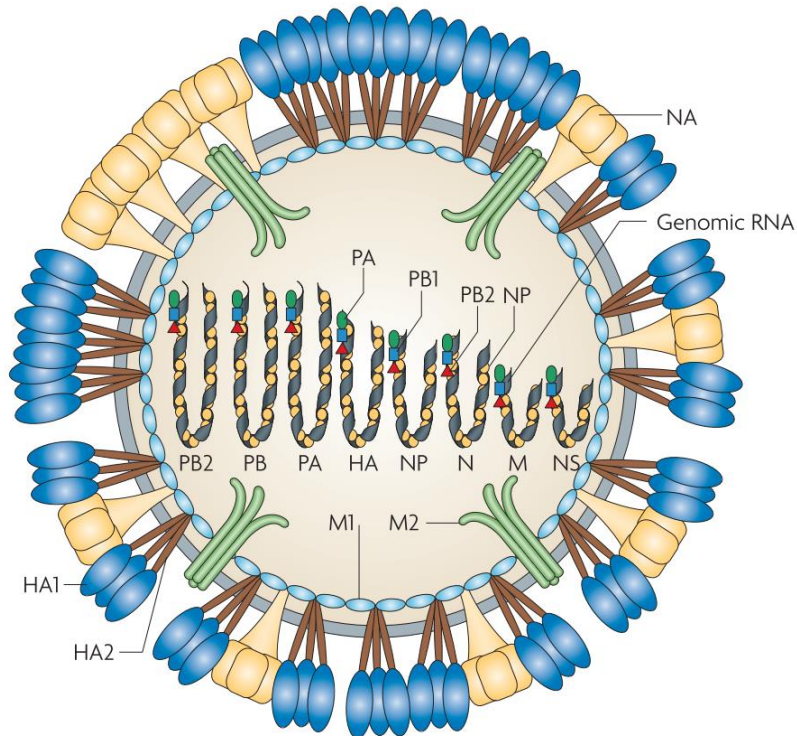


Figure 2-7: **Schematic illustration of IAV.** IAV is surrounded by a lipid membrane embedding the three potential surface glycoproteins hemagglutinin (HA, major antigen), neuraminidase (NA), as well as matrix protein 2 (M2). The matrix protein 1 (M1) protein is located inside the viral membrane; the viral genome consists of eight negative-strand RNA segments packed as a ribonucleoprotein in complex with non-structural protein 1 (NP), polymerase acidic protein (PA), polymerase basic protein 1 (PB1) and polymerase basic protein 2 (PB2). Illustration modified from Hedestam et al. 2008 [123].

2.3.2. Surface antigens – hemagglutinin and neuraminidase

Influenza viruses infect over a billion people around the world and account for approximately half a million deaths per year [124]. Due to antigenic changes, influenza viruses are a major threat for pandemics as well as epidemics [125]. HA is the major antigen in the seasonal influenza virus. It is a homotrimeric surface glycoprotein and plays a key role in terms of pathogen-host interactions. Each protomer of HA contains a receptor-binding site (RBS) in the head domain that binds to sialylated glycan receptors on the host cell surface. Many residues of the RBS are conserved among influenza A and B HAs. However, major structural changes can be observed among natural circulating strains. The specificity of the RBS in humans influenza virus is α 2,6-linked sialic acid and for avian influenza virus α 2,3-linked sialic acid [126].

The glycosylation of HA (Figure 2-8 A) is known to affect the virulence by modulating virus receptor binding [127], stimulating the host immune response [2, 128] and masking antigenic sites [129]. Glycosylation of the homotetrameric surface glycoprotein NA (Figure 2-8 B) influences virus entry [130], release [131] and neurovirulence [132].

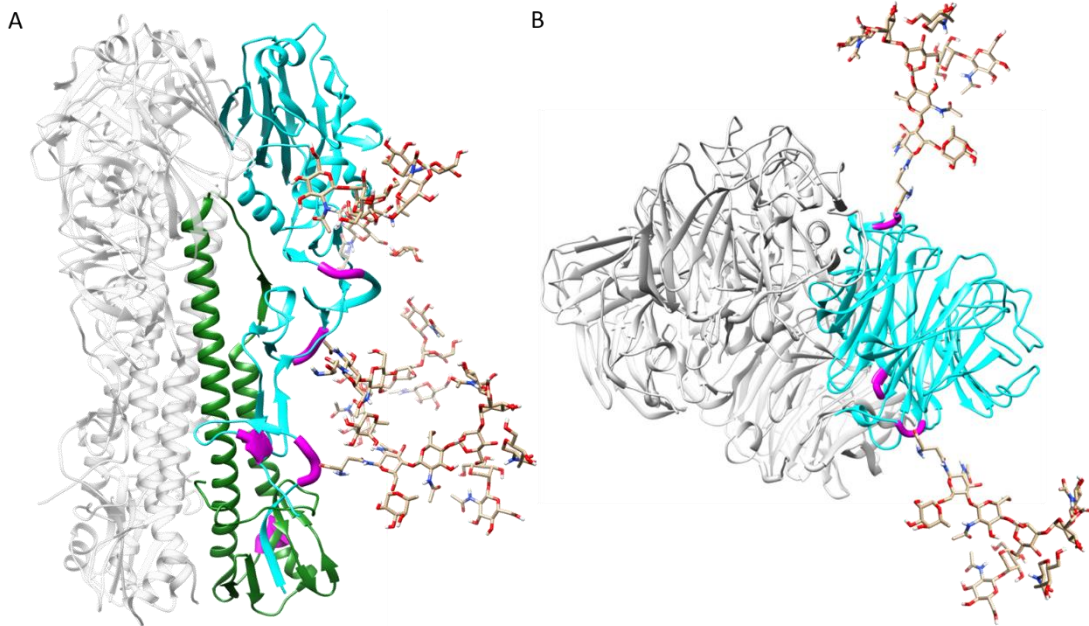


Figure 2-8: **Three dimensional structure of IAV glycoproteins HA (A) and NA (B).** For HA, one monomer is colored in cyan (HA1) and green (HA2). One monomer of NA is colored in cyan. Exemplary, the potential N-glycosylation sites of the colored monomers are highlighted in magenta. Adapted from Butler and Reichl et al. 2019 [133].

The number of potential N-glycosylation sites among different influenza viruses is rather variable for HA and more or less conserved for NA [134]. In this thesis, influenza A/PR/8/34 H1N1 virus was used, which displays six potential N-glycosylation sites at the HA protein (Figure 2-8 A, five at the HA1 (cyan) and one at the HA2 (green) domain) and five at the NA protein (Figure 2-8 B, please note, not all potential N-glycosylation sites of NA are illustrated because of model limitations). Both, HA and NA, have main

functions in the infection of host cells and release of virions, therefore, the replication cycle of IAV is described briefly next.

2.3.3. Replication cycle

IAV infects host cells by attachment of the HA RBS to surface glycoconjugates containing sialylated epitopes [135, 136]. Afterwards, IAV uses the sialidase function of NA to liberate nonproductive HA associations and to identify the proper sialylated receptor on the host cell surface [137]. This receptor remains unknown so far. However, the HA-mediated binding to this receptor triggers endocytosis of the virion into the host cell. Trafficking through the cell to the endosome activates the M2 channels (active at low pH) which causes a change of the three-dimensional structure of HA to expose the fusion peptide [138, 139]. Because of the acidification, vRNPs are released from M1 followed by transfer to the host cytoplasm, followed by HA-mediated fusion [140].

The fusion of the viral-endosomal membrane requires cleavage of HA into its domains HA1 and HA2 [141] and multiple steps, which are comprehensively reviewed by Harrison (2015) [142] and White and Whittaker (2016) [143]. Through cleavage of HA the fusion peptide is exposed in the N-terminus of HA2 and inserts into the endosomal membrane while the C-terminus is anchored in the viral membrane. Folding of the HA2 trimer brings both membranes close to each other, followed by a collapse into a six-helix bundle, initiating the fusion of both membranes. However, the complete process remains to be not fully understood till today.

Next steps are the IAV genome trafficking to the host cell nucleus, followed by replication and transcription of the vRNA as well as the assembling and trafficking of the vRNPs before budding and release [144-147]. According to the focus of this thesis, the assembly of the virus membrane proteins HA, NA and M2 is described in more detail. All membrane proteins are synthesized by ribosomes

associated with the endoplasmic reticulum (ER) membrane. The N-terminus of HA and M2 and the C-terminus of NA are translocated into the ER lumen [148]. At this point, the glycoproteins HA, NA are glycosylated at multiple *N*-glycosylation sites (M2 has also a potential *N*-glycosylation site, which was not described to be glycosylated, so far) [5, 134]. One of the main functions of glycosylation at this stage is to support the folding efficacy of HA and NA by recruiting lectin chaperons and an associated oxidoreductase [149, 150]. HA trimerizes from three individual monomers, while NA tetramerizes from two co-translationally formed dimers [150]. The fusion incompetent form of HA (HA0) is activated by cleavage into its subunits HA1 and HA2 in the trans-Golgi network and plasma membrane [151]. *N*-glycosylation of HA and NA can be (co-)post-transcriptional modified with monosaccharides in the Golgi [5].

2.4. Influenza vaccines

There are three different types of influenza vaccines approved in the United States by the U.S. Food and Drug Administration (FDA): egg-based (e.g. Vaxigrip Tetra®), cell-based (Flucelvax®) and recombinant (FluBlock®) flu vaccine. Recently, the European Medicines Administration (EMA) approved also a recombinant flu vaccine Supemtek® from Sanofi for 2022/2023 besides egg- and cell-based flu vaccines [152]. Candidate vaccine viruses (CVVs) are chosen by authorities (World Health Organization (WHO) and the Committee for Medicinal Products for Human Use (CHMP) to the EMA for each influenza season and either comprise two A and one B strain (trivalent) or two A and two B strains (quadrivalent) [153].

The majority of influenza vaccines is produced in embryonated chicken eggs, representing from the point of view of glycan research a wide mixture of glycan structures from different cell types and cell lineages. Egg-based flu vaccines can be inactivated (killed) or live attenuated (weakened, usually used for nasal spray application). Moreover, influenza vaccines are also produced under more defined conditions by using mammalian cell lines, such as MDCK (e.g. Flucelvax) or African green monkey kidney cells (Vero) (Baxter, not available anymore).

IAV sampled from infected humans can be adapted to propagation in embryonated chicken eggs. Therefore, HA from such egg-adapted IAV can have amino acid mutations close to the receptor-binding site [154] resulting in lower antibody titers. Studies have shown that cell-based flu vaccines provide a better protection compared to standard-dose egg-based vaccines [3]. Furthermore, cell-based flu vaccines production is faster and more flexible compared to egg-based production systems and does not rely on the supply of pathogen-free eggs for vaccine production, being an advantage in the event of a pandemic. Cells used for production are “banked” (ensuring the very same cell is used to avoid changes in cell growth and/or production behavior as well as different product properties) and ensure a stable

supply. Although adherent cells often show higher cell-specific yields (i.e. product molecules produced per cell), scale-up can be technically very challenging and suspension cells are considered the preferred process option. The third group, the recombinant flu vaccines, has only the major antigen HA. Insect cell lines are used to express recombinant HA (rHA) (Flublock) [155]. Via insertion of the specific HA DNA to a host cell using *Baculovirus*, rHA is produced. This process is the fastest compared to egg- and animal cell-based production systems because it does not rely on the adaption of a vaccine virus to eggs or the production cell line. However, because of the genetically modifications, recombinant vaccines undergo more strict critical quality attributes (CQAs), especially in terms of glycoanalysis, compared to egg- and cell-based flu vaccines [156, 157].

Emerging methods for viral vaccine production are based on genome level. SARS-CoV-2 vaccines were recently released in 2020/2021, comprising new technologies like mRNA vaccines (Biontech/Pfizer, Moderna) and viral vector vaccines (e.g. Astrazeneca). Instead of the recombinant antigen, the genetic information to produce the antigen in the human itself is injected. Similar technology is expected to be approved in the future for flu vaccines as well [158, 159].

2.5. The importance of the glycosylation of influenza A virus proteins on vaccine preparations

This paragraph contains parts adapted from Pralow et al. (2021) [160].

After introducing influenza viruses as well as glycomics and glycoproteomics, finally the importance and analysis of glycosylation for influenza virus and its vaccine will be portrayed in this paragraph. Glycosylation of the stalk region appears in all HA from different strains [161] and is described to bind chaperons in the ER and contribute to the HA trimer assembly [162]. Emerging pandemic IAV viruses mostly begin with a low degree of glycosylation of the head region, which increases as the strain circulates seasonally because the AA residues mutate rapidly to evade antibody recognition [163]. However, AAs that are shielded by *N*-glycans seem to mutate at a lower rate compared to AAs of more exposed regions [164]. The increase of *N*-glycosylation sites on the HA head region over time implies a correlation of HA glycosylation with antigenicity and immunogenicity. However, only few information is available regarding the specific IAV glycan fine-structures and the mechanism behind their interaction with antibody responses.

Glycosylation of the HA head domain can interfere with receptor binding and/or membrane fusion [165]. IAV evades the immune response by decreasing the affinity of neutralizing antibodies through allosteric effects caused by mutations in the AA sequence (predominantly in the head domain of HA) [166]. Mutations can increase the avidity of HA for the host sialic acid receptor, which might cause IAV to bind host cells more avidly than competing antibodies [167]. Furthermore, such mutations result in *N*-glycosylation sites. Occupied glycosylation sites in the head domain of HA can also block antibody binding and therefore impact antigenicity [168, 169]. Antigenicity changes based on HA site-specific glycosylation are rarely described because HA inhibition assays are more complicated to be interpreted

due to the lower receptor binding avidity caused by additional glycosylation sites [168]. A mutation that increases receptor avidity can also have a negative impact, as the viral fitness (ability to propagate) relies on the balance of receptor binding as well as membrane fusion and the release of virions from the cell. Increasing the receptor binding avidity could disturb this balance and cause a decrease of released virions. However, HA head domain glycans can support viral fitness by balancing the increased receptor binding avidity of an AA mutation [170]. A drawback of additional HA head domain glycosylation is the increased binding of lectins (e.g. surfactant protein-D (SP-D) and mannose binding lectins) of the innate immune system to glycans linked to HA, compromising the viral fitness through inactivation of IAV [171, 172]. Furthermore, increasing glycosylation can negatively impact the assembly of HA trimers [165]. However, such interactions depend on the specific glycan structures, including a variety of different epitopes exposed to different binding sites, with different binding efficacy and affinity.

The importance of glycosylation changes with a switch from the wild type virus to the vaccine protein glycosylation. In this case, glycosylation can be used to optimize vaccine's efficacy. Increasing interest of cross-reactive IAV vaccines with mono- or non-glycosylated proteins evolves in the scientific community, pointing into the direction of an IAV vaccine, capable of providing protection against different IAV strains [173-175]. As an alternative, several studies also demonstrated the ability to modulate the virulence of IAV or the immunogenicity of influenza vaccines by addition and removal of *N*-glycosylation sites, especially in the head region of HA [169], or by selecting certain host cell systems to produce specific *N*-glycan patterns [2, 176]. Addition of sites to the head region of A/Hong Kong/1/68 (H3N2) HA resulted in increased sensitivity to SP-D *in vitro* and attenuated virulence in mice [177] whereas removal of sites from the globular head region of HA1 from H3N2 (A/Beijing/353/89) and H1N1 (A/Brazil/11/78) led to resistance to neutralization by SP-D and increased virulence in mice [178, 179]. Interestingly, addition of *N*-glycans to the globular head region of HA of the mouse-adapted IAV PR/8/34 (H1N1) virus,

which lacks *N*-glycans on the head region of its HA [166], resulted in sensitivity to SP-D and attenuated virulence in mice [178]. Other studies show differences in antigenicity and immunogenicity of influenza vaccines based on the structure and type of the attached *N*-glycans. Complex-type *N*-glycosylated HA produced in mammalian cells (also MDCK cells [2]) is described to induce higher antibody titers than HA produced in insect cells, which features oligomannose-type *N*-glycans [2, 3]. The Galili-epitope bears the potential to trigger immunogenicity and therefore vaccine efficacy. This was demonstrated on alpha-galactosyltransferase negative mice. The treatment with IAV carrying *in-vitro* alpha-galactosylated complex *N*-glycans (propagated in embryonated chicken eggs) led to an increased immunogenicity compared to the control group treated without alpha-galactosylation [180].

The advent of powerful glycoanalytical techniques has recently enabled investigators to perform comprehensive glycosylation studies including both, the characterization of all glycan compositions present at one glycosylation site (microheterogeneity) and the determination of the glycosylation site occupancy (macroheterogeneity) of influenza virus glycoproteins [181-185]. Furthermore, in-depth structural analysis of the influenza virus *N*-glycome can be performed. xCGE-LIF technology was applied to establish an *N*-glycome “fingerprint” of HA and other viral proteins derived from influenza A/PR/8/34 virus propagated in adherent and suspension MDCK cells [37, 186]. Hussain *et al.* used IM-MS to identify *N*-glycans of two H3N2 and one H1N1 IAV strain propagated in adherent MDCK II cells [187]. Recently, She *et al.* 2020 used nano-PGC-LC-MS(/MS) to identify sulfated complex-type *N*-glycans derived from IAV propagated in different host cells (including MDCK cells) [188]. Harvey *et al.* (2018) showed MS-based glycan and glycopeptide analysis of different IAV strains and characterized rHA produced in different expression systems such as mammalian cells (e.g. MDCK, Madin-Darby bovine kidney (MDBK), human embryonic kidney 293 (HEK293)), insect cells (e.g. High Five, express SF+, sf21), tobacco plants and embryonated chicken eggs [184]. There are a few studies combining the compositional analysis of

the IAV glycome and glycoproteome [181, 182, 189, 190]. However, to this date, no fine structural analysis including information about topology, branching and monosaccharide linkages of released *N*-glycans derived from IAV in combination with the site-specific analysis of IAV glycoproteins has been performed. Without such a detailed structural *N*-glycan analysis, however, it is not possible to elucidate immunogenic epitopes. Based on information about micro- and macroheterogeneity of glycans, the topology of the glycosylated viral protein can be derived and used to elucidate details regarding the access of antibodies to the surface of viral antigens. As all these aspects are strongly determined by the host cell system [35], fine structural *N*-glycan analysis together with corresponding immunogenic studies could help to identify host cell candidates providing advantages regarding antigenicity and immunogenicity of viral antigens [2, 3, 176, 191]. In contrast, compositional analysis only provides information about glycan masses summarizing different amounts of monosaccharides (e.g. Hex, HexNAc, dHex). Therefore, the aim of this work is to create a comprehensive LC-MS-based workflow which allows to obtain site-specific glycopeptide information and information about the fine-structure of glycans. These information can be combined to have site-specific and fine-structural information of a single glycan attached to an IAV glycoprotein.

3. MATERIALS AND METHODS

In this chapter, all materials and methods are described. A detailed list of all used chemicals, consumables and instruments is given in the appendix section (A. Materials). An overview of the main workflow established is illustrated in Figure 3-1. First, the mammalian cell culture and influenza virus infection is described. Furthermore, virus harvest, inactivation and the relative protein quantification are elucidated. The glycomics part of the workflow comprises the *N*-glycan release, enrichment, measurement using nano-PGC-LC-MS(/MS) and data analysis. Regarding the latter, the established post-column make-up flow (PCMF) is described in more detail. In the glycoproteomics part, beside sample preparation, measurement and data analysis, the newly identified specificity of the proteolytic enzyme flavastacin is focused on.

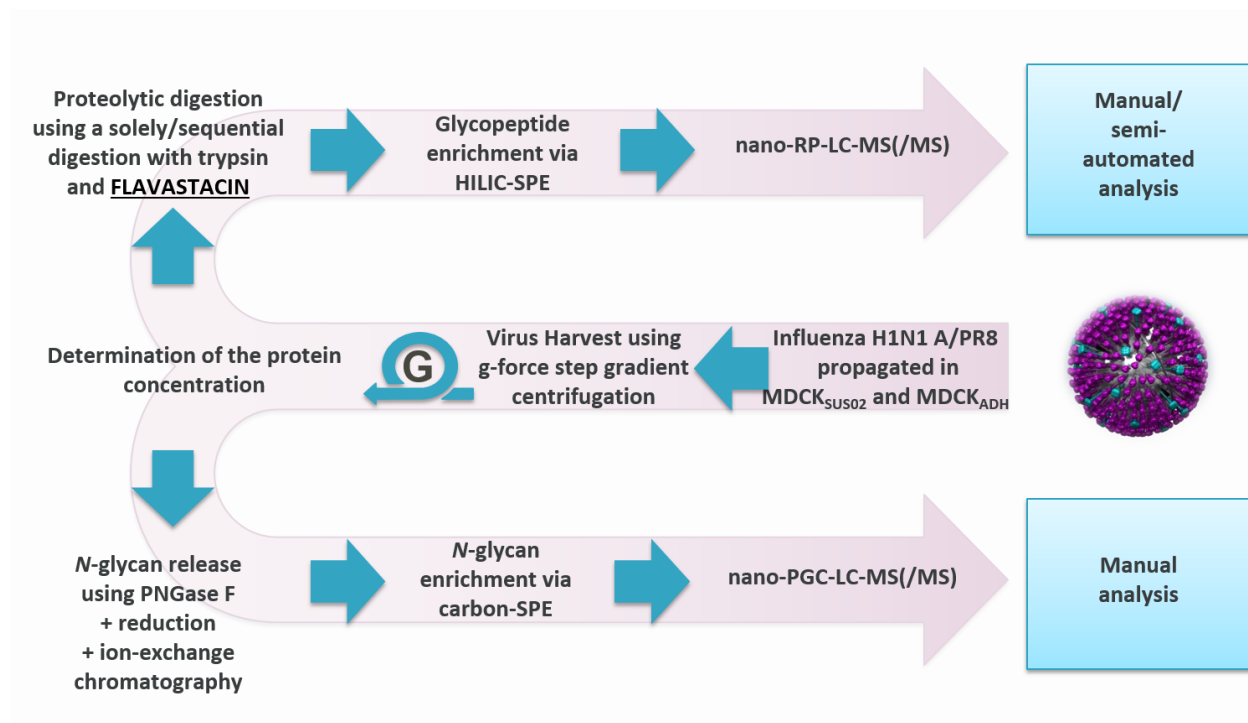


Figure 3-1: **Workflow.** Comprehensive analysis of the glycosylation of hemagglutinin (HA) and neuraminidase (NA) from IAV PR/8/34 (H1N1) propagated in two mammalian cell lines.

This chapter is adapted from Pralow et al. (2017) [81], Nguyen-Khuong and Pralow et al. (2018) [192] and Pralow et al. (2021) [160].

3.1. Cell lines, cell cultivation and virus infection

Cell lines, cell cultivation and virus infection used in this thesis was adapted from previous work [193, 194]. Briefly, MDCK.ADH cells (European collection of authenticated cell cultures (ECACC), No. 84121903) were cultivated in Glasgow Minimum Essential Medium (GMEM, Gibco Invitrogen, No. 22100-093), supplemented with 10 % (v/v) fetal calf serum (FCS)_(aq) and 1 % (v/v) peptone_(aq) (MC33, Lab M Limited). For infections experiments, MDCK.ADH cells were cultivated in triplicates in GMEM supplemented with 5.5 mmol/L glucose_(aq) (7509, Roth), 47 mmol/L NaHCO_{3(aq)} (HN01, Roth) and 0.2% (v/v) peptone_(aq). The MDCK.SUS2 cell line was generated by adaptation of MDCK.ADH cells (kindly provided by Prof. K. Scharfenberg, University of Applied Sciences Emden/Leer) and cultivated in biological duplicates in chemical defined medium SMIF8 PGd 2x (protein- and peptide-free; Gibco, through contact with Prof. K. Scharfenberg) with addition of 23.8 mmol/L NaHCO_{3(aq)}, 85.6 mmol/L NaCl_(aq), 0.1% (w/v) Pluronic-F68, 0.001‰ (w/v) ethanolamine (98%), 1.6 mmol/L L-glutamic acid, 20.3 mmol/L D-(+)-glucose, 4 mmol/L glutamine and 4 mmol/L pyruvate [193]. Cells were infected with human IAV PR/8/34 (H1N1) (no. 3138, Robert Koch Institute), in-house-generated adherent MDCK-derived viral stock, 5.17×10^8 TCID₅₀ /mL, HA titer of 2.63 log₁₀ HA units/100 µL) with a multiplicity of infection of 0.025 and 2×10^{-6} units trypsin (27250-018, Gibco) per cell. Vented shaker flasks with a volume of 250 mL (MDCK.SUS2) or 850 cm² roller bottles (MDCK.ADH; both VWR) were used for all infection experiments. 72 h post infection, supernatants containing released virus particles were harvested and subsequently purified.

3.2. Virus harvesting, purification and inactivation

Virus particles were harvested from the cells via *g*-force step-gradient centrifugation as described before [36, 195]. This involved pelleting of the cell culture supernatant with three centrifugation steps: (I. 300 x *g*, 4°C, 20 min; II. 4.000 x *g*, 4°C, 35 min; III. 10.000 x *g*, 4°C, 45 min). After each step, the supernatant was collected for further centrifugation. After a final ultracentrifugation step (100.000 x *g*, 4°C, 90 min), the pellet was resuspended in 50 µL 100 mM Tris-HCL_(aq) buffer with 5% SDS (w/v) and incubated for 10 min at 56°C for inactivation using a thermomixer (Eppendorf).

3.3. Protein concentration assay

The protein concentration of the purified IAV harvest was determined using the QuantIT protein assay (Q33210, Life technologies), following the assay instructions from the supplier.

3.4. *N*-glycan analysis: Sample preparation, measurement and data analysis

3.4.1. Chemicals, solvents and buffers

Bovine fetuin (F2379-100MG), PNGase F (P7367), potassium hydroxide (KOH, 484016-1KG), dithiothreitol (DTT, D5545-5G), iodoacetamide (IAA, I1149-25G), ammonium bicarbonate (ABC, 09830-500G), sodium borohydride (NaBH₄, 71320-100G), acetic acid (49199-50ML-F), ammonium acetate (A-1542) and acetonitrile (ACN, LC-MS Grade ≥ 99.5%, 34967) were all purchased from Sigma-Aldrich. Urea was obtained from AppliChem (A1049). Methanol LC/MS-grade was purchased from Fisher Scientific (10031094). All buffers and solutions were prepared with deionized and purified water (dH₂O) using a Milli-Q water purification system (18.2 MΩ·cm⁻¹ at 25°C, total organic carbon of 3 ppb) from Merck Millipore. For LC-MS solvents, water was further purified using an LC-Pak Polisher from Merck Millipore.

3.4.2. *N*- and *O*-glycan release

N- and *O*-linked glycan release for MS analysis was adapted from methods previously described [40, 43, 196]. One hundred microgram (100 µg) of bovine fetuin (Chapter 4.2), as well as 100 µg of IAV proteins (Chapter 4.3) were dissolved in 200 µL of 8 M urea buffer (0.05 M ABC_(aq) pH 8.5; urea_(aq)). The denatured and dissolved sample was transferred to a filter unit (Nanosep® Omega™ with polyethersulfone membrane, MWCO 10 kDa; PALL Life Sciences) [92]. The sample was centrifuged at 14.000 x *g* for 10 min at standard room temperature (RT). Afterwards, 30 µL of 10 mM DTT (in urea_(aq)) was added to the sample and incubated at 56°C for 20 minutes using an Eppendorf ThermoMixer). The solution was then centrifuged (17.000 x *g*, 10 min, RT) before 30 µL of 55 mM IAA (in urea_(aq)) was added and incubated at RT for 20 min in the dark. The sample was then washed three times with 100 µL 50 mM ABC_(aq) followed by centrifugation (17.000 x *g*, 10 min RT). The denatured, reduced and alkylated protein on top of the filter membrane was finally reconstituted in 50 µL of 50 mM ABC_(aq).

The *N*-linked glycans were released by adding 1U of PNGase F reconstituted in PBS (pH 8.0) to the filter unit containing the reduced and alkylated protein. Samples were incubated overnight at 37°C using a temperature-controlled incubator (Titramax 1000 with Incubator 1000, Heidolph Instruments). The released *N*-glycans were harvested by washing and centrifuging the filter unit at 14,000 x *g* for 10 min at RT in three steps as followed: Centrifuging the supernatant, washing and centrifuging with 50 µL 50 mM ABC_(aq) + 5% ACN (v/v) and washing and centrifuging with 50 µL 50 mM ABC_(aq). The flow-through was collected and pooled. Glycosylamines were removed from the reducing terminus of the glycans by adding 10 µL of 100 mM ammonium acetate at pH 5 (final concentration 15 mM) for 60 min at RT. Approximately 150 µL flow-through was collected and dried in a vacuum centrifuge (RVC 2–33 CDplus, ALPHA 2–4 LDplus, Martin Christ). The *N*-glycans were reduced with 20 µL of 1 M NaBH₄ in 50 mM

KOH_(aq) at 50°C for 3 h. The reduction was quenched with 1 µL acetic acid and the *N*-linked glycan alditols were desalted as described below.

Following the *N*-glycan release, *O*-linked glycans were released by aspirating the glycoproteins on top of the filter membrane and then conducting reductive β-elimination by incubating overnight with 20 µL of 0.5 M NaBH₄ in 50 mM KOH_(aq) at 50°C. Again, the reduction was quenched with 1 µL acetic acid and also the *O*-linked glycan alditols were desalted as described below.

3.4.3. Desalting and enrichment of glycans

The separate fractions of *N*- and *O*-glycans were desalted using cation exchange columns comprising 30 µL of a methanolic slurry of AG50W-X8 cation-exchange resin (142-1431, BioRad) packed on top of a 20 µL C18 StageTip Frit (SP201, Thermo Fisher Scientific). Residual borate was removed by addition of 50 µL methanol followed by drying under vacuum. This was repeated 5 times.

Further clean-up of IAV samples was achieved via PGC-SPE using 5 µL of a 80% ACN_{aq}+0.1% TFA_{aq} slurry of Carbograph (1769, Grace) packed on top of a 20 µL C18 StageTip Frit. The SPE was performed in the following steps, each three times using 10 µL, respectively, followed by centrifugation using a table centrifuge (Eppendorf): I. Wash with 50% ACN_{aq}+0.1% TFA_{aq}; II. Equilibration with 0.1% TFA_{aq}; III. Load 10 µL (≈1 µg/µL) *N*-glycans resuspended in 0.1% TFA_{aq} and repeat in total three times by pipetting the permeate on top of the tip, IV. Wash with 0.1% TFA_{aq}, V. Elute with 50% ACN_{aq}+0.1% TFA_{aq}. The eluate was dried in a vacuum centrifuge (RVC 2–33 CDplus, ALPHA 2–4 LDplus, Martin Christ GmbH) and resuspended in 10 µL 0.1% TFA_{aq}.

For bovine fetuin, each separate *N*- and *O*-glycan fraction was dissolved in 25 µL dH₂O, respectively. Finally, both fractions were pooled. Ten microliters (10 µL) of released glycans (equivalent of 20 µg of bovine fetuin) were aliquoted (final concentration 2 µg/µL) and stored at -20°C.

3.4.4. *N*- and *O*-glycan measurement by nano-PGC-LC-MS(/MS)

For LC separation an UltiMate 3000 RSLCnano system from Thermo Fisher Scientific was used. Five microliters (5 μ L) of a glycan sample were loaded isocratically through a nanoViper 20 μ L FS/PEEK-sheathed sample loop on a Hypercarb PGC (Hypercarb Kappa, 320 μ m x 3 cm, 35005-030315, Thermo Fisher Scientific) trap column via 100% loading buffer A (10 mM ABC_(aq)) with a flow rate of 7 μ L/min for 5 min. Afterwards, the loaded trap column was switched in-line with the analytical nano-column (Hypercarb Kappa, 75 μ m x 10 cm, 35003-100065, Thermo Fisher Scientific). The separation was performed at a flow of 0.9 μ L/min by using a gradient of buffer A (10 mM ABC_(aq)) and buffer B (10 mM ABC_(aq) + 90% ACN (v/v)). Specifically, separation was performed using the following multi-step binary gradient: 5-9% buffer B till 17 min, 9-56% buffer B till 65 min, 56-90% buffer B till 90 min and 5% buffer B 90-100 min.

To improve the electrospray stability, both the nano-PGC-LC gradient flow and the isocratic PCMF (100% ACN) were combined using a T-piece (SC901, Thermo Fisher Scientific). The PCMF was pumped from a glass syringe (500 μ L; Hamilton) with 2 μ L/min using the integrated syringe pump of the LTQ Orbitrap Elite mass spectrometer (Thermo Fisher Scientific). Both, the nano-LC flow and the PCMF were transferred to the T-piece, each via a NanoViper capillary (20 μ m inner diameter (InD), 1/32" outer diameter (OD), 650 mm length, 6041.5275, Thermo Fisher Scientific) and the resultant mixed flow was directed to a electropolished stainless steel spray needle (Nano-bore emitter, 30 μ m InD, 1/32" OD, 40 mm length, ES542, Thermo Fisher Scientific) via a NanoViper capillary (75 μ m InD, 1/32" OD, 550 mm length, 6041.5760, Thermo Fisher Scientific). The sheme for this setup is illustrated in Figure 3-2.

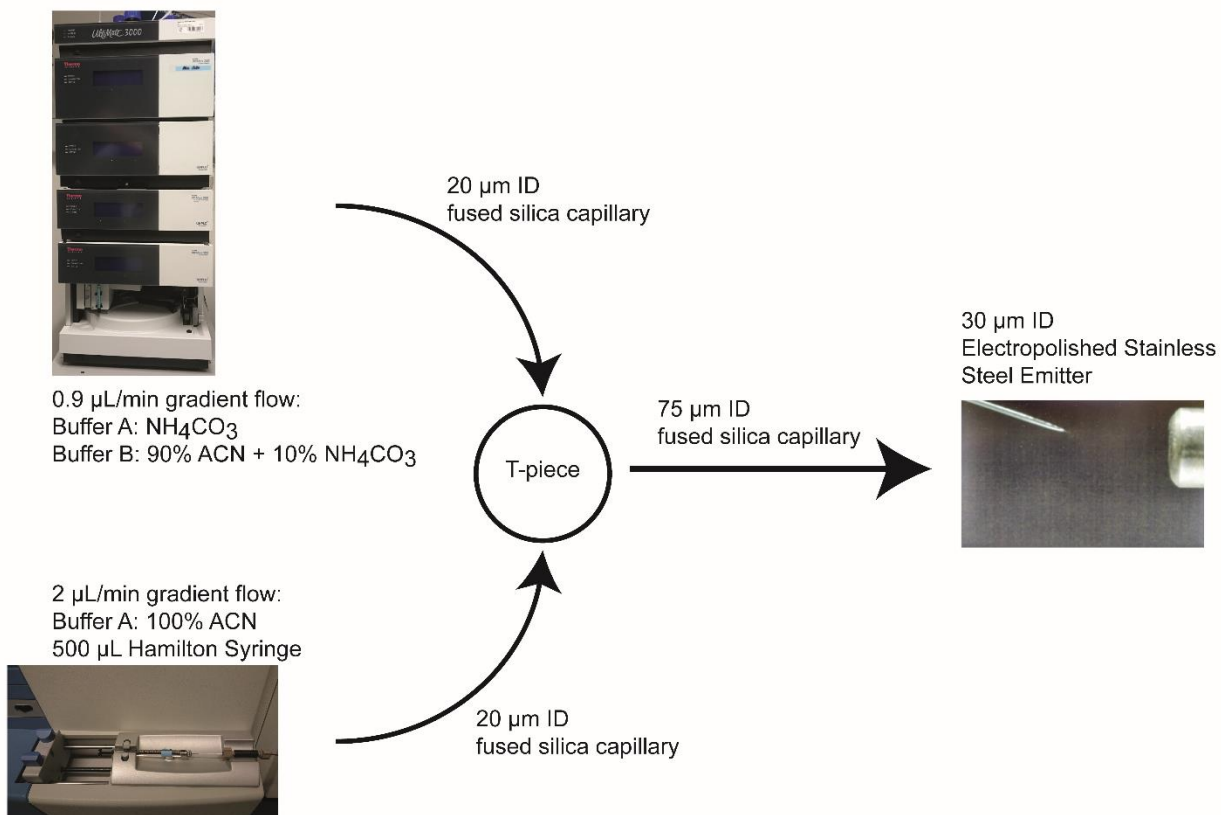


Figure 3-2: **Schematic of the PCMF setup.** The post-column nano-flow is supplemented with an additional capillary flow of 100% ACN via a T-piece. The combined flow is directed to a 30 μm ID stainless steel electrospray needle via a 75 μm ID fused silica capillary. Adapted from Nguyen-Khuong and Pralow et al. (2018) [192].

The eluting glycans were measured on an LTQ Orbitrap Elite mass spectrometer from Thermo Fisher Scientific using a Nanospray Flex™ source (Thermo Fisher Scientific) in negative ion mode with a capillary voltage of 2.7 kV. Glycans were fragmented using HCD at a NCE of 75 with an activation time of 0.1 ms. The top 5 most intensive precursor ions with a charge state >1 were chosen for fragmentation. The mass range for MS and MS/MS was 150-2000 m/z. For bovine fetuin, these settings were used in both setups, with and without PCMF.

3.4.5. Analysis of PGC-LC-MS(-MS) data from releases glycans

Data files of bovine fetuin were analyzed manually using Xcalibur (Version 2.2, Qual Browser, Thermo Fisher Scientific). Acquired MS scans were analyzed and a list of monoisotopic masses corresponding to glycans was sent to ExPASy GlycoMod (<http://web.expasy.org/glycomod>) to retrieve a list of possible matching glycan compositions. The compositions and glycan structures were confirmed by manually annotating the MS/MS fragment ion spectra of each glycan structure, also considering diagnostic ions that are characteristic for certain glycan epitopes.

MS datasets obtained from the IAV *N*-glycan analysis of each cell line and their biological replicates (three MDCK.ADH and two MDCK.SUS2) were inspected manually. Specifically, deconvoluted precursor masses were submitted to ExPASy GlycoMod (free online tool: <https://web.expasy.org/glycomod/>) and UniCarb-DB to calculate possible *N*-glycan compositions. The fine-structure of *N*-glycans was identified based on the manual annotation of specific fragment ion spectra with diagnostic cross-ring fragment ions [41, 75]. Identified structures of different *N*-glycan isomers were additionally secured based on their retention behavior [76].

3.5. Glycopeptide analysis: Sample preparation, measurement and data analysis

3.5.1. Chemicals, solvents and buffers

The proteins bovine serum albumin (BSA; A3912-100G) and lactotransferrin from human milk (hLTF; L4894-5MG) were purchased from Sigma-Aldrich. Enzymes used for proteolytic digestion were trypsin (trypsin sequencing grade modified, V5111) from Promega, endoproteinase AspN [flavastacin [81]] (AspN, P8104S) from New England Biolabs and PNGase F (P7367) from Sigma-Aldrich. KOH (484016-1KG), DTT (D5545-5G), IAA (I1149-25G), ABC (09830-500G) and ACN, LC-MS Grade \geq 99.5%, 34967) were all purchased from Sigma-Aldrich. Urea (A1049), Tris-HCL (A3452) and calcium chloride (A46899)

were from AppliChem and trifluoroacetic acid (TFA, 28904) was purchased from Thermo Fisher Scientific. All chemicals were in the highest grade available. All solvents for LC-MS were MS grade. All buffers and solutions were prepared with deionized and purified water using a Milli-Q water system (18.2 M Ω ·cm⁻¹ at 25°C, total organic carbon of 3 ppb) from Merck Millipore. For LC-MS solvents, water was further purified using the LC-Pak Polisher from Merck Millipore.

3.5.2. Proteolytic digestion via filter-aided sample preparation

BSA and hLTF (Chapter 4.1), as well as proteins from IAV (Chapter 4.3) were digested proteolytically using a modified version of the FASP method [92]. Briefly, 100 μ g of each protein were applied to a filter unit (Nanosep[®] Omega[™] with polyethersulfone membrane, MWCO 10 kDa; PALL Life Sciences). Samples were treated with urea buffer_(Tris-HCl) (8 M urea in 0.1 M Tris-HCl_(aq) pH 8.5; AppliChem), followed by reduction with 40 mM DTT_(aq) and alkylation with 55 mM IAA_(aq) – each dissolved in 50 mM ABC buffer_(aq). Each filter unit was washed three times with urea buffer_(Tris-HCl) and three times with ABC buffer_(aq). Proteins were digested proteolytically with trypsin using an enzyme/protein ratio of 1:30 (w/w). Samples were incubated overnight at 37°C and 350 rpm using a temperature controlled incubator (Titramax 1000 + Inkubator 1000, Heidolph). Digests were collected by centrifugation. Filter units were washed twice, first using 50 μ L ABC buffer_(aq) with 5%_(v/v) ACN, then using 50 μ L dH₂O; in between samples were centrifuged. The flow through was kept along with the digest, in order to be dried by vacuum centrifugation.

After tryptic digestion, approximately 20 μ g peptides were reconstituted in 20 μ l 1x AspN reaction buffer (New England Biolabs; 50 mM Tris-HCL, 2.5 mM Zinc Sulfate, pH 8.0). Afterwards AspN (Flavastacin) was added (enzyme/protein ratio 1:20) to the peptide solution and incubated overnight at 37°C as recommended by the supplier. The enzyme reaction was stopped via centrifugation through a filter unit (same as described above). The flow through (tryptic digests, as well as the sequential digests

with trypsin and AspN (Flavastacin) were dried by vacuum centrifugation and reconstituted in 0.1%_(v/v) TFA_(aq) prior to LC-MS(/MS) measurements.

3.5.3. Enrichment of glycopeptides

Glycopeptides derived from IAV were enriched using a HILIC-SPE method developed by Selman et al. (2011) [197] and modified by Hoffmann et al. (2018) [97]. A 20 μ L pipette tip was stuffed with cotton to the 10 μ L marking. The SPE was performed in the following steps, each three times using 10 μ L, respectively, followed by centrifugation using a table centrifuge (Eppendorf): I. Wash with 0.1% TFA_{aq}; II. Equilibration with 83% ACN_{aq}+0.1% TFA_{aq}; III. Load 10 μ L (\approx 1 μ g/ μ L) *N*-glycans resuspended in 83% ACN_{aq}+0.1% TFA_{aq} and repeat in total three times by pipetting the permeate on top of the tip, IV. Wash with 83% ACN_{aq}+0.1% TFA_{aq}, V. Elute with 0.1% TFA_{aq}. The eluate was dried in a vacuum centrifuge (RVC 2–33 CDplus, ALPHA 2–4 LDplus, Martin Christ GmbH) and resuspended in 10 μ L 0.1% TFA_{aq}.

3.5.4. *N*-glycopeptide measurement by RP-LC-MS(/MS)

For all measurements, an UltiMate 3000 Rapid Separation LC system from Thermo Fisher Scientific was used. Samples (\approx 500 ng) were loaded isocratically on a trap column (Acclaim PepMap[®]100, 100 μ m x 2 cm nanoViper C18, 5 μ m, 100 Å, Thermo Fisher Scientific) via 100% loading buffer A (98%_(v/v) dH₂O, 2%_(v/v) ACN, 0.05%_(v/v) TFA) with a flow rate of 7 μ L/min within the first 5 min. Afterwards, the loaded trap column was switched in line with the separation column (Acclaim PepMap[®]RSLC, 75 μ m x 25 cm nanoViper C18, 2 μ m, 100 Å, Thermo Fisher Scientific), with a nano-flow rate of 0.3 μ L/min of 4% nano buffer B (10%_(v/v) dH₂O, 10%_(v/v) trifluoroethanol (TFE), 80%_(v/v) ACN, 0.1%_(v/v) formic acid (FA)) and nano buffer A (98%_(v/v) dH₂O, 2%_(v/v) ACN, 0.1 %_(v/v) FA). The separation was performed by a multi-step binary nano A/B gradient: 4-55% nano buffer B till 80 min, 55-90% nano buffer B till 100 min, 90% nano buffer B till 110 min and 4% nano buffer B 110-150 min.

The eluting peptides were measured on an LTQ Orbitrap Elite mass spectrometer from Thermo Fisher Scientific using a Nanospray Flex™ source in positive ionization mode with a capillary voltage of -2.7 kV. Peptides and glycopeptides were fragmented using HCD with NCE of 35 with an activation time of 0.1 ms. The five most intense precursor ions with a charge state >1 were chosen for fragmentation. The recorded mass range for MS was 350-2000 m/z and for MS/MS 150-2000 m/z.

3.5.1. Analysis of RP-LC-MS(/MS) data from *N*-glycopeptides

The MS(/MS) data from BSA and hLTF were imported into *Proteome Discoverer* (Version 1.4, Thermo Fisher Scientific) and searched against UniProt-KB/SwissProt database (542258 sequences; downloaded January, 2014) using *MASCOT* (Version 2.5, Matrix Science). The MS(/MS) data were screened against an unspecific *in-silico* digestion of the mammalian taxonomy database with the fixed modification of cysteine with carbamidomethyl, variable deamidation of asparagine and variable oxidation of methionine. The precursor ion mass tolerance was set to 5 ppm and the fragment ion mass tolerance to 0.02 Da. The protein relevance threshold was set to 20 and the peptide cut off score to 10. The target false discovery rate for peptide hits was set to 0.01 (strict setting and to 0.05 as relaxed setting).

Manual glycopeptide analysis from hLTF and IAV glycoproteins was performed according to Pralow et al. (2017) [81] and Hoffmann et al. (2018) [97]. All acquired MS/MS spectra were manually screened for specific oxonium ions ($[M+H]^+$: HexNAc m/z 204.08, HexNAc₁Hex₁ m/z 366.14, HexNAc₁Hex₁dHex₁ m/z 512.20, HexNAc₁Hex₂ m/z 528.19). Oxonium ions of sialic acid are not supposed to be detected in IAV glycoprotein samples, because of the presence of NA in the samples. Therefore, the oxonium ion $[M + H]^+$ NeuAc-H₂O m/z 274.10 was only used for hLTF. After the classification of each glycopeptide fragment ion spectrum, the mass of the peptide moiety (Mp) for each fragment ion spectrum was determined based on a specific fragment ion pattern: Peptide mass minus ammonia $[Mp + H^+ - NH_3]^+$, peptide mass $[Mp + H]^+$, peptide mass plus ^{0,2}X-ring cleavage of the innermost GlcNAc $[Mp + H + ^{0,2} X$

GlcNAc]⁺ and peptide mass plus innermost GlcNAc [Mp + H + GlcNAc]⁺ (only for *N*-glycopeptides, for *O*-glycopeptides [Mp + H + ^{0.2} X GlcNAc]⁺ is missing) [97]. The putative peptide masses from IAV glycoproteins were searched against an *in-silico* digestion of HA (Uniprot accession number P03452), NA (Uniprot accession number P03468) and M2 (Uniprot accession number P06821), using ExPASy FindPept (free online tool <https://web.expasy.org/findpept/>). For hLTF, the Uniprot accession number P02788 was used. Variable modifications were the carbamidomethylation of cysteine and the oxidation of methionine. N-terminal carbamidomethylation was imitated by subtracting 57 Da from unknown predicted peptide masses. The precursor mass error was set to 10 ppm and an unspecific digestion was performed to take missed cleavages as well as unspecific cleavage products into account. The resulting peptide sequences were screened for possible *N*-glycosylation consensus sequences. To confirm the predicted peptide sequence, an *in-silico* fragmentation of the predicted peptide sequences was performed using MS-Product from ProteinProspector (v 5.16.0) (free online tool: <http://prospector.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msproduct>). Specific b- and y-ions of the *in-silico* fragmented predicted peptide sequence were compared with MS/MS fragment ions derived from the measured spectra and manually annotated with a maximum mass error of 10 ppm. By subtracting the determined peptide mass from the monoisotopic singly charged glycopeptide precursor mass, the mass of the putative glycan moiety was calculated. This mass was then screened for possible *N*-glycan compositions using ExPASy GlycoMod and UniCarb-DB. The mass error was set to 10 ppm and variable modifications were the carbamidomethylation of cysteine and the oxidation of methionine.

Based on the results of the manually annotated reference data set of IAV glycoproteins (first replicate of each sample type), other biological replicates were analyzed using an in-house-developed software glyXtool^{MS} for the semi-automated analysis of glycopeptide MS data. This software was developed by Pioch et al. (2018) and is available online on GitHub (<https://github.com/glyXera/glyXtoolMS>) [83]. The

mass error was set to 10 ppm and an unspecific digestion was performed to take missed cleavages as well as unexpected cleaved products into account. Variable modifications were the carbamidomethylation of cysteine and oxidation of methionine. The software compares defined possible *N*-glycan/peptide combinations with the identified monoisotopic glycopeptide precursor masses. MS/MS fragment ion spectra of matching glycopeptide precursor are further automatically annotated with b- and y-ions of possible glycopeptide sequences. It also performs a glycopeptide scoring based on the recognition of glycan-specific oxonium ions in the MS/MS fragment ion spectra. After data processing, results of identified glycopeptides were manually confirmed or discarded based on the coverage of specific b- and y-ions in the corresponding fragment ion spectra

3.5.2. Relative quantification of *N*-glycopeptides

Site-specific relative quantification of *N*-glycopeptides derived from IAV glycoproteins was performed using Byonic and Byologic (ProteinMetrics). A reference dataset of each sample was analyzed (first replicate of each sample type). The mass error was set to 10 ppm and 30 ppm for precursors and fragment ions, respectively. An unspecific digestion was performed to take missed cleavages as well as unexpected cleavage products into account. Variable modification was oxidation of methionine. Carbamidomethylation of cysteine was set as fixed modification. The mammalian *N*-glycome database provided from Byonic, which includes all *N*-glycan compositions identified from the *N*-glycomic analysis, was used for variable modification of *N*-glycosylation sites. Based on the *N*-glycoproteomic analysis, the percentage number of all individual *N*-glycan compositions of each HA and NA *N*-glycosylation site was determined from the sum of the extracted ion chromatogram (EIC) peak area of each multiply charged (>1) precursor signal relative to the sum of all *N*-glycopeptide signals of the respective *N*-glycosylation site.

3.6. Graphical illustration

The molecular structure of HA was modeled using the protein data bank (PDB) entry number 1RU7. To model the molecular structure of neuraminidase, the Swiss-Model template ID 5hug.1 was used. For model processing and design the open source software UCSF Chimera Version 1.10.2 was utilized [198].

4. RESULTS AND DISCUSSION

This chapter consists of three parts. In the first part, the finding of a new specificity of a proteolytic enzyme, as well as its evaluation for *N*-glycoproteomic analysis is demonstrated. Next, the development of a PCMF for stable and efficient nano-PGC-LC-MS(/MS) analysis of released glycans is described. Finally, both methods are applied to perform a comprehensive fine-structural and site-specific *N*-glycan analysis of IAV propagated in two different MDCK cell lines.

4.1. Improvement of the glycoproteomic toolbox with the discovery of a unique C-terminal cleavage specificity of flavastacin for *N*-glycosylated asparagine

This chapter is adapted from Pralow et al. (2017) [81].

4.1.1. Introduction

For every *N*-glycoproteomic analysis workflow, numerous parameters have to be adjusted optimally. In particular, proteolytic digestion using sequential steps and the selection of specific enzymes is important to overcome common problems, e.g. the analysis of large *N*-glycopeptides (with low charge density and/or sequence constraints) or *N*-glycopeptides with multiple glycosylation sites. Here, a new approach for the proteolytic digestion of glycoproteins by using flavastacin, a protease that was found to cleave specifically at the C-terminus of *N*-glycosylated asparagine is shown. The glycoprotein hLTF and the non-glycosylated protein BSA were used as model proteins to demonstrate this newly identified cleavage specificity. According to the manufacturer's recommendation, flavastacin works only on peptides smaller than 50 AAs. Therefore, hLTF and BSA were first treated with trypsin before flavastacin was added (see Materials and Methods).

The hLTF is a well-characterized glycoprotein present in human milk, saliva, tears, nasal secretions and other body fluids [199], which contains three potential *N*-glycosylation sites (N₁₅₆, N₄₉₇ and N₆₄₂). The sites N₁₅₆ and N₄₉₇ were described to carry complex-type mostly core-fucosylated and sialylated *N*-glycans and also N-acetylactosamine (LacNAc) extensions [200]. For site N₆₄₂, *N*-glycosylation was recently confirmed [97]. The theoretical tryptic *N*-glycopeptides of hLTF are (R)PFLN₁₅₆WTGPPEIEAAVAR(F) (1964.0155 Da) (with a commonly missed cleavage site due to proline: (R)TAGWNVPIGTLRPFLN₁₅₆WTGPPEIEAAVAR(F) (3229.7036 Da)), (R)TAGWNIPMGLLFN₄₉₇QTGSK(F) (2036.9812 Da) and (R)N₆₄₂GSDCPDK(F) (834.3178 Da). These peptides consist of 18/(30), 19 and 8 AAs, respectively. The large peptide moieties of the sites N₁₅₆ and N₄₉₇ in combination with an *N*-glycosylation can reach masses, which are unsuitable for *N*-glycoproteomic analysis [91].

4.1.2. Results and discussion

N-glycoproteomic analysis of flavastacin-generated hLTF *N*-glycopeptides revealed that all *N*-glycopeptides feature the *N*-glycosylated asparagine at the C-terminus, while the N-terminus was either a tryptic or an unspecific cleavage site. In addition, the generated *N*-glycopeptides were shorter compared to a solely tryptic digest. The identified *N*-glycopeptides for the *N*-glycosylation sites N₁₅₆ and N₄₉₇ were 4–16 and 4–13 AAs long, respectively (see Table 4-1). No *N*-glycopeptide was detected for the *N*-glycosylation site N₆₄₂ [200]. No further enrichment of glycopeptides prior to MS analysis was performed, since the digestion strategy together with a long separation gradient and a high-resolution MS measurement resulted in a comprehensive coverage of glycopeptides as shown in Figure 4-1. The eluting *N*-glycopeptides are depicted by extracted ion chromatograms of MS/MS spectra containing glycan-specific oxonium ions: *m/z* 204.087 HexNAc [M+H]⁺, *m/z* 274.093 (NeuAc-H₂O [M+H]⁺) and *m/z* 366.141 Hex+HexNAc [M+H]⁺. The unspecific cleavage specificity due to flavastacin resulted in redundant *N*-glycopeptide signals over the elution time range with peptide moieties of different length. This can be a disadvantage for the analysis of more complex glycopeptide samples.

Table 4-1: **Manually annotated *N*-glycopeptide sequences of human lactotransferrin (hLTF) from MS(/MS) spectra of nano-RP-LC-ESI-OT-MS²(HCD) measurements after sequential digestion of hLTF with trypsin and flavastacin. CAM - carbamidomethylation; MSO - methionine S-oxidation. Adapted from Pralow et al. (2017) [81].**

N-Glycosylation Site	Experimental/Observed mass [Da]	Theoretical mass [Da]	Δ Mass [Da]	Peptide	Position	Modifications	tR [min]
156	490.2640	490.2660	0.002	(R)PFLN(W)	153-156		24.73
	860.4960	860.4990	0.003	(G)TLRPFLN(W)	150-156		25.07
	1127.6620	1127.6570	-0.005	(V)PIGTLRPFLN(W)	147-156		40.09
	1226.7230	1226.7260	0.003	(N)VPIGTLRPFLN(W)	146-156		41.54
	1755.9520	1755.9540	0.002	(R)TAGWNVPIGTLRPFLN(W)	141-156		42.17
	1340.7680	1340.7680	0.000	(W)NVPIGTLRPFLN(W)	145-156		42.93
497	856.4950	856.4961	-0.001	(N)IPM(-48)GLLFN(Q)	490-497	M(CAM) (-105)	30.11
	506.2990	506.2970	-0.002	(G)LLFN(Q)	494-497		32.10
	563.3210	563.3190	-0.002	(M)GLLFN(Q)	493-497		36.07
	1385.7223	1385.7246	-0.002	(R)TAGWNIPM(-48)GLLFN(Q)	485-497	M(CAM) (-105)	39.58
	920.4930	920.4910	-0.002	(N)IPMGLLFN(Q)	490-497	MSO	39.95
	904.4950	904.4960	0.001	(N)IPMGLLFN(Q)	490-497		45.56
	791.4110	791.4120	0.001	(I)PMGLLFN(Q)	491-497		47.45
	694.3590	694.3590	0.000	(P)MGLLFN(Q)	492-497		53.33

RESULTS AND DISCUSSION

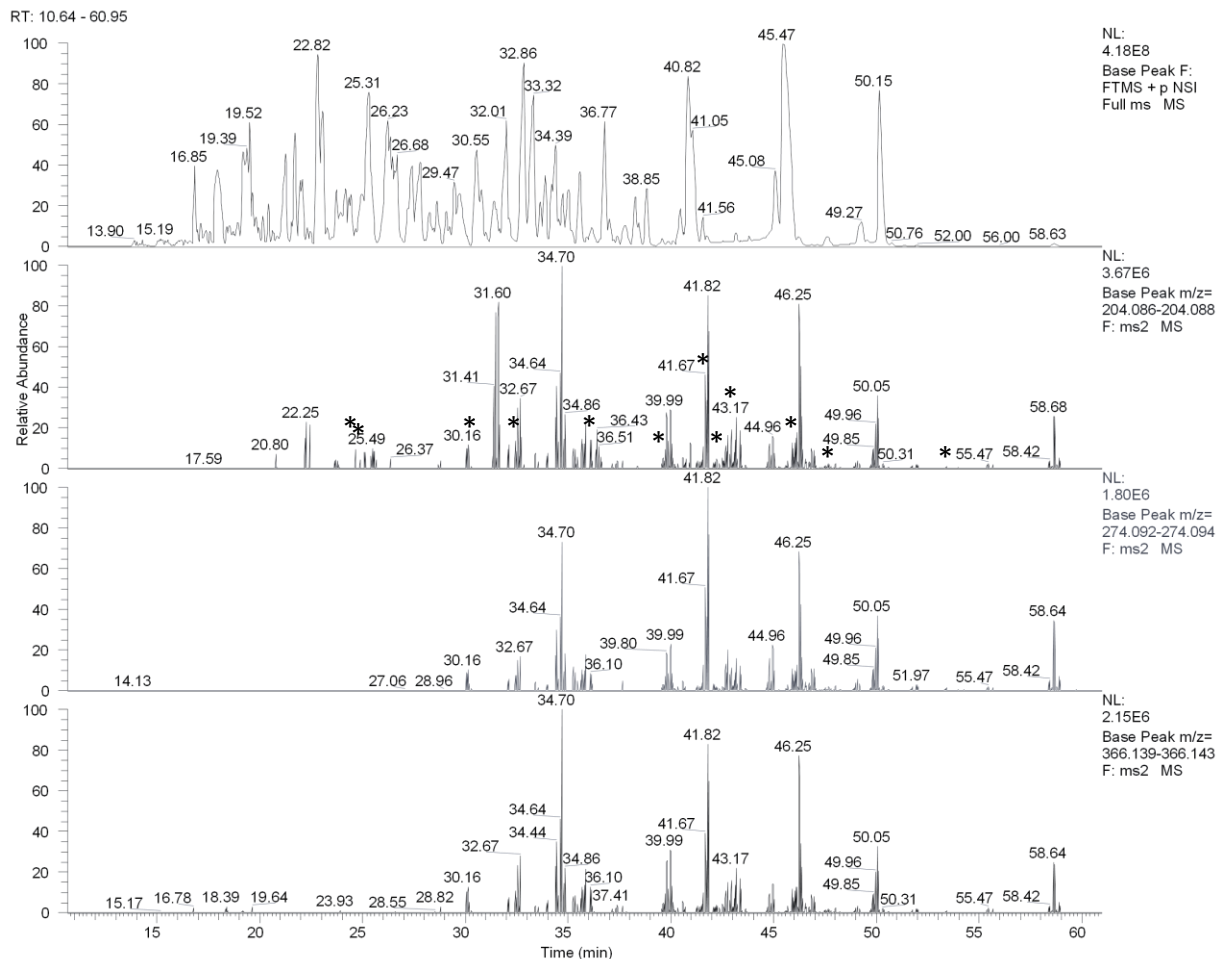


Figure 4-1: **Base peak ion-chromatogram (BPC) and oxonium ion related extracted ion-chromatograms (EIC) of MS/MS spectra of nanoRP-LC-ESI-OT-MS²(HCD) measured hLTF after sequential digestion with trypsin and flavastacin.** Starting from the top: BPC of MS spectra (grey), EIC of MS/MS spectra of HexNAc within the m/z range 204.086-204.088 $[M+H]^+$ (red), EIC of MS/MS spectra of NeuAc-H₂O within the m/z range 274.092-274.094 (green), EIC of MS/MS spectra of HexHexNAc within the m/z range 366.139-366.143 (blue). The ion-chromatograms are illustrated in the time range 10.64-60.95 min. The accepted mass error of the EIC of the specific oxonium ions is 5 ppm. Signals listed in Table 1 are marked using asterisks. Adapted from Pralow et al. (2017) [81].

To illustrate the unique cleavage specificity of flavastacin, two examples for annotated MS/MS spectra of the *N*-glycopeptide containing sites N₄₉₇ (Figure 4-2A) and N₁₅₆ (Figure 4-2B) are shown. Both *N*-glycopeptides have the *N*-glycosylated asparagine at the C-terminus and the tryptic cleavage site at the N-terminus. For the site N₄₉₇, the fragment ion spectrum of the *N*-glycopeptide TAGWNIPM*GLLFN₄₉₇ with the *N*-linked glycan Hex₅HexNAc₄dHex₁NeuAc₁ (precursor ion: m/z 1184.1659 $[M+3H]^{3+}$) is depicted, along with the corresponding peptide ion 1385.7266 $[M+H]^+$ (see

Figure 4-2A). The example for the fragment ion spectrum of site N₄₉₇ shows a neutral loss of 105 Da, related to a carbamidomethyl-methionine residue (see Figure 4-2A). This neutral loss from a carbamidomethylated methionine has been described only rarely in literature [201]. Without the awareness of a carbamidomethylation of methionine, only a neutral loss of 48 Da would be recognized, which can mistakenly also be interpreted as a side chain loss of methionine sulfoxide. However, due to the specific digestion strategy involving both trypsin and flavastacin, in combination with high-resolution LC-MS, the identification of such unlikely modifications is also possible. For site N₁₅₆, the fragment ion spectrum of the *N*-glycopeptide TAGWNVPIGTLRPFLN₁₅₆ with the *N*-linked glycan Hex₅HexNAc₄dHex₂NeuAc₁ (precursor ion: m/z 1321.2479 [M+3H]³⁺) is depicted (see Figure 4-2B) as well as the corresponding peptide ion 1755.9515 [M+H]⁺. Both fragment ion spectra are dominated by the b-ion series, oxonium ions (B-ions) and the aforementioned fragmentation pattern: [peptide - NH₃ + H⁺]; [peptide + H⁺] and [peptide + GlcNAc + H⁺] (Figure 4-2).

RESULTS AND DISCUSSION

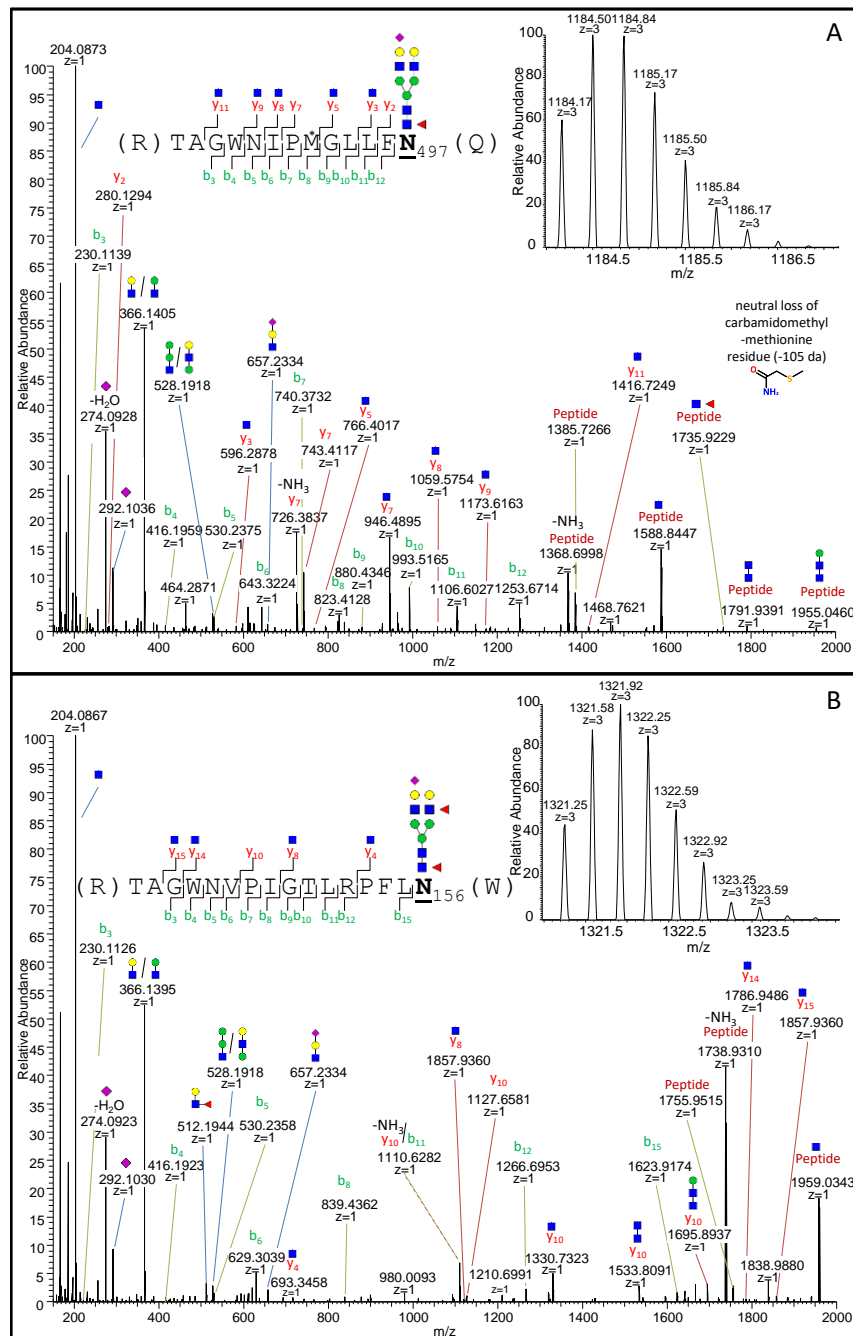


Figure 4-2: **Fragment ion spectra of nanoRP-LC-ESI-OT-MS²(HCD) measured hLTF N-glycopeptides after sequential digestion with trypsin and flavastacin.** (A) For site N₄₉₇ the fragment ion spectrum of the N-glycopeptide sequence TAGWNIPM*GLLFN₄₉₇ with the N-linked glycan Hex₅HexNAC₄Hex₁NeuAc₁ and the corresponding precursor ion m/z 1184.1659 [M+3H]³⁺ is shown. The * indicates the carbamidomethylation of methionine. (B) For site N₁₅₆ the fragment ion spectrum of the N-glycopeptide sequence TAGWNVPIGTLRPFLL₁₅₆ with the N-linked glycan Hex₅HexNAC₄Hex₂NeuAc₁ and the corresponding precursor ion m/z 1321.2479 [M+3H]³⁺ is shown. N-glycan structures and oxonium ions are illustrated according to the Symbol Nomenclature for Graphical Representations of Glycans [6]. The b-ion series are highlighted in green, the y-ion series in red and the B-ions in blue. The isotopic patterns of the precursors are shown at the upper right corner. Adapted from Pralow et al. (2017) [81].

Analysis of flavastacin-generated BSA peptides revealed primarily tryptic cleavage at the C-terminus (34 of 37 peptides) and tryptic and unspecific cleavage at the N-terminus (13 tryptic, 24 unspecific cleavages). Twelve peptides had tryptic cleavages at both termini (Table 4-2). For the non-glycosylated BSA no flavastacin-generated peptides were detected with an asparagine at the C-terminus, which correlates to observations for non-glycosylated asparagines of the *N*-glycosylated hLTF (see Supplementary Table A1-1). To check for possible unspecific cleavages of the tryptic digest and its influence on the flavastacin digest, identified peptides of solely tryptic digested BSA (Supplementary Table A1-2) and hLTF (Supplementary Table A1-3) were examined. Here, almost exclusively, specific cleavages were identified in the tryptic digests of BSA and hLTF. This strongly suggests that the observed *N*-glyco-specific cleavage of hLTF, as well as the unspecific cleavage of BSA and hLTF of the combined digest (trypsin and flavastacin) can only be linked to the activity of flavastacin.

Whilst it has been described that flavastacin has specificities towards the N-terminus of aspartic acid, glutamic acid and cysteine, a unique cleavage specificity of flavastacin for the C-terminus of *N*-glycosylated asparagine was found, which was not explored up to now. All manually annotated hLTF *N*-glycopeptide related peptide sequences are listed in Table 4-1 (other identified non-glycosylated peptides are listed in Supplementary Table A1-1). Every single hLTF *N*-glycopeptide sequence has been cleaved at the *N*-glycosylated asparagine at the C-terminus – independent of the *N*-glycoform attached to the respective *N*-glycosylation site. In addition, the N-terminus is a tryptic or an unspecific cleavage site. Based on the manually annotated peptide sequences and the database-assisted MASCOT search, no strict N-terminal cleavage of aspartic acid could be observed, neither for BSA nor for hLTF (Table 4-1, Table 4-2 and Supplementary Table A1-1).

RESULTS AND DISCUSSION

Table 4-2: *Proteome Discoverer results of a nano-RP-LC-ESI-OT-MS²(HCD) measurement after sequential digestion of BSA with trypsin and flavastacin. MASCOT search against unspecific in-silico digestion of mammalian taxonomy (UniProt-KB/SwissProt database). Peptide sequences with the N-terminal cleavage of aspartic acid or deamidated asparagine (as indicated from the supplier of flavastacin) are highlighted in red. Adapted from Pralow et al. (2017) [81].*

Sequence	# PSM	Modifications	MH+ [Da]	IonScore	ΔM [Da]
(K)TVMENFVAFVDK(C)	2	N5(Deamidated)	1400.67510	107	-0.002
(R)MPCTEDYLSLILNR(L)	2	C3(Carbamidomethyl)	1724.82927	98	-0.005
(K)LGEYGFQNALIVR(Y)	42	N8(Deamidated)	1480.77873	98	0.004
(K)LGEYGFQNALIVR(Y)	13		1479.79204	97	0.000
(K)TVMENFVAFVDK(C)	9		1399.69121	93	-0.001
(K)TVMENFVAFVDK(C)	1	M3(Oxidation)	1415.68498	93	-0.003
(K)VPQVSTPTLVEVSR(S)	4		1511.84062	86	-0.002
(P)CTEDYLSLILNR(L)	3	C1(Carbamidomethyl)	1496.73870	83	-0.002
(R)RHPEYAVSVLLR(L)	6		1439.81028	82	-0.002
(M)ENFVAFVDK(C)	7		1068.53703	77	0.001
(R)KVPQVSTPTLVEVSR(S)	2		1639.93723	76	-0.001
(K)DAFLGSFLYEYSR(R)	2		1567.73613	76	-0.007
(E)YGFQNALIVR(Y)	6	N5(Deamidated)	1181.63017	75	-0.001
(G)EYGFQNALIVR(Y)	5	N6(Deamidated)	1310.67229	74	0.000
(M)ENFVAFVDK(C)	2	N2(Deamidated)	1069.51872	71	-0.001
(K)LGEYGFQNAL(I)	2		1111.54180	68	0.000
(K)KQTALVELLK(H)	13		1142.71416	68	-0.000
(Y)FYAPPELLYANK(Y)	1		1491.75029	68	-0.002
(K)LVNELTEFAK(T)	1	N3(Deamidated)	1164.61382	65	-0.001
(E)DYLSLILNR(L)	4		1106.62212	63	0.002
(K)DAIPENLPLTADFAEDK(D)	2		1955.95671	60	-0.003
(S)TPTLVEVSR(S)	1		1001.56133	57	-0.001
(K)HLVDEPQNLK(Q)	2		1305.71526	57	-0.001
(R)HPEYAVSVLLR(L)	2		1283.70691	56	-0.004
(P)EYAVSVLLR(L)	3		1049.59648	56	-0.003
(E)YAPPELLYANK(Y)	1		1344.68120	55	-0.002
(Q)VSTPTLVEVSR(S)	1		1187.66106	55	-0.002
(F)LGSFLYEYSR(R)	3		1234.60918	55	-0.001
(N)FVAFVDK(C)	2		825.44951	55	-0.001
(N)LPPLTADFA(E)	7		944.50633	54	-0.001
(A)IPENLPLTA(D)	5		1064.59661	54	-0.002
(Y)APPELLYANK(Y)	1		1181.61870	54	-0.001
(T)ALVELLK(H)	1		785.51219	54	-0.001
(E)YAVSVLLR(L)	2		920.55394	52	-0.002
(K)LVNELTEFAK(T)	1		1163.62810	52	-0.003
(Q)TALVELLK(H)	3		886.55974	48	-0.001
(G)FQNALIVR(Y)	1		960.56017	47	-0.002

Flavastacin shows a clear specificity for the C-terminus of *N*-glycosylated asparagine N₁₅₆ and N₄₉₇ in hLTF (illustrated in Figure 4-3). Due to the presence of multiple *N*-glycosylation sites and the well-described complex-type *N*-glycan structures, hLTF is a very suitable glycoprotein to demonstrate this newly found specificity of flavastacin.

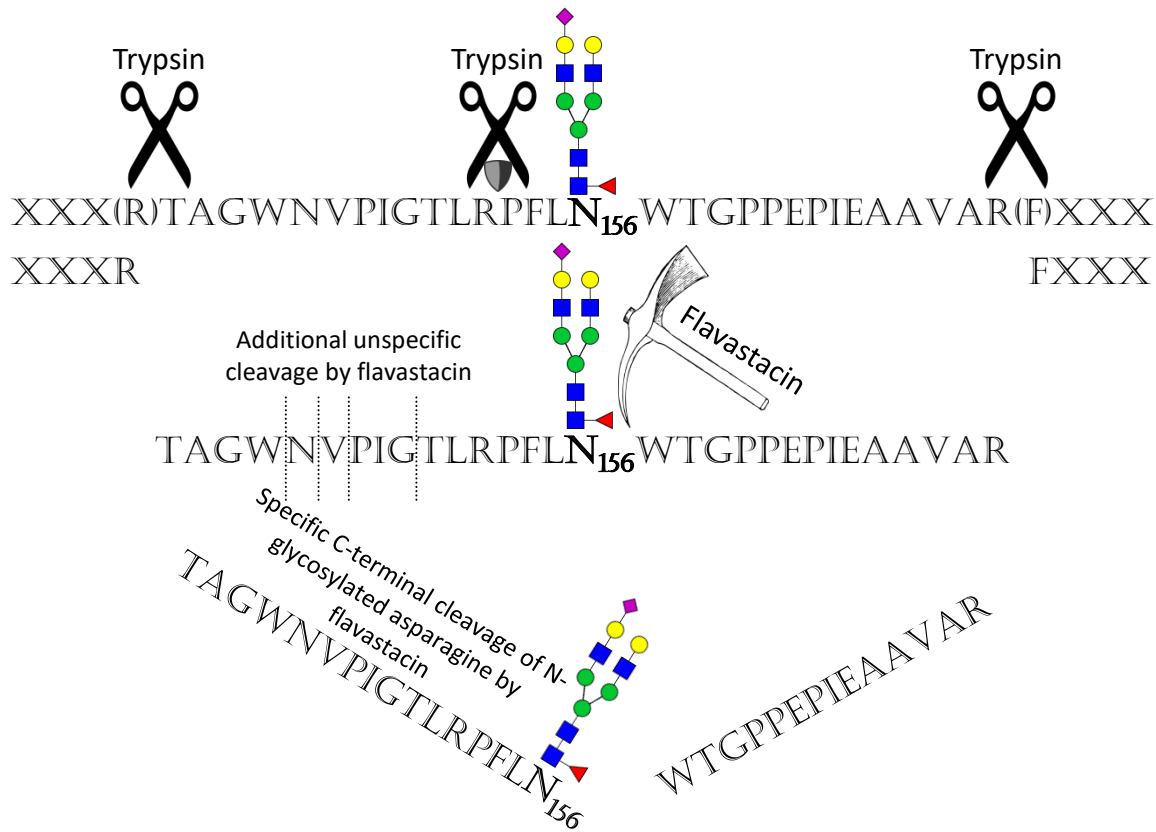


Figure 4-3: **Specificity of flavastacin for the C-terminus of *N*-glycosylated asparagine.** The amino acid sequence of the tryptic *N*-glycopeptide of hLTF with the *N*-glycan Hex₅HexNAc₄dHex₁NeuAc₁ linked to *N*-glycosylation site N₁₅₆ is shown. The scissors symbolize trypsin; the shield stands for cleavage inhibition of trypsin due to proline (P). Flavastacin is symbolized by a pick and its specific C-terminal cleavage of *N*-glycosylated asparagine, as well as its unspecific N-terminal cleavages are shown. The *N*-glycan structure is illustrated according to the Symbol Nomenclature for Graphical Representations of Glycans [6]. Adapted from Pralow et al. (2017) [81].

In contrast to previous work [88, 89], specificity of flavastacin for the N-terminus of aspartic acid could not be confirmed, neither for hLTF nor for BSA. However, it was demonstrated, that the sequential combination of trypsin and flavastacin for protein digestion successfully cleaves the *N*-glycoprotein hLTF

in well annotatable *N*-glycopeptide sequences. Interestingly, in contrast to unspecific digestion strategies using proteinase K or pronase, flavastacin works as an “*N*-glyco-specific” proteolytic enzyme (specific for *N*-glycosylated asparagine at the C-terminus). This property improves data quality as well as data analysis and therefore facilitates *N*-glycoproteomics significantly. However, the unspecific cleavage due to flavastacin at the N-terminus results in the distribution of redundant *N*-glycopeptide signals with peptide moieties of different length.

Overall, this finding improves the glycoproteomic toolbox and helps to overcome common problems in *N*-glycoproteomics, i.e. the presence of too large *N*-glycopeptides with too many AAs and/or too many *N*-glycosylation sites for LC-MS analysis. Despite the fact that this specificity of flavastacin and its cleaving mechanism need to be examined also for other glycoproteins (as well as for more complex (glyco-) protein mixtures like IAV glycoproteins (Chapter 4.3)) and in particular for other types of glycosylation (like high-mannose-type, hybrid-type and O-glycosylation), the use of flavastacin will already be beneficial for glycoscience now, as it allows researchers to dig faster and deeper into *N*-glycoproteomes.

4.2. Improvement of electrospray stability in negative ion mode for nano-PGC-LC-MS glycoanalysis via post-column make-up flow

This chapter is adapted from Nguyen-Khuong and Pralow et al. (2018) [192].

4.2.1. Introduction

With the ever-increasing interest to develop methods with higher sensitivity, there is significant interest in transitioning MS-based glycomics analysis to nano-scale instrument architecture (nano-LC and nano-ESI). This is due to the significantly lower flow rate and sample dilution within nano-LC columns and the superior ionization and sampling rate of nano-ESI [202, 203]. Unlike capillary flow-adapted source sprayers, the architecture of standard nano-source sprayers does not rely on a nebulizer to facilitate the desolvation of the liquid. Instead, electrospray efficiency is functionally dependent on the potential gradient between the electrospray tip and the orifice [202, 203].

In this section, an easy-to-use setup is described, which enables the analysis of reduced native glycans via nano-PGC-LC-ESI-MS(/MS) in negative ion mode. From our observations, periodical droplet forming attended by intermittent corona discharge events occurred at the nano-electrospray needle. This was particularly the case during periods of high aqueous buffer content of the nano-PGC-LC elution gradient. In an effort to improve the spray, an organic PCMF is introduced to enhance the electrospray stability and to improve the overall performance of the analyses.

4.2.2. Results

The motivation of this work is based on the observation that during the periods of high aqueous content of the standard nano-PGC-LC elution gradient, the spray stability was compromised. As depicted in Figure 4-4, instances of inconsistent nano-ESI occurred whereby periodic droplets formed on the spray needle associated with intermittent corona discharge events. These droplets dissipated away from the

orifice of the mass spectrometer, causing signal drop-outs in ionization- and sampling-rate, i.e. a breakdown of the ion current (see zoom-in Figure 4-5). As a result of the unstable nano-ESI, detection and fragmentation of the analytes was impaired. This is demonstrated in the BPC of pooled fetuin *N*- and *O*-glycans as displayed in Figure 4-5A. It is further emphasized in Figure 4-6A, where the EIC of all selected (glycan) ion traces are repeatedly interrupted by such signal drop-outs. These signal drop-outs were in sync with the formation of the droplets. In particular, during the early periods of the nano-LC gradient (with high aqueous content) the signal was deleteriously affected. This is demonstrated during the elution of small *O*-glycans at the beginning of the gradient (Figure 4-5A, EIC 675.25¹⁻ and EIC 966.34¹⁻). It should be emphasized that for both nano-PGC-LC-ESI-MS(/MS) setups (with and without PCMF), there was an efficient separation of the glycans under the given elution gradient, but without PCMF, the eluting glycans accumulated into the droplets growing at the needle tip. Thus, without PCMF, in addition to the spray instabilities, also the chromatographic separation was affected post-column. At lower extent, this occurred also during the later periods of the nano-LC gradient (with higher organic content; see Figure 4-6A, EIC 1111.39²⁻ and EIC 1439.51²⁻).

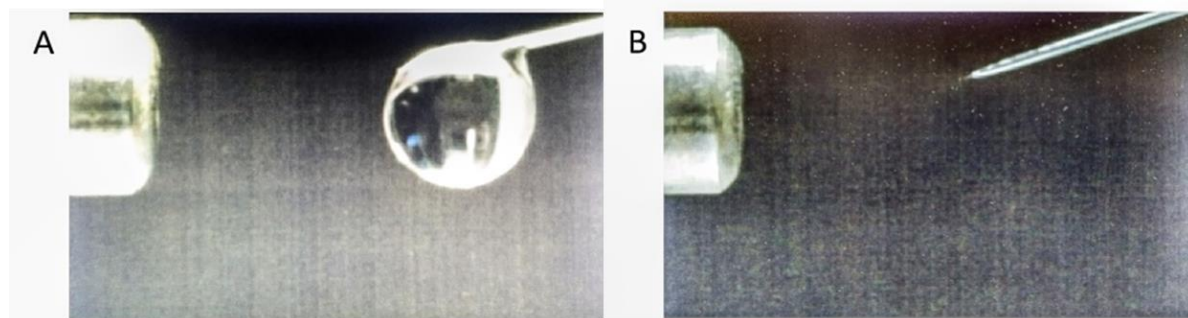


Figure 4-4: **Comparison of the spray stability during elution when using nano-PGC-LC-ESI-MS(/MS) in negative ion mode.** A) Image of the needle and the nano-source. A droplet forms and is released periodically. B) Due to the supplementation with a PCMF of 100% ACN, no droplet is forming anymore and a stable electro-spray is achieved. Adapted from Nguyen-Khuong and Pralow et al. (2018) [192].

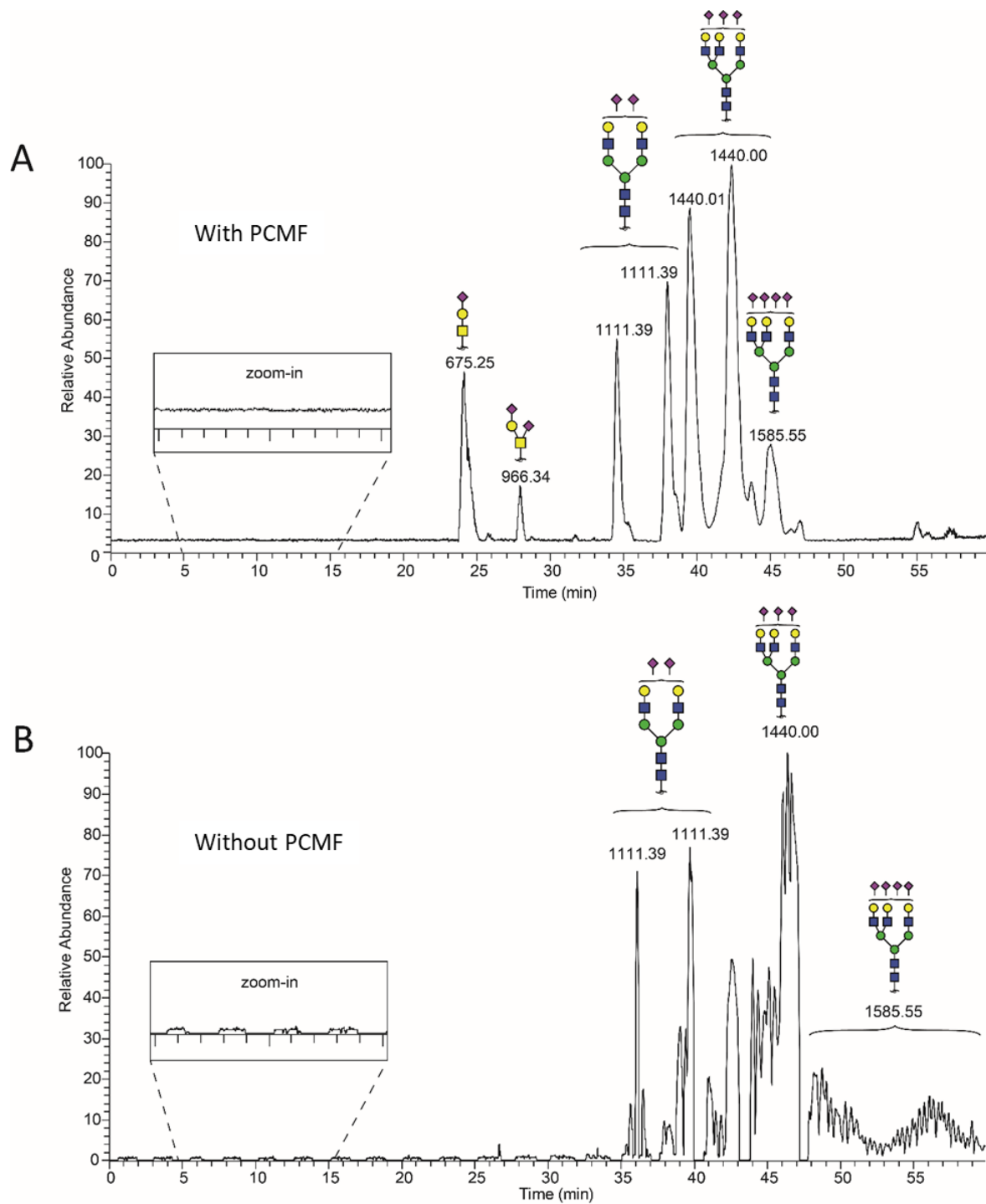


Figure 4-5: **Comparison of the BPC of eluted N- and O-glycans released from bovine fetuin with (A) and without (B) PCMF supplementation.** The recovery of smaller alditols such as O-glycans (like, $\text{HexNAc}_1\text{Hex}_1\text{NeuAc}_1$ and $\text{HexNAc}_1\text{Hex}_1\text{NeuAc}_2$) can be observed, while still the larger N-glycans (like, $\text{HexNAc}_4\text{Hex}_5\text{NeuAc}_2$, $\text{HexNAc}_5\text{Hex}_6\text{NeuAc}_3$, $\text{HexNAc}_5\text{Hex}_6\text{NeuAc}_4$) are recovered after the PCMF implementation, as well. Using PCMF, no breakdown of the BPC ion trace occurs anymore (zoom-in). Symbolic representations of N-glycan structures were drawn using GlycoWorkbench Version 1.1, following the Symbol Nomenclature for Graphical Representations of Glycans [6]. Adapted from Nguyen-Khuong and Pralow et al. (2018) [192].

To decrease the aqueous content throughout the analysis without compromising the chromatographic separation of the glycans, a PCMF was implemented in between column and sprayer, as explained in the Materials and Methods section. Three different capillary dimensions (20, 50 and 75 μm) were tested as the capillary to deliver the flow from the T-piece to the electrosprayer. It was observed, that using the 20 μm and 50 μm ID capillaries for the outlet (Figure 4-4) would still produce signal breakdowns caused by the formation of a droplet that would not desolvate properly (data not shown). A 75 μm ID capillary was the smallest capillary which would allow for a proper mixing of the column flow and the PCMF prior to electrospray. The slightly bigger ID of the electropolished stainless steel spray needle (30 μm , for capillary and tip) compared to the ID of common fused silica needles (20 μm for capillary and 10 μm at the tip) fitted better to the 3-fold increased flow rate.

With the application of a PCMF of neat organic solvent (100% ACN), the spray was stable throughout the entire nano-LC run (see Figure 4-4B), i.e. from high aqueous to high organic content of the gradient (Figure 3B & zoom-in). With the application of the PCMF, a better signal stability and better peak shapes and hence better BPC and EIC were achieved (see Figure 4-5B & Figure 4-6B).

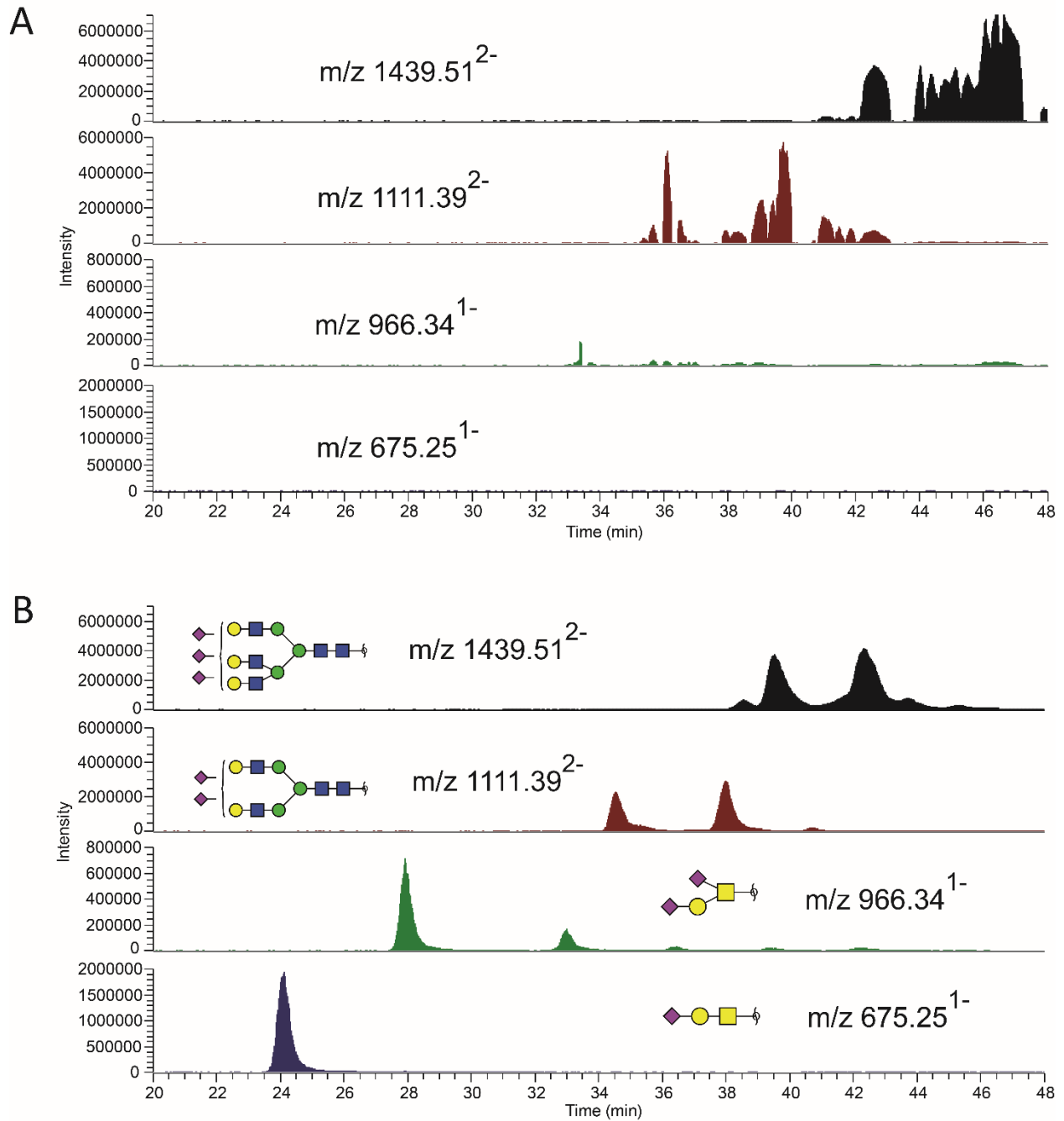


Figure 4-6: **Comparison of the EIC of selected masses corresponding to N- and O-glycans released from bovine fetuin before (A) and after (B) PCMF supplementation.** With PCMF, the EIC ion traces of the different N- (m/z 1439.51²⁻ and 1111.39²⁻) and O-glycans (966.34¹⁻ and 675.25¹⁻) are significantly improved with good nano-PGC-LC separation of their isoforms. Adapted from Nguyen-Khuong and Pralow et al. (2018) [192].

Table 4-3 reports the glycan compositions, their spectra copy numbers and tR, which were observed and confirmed via MS/MS interrogation with and without PCMF. Without PCMF not a single spectrum could be counted for *O*-glycans, eluting at early time points of the gradient when the aqueous content was high. For the *N*-glycans, eluting at later time points of the gradient with lower aqueous content, a significantly lower spectral copy number was obtained compared to the results achieved with PCMF. In fact, with PCMF, the spectral copy numbers of the glycans were improved by a factor of at least two for the majority of the fetuin glycans. Taken with the results in Table 4-3, one can summarize that at nano-flow, the detection of glycans, particularly those at low abundance, as well as the resolution of such peaks is severely compromised without using PCMF.

RESULTS AND DISCUSSION

Table 4-3: *N- and O-glycan compositions released from bovine fetuin detected via nano-PGC-LC-ESI-MS(/MS) with and without the use of a PCMF. Comparison of the spectra copy numbers (count of MS/MS fragment spectra of each precursor (glycan) ion $[M-2H]^2$) corresponding to the glycan compositions found between the two analytical measurements. Note that there is no spectral copy number for the O-glycan 675.25⁺, because of the missing doubly charged precursor. Adapted from Nguyen-Khuong and Pralow et al. (2018) [192].*

N-glycan compositions	spectra copy numbers		[M-H]-	MS/MS confirmed	tR [min] (with PCMF)
	NO PCMF	PCMF			
(Hex)3 (HexNAc)3 (NeuAc)3 + (Man)3(GlcNAc)2	57	299	2880.0205	YES	38.59, 39.53, 42.35, 43.67
(Hex)2 (HexNAc)2 (NeuAc)2 + (Man)3(GlcNAc)2	52	156	2223.7788	YES	34.57, 37.99, 40.73
(Hex)3 (HexNAc)3 (NeuAc)2 + (Man)3(GlcNAc)2	33	70	2588.91724	YES	36.52, 38.24, 39.88
(Hex)2 (HexNAc)2 (NeuAc)3 + (Man)3(GlcNAc)2	20	41	2514.86639	YES	35.48, 36.92, 38.54, 42.03
(Hex)3 (HexNAc)3 (Deoxyhexose)1 (NeuAc)2 + (Man)3(GlcNAc)2	18	33	2734.97094	YES	36.27, 37.63, 38.98, 40.51
(Hex)2 (HexNAc)2 (Deoxyhexose)1 (NeuAc)2 + (Man)3(GlcNAc)2	15	19	2369.83618	YES	34.54, 36.47, 39.83, 45.93
(Hex)2 (HexNAc)2 (NeuAc)1 + (Man)3(GlcNAc)2	13	28	1932.68396	YES	32.84, 36.42, 39.36, 42.11
(Hex)2 (HexNAc)3 (NeuAc)2 + (Man)3(GlcNAc)2	6	5	2426.86084	YES	37.8
(Hex)2 (HexNAc)2 (Deoxyhexose)1 (NeuAc)1 + (Man)3(GlcNAc)2	3	2	2078.73754	YES	32.32, 34.99, 38.26
(Hex)4 (HexNAc)2 (Deoxyhexose)1 + (Man)3(GlcNAc)2	2	5	2111.7461	YES	36.9, 41.75, 45.02
(Hex)3 (HexNAc)3 (NeuAc)4 + (Man)3(GlcNAc)2	1	232	3171.07884	YES	41.72, 44.6, 46.64
(Hex)3 (HexNAc)3 (Deoxyhexose)1 (NeuAc)3 + (Man)3(GlcNAc)2	-	31	3026.06762	YES	40.98, 43.97
(Hex)2 (HexNAc)3 (Deoxyhexose)1 (NeuAc)1 + (Man)3(GlcNAc)2	-	8	2281.73242	YES	34.46, 37.96, 44.8, 51.27
(Hex)1 (HexNAc)2 (Deoxyhexose)1 + (Man)3(GlcNAc)2	-	7	1625.59606	YES	33.13
(Hex)1 (HexNAc)1 (NeuAc)1 + (Man)3(GlcNAc)2	-	6	1567.55468	YES	31.97, 35.43
(Hex)2 (HexNAc)2 (Deoxyhexose)1 + (Man)3(GlcNAc)2	-	6	1787.64942	YES	33.97
(HexNAc)2 (Deoxyhexose)1 + (Man)3(GlcNAc)2	-	5	1463.54456	YES	30.95, 32.25
(Hex)1 (HexNAc)2 (NeuAc)1 + (Man)3(GlcNAc)2	-	2	1770.63586	YES	32.26, 35.67
(Hex)3 (HexNAc)2 (Deoxyhexose)1 + (Man)3(GlcNAc)2	-	2	1949.70092	YES	32.87, 35.45
(Hex)2 (HexNAc)2 + (Man)3(GlcNAc)2	-	1	1641.5879	YES	31.7
O-glycan compositions	spectra copy numbers		[M-H]-	MS/MS confirmed	tR [min] (with PCMF)
	NO PCMF	PCMF			
(Hex)1 (HexNAc)1 (NeuAc)2	-	17	966.35	YES	27.93
(Hex)1 (HexNAc)1 (NeuAc)1	-	-	675.25	NO	24.12

4.2.3. Discussion

The identification of glycans using nano-PGC-LC-ESI-MS(/MS) with subsequent detection of the mass and fragmentation in negative ion mode is a powerful glycoanalytical tool. This is centered around the ability of the technique to not only allow to characterize both *N*- and *O*-glycans, but also to separate isobaric structures using PGC and to detect native sialylated and neutral structures in one run. Negative ion polarity detection of fragment ions is diagnostic for certain types of branching, linkage and thus central to identifying glyco-epitopes. Central to the detection of glycans in negative polarity is the use of basic buffers, which deprotonate the analytes and increase their propensity for MS-detection in negative ion mode. This methodology for glycan analysis has already been described in detail, for analytical- [204] and capillary- [40, 65, 76, 205, 206], but not much for nano-scale PGC-LC-ESI-MS(/MS) [43].

For analytical- and capillary-scale LC-MS, a coaxial addition of sheath gas and/or liquid is standard, as the gas assists nebulization and evaporation/desolvation of the droplets and the added organic liquid reduces their surface tension [207-211]. For nano-scale LC-MS, usually flow rates and spray-tip diameters are much smaller, hence no sheath gas/liquid is required to form a stable nano-electrospray [212-216].

For nano-PGC-LC, the only commercially available columns are from Thermo Fisher Scientific (see Materials and Methods). Compared to common reversed-phase nano-LC columns with a typical maximum flow rate of 0.3 μL , the nano-PGC-LC columns are usually operated at about a three times higher flowrate. When coupling these columns at 0.9 $\mu\text{L}/\text{min}$ to MS in negative ion mode with the nano-spray source from Thermo Fisher Scientific (see Materials and Methods), repeating spray instability was observed, which is described to occur at higher propensity in negative ion mode [217, 218]. As illustrated in the results section, the ion traces of the analytes are detrimentally compromised without augmentation of the spray with an appropriate organic solvent. Most significant improvement of the

electrospray was reached during the early gradient periods of high aqueous content (Figure 4-5 and Figure 4-6).

This sheath/make-up flow ESI-setup is commonly used in analytical and capillary scale LC-MS solutions, yet to our knowledge has not been applied to a nano-PGC-LC-ESI-MS(/MS) approach, operated in negative ion mode. A similar solution to improve ESI for nano-scale LC-MS, is for example, the CaptiveSpray nanoBooster™ (Bruker Daltonics, Germany) that is also used for glycomic investigations with nano-PGC-LC-ESI-MS(/MS) in negative ion mode [43]. Staples et al. (2010) has described another solution, but for a nano-Chip-based format applied to HILIC-ESI-MS(/MS), analyzing glycosaminoglycans [219]. Ni et al. (2013) used a similar approach using fluorid-mediated negative ion-mode microchip PGC-LC-MS to analyze released *N*-glycans [220].

The demonstrated approach can be easily applied to any conventional nano-U/HPLC-ESI-MS(/MS) setup with an additional (syringe) pump, a T-piece and a respective spray needle, adjusted to the increased flow rate. The approach is modular and thus allows the user to adjust and replace any part as well as to individually optimize the setup to the given instrumentation. Based on this pilot study, further work is necessary to fully elaborate the robustness and sensitivity of the approach. Following the trend within the analytical community towards more sensitive nano-flow and nano-spray LC-MS technologies, the use of complementary organic solvents for PCMF might be useful to any analyte class that requires negative ion detection, for example (oligo-)nucleotides, metabolites or (glyco-)lipids.

4.3. Comprehensive *N*-glycosylation analysis of the IAV proteins HA and NA from adherent and suspension MDCK cells

This chapter is adapted from Pralow et al. (2021) [160].

4.3.1. Introduction

In this section, a comprehensive MS-based workflow was established, i.e. a combination of glycoproteomic and glycomic data allowing to link compositions to possible glycan structures. This enabled, in some cases, a site-specific structural analysis of glycoproteins. For glycomic analysis, nano-PGC-LC-MS(/MS) equipped with a PCMF (Chapter 4.2) was used. This method can ascertain information-rich data as it makes use of information derived from the chromatographic separation (which even allows separation of isobaric *N*- and *O*-glycans on the PGC column [76]) and information derived from MS/MS fragmentation patterns acquired in negative ion mode (unique and characteristic patterns that provide structural glycan information). The latter contain cross-ring fragments, which can be used to confirm the topology, branching and monosaccharide linkages of a glycan [72, 74, 75]. *N*-glycopeptides (derived from digestion with trypsin and flavastacin (Chapter 4.1)) were measured on a nano reversed-phase LC (nano-RP-LC) coupled to MS/MS. Data acquisition was followed by manual and semi-automated analysis of HCD-generated *N*-glycopeptide fragment ion spectra using glyXtool^{MS} [83], an in-house-developed glycopeptide analysis software and the search engine Byonic [113].

This workflow was challenged by analyzing both, the “fine-structure of the whole IAV *N*-glycome” (*N*-glycomics) and the “site-specific glycan compositions of the IAV glycoproteins” (*N*-glycoproteomics, mainly HA and NA) propagated in two closely related cell lines – MDCK.ADH, an adherent cell line and MDCK.SUS2, a suspension cell line derived thereof. Furthermore, we investigated whether adaptation to

growth in suspension would impact the glycan pattern of the viral envelope proteins HA and NA regarding alterations of *N*-glycan micro- and macroheterogeneity.

4.3.2. Results

In Figure 3-1, the established workflow is illustrated. Virus harvests were purified via *g*-force step-gradient centrifugation to obtain whole virus particle fractions without contaminating cell debris as described previously [36, 195]. In the next step, the virus was inactivated using SDS and the protein concentration is determined. For glycomic analysis (bottom, Figure 3-1), the *N*-glycans were released using PNGase F, followed by reduction of the reducing end of the *N*-glycans and clean-up via ion-exchange-SPE. Furthermore, *N*-glycans were enriched via PGC-SPE and analyzed using nano-PGC-LC-MS(/MS) equipped with a PCMF [192] followed by manual data analysis. For glycoproteomic analysis (top, Figure 3-1), viral proteins were digested using trypsin and flavastacin [81] followed by glycopeptide enrichment using hydrophilic interaction chromatography (HILIC)-SPE [97]. Subsequently, glycopeptides were measured using nano-RP-LC-MS(/MS) [97] and analyzed manually as well as semi-automatically.

4.3.2.1. Characterization of influenza *N*-glycans fine-structure

First, a list of IAV protein *N*-glycan structures of all samples was established to compare the *N*-glycome of IAV proteins propagated in MDCK.ADH and MDCK.SUS2 cells. In a next step, this list was used for subsequent *N*-glycoproteomic analysis. Analysis of the viral *N*-glycome was performed using nano-PGC-LC-MS(/MS) [40, 43, 72]. Here, nano-PGC-LC-MS(/MS) was used for the analysis of *N*-glycans derived from virus glycoproteins to confirm rich data sets derived from chromatographic separation (separation of isobaric glycans) [76] and MS/MS fragmentation (i.e. specific cross-ring fragment ions) [72, 74]. Please note, that despite virus particle purification and concentration, detection of a small amount of glycans from host cell glycoproteins was to be expected, due to attachment and/or incorporation of host cell proteins into the viral membrane [221].

To demonstrate the performance of this method, exemplarily, isomeric *N*-glycans that were found for virus particles expressed in both cell lines are displayed in Figure 4-7. To facilitate understanding and assignments in figures and tables, *N*-glycan structures are labelled using a unique glycan-composition ID in Supplementary Table A2-1). The EIC of the *N*-glycan (Hex)₂ (HexNAc)₃ (dHex)₁ + (Man)₃(GlcNAc)₂ (m/z 994.86²) indicated the presence of four peaks - each an isomeric structure: (i) core fucosylated, galactosylated biantennary with a bisecting GlcNAc (30 min, C21), (ii) galactosylated, biantennary with an outer arm Fuc (3-arm), terminated with an GalNAc (3-arm) (31 min, C22), (iii) core fucosylated, alpha-galactosylated, triantennary glycan (35 min, C23) and (iv) core fucosylated, galactosylated, triantennary glycan (35.5 min, C24). In this example, C21, C22 and C23 were detected in both samples (Figure 4-7A and B). The fourth structure (C24), however, was only detected for virions produced in MDCK.SUS2 cells (Figure 4-7B). Using only compositional or partly structural information obtained from conventional glycomic measurement methods, i.e. HILIC with fluorescence detection optional coupled to MS (HILIC-FLD-MS) and MALDI-TOF-MS, this qualitative difference would have been missed. Only the combination of a suitable *N*-glycan separation method that includes isobaric structures and the online detection of fine-structural information using nano-PGC-LC-MS(/MS), allowed such a fine-structural *N*-glycome analysis in a single measurement.

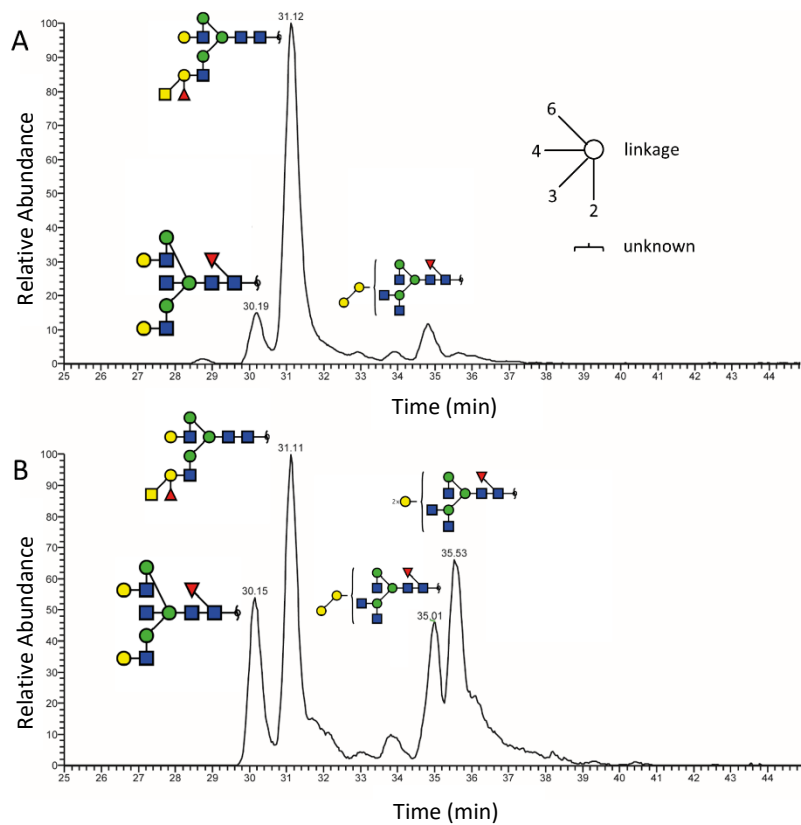


Figure 4-7: **Separation of *N*-glycan isomers using nano-PGC-LC-MS(MS)**. EIC of the *N*-glycan composition (Hex)₂(HexNAc)₃(dHex)₁ + (Man)₃(GlcNAc)₂ (m/z 994.86²⁻) derived from IAV propagated in (A) MDCK.ADH and (B) MDCK.SUS2 cells. Symbolic representations of *N*-glycan structures were drawn with GlycoWorkbench Version 1.1, following the Symbol Nomenclature for Graphical Representations of Glycans [6]. Linkage legend is given in the upper right corner. Adapted from Pralow et al. (2021) [160].

For the annotation of the fine-structure of the detected *N*-glycans, the identity of isomers was confirmed by manual annotation of their MS/MS fragment ion spectra. Specific (cross-ring-) fragment ions were used to confirm the fine-structure of the annotated *N*-glycan isomers from Figure 4-7 as illustrated exemplarily in Figure 4-8. In this instance, a combination of different fragment ions found in the MS/MS fragment ion spectra were used to identify e.g. the presence of either core or antenna fucosylation, the presence of Gal- α -(1,3)-Gal (Galili)-epitopes or bisecting structures. For example, the fragment ion m/z 350¹⁻ can be found in the core-fucosylated *N*-glycan structures depicted in Figure 4-8 (A, C and D), but is missing in the antenna-fucosylated structure depicted in Figure 4-8 (B), thus

serving as a diagnostic fragment ion to identify core fucosylation. The detection of (cross-ring-) fragments also facilitated identification of the architecture of glyco-epitopes, such as the combination of the fragment ions m/z 383¹⁻ and 586¹⁻, which together indicated the presence of a Galili-epitope (Figure 4-8C). Furthermore, structural differentiation of the *N*-glycan C22 was achieved due to the observation of the specific diagnostic fragment ion for GalNAc-Gal(Fuc)-GlcNac (m/z 773¹⁻, Figure 4-8B). In addition, antenna fucosylation of the 3-arm Gal was identified using the fragment ions m/z 325¹⁻, 570¹⁻ and 773¹⁻ (Figure 4-8B). In addition, the D-ion could be used to distinguish the arm specificity of the glyco-epitope. For example, the 688¹⁻ ion enabled to distinguish the presence of a terminal Gal on the 6-arm of the C22 glycan, as opposed to the 3-arm epitope GalNAc-Gal(Fuc)-GlcNac (Figure 4-8B).

RESULTS AND DISCUSSION

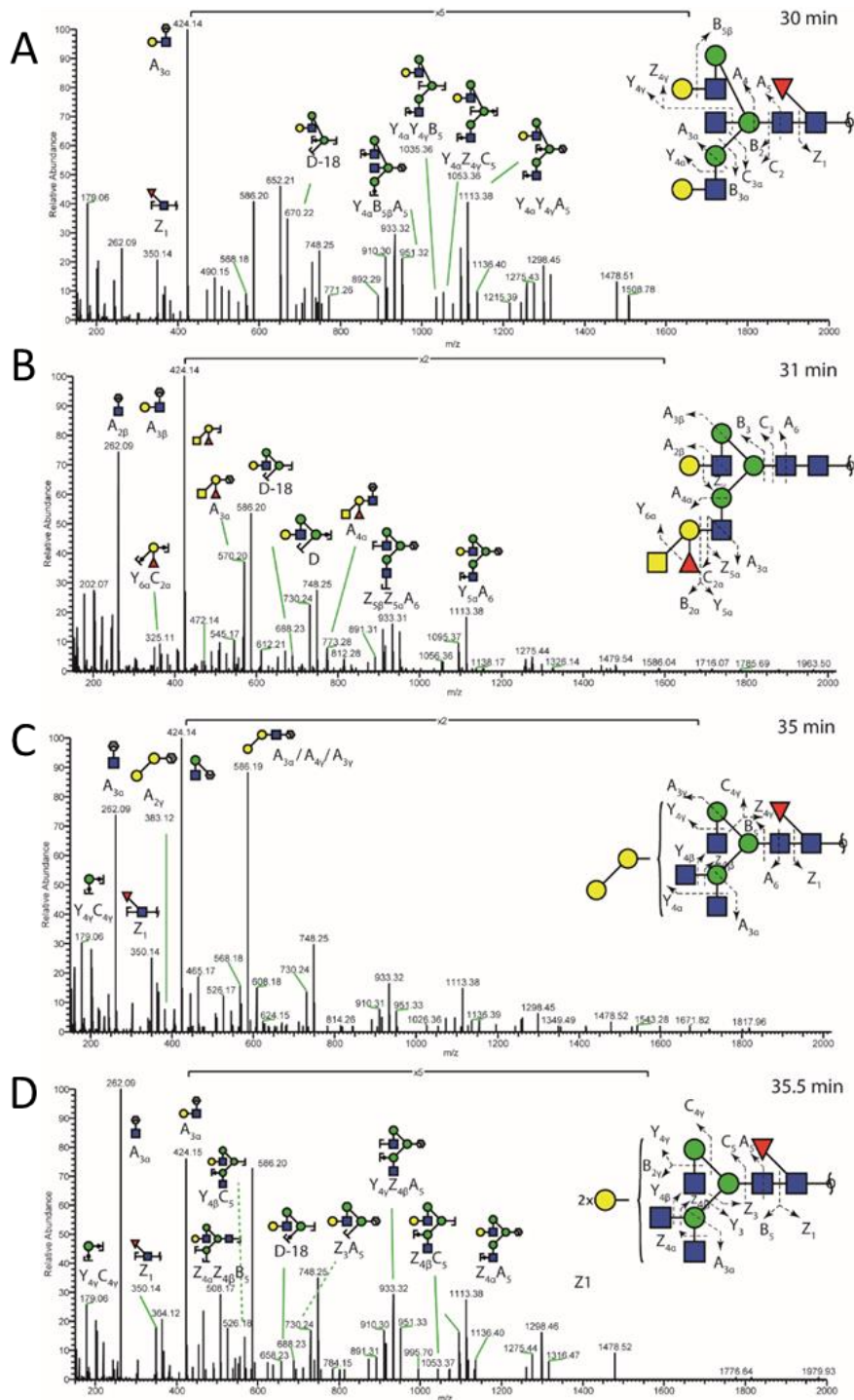


Figure 4-8: **Negative ion mode HCD fragmentation of different N-glycan isomers.** MS/MS fragment ion spectra of different N-glycan isomers with the mutual glycan composition (Hex)₂ (HexNAc)₃ (dHex)₁ + (Man)₃(GlcNAc)₂ (m/z 994.86²). Diagnostic cross-ring fragment ions for the identification of specific N-glycan structures are illustrated. Symbolic representations of N-glycan structures were drawn with GlycoWorkbench Version 1.1, following the Symbol Nomenclature for Graphical Representations of Glycans [6]. Adapted from Pralow et al. (2021) [160].

All identified *N*-glycan structures of the IAV *N*-glycome propagated in MDCK.ADH and MDCK.SUS2 cells (all replicates) are listed in Supplementary Table A2-1. In total, 85 unique *N*-glycan structures from IAV glycoprotein samples were identified based on tR and MS/MS fragment ion spectra analysis. Nine of these structures did not have sufficient MS/MS information for structural annotation, therefore only compositions are reported in this case.

There were 62 unique *N*-glycans detected in samples from IAV glycoproteins propagated in MDCK.ADH cells and 70 in MDCK.SUS2 cells. All *N*-glycans detected were neutral, i.e. no sialylated structures were found. The majority of *N*-glycans of both cell lines were complex-type with and without core- and antenna-fucosylation. Interestingly, antenna fucosylation was mostly identified to be the glyco-epitope Gal(Fuc)-GlcNAc (blood group H type 2, identified via the diagnostic fragment ions 409¹⁻ and 427¹⁻ [41, 64, 72]). In a lower abundance, also the glyco-epitope GalNAc-Gal(Fuc)-GlcNAc (blood group A, identified by the diagnostic fragment ion *m/z* 773¹⁻ [41, 72]), was detected in IAV derived from both cell lines. In particular, the latter was found with a high relative abundance (Figure 4-9) and frequency (Supplementary Table A2-1), especially in samples derived from MDCK.ADH cells. Apart from this, many complex-type *N*-glycans with terminal alpha-galactosylation could be identified in IAV derived from both cell lines (fragment ions *m/z* 341¹⁻, 383¹⁻ and 586¹⁻ [41, 222]), with a slightly higher occurrence in MDCK.ADH cells. Glycans eluting at earlier tRs are annotated as complex-type *N*-glycans carrying a bisecting GlcNAc, again with a relative higher abundance in samples derived from MDCK.ADH cells. The major structure identified in all the samples was the core-fucosylated three antennary complex-type *N*-glycan C31, with a slightly higher abundance in samples derived from MDCK.SUS2 cells (Figure 4-9, tR 36-37 min). Besides complex-type *N*-glycans, also hybrid-type *N*-glycans were detected, partly carrying core-/antenna-fucosylation (blood group H type 2) but also with the blood group A glyco-epitope GalNAc-Gal(Fuc)-GlcNAc (Supplementary Table A2-1, Figure 4-7, Figure 4-8 & Figure 4-9). Oligomannose-

type *N*-glycans were found in all samples, but with a much higher relative abundance in samples derived from MDCK.ADH cells (Figure 4-9).

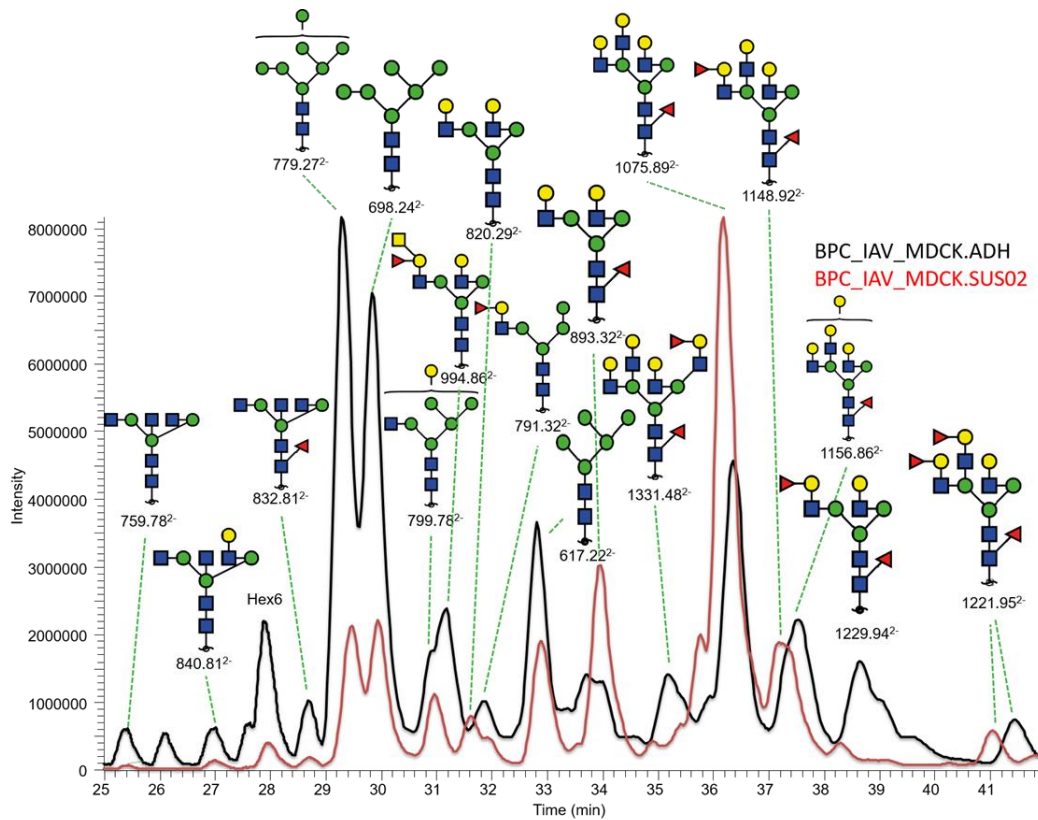


Figure 4-9: **Comparison of the most abundant *N*-glycan structures.** Overlay of the base peak chromatogram (BPC) of the nano-PGC-LC-MS(/MS) measurements of *N*-glycans derived from IAV propagated in MDCK.ADH (black) and MDCK.SUS2 cells (red). The *t*R frame from 25-42 min is illustrated. The highest abundant *N*-glycan structure and the corresponding double negatively charged molecular mass under each peak is depicted. Hex6 marks a signal from a polymer. Symbolic representations of *N*-glycan structures were drawn with GlycoWorkbench Version 1.1, following the Symbol Nomenclature for Graphical Representations of Glycans [6]. Adapted from Pralow et al. (2021) [160].

Overall, the results show a similar *N*-glycan pattern of the two surface proteins HA and NA from IAV propagated in MDCK.ADH and MDCK.SUS2 cells. However, *N*-glycan analysis of both samples via xCGE-LIF demonstrated that IAV samples derived from MDCK.ADH cells had a higher relative abundance of signals at higher MTU'' (referring to larger *N*-glycans). In contrast, IAV samples derived from MDCK.SUS2 had a higher relative abundance of signals referring to smaller *N*-glycan structures (Supplementary Figure A2-1).

4.3.2.2. Characterization of influenza glycopeptides

Besides a structural elucidation of the IAV *N*-glycome, a site-specific glycoproteomic analysis of HA and NA from IAV propagated in MDCK.ADH and MDCK.SUS2 cells was performed. For in-depth characterization of the micro- and macroheterogeneity of both proteins the in-house developed software glyXtool^{MS} was employed.

Figure 4-10 shows an MS/MS fragment ion spectrum of an *N*-glycopeptide derived from HA with the precursor mass m/z 1036.4249³⁻ after a sequential digestion using trypsin and flavastacin [81], respectively. The putative peptide mass m/z 973.4782¹⁻ was identified based on the annotated conserved fragment ion pattern (Hoffmann et al. 2018) [97]. The peptide sequence was verified by annotation of specific b- and y-ions. Furthermore, glycan moiety-derived oxonium ions (B-ions) and neutral-loss fragment ions (Y-ions) were annotated (nomenclature Domon and Costello [223]). The mass of the *N*-glycan moiety was determined (neutral loss of 2133.7965 Da) and allowed the annotation of the putative *N*-glycan composition (Hex)₃ (HexNAc)₃ (dHex)₁ + (Man)₃(GlcNAc)₂.

RESULTS AND DISCUSSION

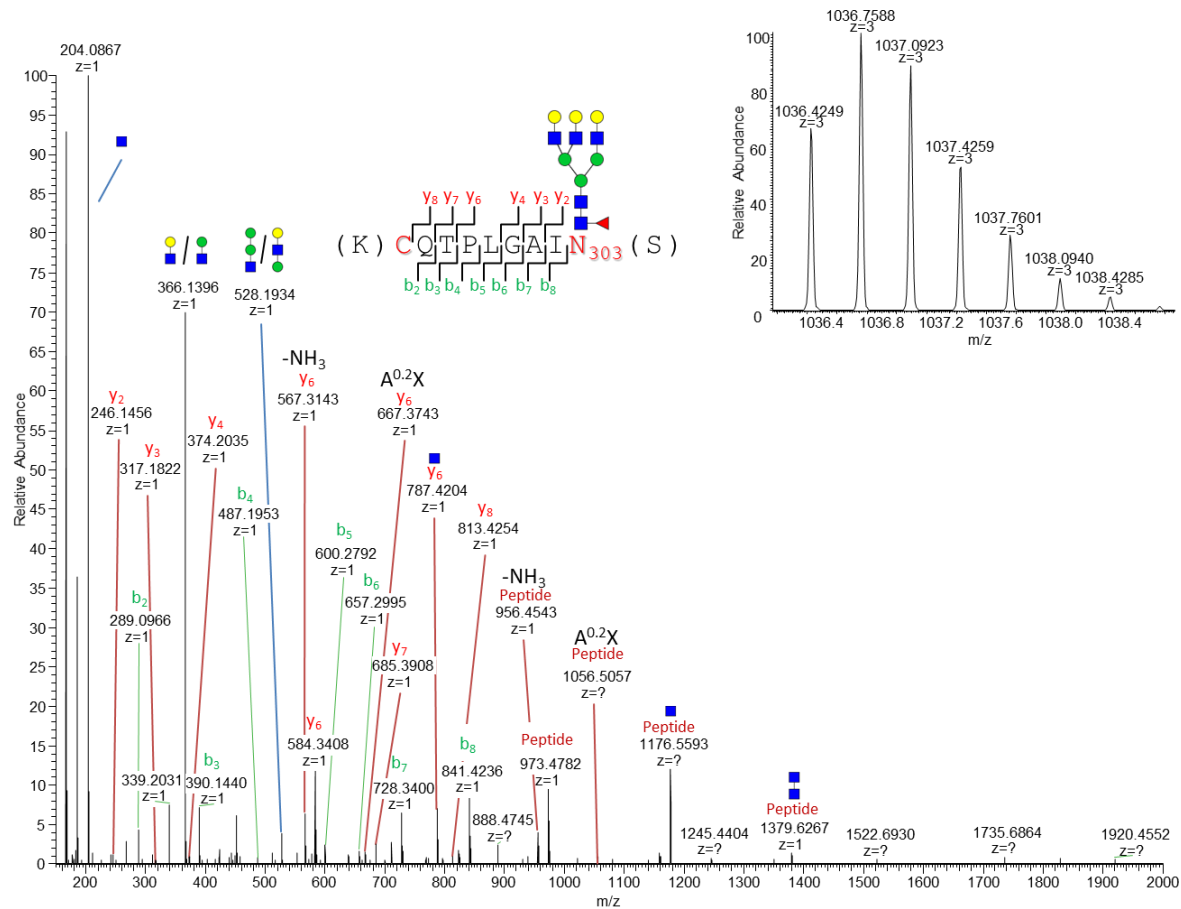


Figure 4-10: **Site-specific glycopeptide analysis using nano-RP-LC-MS(/MS)**. Fragment ion spectrum of a HA N-glycopeptide (from IAV propagated in MDCK.SUS2 cells), derived from a sequential proteolytic digestion using trypsin and flavastacin. The N-glycopeptide sequence is (K)CQTPLGAIN₃₀₃(S) with the linked N-glycan composition Hex₃HexNAc₃dHex₁+Man₃GlcNAc₂ (precursor mass: 1036.4249³⁺). Amino acids highlighted in red indicate the carbamidomethylation of cysteine and the N-glycosylation of asparagine. For the validation of the (glyco-)peptide sequence, specific b- (green) and y- (red), as well as B-(oxonium-, blue) and Y-ions (peptide + glycan moiety) are annotated. The isotopic pattern of the precursor is located at the upper right corner. Symbolic representations of N-glycan structures were drawn with GlycoWorkbench Version 1.1, following the Symbol Nomenclature for Graphical Representations of Glycans [6]. Adapted from Pralow et al. (2021) [160].

A list of all identified peptide moieties for each N-glycosylation site of HA and NA is given in Table 4-4.

While the diversity of peptide moieties was rather low, we observed the occurrence of N-terminal carbamidomethylation of peptides derived from HA as well as NA predominantly from IAV propagated in MDCK.SUS2 cells. Using solely tryptic digestion resulted in the identification of N-glycopeptides only from HA2 at the site N497 (not shown). In contrast, the sequential digestion with trypsin and flavastacin enabled the identification of nearly all potential N-glycosylation sites of HA (N27/28, N40, N285, N303)

and one site from NA (N73) (Table 4-4). However, site N497 was not detected. Please note that all detected *N*-glycopeptide fragment-ion spectra are related to HA or NA. No host cell-related glycopeptides were detected, except one *O*-glycopeptide belonging to bovine fetuin (medium component).

Table 4-4: **Identified (glyco)-peptide sequences of IAV surface glycoproteins hemagglutinin and neuraminidase.** For each *N*-glycosylation site, all identified (glyco)-peptide-moieties are listed. Furthermore, modified amino acids are underlined and the *N*-glycosylation site is highlighted in bold. Adapted from Pralow et al. (2021) [160].

Protein	Site	Peptide Sequence	Modification	Spectrum
Hemagglutinin	N27/28	(A)DTIC <u>I</u> GYHANN(S)	(C) Carbamidomethylation	Supplementary Fig. A2-2
		(A) <u>D</u> TIC <u>I</u> GYHANN(S)	(C) Carbamidomethylation (D) Carbamidomethylation	
	N40	(N)STDTVDTVLEKN(V)		Supplementary Fig. A2-3
	N285	(R)GFGSGIITSN(A)		Supplementary Fig. A2-4
		(R) <u>G</u> FSGIITSN(A) (G)SGIITSN(A)	(N-terminal) Carbamidomethylation	
N303	(K) <u>C</u> QTPLGAIN(S)	(C) Carbamidomethylation	Figure 4-10	
	(K) <u>C</u> QTPLGAIN(S)	(C) Carbamidomethylation (Q) NH ₃ loss		
	(Q)TPLGAIN(S)			
N497	(R)NGTYDYPK(Y)		Supplementary Fig. A2-5	
Neuraminidase	N73	(K)DTTSVILTGN(S)		Supplementary Fig. A2-6
		(K) <u>D</u> TTSVILTGN(S)	(D) Carbamidomethylation	

In Supplementary Table A2-2, all *N*-glycopeptides detected for HA and NA are listed. All detected site-specific *N*-glycan compositions were putative complex- or hybrid-type *N*-glycan structures (see *N*-glycan

analysis). The glycoproteomic analyses of the biological replicates showed high similarity in their base-peak chromatograms (BPC), with only moderate differences in the relative abundance of selected peaks. IAV propagated in MDCK.ADH showed even less variation in relative peak abundance (Supplementary Figure A2-7) compared to IAV propagated in MDCK.SUS2 (Supplementary Figure A2-8). *N*-glycopeptides were only considered if they were verified on MS as well as MS/MS level; missing MS/MS information for low-abundant *N*-glycopeptides can explain small discrepancies between the biological replicates.

In Figure 4-11, the structure of HA, its potential glycosylation sites, as well as the *N*-glycan compositions detected is visualized. As mentioned before, HA is the major antigen of IAV and each monomer comprises the two domains, HA1 and HA2. Whereby HA1 forms mainly the head region with the receptor binding site and HA2 forms the stem region (Figure 4-11). The HA1 of IAV strain PR/8/34 (H1N1) has five potential *N*-glycosylation sites (N27/28, N40, N285 and N303), all located at the stem region. Starting from the bottom of the stem region, N27/28 in IAV samples derived from MDCK.SUS2 and MDCK.ADH cells carried potentially three- and four-antennary complex-type *N*-glycans, or *N*-glycans with LacNAc extensions (based on the structural *N*-glycan analysis shown before). In addition, IAV propagated in MDCK.SUS2 cells displayed some compositions referring to di-antennary *N*-glycan structures. Furthermore, IAV from MDCK.SUS2 cells had a remarkable high amount of the (Hex)₃ (HexNAc)₃ (dHex)₁ + (Man)₃(GlcNAc)₂ composition at N27/28. All structures were exclusively fucosylated (core- and/or antenna-fucosylated). To some extent, possible alpha-galactosylated structures could be detected (higher variability in IAV from MDCK.ADH cells). Please note that the vicinity of the two *N*-glycosylation sites N27 and N28 does typically not allow to specifically link an *N*-glycan moiety to one or the other site. Therefore, both sites were referred to as a single locus. Nevertheless, due to the specific cleavage behavior of the proteolytic enzyme flavastacin, our results suggest N28 to be exclusively *N*-glycosylated but N27 not carrying any *N*-glycans. For N40, we could not

identify any *N*-glycosylation in IAV from MDCK.ADH cells and only one *N*-glycan composition (Hex)₃ (HexNAc)₃ (dHex)₁ + (Man)₃(GlcNAc)₂ in IAV from MDCK.SUS2 cells. Compared to the high microheterogeneity of the other *N*-glycosylation sites of HA1 (Figure 4-11), this finding might indicate a poor accessibility of N40 in the three dimensional space of HA. At the site N303 tri-, as well as tetra-antennary structures (or structures with LacNAc extensions) dominated. IAV derived from MDCK.SUS2 cells also displayed various compositions referring to di-antennary *N*-glycan structures. All corresponding structures carried an antenna and/or core-fucosylation and to some extent Galili-epitopes (especially in IAV from MDCK.ADH cells). The closest site to the HA1 head region, N285, carried similar *N*-glycan compositions to N303. For IAV derived from MDCK.ADH cells the variability of different *N*-glycan compositions, especially potentially alpha-galactosylated complex type (or hybrid-type structures based on the aforementioned structural *N*-glycan analysis) *N*-glycans could be observed. In IAV derived from MDCK.SUS2 cells, a higher variability of potential di-antennary *N*-glycan structures was found.

HA2 has only one potential *N*-glycosylation site at N497. Our analysis revealed only the composition (Hex)₃ (HexNAc)₄ (dHex)₁ + (Man)₃(GlcNAc)₂ for IAV from MDCK.SUS2 cells and the compositions (Hex)₃ (HexNAc)₃ (dHex)₁ + (Man)₃(GlcNAc)₂ and (Hex)₃ (HexNAc)₃ (dHex)₂ + (Man)₃(GlcNAc)₂ for IAV from MDCK.ADH cells.

Overall, *N*-glycopeptide analysis of HA propagated in MDCK.SUS2 and MDCK.ADH cells showed an increased variability, as well as complexity of *N*-glycosylation at sites closer to the head region of the antigen. MDCK.ADH cells seem to produce IAV particles with HA displaying a higher variability and complexity of *N*-glycans as MDCK.SUS2 cells that is in accordance with Hämmerling et al. 2017 [186]. In general, all *N*-glycan structures exhibited antenna and/or core fucosylation. In IAV from MDCK.ADH cells, many observed compositions might refer to alpha-galactosylated structures.

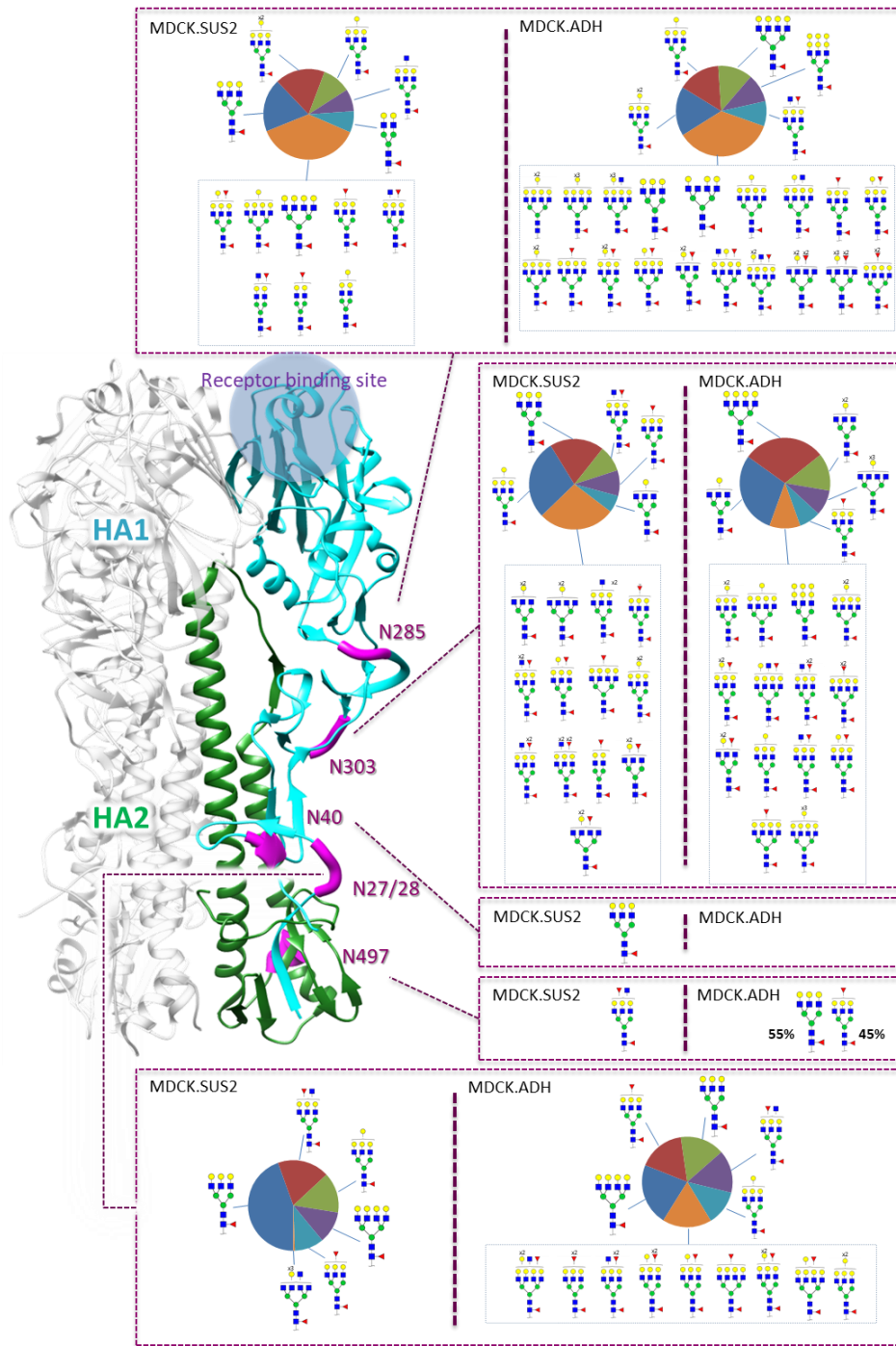


Figure 4-11: **Graphical illustration of the trimeric IAV antigen HA.** The molecular structure of HA was modeled using the PDB entry number 1RU7. For model processing and design the open source software UCSF Chimera Version 1.10.2 was utilized. HA1 (cyan) and HA2 (green) of an HA monomer are highlighted. N-glycosylation consensus sequences are shown in magenta and labeled with the corresponding N-glycosylation site. The site-specific N-glycan compositions of HA propagated in MDCK.SUS2 and MDCK.ADH cells are illustrated on the right site. For each site, the five most abundant N-glycan compositions are relatively quantified in a pie diagram. The remaining low abundant N-glycans are summed in a 6th piece. Symbolic representations of N-glycan structures were drawn with GlycoWorkbench Version 1.1, following the Symbol Nomenclature for Graphical Representations of Glycans [6]. Adapted from Pralow et al. (2021) [160].

In Figure 4-12, the structure of NA is shown. NA is a tetrameric antigen carrying five potential *N*-glycosylation sites. N44, N58 and N73 in the hypervariable stalk region and N131 and N220 in the head region (please note that the hypervariable stalk region is not completely visualized in Figure 4-12 because of model limitations due to missing protein sequences). Our analysis revealed only N73 to be *N*-glycosylated. In contrast to HA, possible di-antennary complex structures were the most dominant *N*-glycans. In IAV derived from MDCK.ADH cells, a possible hybrid type *N*-glycan structure had a high abundance. Comparable to HA, most of the *N*-glycan compositions at NA were antenna- and/or core-fucosylated and to some extent α -galactosylated. As for HA, MDCK.ADH cells seemed to provide a higher variability and complexity of *N*-glycan compositions compared to MDCK.SUS2 cells. Finally, for the third IAV glycoprotein M2, no *N*-glycosylation could be detected. In addition, no *O*-glycopeptides derived from HA or NA could be detected.

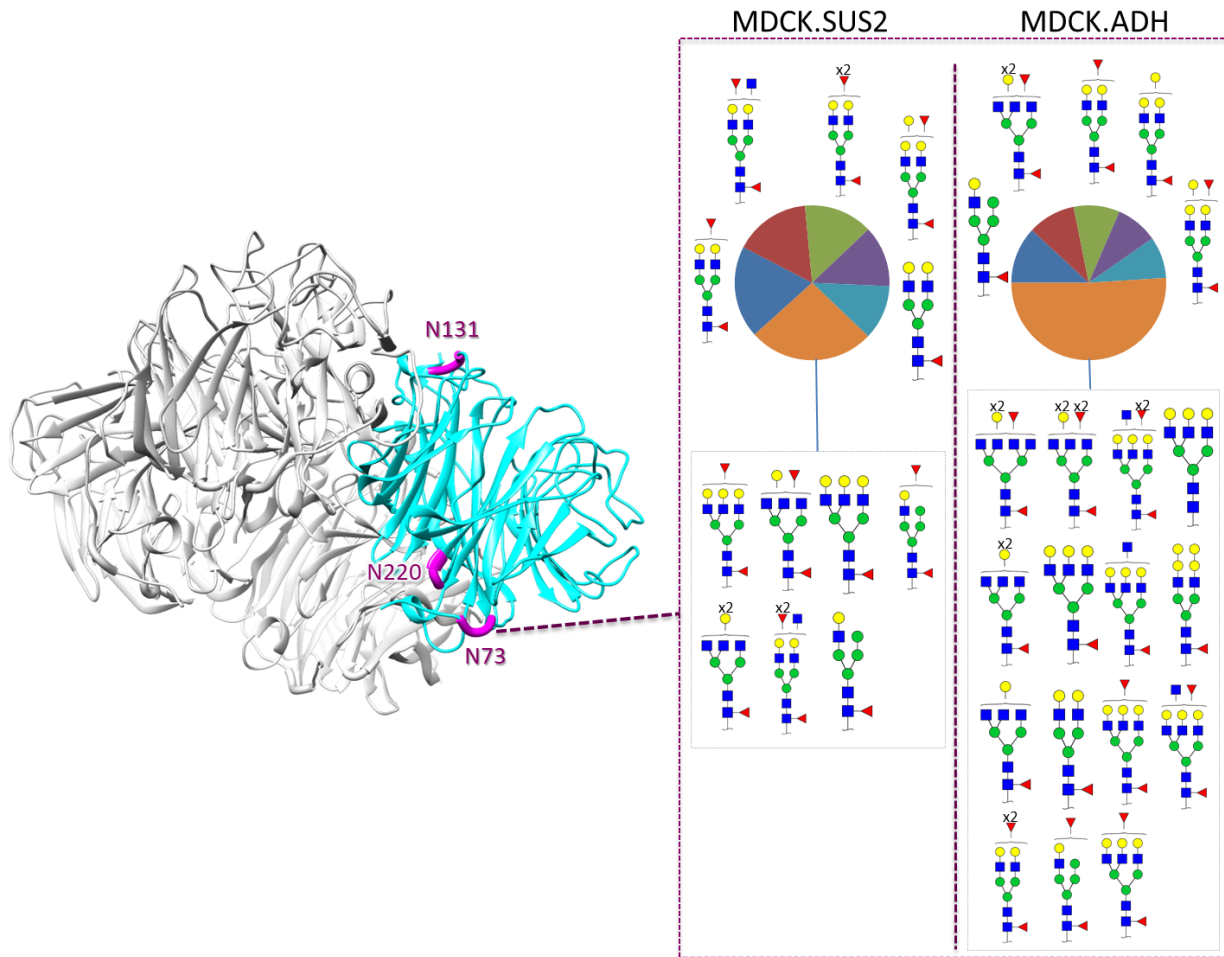


Figure 4-12: **Graphical illustration of the tetrameric IAV antigen NA.** The molecular structure of NA was modeled using the Swiss-Model template ID 5hug.1. For model processing and design the open source software UCSF Chimera Version 1.10.2 was utilized. An NA monomer is highlighted in cyan. N-glycosylation consensus sequences are shown in magenta and labeled with the corresponding N-glycosylation site. The site-specific N-glycan compositions of NA propagated in MDCK.SUS2 and MDCK.ADH cells are illustrated on the right site. For each site, the five most abundant N-glycan compositions are relatively quantified in a pie diagram. The remaining low abundant N-glycans are summed in a 6th piece. Symbolic representations of N-glycan structures were drawn with GlycoWorkbench Version 1.1, following the Symbol Nomenclature for Graphical Representations of Glycans [6]. Adapted from Pralow et al. (2021) [160].

4.3.2.3. Assembling of structural and site-specific *N*-glyco analysis results

Datasets from the fine-structural *N*-glycan and the site-specific *N*-glycopeptide analysis can be assembled to achieve comprehensive structural and site-specific *N*-glycan information. Exemplarily, in Figure 4-11 at the site N285 (closest to the head region) of HA propagated in MDCK.ADH cells, (Hex)₅ (HexNAc)₃ (dHex)₁ + (Man)₃(GlcNAc)₂ (dark blue) is a *N*-glycan composition with a high relative abundance. When looking up the *N*-glycan list (Supplementary Table A2-1), we see for the putative *N*-glycan composition only one fine-structural entry providing evidence of a three antennary double alpha-galactosylated core-fucosylated complex-type *N*-glycan (C45). Another example is the high abundant potential hybrid-type *N*-glycan (Hex)₂ (HexNAc)₁ (dHex)₁ + (Man)₃(GlcNAc)₂ located at N73 at NA propagated in MDCK.ADH cells (Figure 4-12). As before, in the fine-structural *N*-glycan list (Supplementary Table A2-1), only one entry for this composition exists: A hybrid-type *N*-glycan carrying a blood group H type 2 epitope at the 3-antenna (H01). In case of (Hex)₂ (HexNAc)₂ (dHex)₂ + (Man)₃(GlcNAc)₂, the highest abundant *N*-glycan composition located on N73 on NA propagated in MDCK.SUS2 cells, two entries exist in the IAV *N*-glycan list (Supplementary Table A2-1): The complex-type *N*-glycan with two antennary, a core-fucosylation and a blood group H type 2 epitope at the 3-antenna (C17), as well as the hybrid-type *N*-glycan with a core-fucosylation and a blood group A epitope at the 3-antenna (H09). Both structures were annotated in IAV from MDCK.SUS2 cells. For IAV from MDCK.ADH cells, only the complex-type *N*-glycan structure was detected.

4.3.3. Discussion

Most reports dealing with *N*-glycomic/*N*-glycoproteomic analysis of IAV antigens up to now lack a fine-structural analysis of *N*-glycans and provide only compositions or partly structural information.

In this thesis, I report about a comprehensive site-specific (*N*-glycoproteomics) and fine-structural (*N*-glycomics) analysis of *N*-glycans derived from the proteins of IAV PR/8/34 (H1N1) propagated in

MDCK.ADH and MDCK.SUS2 cells. The viral proteins studied comprise the whole IAV proteome, with HA and NA as the most relevant IAV surface glycoproteins in influenza vaccine production. For the first time, a comprehensive fine-structural analysis of the IAV whole *N*-glycome using nano-PGC-LC-MS(/MS), in combination with site-specific *N*-glycopeptide analysis was performed that enabled the establishment of a fine structural list of the whole virus *N*-glycome. This list was subsequently used as a starting point to link these structures to *N*-glycan compositions from the site-specific *N*-glycoproteomic analysis. Finally, site-specific *N*-glycan microheterogeneity was mapped to the three-dimensional structure of the HA, as well as the NA molecule. Unexpectedly, *N*-glycans decorated with potentially immunogenic epitopes (Galili-epitope, blood group H type 2 and blood group A) could be identified.

4.3.3.1. Novel methods used for influenza *N*-glycan analysis

In previous work only xCGE-LIF was used for IAV *N*-glycomic analysis [2, 35-37, 47]. However, due to complexity of the IAV glycosylation, additional experiments such as sequential exoglycosidase digestion followed by repeated xCGE-LIF measurement are required for fine-structural analysis.

In this thesis, released *N*-glycans from IAV antigens propagated in MDCK.ADH and MDCK.SUS2 cells were analyzed by nano-PGC-LC-MS(/MS). Nano-PGC-LC-MS(/MS) allows fine-structural analysis of glycans by the separation of isobaric structures followed by additional structural analysis via negative ion mode fragmentation generated marker ions [72, 224]. Compared to xCGE-LIF-based glycoanalysis, nano-PGC-LC-MS(/MS) has limitations in terms of throughput and reproducibility [76, 80] (see, for example, the tR difference of structure C43 in IAV derived from both cell lines, Figure 4-9). However, the glycan nano-PGC-LC-MS(/MS) BPCs of the replicates of IAV propagated in MDCK.ADH cells (Supplementary Figure A2-9) as well as MDCK.SUS2 cells (Supplementary Figure A2-10) show similar peak profiles with only moderate differences in the relative peak abundances.

Furthermore, interpretation of nano-PGC-LC-MS(/MS) data is rather challenging. Established glycomic methods are way more standardized and are provided with suitable databases. On the other hand, for example, xCGE-LIF measurements are more sensitive compared to MS and can detect *N*-glycans of very low abundancies.

Nevertheless, glycans have different ionization and fragmentation efficacies and when using MS-based analysis, preferable high abundant signals are chosen for fragmentation. Therefore, it is still possible to miss some unknown glycan structures in the analysis.

Monitoring micro- and macroheterogeneity of glycoproteins is of increasing interest. The trend in the biopharmaceutical industry towards manufacturing of more complex recombinant glycoproteins with multiple glycosylation sites (compared to the less complex mAbs with a single glycosylation site) requires not only a glycoanalysis on a global level (e.g. HILIC-FLD-MS, xCGE-LIF, MALDI-TOF-MS, PGC-LC-MS), but also on a protein structure-related site-specific level. In recent years enormous progress took place in optimizing the sample preparation and pushing LC-MS instrument-specific parameters to an optimal outcome [184]. Apart from the workflow, the “design” of the proteolytic digestion of the glycoproteins is essential. Normally, a tryptic digestion is preferred because of the high specificity. But sometimes *N*-glycan consensus sequences are located in peptide sequences, which carry multiple glycosylation sites or contain too many AAs to be readily analyzed by MS when using trypsin. Here, using unspecific enzymes like proteinase K or pronase allows to uncover such glycosylation sites, but results in a high redundancy of unique glycopeptides with different peptide-moieties.

The *N*-glycoproteomic analysis in this thesis features a sequential digestion workflow using the proteases trypsin and flavastacin [81, 97]. Because of the missing vicinity of a tryptic cleavage site to most of the potential *N*-glycosylation sites of HA and NA, using only trypsin was not efficient enough to perform a site-specific *N*-glycopeptide analysis, since the peptide moiety had too many AAs and/or

multiple consensus sequences for *N*-glycosylation. Flavastacin is cleaving the C-terminus of *N*-glycosylated asparagine, generating well analyzable *N*-glycopeptides, enabling the identification of all IAV HA1 *N*-glycosylation sites. In contrast to other unspecific enzymes used for the glycoproteomic analysis of IAV antigens (e.g. chymotrypsin, endoproteinase AspN, proteinase K, pepsin, Glu-C) [181, 182, 184], the “glyco”-specific cleavage behavior of flavastacin reduce the amount of redundant peptide moieties and increase confidence for peptide sequence identification.

Losing *N*-glycopeptide signals referring to site N497 of HA after flavastacin treatment can be explained by the tryptic cleavage site N-terminal to the glycosylated asparagine that was generated. The peptide sequence after tryptic digestion was (R)NGTYDYPK(Y) resulting in (R)N(G) after flavastacin treatment. Because of their low hydrophobicity, *N*-glycopeptides with only asparagine as peptide moiety might have been lost in our LC setup (hydrophobic loading onto the RP trap-column with the flow-through directed to waste).

Despite the availability of commercialized software (e.g. Byonic (ProteinMetrics), ProteinScape (Bruker Daltonic)) that enable a fully automated *N*-glycoproteomic analysis, it is still necessary to validate automated identifications manually. Because of the “black box” nature of such software, possible identifications can be lost when the search space is not properly defined. The recently published software glyXtool^{MS} performs a semi-automated analysis, providing not only tools for manual validation, but also enabling the manual annotation of not or wrongly identified MS/MS spectra [83]. Therefore, in this thesis, glyXtool^{MS} was used for the *N*-glycoproteomic analysis of the biological replicates. To challenge this analytical workflow we used IAV particles propagated in two closely related cell lines.

Some *N*-glycan structures/compositions were detected in the *N*-glycoanalysis, but were not detected as an *N*-glycopeptide, neither derived from IAV glycoproteins HA, NA, or M2, nor derived from glycoproteins of the cell line or the medium (e.g. from bovine origin). This might be due to the fact that

the abundance of various *N*-glycopeptides with the same *N*-glycan composition was too low for detection and fragmentation in the glycopeptide analysis. After *N*-glycan release, however, they assembled to one single high abundant *N*-glycan signal that can be detected in the *N*-glycomic analysis. Furthermore, we cannot exclude that a small amount of low abundant *N*-glycopeptides of IAV (or remaining host cell glycoproteins) were not identified at all by our workflow.

Recently published work on IAV glycosylation used only LC-MS-based site-specific glycopeptide analysis [183, 189, 225, 226], or in combination with MALDI-TOF-MS-, [181, 190, 227] respectively, HILIC-MS-based [182] analysis of released *N*-glycans. In addition, more complex viral glycoproteins like the envelope glycoprotein of HIV [228-230] with up to 30 glycosylation sites and the SARS-CoV-2 spike protein [231] were recently characterized. The HIV glycoprotein was investigated by site-specific glycopeptide analysis via LC-MS and compositional analysis of its released glycans using HILIC-FLD and IM-MS. The SARS-CoV-2 spike-protein was analyzed using only LC-MS-based site-specific glycopeptide analysis. As LC-MS-based site-specific analysis of glycoproteins lacks structural information of glycans (except for some partial glycan structural information that can be gained by using IM-MS [98] or by analyzing the ratio of specific fragment-ion intensities of glycopeptide MS/MS spectra [97]), additional glycomic analysis is necessary for the fine-structural elucidation of the glycans, as performed in this thesis. In contrast to standard MS-based glycomic approaches like MALDI-TOF-MS and HILIC-MS, (nano-)PGC-LC-MS(/MS) analysis allows to obtain more structural information of glycans but is more laborious. In particular, as data analysis is mainly manual, (nano-)PGC-LC-MS(/MS) requires quite some expertise and has limitations to identify large glycan structures with very high molecular masses (e.g. glycan structures beyond the xCGE-LIF peaks >430 MTU'' in the IAV *N*-glycan fingerprints in Supplementary Figure A2-1). Performing additional glycomic analysis by nano-PGC-LC-MS(/MS) enabled to differentiate between complex- and hybrid-type *N*-glycans as well as between antennary and bisecting GlcNAc. More

important, *N*-glycan structures with terminal blood group epitopes (A, B and H, Supplementary Table A2-1)) and non-human epitopes like the Galili-epitope could be structurally elucidated and identified. Without additional fine-structural *N*-glycan analysis, such information would have been missed.

4.3.3.2. Influenza antigen glycosylation

Comparing glycome-/glycoproteome data among influenza virus-related studies is problematic for several reasons: I.) Proteins of different IAV strains are hard to compare with each other because of their variable structure, as well as the number and location of potential *N*-glycosylation sites. II.) Besides embryonated chicken eggs (undefined mixture of different cells), a high number of possible host cells are used for IAV propagation, with a significant impact on IAV *N*-glycan structures. III.) Recombinant IAV antigens derived from mammalian or insect cell lines originate from a totally different environment compared to antigens incorporated into a lipid bilayer membrane of an intact IAV (e.g. no inclusion into a lipid bilayer, no vicinity of NA). As a result, the *N*-glycome of the recombinant antigens can include *N*-glycan structures that are biological impossible to obtain for intact IAV (e.g. sialylated epitopes [176]). IV.) The choice of cultivation medium and the number of passages and adaption time of the cell line can affect the glycosylation machinery of the host cell system and therefore the *N*-glycosylation pattern of IAV antigens. V.) Different analytical approaches in glycomics and glycoproteomics bear different information content regarding IAV glycosylation. VI.) Despite virus purification, cellular (host cell (glyco-) proteins) as well as animal-derived (medium components) contaminants can be found.

To give a short overview of the results of similar studies: An et al. (2019) analyzed the *N*-glycome of IAV H1N1 propagated in MDCK cells using MALDI-TOF-MS and observed mainly oligomannose-type (84%) and smaller proportions of complex-type (15%) and hybrid-type (1%) *N*-glycans [181]. Beside complex-type *N*-glycans, site-specific glycopeptide analysis revealed high abundant oligomannose-type *N*-glycans linked to the stalk region of HA from IAV H1N1 [181]. Khatri et al. (2016) analyzed the *N*-glycome of IAV

H1N1 propagated in embryonated chicken eggs using HILIC-MS. They identified a low degree of oligomannose-type (<6%) and hybrid-type ($\approx 9\%$) *N*-glycans, but a high amount of highly complex-type *N*-glycans ($\approx 85\%$). The site-specific glycopeptide analysis showed predominantly highly complex-type *N*-glycans linked to the stalk region, except for site N303, which was mostly decorated with oligomannose-type *N*-glycans [182]. Similar findings were already published earlier [2, 35]. She et al. (2017) performed a site-specific glycopeptide analysis of two high-yield candidate reassortant vaccines (NIBRG-121xp and NYMC-X181A) propagated in eggs [183]. Based on compositional data, the authors could identify complex- and oligomannose-type *N*-glycans. Beside this, also sulfated complex-type *N*-glycans were detected at HA and NA (but not in our study). Very recently, the authors did nano-PGC-LC-MS(/MS) analysis to confirm the sulfated glycan structures. Finally, they could identify sulfated complex *N*-glycans on glycoproteins derived from IAV propagated in eggs and MDCK cells [188].

In this study, *N*-glycan analysis of the purified whole virus lysate revealed oligomannose-, hybrid-, as well as complex-type *N*-glycan structures (with and without antenna- and/or core-fucosylation). Structures with antenna fucosylation were identified to consist of the epitopes blood-group H type 2 (in accordance with the annotation of An et. al (2019), who refers the partly structural annotation to the known *N*-glycan processing pathways in MDCK cells [181]) and blood-group A. To some extent structures with bisecting GlcNAc were identified. Compositions suggesting hybrid-type structures but also complex-type structures with Galili-epitopes could be successfully analyzed and confirmed. MDCK cells (especially MDCK.ADH) are of high interest for the cell based vaccine production. Adherent cells are difficult to scale up and to handle in high cell density cultures; however cell lines growing in suspension can overcome such limitations. To stress the glycomic/glycoproteomic workflow and to obtain further insights into the impact of cell line adaptation to growth in suspension on viral antigen composition, a detailed comparison of the site-specific and structural glycosylation pattern of viral proteins propagated

in MDCK.ADH and MDCK.SUS2 cells was performed. As expected, both cell lines showed a similar qualitative *N*-glycan pattern. However, relative abundances of individual *N*-glycan structures differed. Notable, the amount of oligomannose-type *N*-glycans was reduced in IAV from the serum-free MDCK.SUS2 cell cultures compared to IAV propagated in MDCK.ADH cells. In part, this could result from animal-derived medium components (fetal calf serum (FCS), lab M peptone), used to supplement the GMEM medium for cultivation of MDCK.ADH cells.

All potential *N*-glycosylation sites of HA were identified to be glycosylated with core-fucosylated complex and/or hybrid-type *N*-glycans. Complex-type *N*-glycosylated HA produced in mammalian cells is described to induce higher antibody titers than HA produced in insect cells, which features oligomannose-type *N*-glycans [3]. HA of IAV PR/8/34 (H1N1) has all *N*-glycosylation sites at the stem region and no *N*-glycosylation site at the head region close to the receptor binding site. However, we observed an increase of *N*-glycan variability and complexity at sites closer to the head region of HA (N285 and N303). The compositions of the *N*-glycans suggest multiantennary and/or LacNAc extended epitopes (based on the fine-structure *N*-glycan analysis) on these sites. To some extent, their high molecular mass *N*-glycans might have the ability to shield or cover the head region of HA and might be more exposed and recognizable to the immune system. Therefore, differences in antigenicity and immunogenicity of HA derived from both cell lines cannot be excluded.

For NA, we found only N73, one out of five potential *N*-glycosylation sites, to be glycosylated. This site is located at the stem region of NA. The site-specific *N*-glycopeptide analysis revealed a similar picture as the results from HA. *N*-glycans attached to NA, however, had lower complexity (mostly two antennary structures). Again, mainly antenna- and/or core-fucosylated complex-/hybrid-type *N*-glycans were identified.

Based on data, *N*-glycans with immunogenic epitopes like Galili-epitope, blood-group H type 2 and blood-group A are also suggested to be more abundant close to the head region (more dominant in IAV from MDCK.ADH cells). For vaccines, the Galili-epitope bears the potential to trigger immunogenicity and therefore vaccine efficacy. This was demonstrated on alpha-galactosyltransferase negative mice. The treatment with IAV carrying *in-vitro* alpha-galactosylated complex *N*-glycans (propagated in embryonated chicken eggs) led to an increased immunogenicity compared to the control group treated without alpha-galactosylation [180]. Recently, Galili et al. (2020) addressed the potential to amplify immunogenicity of prospective SARS-CoV-2 vaccines by glycoengineering the coronavirus glycan shield to present Galili-epitopes [232]. The different blood group epitopes detected on IAV antigens (HA and NA) might also have an effect on the immunogenicity or vaccine efficacy. Unfortunately, so far, no immunogenic study for such specific blood-group epitopes on IAV vaccines is available.

The identification of non-human glyco-epitopes like α -Gal, might challenge the interpretation of animal trials. Exposing *N*-glycans is part of the main strategies of IAV to evade the host's immune system [129]. Most animal studies for vaccines are performed in mice and ferrets, which encode, in contrast to humans, the alpha-galactosyltransferase. Therefore, depending on the IAV vaccine production system, the outcome of immunological studies might not reflect the immunogenicity of the IAV vaccine in humans.

As described in the introduction, several studies discuss the manipulation of IAV vaccine glycosylation (by depleting or shortening the glycans [173-175], adding specific glycan structures [176-179] or by adding/removing specific glycosylation sites [169]). The trend to produce rHA vaccines (e.g. Flublock [233, 234]) could take advantage of such glycoengineering ideas. The production of specific immunogenic epitopes in combination with addition or removal of *N*-glycosylation sites of HA plus the lack of membrane and NA (enabling sialylated eptiopes), might contribute to the efficacy of a vaccine.

Overall, comprehensive glycoprofiling of IAV proteins might be more requested to interpret findings of animal trials and might gain more importance in the future to better understand antigenicity and immunogenicity of IAV strains expressed in different host systems.

5. CONCLUSIONS

Not only glycosylated biopharmaceuticals such as recombinant proteins but also viral antigens with multiple *N*-glycosylation sites and different types of glycosylation are of significant importance in the pharmaceutical market. Although regulatory authorities do not ask for a glycan analysis of antigens from virus particles used for vaccine manufacturing yet, this might come one day. In particular, as the number of licensed viral vaccines and the number of manufacturing technologies increases.

Well-integrated technologies for *N*-glycan analysis (e.g. xCGE-LIF, HILIC-FLD, MALDI-TOF-MS) are optimized for the profiling of mAbs, other recombinant proteins (e.g. hormones) and vaccines. Those methods have a high resolution and a very low detection limit. Furthermore, they allow the establishment of high(er) throughput measurement platforms and the identification and quantification of *N*-glycan isomers. While LC-MS is the method of choice when it comes to site-specific glycopeptide analysis, it lacks the option to separate glycan isomers and provide structural information of glycans.

The aim of this thesis was to establish a comprehensive LC-MS workflow, capable to uncover both, the fine structural information of a glycan, as well as its location on a protein. This workflow was used to analyse the glycosylation of glycoproteins (mainly HA and NA) from IAV propagated in two closely related mammalian cell lines (MDCK.ADH and MDCK.SUS2 cells).

For the site-specific glycopeptide analysis, the choice of enzyme used for glycoprotein cleavage into suitable glycopeptides is crucial. Specific enzymes like trypsin tend to result in glycopeptides with a peptide moiety with too many amino acids or multiple glycosylation sites. This results into the loss of possible glycopeptide signals and/or can make the data analysis more complicated. On the other hand, unspecific enzymes like pronase or proteinase K generate smaller glycopeptides, but do not follow strict cleavage rules, which results in longer data analysis time and redundant glycopeptide signals (because of

multiple peptide moieties of a single glycosylation site), which could also lead to the loss of smaller signals. Screening for other enzymes for an optimized cleavage of IAV glycoproteins, Flavastacin was identified to cleave specifically C-terminal *N*-glycosylated asparagine, while the N-terminus is cleaved unspecifically. The enzyme was evaluated using the glycoprotein hLTF. Using this enzyme for the analysis of IAV glycoproteins resulted in a much higher amount of glycopeptide signals from very different glycosylation sites compared to trypsin alone. Overall, this protease helps to digest proteins with no tryptic cleavage sites close to the glycosylation site and assists (semi-)automated glycopeptide analysis software because of the “glyco-specific” cleavage behavior.

Glycomic analysis using nano-PGC-LC instrumentation was disturbed because of low spray efficacy through droplet formation, leading to instable measurements. The droplet formation was caused because by a high aquos content in the LC buffer system. Therefore, a PCMF with 100% ACN was attached to the system to increase the organic content of the buffer without affecting the PGC separation efficacy. This decreased the surface tension and resulted in a stable nano-spray, which improved the signal intensity, fragment ion spectra quality as well as identification rate of detected glycans (especially smaller glycans at early tR) tremendously.

Previous studies analyzed different IAV in several expression systems. However, a combination of site-specific glycopeptide analysis and fine structural glycan analysis was still missing. Combining the aforementioned established methods, a comprehensive workflow was designed that allowed to combine site-specific glycopeptide information with fine structural glycan information. While the *N*-glycosylation pattern of the IAV glycoproteins HA and NA propagated in MDCK.ADH and MDCK.SUS2 cells was similar, it displayed differences in relative abundances of individual *N*-glycan structures. Fine structural glycan analysis revealed specific glycan epitopes like the blood group H type 2 and the blood

group A epitope. Furthermore, a high degree of the immunogenic Galili-epitope was found on HA and NA of IAV produced in both cell lines (especially in MDCK.ADH).

All potential *N*-glycosylation sites of HA were shown to be glycosylated with an increasing complexity and variability of glycosylation at sites closer to the head region (antigen binding site). HA from both cell lines was mainly decorated with high molecular mass *N*-glycan structures with up to four antennaries. For NA, only one potential *N*-glycosylation site (N73) was identified. In contrast to HA, *N*-glycan structures were smaller with mostly two antennaries.

Combining both, site-specific as well as fine structural glycan analysis, allowed to analyse the specific glycan structure at a specific glycosylation site. But this was only the case for glycan compositions (linked to the specific molecular mass of a glycan moiety) with only a single identified mass isomer. Not for compositions with multiple glycan mass isomers. However, for the ladder, this workflow provided the potential glycan structures located at a specific glycosylation site of a protein. Therefore, this work shows successfully the combination of LC-MS based methods for a site-specific and fine structural glycan analysis of viral glycoproteins.

6. OUTLOOK

Glycosylation analysis of virus particle-derived antigens (not recombinant proteins) is not a CQA in the biopharmaceutical industry, yet. Recent studies might indicate that the glycosylation of viral antigens influences the antibody response [164]. This might result in less vaccine efficacy and more severe diseases like in young adults infected with the 2009 H1N1 [235]. Therefore, the monitoring of vaccine antigen glycosylation might be important in future vaccine production.

The workflow established in the scope of this thesis enables a very complex in-depth explorative glycosylation analyses, but demand a high level of expertise and cost intensive instrumentation. Such complex glycan analysis workflow might not be feasible for routine production, but is very important for research and development and in-depth antigen characterization. In addition, interpretation of antigenicity and immunogenicity of vaccines, based on their antigen glycosylation, will be facilitated.

There are different possibilities to further optimize the presented workflow. One option is to add an additional separation step prior LC-MS measurement. Using a standard SDS-gel, sufficient to separate the IAV glycoproteins by their molecular mass, could isolate specific glycoproteins. These proteins could be transferred to a western blot membrane followed by an *N*-glycan release [40]. Therefore, specific glycan fine structures could be linked to specific IAV glycoproteins. However, such additional sample preparation step could potentially decrease the signal intensity in LC-MS.

Another option would be using an unspecific digestion of the IAV glycoproteins using pronase as described by Stavenhagen et al. 2015 [236]. Therefore, the amount of AAs attached to the glycan moiety could be small enough to still pass the PGC column, but large enough to provide site-specific information. Measurement and fragmentation in negative ion mode could simultaneously provide fine-structural information. Accordingly, using LC-MS instrumentation in combination with different

fragmentation techniques as well as different NCEs would allow providing fine-structural as well as site-specific information within a single LC-MS measurement.

The finding that α -galactosylated *N*-glycan structures in HA as well as NA of IAV are found for both, MDCK.ADH and MDCK.SUS2 cells, might harbor the potential of modulating antigenicity and immunogenicity of cell-derived compared to eggs-derived vaccines [180, 182, 237]. To elucidate possible effects of non-human glycan epitopes identified for IAV antigens in terms of vaccine efficacy, further structure-function-relation as well as animal studies of influenza virus produced in different host cells need to be performed. For animal studies, the individual potential glycome of the animal as well as the potential glycome of the tested vaccine need to be taken into account, as specific glycan epitopes on viruses could influence the immunogenicity in the different animals.

More method and technology developments will be seen in the future with the potential to analyze glycan fine structure, site-specifically, in a single run, without the need of a comprehensive orthogonal analysis. Such measurements would provide specific fine structural information for each site-specific glycopeptide signal, independent of the amount of different glycan isomers in the sample. Especially IM-MS was demonstrated to bear the potential to identify epitopes on small glycopeptides [20, 238]. However, IM-MS lacks the potential to provide detailed fine structural information of the glycan moiety from glycopeptides. Nevertheless, further development of IM-MS instrumentation in combination with different fragmentation techniques and pre-separation techniques like PGC-LC might meet the requirement of a real fine structural site-specific glycan analysis.

REFERENCES

1. Holmes, K. K., Bertozzi, S., Bloom, B. R., Jha, P., Gelband, H., DeMaria, L. M. & Horton, S. (2017) Major infectious diseases: key messages from disease control priorities, *Major infectious diseases*.
2. Hutter, J., Rodig, J. V., Hoper, D., Seeberger, P. H., Reichl, U., Rapp, E. & Lepenies, B. (2013) Toward animal cell culture-based influenza vaccine design: viral hemagglutinin N-glycosylation markedly impacts immunogenicity, *Journal of immunology*. **190**, 220-30.
3. de Vries, R. P., Smit, C. H., de Bruin, E., Rigter, A., de Vries, E., Cornelissen, L. A., Eggink, D., Chung, N. P., Moore, J. P. & Sanders, R. W. (2012) Glycan-dependent immunogenicity of recombinant soluble trimeric hemagglutinin, *Journal of virology*, JVI. 01084-12.
4. York, I. A., Stevens, J. & Alymova, I. V. (2019) Influenza virus N-linked glycosylation and innate immunity, *Biosci Rep*. **39**.
5. Varki, A. & Hart, G. W. (2017) *Essentials of Glycobiology*, Cold Spring Harbor Laboratory Press.
6. Varki, A., Cummings, R. D., Aebi, M., Packer, N. H., Seeberger, P. H., Esko, J. D., Stanley, P., Hart, G., Darvill, A., Kinoshita, T., Prestegard, J. J., Schnaar, R. L., Freeze, H. H., Marth, J. D., Bertozzi, C. R., Etzler, M. E., Frank, M., Vliegenthart, J. F., Lutteke, T., Perez, S., Bolton, E., Rudd, P., Paulson, J., Kanehisa, M., Toukach, P., Aoki-Kinoshita, K. F., Dell, A., Narimatsu, H., York, W., Taniguchi, N. & Kornfeld, S. (2015) Symbol Nomenclature for Graphical Representations of Glycans, *Glycobiology*. **25**, 1323-4.
7. Taylor, M. E. & Drickamer, K. (2011) *Introduction to glycobiology*, Oxford university press.
8. Hennig, R. (2009) Development of high-throughput method for the analysis of total plasma N-glycans by CGE-LIF. Diploma thesis, Otto von Guericke Universität, Magdeburg, Deutschland,
9. Christiansen, M. N., Chik, J., Lee, L., Anugraham, M., Abrahams, J. L. & Packer, N. H. (2014) Cell surface protein glycosylation in cancer, *Proteomics*. **14**, 525-46.
10. Rhodes, J., Campbell, B. J. & Yu, L.-G. (2010) Glycosylation and Disease, *Encyclopedia of Life Sciences (ELS) John Wiley & Sons, Ltd: Chichester*.
11. Defaus, S., Gupta, P., Andreu, D. & Gutierrez-Gallego, R. (2014) Mammalian protein glycosylation--structure versus function, *The Analyst*. **139**, 2944-67.
12. Bergstrom, J. P. & Helander, A. (2008) Clinical characteristics of carbohydrate-deficient transferrin (disialotransferrin) measured by HPLC: sensitivity, specificity, gender effects, and relationship with other alcohol biomarkers, *Alcohol and alcoholism*. **43**, 436-41.
13. van Kooyk, Y., Kalay, H. & Garcia-Vallejo, J. J. (2013) Analytical tools for the study of cellular glycosylation in the immune system, *Frontiers in immunology*. **4**, 451.
14. Lingg, N., Zhang, P., Song, Z. & Bardor, M. (2012) The sweet tooth of biopharmaceuticals: importance of recombinant protein glycosylation analysis, *Biotechnology journal*. **7**, 1462-72.
15. Aziz, F. & Qiu, Y. (2014) The role of anti-LeY antibody in the downregulation of MAPKs/COX-2 pathway in gastric cancer, *Current drug targets*. **15**, 469-76.
16. Ravn, V. & Dabelsteen, E. (2000) Tissue distribution of histo-blood group antigens, *Apmis*. **108**, 1-28.
17. Greenwell, P. (1997) Blood group antigens: molecules seeking a function?, *Glycoconjugate journal*. **14**, 159-173.
18. Thaysen-Andersen, M. & Packer, N. H. (2014) Advances in LC-MS/MS-based glycoproteomics: getting closer to system-wide site-specific mapping of the N- and O-glycoproteome, *Biochimica et biophysica acta*. **1844**, 1437-52.

19. Huffman, J. E., Pucic-Bakovic, M., Klaric, L., Hennig, R., Selman, M. H., Vuckovic, F., Novokmet, M., Kristic, J., Borowiak, M., Muth, T., Polasek, O., Razdorov, G., Gornik, O., Plomp, R., Theodoratou, E., Wright, A. F., Rudan, I., Hayward, C., Campbell, H., Deelder, A. M., Reichl, U., Aulchenko, Y. S., Rapp, E., Wuhrer, M. & Lauc, G. (2014) Comparative performance of four methods for high-throughput glycosylation analysis of immunoglobulin G in genetic and epidemiological research, *Molecular & cellular proteomics : MCP*. **13**, 1598-610.
20. Pralow, A., Cajic, S., Alagesan, K., Kolarich, D. & Rapp, E. (2020) State-of-the-Art Glycomics Technologies in Glycobiotechnology, *Advances in Glycobiotechnology*, Springer, Cham, 379-411.
21. Helenius, A. & Aebi, M. (2001) Intracellular functions of N-linked glycans, *Science*. **291**, 2364-2369.
22. Varki, A. (2017) Biological roles of glycans, *Glycobiology*. **27**, 3-49.
23. Moremen, K. W., Tiemeyer, M. & Nairn, A. V. (2012) Vertebrate protein glycosylation: diversity, synthesis and function, *Nat Rev Mol Cell Biol*. **13**, 448-62.
24. Varki, A. (1993) Biological roles of oligosaccharides: all of the theories are correct, *Glycobiology*. **3**, 97-130.
25. Woods, R. J. (1995) Three-dimensional structures of oligosaccharides, *Current opinion in structural biology*. **5**, 591-598.
26. Wang, J.-R., Gao, W.-N., Grimm, R., Jiang, S., Liang, Y., Ye, H., Li, Z.-G., Yau, L.-F., Huang, H. & Liu, J. (2017) A method to identify trace sulfated IgG N-glycans as biomarkers for rheumatoid arthritis, *Nature communications*. **8**, 631.
27. Pomin, V. H. (2015) Sulfated glycans in inflammation, *European journal of medicinal chemistry*. **92**, 353-369.
28. Yoshimura, T., Hayashi, A., Handa-Narumi, M., Yagi, H., Ohno, N., Koike, T., Yamaguchi, Y., Uchimura, K., Kadomatsu, K. & Sedzik, J. (2017) GlcNAc6ST-1 regulates sulfation of N-glycans and myelination in the peripheral nervous system, *Scientific reports*. **7**, 42257.
29. Takashiba, M., Chiba, Y. & Jigami, Y. (2006) Identification of phosphorylation sites in N-linked glycans by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, *Analytical chemistry*. **78**, 5208-5213.
30. Tarentino, A. L., Gomez, C. M. & Plummer Jr, T. H. (1985) Deglycosylation of asparagine-linked glycans by peptide: N-glycosidase F, *Biochemistry*. **24**, 4665-4671.
31. Guile, G. R., Rudd, P. M., Wing, D. R., Prime, S. B. & Dwek, R. A. (1996) A rapid high-resolution high-performance liquid chromatographic method for separating glycan mixtures and analyzing oligosaccharide profiles, *Anal Biochem*. **240**, 210-26.
32. Rudd, P. M. & Dwek, R. A. (1997) Rapid, sensitive sequencing of oligosaccharides from glycoproteins, *Current opinion in biotechnology*. **8**, 488-97.
33. Ahn, J., Bones, J., Yu, Y. Q., Rudd, P. M. & Gilar, M. (2010) Separation of 2-aminobenzamide labeled glycans using hydrophilic interaction chromatography columns packed with 1.7 microm sorbent, *Journal of chromatography B, Analytical technologies in the biomedical and life sciences*. **878**, 403-8.
34. Callewaert, N., Geysens, S., Molemans, P. & Contreras, R. (2001) Ultrasensitive profiling and sequencing of N-linked oligosaccharides using standard DNA-sequencing equipment, *Glycobiology*. **11**, 275-281.
35. Schwarzer, J., Rapp, E., Hennig, R., Genzel, Y., Jordan, I., Sandig, V. & Reichl, U. (2009) Glycan analysis in cell culture-based influenza vaccine production: influence of host cell line and virus strain on the glycosylation pattern of viral hemagglutinin, *Vaccine*. **27**, 4325-36.
36. Schwarzer, J., Rapp, E. & Reichl, U. (2008) N-glycan analysis by CGE-LIF: profiling influenza A virus hemagglutinin N-glycosylation during vaccine production, *Electrophoresis*. **29**, 4203-14.

-
37. Hennig, R., Rapp, E., Kottler, R., Cajic, S., Borowiak, M. & Reichl, U. (2015) *N*-Glycosylation fingerprinting of viral glycoproteins by xCGE-LIF.
38. Ruhaak, L. R., Hennig, R., Huhn, C., Borowiak, M., Dolhain, R. J., Deelder, A. M., Rapp, E. & Wuhrer, M. (2010) Optimized workflow for preparation of APTS-labeled *N*-glycans allowing high-throughput analysis of human plasma glycomes using 48-channel multiplexed CGE-LIF, *Journal of proteome research*. **9**, 6655-6664.
39. Wuhrer, M., de Boer, A. R. & Deelder, A. M. (2009) Structural glycomics using hydrophilic interaction chromatography (HILIC) with mass spectrometry, *Mass spectrometry reviews*. **28**, 192-206.
40. Jensen, P. H., Karlsson, N. G., Kolarich, D. & Packer, N. H. (2012) Structural analysis of *N*- and *O*-glycans released from glycoproteins, *Nature protocols*. **7**, 1299-310.
41. Everest-Dass, A. V., Kolarich, D., Campbell, M. P. & Packer, N. H. (2013) Tandem mass spectra of glycan substructures enable the multistage mass spectrometric identification of determinants on oligosaccharides, *Rapid communications in mass spectrometry : RCM*. **27**, 931-9.
42. Reiding, K. R., Blank, D., Kuijper, D. M., Deelder, A. M. & Wuhrer, M. (2014) High-throughput profiling of protein *N*-glycosylation by MALDI-TOF-MS employing linkage-specific sialic acid esterification, *Analytical chemistry*. **86**, 5784-93.
43. Kolarich, D., Windwarder, M., Alagesan, K. & Altmann, F. (2015) Isomer-Specific Analysis of Released *N*-Glycans by LC-ESI MS/MS with Porous Graphitized Carbon, *Methods in molecular biology*. **1321**, 427-35.
44. Wuhrer, M., Deelder, A. M. & Hokke, C. H. (2005) Protein glycosylation analysis by liquid chromatography-mass spectrometry, *Journal of chromatography B, Analytical technologies in the biomedical and life sciences*. **825**, 124-33.
45. Reiding, K. R., Lonardi, E., Hipgrave Ederveen, A. L. & Wuhrer, M. (2016) Ethyl Esterification for MALDI-MS Analysis of Protein Glycosylation, *Methods in molecular biology*. **1394**, 151-162.
46. Shubhakar, A., Kozak, R. P., Reiding, K. R., Royle, L., Spencer, D. I., Fernandes, D. L. & Wuhrer, M. (2016) Automated High-Throughput Permethylolation for Glycosylation Analysis of Biologics Using MALDI-TOF-MS, *Analytical chemistry*. **88**, 8562-9.
47. Rodig, J. V., Rapp, E., Bohne, J., Kampe, M., Kaffka, H., Bock, A., Genzel, Y. & Reichl, U. (2013) Impact of cultivation conditions on *N*-glycosylation of influenza virus a hemagglutinin produced in MDCK cell culture, *Biotechnology and bioengineering*. **110**, 1691-703.
48. Hennig, R., Cajic, S., Borowiak, M., Hoffmann, M., Kottler, R., Reichl, U. & Rapp, E. (2016) Towards personalized diagnostics via longitudinal study of the human plasma *N*-glycome, *Biochimica et biophysica acta*. **1860**, 1728-38.
49. Konze, S. A., Cajic, S., Oberbeck, A., Hennig, R., Pich, A., Rapp, E. & Buettner, F. F. (2017) Quantitative assessment of sialo-glycoproteins and *N*-glycans during cardiomyogenic differentiation of human induced pluripotent stem cells, *Chembiochem*. **18**, 1317-1331.
50. Thiesler, C. T., Cajic, S., Hoffmann, D., Thiel, C., van Diepen, L., Hennig, R., Sgodda, M., Weißmann, R., Reichl, U. & Steinemann, D. (2016) Glycomic characterization of induced pluripotent stem cells derived from a patient suffering from phosphomannomutase 2 congenital disorder of glycosylation (PMM2-CDG), *Molecular & Cellular Proteomics*. **15**, 1435-1452.
51. Kottler, R., Mank, M., Hennig, R., Muller-Werner, B., Stahl, B., Reichl, U. & Rapp, E. (2013) Development of a high-throughput glycoanalysis method for the characterization of oligosaccharides in human milk utilizing multiplexed capillary gel electrophoresis with laser-induced fluorescence detection, *Electrophoresis*. **34**, 2323-36.
52. Ciucanu, I. (2006) Per-O-methylation reaction for structural analysis of carbohydrates by mass spectrometry, *Anal Chim Acta*. **576**, 147-55.
-

-
53. Morelle, W. & Michalski, J. C. (2007) Analysis of protein glycosylation by mass spectrometry, *Nature protocols*. **2**, 1585-602.
54. Harvey, D. J. (2000) Electrospray mass spectrometry and fragmentation of N-linked carbohydrates derivatized at the reducing terminus, *Journal of the American Society for Mass Spectrometry*. **11**, 900-15.
55. Melmer, M., Stangler, T., Schiefermeier, M., Brunner, W., Toll, H., Rupprechter, A., Lindner, W. & Premstaller, A. (2010) HILIC analysis of fluorescence-labeled N-glycans from recombinant biopharmaceuticals, *Analytical and bioanalytical chemistry*. **398**, 905-14.
56. Stockmann, H., O'Flaherty, R., Adamczyk, B., Saldova, R. & Rudd, P. M. (2015) Automated, high-throughput serum glycoprofiling platform, *Integr Biol (Camb)*. **7**, 1026-32.
57. Anumula, K. R. & Dhume, S. T. (1998) High resolution and high sensitivity methods for oligosaccharide mapping and characterization by normal phase high performance liquid chromatography following derivatization with highly fluorescent anthranilic acid, *Glycobiology*. **8**, 685-94.
58. Klapoetke, S., Zhang, J., Becht, S., Gu, X. & Ding, X. (2010) The evaluation of a novel approach for the profiling and identification of N-linked glycan with a procainamide tag by HPLC with fluorescent and mass spectrometric detection, *J Pharm Biomed Anal*. **53**, 315-24.
59. Lauber, M. A., Yu, Y. Q., Brousmiche, D. W., Hua, Z., Koza, S. M., Magnelli, P., Guthrie, E., Taron, C. H. & Fountain, K. J. (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**, 5401-9.
60. Tomiya, N., Awaya, J., Kurono, M., Endo, S., Arata, Y. & Takahashi, N. (1988) Analyses of N-linked oligosaccharides using a two-dimensional mapping technique, *Analytical biochemistry*. **171**, 73-90.
61. Marino, K., Bones, J., Kattla, J. J. & Rudd, P. M. (2010) A systematic approach to protein glycosylation analysis: a path through the maze, *Nat Chem Biol*. **6**, 713-723.
62. Nakano, M., Saldanha, R., Gobel, A., Kavallaris, M. & Packer, N. H. (2011) Identification of glycan structure alterations on cell membrane proteins in desoxyepothilone B resistant leukemia cells, *Molecular & cellular proteomics : MCP*. **10**, M111 009001.
63. Everest-Dass, A. V., Jin, D. Y., Thaysen-Andersen, M., Nevalainen, H., Kolarich, D. & Packer, N. H. (2012) Comparative structural analysis of the glycosylation of salivary and buccal cell proteins: innate protection against infection by *Candida albicans*, *Glycobiology*. **22**, 1465-1479.
64. Nguyen-Khuong, T., Everest-Dass, A. V., Kautto, L., Zhao, Z. J., Willcox, M. D. P. & Packer, N. H. (2015) Glycomic characterization of basal tears and changes with diabetes and diabetic retinopathy, *Glycobiology*. **25**, 269-283.
65. Ashwood, C., Lin, C. H., Thaysen-Andersen, M. & Packer, N. H. (2018) Discrimination of Isomers of Released N- and O-Glycans Using Diagnostic Product Ions in Negative Ion PGC-LC-ESI-MS/MS, *Journal of the American Society for Mass Spectrometry*. **29**, 1194-1209.
66. Estrella, R. P., Whitelock, J. M., Packer, N. H. & Karlsson, N. G. (2007) Graphitized Carbon LC- MS Characterization of the Chondroitin Sulfate Oligosaccharides of Aggrecan, *Analytical chemistry*. **79**, 3597-3606.
67. Schulz, B. L., Sloane, A. J., Robinson, L. J., Prasad, S. S., Lindner, R. A., Robinson, M., Bye, P. T., Nielson, D. W., Harry, J. L. & Packer, N. H. (2007) Glycosylation of sputum mucins is altered in cystic fibrosis patients, *Glycobiology*. **17**, 698-712.
68. Thomsson, K. A., Schulz, B. L., Packer, N. H. & Karlsson, N. G. (2005) MUC5B glycosylation in human saliva reflects blood group and secretor status, *Glycobiology*. **15**, 791-804.
-

-
69. Karlsson, N. G., Schulz, B. L. & Packer, N. H. (2004) Structural determination of neutral O-linked oligosaccharide alditols by negative ion LC-electrospray-MSn, *Journal of the American Society for Mass Spectrometry*. **15**, 659-72.
70. Schulz, B. L., Oxley, D., Packer, N. H. & Karlsson, N. G. (2002) Identification of two highly sialylated human tear-fluid DMBT1 isoforms: the major high-molecular-mass glycoproteins in human tears, *Biochemical Journal*. **366**, 511-520.
71. Melmer, M., Stangler, T., Premstaller, A. & Lindner, W. (2011) Comparison of hydrophilic-interaction, reversed-phase and porous graphitic carbon chromatography for glycan analysis, *J Chromatogr A*. **1218**, 118-23.
72. Everest-Dass, A. V., Abrahams, J. L., Kolarich, D., Packer, N. H. & Campbell, M. P. (2013) Structural feature ions for distinguishing N- and O-linked glycan isomers by LC-ESI-IT MS/MS, *Journal of the American Society for Mass Spectrometry*. **24**, 895-906.
73. Harvey, D. J. (2005) Fragmentation of negative ions from carbohydrates: part 2. Fragmentation of high-mannose N-linked glycans, *Journal of the American Society for Mass Spectrometry*. **16**, 631-46.
74. Harvey, D. J. (2005) Fragmentation of negative ions from carbohydrates: part 3. Fragmentation of hybrid and complex N-linked glycans, *Journal of the American Society for Mass Spectrometry*. **16**, 647-59.
75. Harvey, D. J., Royle, L., Radcliffe, C. M., Rudd, P. M. & Dwek, R. A. (2008) Structural and quantitative analysis of N-linked glycans by matrix-assisted laser desorption ionization and negative ion nanospray mass spectrometry, *Anal Biochem*. **376**, 44-60.
76. Pabst, M., Bondili, J. S., Stadlmann, J., Mach, L. & Altmann, F. (2007) Mass + retention time = structure: a strategy for the analysis of N-glycans by carbon LC-ESI-MS and its application to fibrin N-glycans, *Analytical chemistry*. **79**, 5051-7.
77. Pabst, M. & Altmann, F. (2008) Influence of electrosorption, solvent, temperature, and ion polarity on the performance of LC-ESI-MS using graphitic carbon for acidic oligosaccharides, *Analytical chemistry*. **80**, 7534-7542.
78. Aldredge, D., An, H. J., Tang, N., Waddell, K. & Lebrilla, C. B. (2012) Annotation of a serum N-glycan library for rapid identification of structures, *Journal of proteome research*. **11**, 1958-1968.
79. Abrahams, J. L., Campbell, M. P. & Packer, N. H. (2018) Building a PGC-LC-MS N-glycan retention library and elution mapping resource, *Glycoconjugate journal*. **35**, 15-29.
80. Ashwood, C., Pratt, B., MacLean, B. X., Gundry, R. L. & Packer, N. H. (2019) Standardization of PGC-LC-MS-based glycomics for sample specific glycotyping, *The Analyst*.
81. Pralow, A., Hoffmann, M., Nguyen-Khuong, T., Rapp, E. & Reichl, U. (2017) Improvement of the glycoproteomic toolbox with the discovery of a unique C-terminal cleavage specificity of flavastacin for N-glycosylated asparagine, *Scientific reports*. **7**, 11419.
82. Barrett, A. J., Woessner, J. F. & Rawlings, N. D. (2012) *Handbook of proteolytic enzymes*, Elsevier.
83. Pioch, M., Hoffmann, M., Pralow, A., Reichl, U. & Rapp, E. (2018) glyXtool(MS): An Open-Source Pipeline for Semiautomated Analysis of Glycopeptide Mass Spectrometry Data, *Analytical chemistry*.
84. Drapeau, G. R. (1980) Substrate specificity of a proteolytic enzyme isolated from a mutant of *Pseudomonas fragi*, *The Journal of biological chemistry*. **255**, 839-40.
85. Tetaz, T., Morrison, J. R., Andreou, J. & Fidge, N. H. (1990) Relaxed specificity of endoproteinase Asp-N: this enzyme cleaves at peptide bonds N-terminal to glutamate as well as aspartate and cysteic acid residues, *Biochem Int*. **22**, 561-6.
-

-
86. Ingrosso, D., Fowler, A. V., Bleibaum, J. & Clarke, S. (1989) Specificity of endoproteinase Asp-N (*Pseudomonas fragi*): cleavage at glutamyl residues in two proteins, *Biochemical and biophysical research communications*. **162**, 1528-34.
87. Kanie, Y., Yamamoto-Hino, M., Karino, Y., Yokozawa, H., Nishihara, S., Ueda, R., Goto, S. & Kanie, O. (2009) Insight into the regulation of glycan synthesis in *Drosophila* chaoptin based on mass spectrometry, *PLoS One*. **4**, e5434.
88. Tarentino, A. L., Quinones, G., Grimwood, B. G., Hauer, C. R. & Plummer, T. H., Jr. (1995) Molecular cloning and sequence analysis of flavastacin: an O-glycosylated prokaryotic zinc metalloendopeptidase, *Arch Biochem Biophys*. **319**, 281-5.
89. Grimwood, B. G., Plummer, T. H., Jr. & Tarentino, A. L. (1994) Purification and characterization of a neutral zinc endopeptidase secreted by *Flavobacterium meningosepticum*, *Arch Biochem Biophys*. **311**, 127-32.
90. Ebeling, W., Hennrich, N., Klockow, M., Metz, H., Orth, H. D. & Lang, H. (1974) Proteinase K from *Tritirachium album limber*, *European Journal of Biochemistry*. **47**, 91-97.
91. Kolarich, D., Jensen, P. H., Altmann, F. & Packer, N. H. (2012) Determination of site-specific glycan heterogeneity on glycoproteins, *Nature protocols*. **7**, 1285-98.
92. Wisniewski, J. R., Zougman, A., Nagaraj, N. & Mann, M. (2009) Universal sample preparation method for proteome analysis, *Nature methods*. **6**, 359-62.
93. Rosenfeld, J., Capdevielle, J., Guillemot, J. C. & Ferrara, P. (1992) In-gel digestion of proteins for internal sequence analysis after one-or two-dimensional gel electrophoresis, *Analytical biochemistry*. **203**, 173-179.
94. Reusch, D., Habegger, M., Falck, D., Peter, B., Maier, B., Gassner, J., Hook, M., Wagner, K., Bonnington, L., Bulau, P. & Wuhrer, M. (2015) Comparison of methods for the analysis of therapeutic immunoglobulin G Fc-glycosylation profiles-Part 2: Mass spectrometric methods, *Mabs*. **7**, 732-742.
95. Selman, M. H., Hoffmann, M., Zauner, G., McDonnell, L. A., Balog, C. I., Rapp, E., Deelder, A. M. & Wuhrer, M. (2012) MALDI-TOF-MS analysis of sialylated glycans and glycopeptides using 4-chloro-alpha-cyanocinnamic acid matrix, *Proteomics*. **12**, 1337-48.
96. Watson, J. T. & Sparkman, O. D. (2007) *Introduction to mass spectrometry: instrumentation, applications, and strategies for data interpretation*, John Wiley & Sons.
97. Hoffmann, M., Pioch, M., Pralow, A., Hennig, R., Kottler, R., Reichl, U. & Rapp, E. (2018) The Fine Art of Destruction: A Guide to In-Depth Glycoproteomic Analyses—Exploiting the Diagnostic Potential of Fragment Ions, *Proteomics*.
98. Hinneburg, H., Hofmann, J., Struwe, W. B., Thader, A., Altmann, F., Varon Silva, D., Seeberger, P. H., Pagel, K. & Kolarich, D. (2016) Distinguishing N-acetylneuraminic acid linkage isomers on glycopeptides by ion mobility-mass spectrometry, *Chem Commun (Camb)*. **52**, 4381-4.
99. Ducret, A., Oostveen, I. V., Eng, J. K., Yates, J. R. & Aebersold, R. (1998) High throughput protein characterization by automated reverse-phase chromatography/electrospray tandem mass spectrometry, *Protein Science*. **7**, 706-719.
100. Thermo Scientific (2014) Trap Columns in
101. 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**, 282574.
102. Domon, B. & Costello, C. E. (1988) A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates, *Glycoconjugate journal*. **5**, 397-409.
103. Huddleston, M. J., Bean, M. F. & Carr, S. A. (1993) Collisional fragmentation of glycopeptides by electrospray ionization LC/MS and LC/MS/MS: methods for selective detection of glycopeptides in protein digests, *Analytical chemistry*. **65**, 877-884.
-

-
104. Biemann, K. (1990) Appendix 5. Nomenclature for peptide fragment ions (positive ions), *Methods in enzymology*. **193**, 886-887.
105. Pasing, Y., Sickmann, A. & Lewandrowski, U. (2012) N-glycoproteomics: mass spectrometry-based glycosylation site annotation, *Biological chemistry*. **393**, 249-58.
106. Wuhrer, M., Catalina, M. I., Deelder, A. M. & Hokke, C. H. (2007) Glycoproteomics based on tandem mass spectrometry of glycopeptides, *Journal of chromatography B, Analytical technologies in the biomedical and life sciences*. **849**, 115-28.
107. Hinneburg, H., Stavenhagen, K., Schweiger-Hufnagel, U., Pengelley, S., Jabs, W., Seeberger, P. H., Silva, D. V., Wuhrer, M. & Kolarich, D. (2016) The Art of Destruction: Optimizing Collision Energies in Quadrupole-Time of Flight (Q-TOF) Instruments for Glycopeptide-Based Glycoproteomics, *Journal of the American Society for Mass Spectrometry*. **27**, 507-19.
108. Scott, N. E., Parker, B. L., Connolly, A. M., Paulech, J., Edwards, A. V., Crossett, B., Falconer, L., Kolarich, D., Djordjevic, S. P., Hojrup, P., Packer, N. H., Larsen, M. R. & Cordwell, S. J. (2011) Simultaneous glycan-peptide characterization using hydrophilic interaction chromatography and parallel fragmentation by CID, higher energy collisional dissociation, and electron transfer dissociation MS applied to the N-linked glycoproteome of *Campylobacter jejuni*, *Molecular & cellular proteomics : MCP*. **10**, M000031-MCP201.
109. Alagesan, K., Hoffmann, M., Rapp, E. & Kolarich, D. (2020) Glycoproteomics Technologies in Glycobiotechnology.
110. Klink (2011) MS Solutions #9: Peptide Sequencing with Electrospray LC/MS Part 1: Ion Types and Nomenclature in
111. Vékey, K., Ozohanics, O., Tóth, E., Jekó, A., Révész, Á., Krenyácz, J. & Drahos, L. (2013) Fragmentation characteristics of glycopeptides, *International Journal of Mass Spectrometry*. **345-347**, 71-79.
112. Lee, L. Y., Moh, E. S., Parker, B. L., Bern, M., Packer, N. H. & Thaysen-Andersen, M. (2016) Towards Automated N-Glycopeptide Identification in Glycoproteomics, *Journal of proteome research*.
113. Bern, M., Kil, Y. J. & Becker, C. (2012) Byonic: advanced peptide and protein identification software, *Current protocols in bioinformatics*. **40**, 13.20. 1-13.20. 14.
114. Ongay, S., Boichenko, A., Govorukhina, N. & Bischoff, R. (2012) Glycopeptide enrichment and separation for protein glycosylation analysis, *Journal of separation science*. **35**, 2341-72.
115. Badr, H. A., Alsadek, D. M., Darwish, A. A., Elsayed, A. I., Bekmanov, B. O., Khussainova, E. M., Zhang, X., Cho, W. C., Djansugurova, L. B. & Li, C. Z. (2014) Lectin approaches for glycoproteomics in FDA-approved cancer biomarkers, *Expert review of proteomics*. **11**, 227-36.
116. Lee, A., Nakano, M., Hincapie, M., Kolarich, D., Baker, M. S., Hancock, W. S. & Packer, N. H. (2010) The lectin riddle: glycoproteins fractionated from complex mixtures have similar glycomic profiles, *Omics: a journal of integrative biology*. **14**, 487-499.
117. Thaysen-Andersen, M., Mysling, S. & Højrup, P. (2009) Site-specific glycoprofiling of N-linked glycopeptides using MALDI-TOF MS: strong correlation between signal strength and glycoform quantities, *Analytical chemistry*. **81**, 3933-3943.
118. Paules, C. & Subbarao, K. (2017) Influenza, *Lancet*. **390**, 697-708.
119. Su, S., Fu, X. L., Li, G. R., Kerlin, F. & Veit, M. (2017) Novel Influenza D virus: Epidemiology, pathology, evolution and biological characteristics, *Virulence*. **8**, 1580-1591.
120. Biggerstaff, M., Cauchemez, S., Reed, C., Gambhir, M. & Finelli, L. (2014) Estimates of the reproduction number for seasonal, pandemic, and zoonotic influenza: a systematic review of the literature, *Bmc Infect Dis*. **14**.
-

-
121. Vajda, J., Weber, D., Brekel, D., Hundt, B. & Muller, E. (2016) Size distribution analysis of influenza virus particles using size exclusion chromatography, *Journal of Chromatography A*. **1465**, 117-125.
122. Stanley, W. M. (1944) The Size of Influenza Virus, *J Exp Med*. **79**, 267-283.
123. Hedestam, G. B. K., Fouchier, R. A. M., Phogat, S., Burton, D. R., Sodroski, J. & Wyatt, R. T. (2008) The challenges of eliciting neutralizing antibodies to HIV-1 and to influenza virus, *Nature Reviews Microbiology*. **6**, 143-155.
124. Lambert, L. C. & Fauci, A. S. (2010) Influenza vaccines for the future, *N Engl J Med*. **363**, 2036-44.
125. Hay, A. J., Gregory, V., Douglas, A. R. & Lin, Y. P. (2001) The evolution of human influenza viruses, *Philos Trans R Soc Lond B Biol Sci*. **356**, 1861-70.
126. Wu, N. C. & Wilson, I. A. (2020) Influenza hemagglutinin structures and antibody recognition, *Cold Spring Harbor perspectives in medicine*. **10**, a038778.
127. de Vries, R. P., de Vries, E., Bosch, B. J., de Groot, R. J., Rottier, P. J. & de Haan, C. A. (2010) The influenza A virus hemagglutinin glycosylation state affects receptor-binding specificity, *Virology*. **403**, 17-25.
128. Wei, C.-J., Boyington, J. C., Dai, K., Houser, K. V., Pearce, M. B., Kong, W.-P., Yang, Z.-y., Tumpey, T. M. & Nabel, G. J. (2010) Cross-neutralization of 1918 and 2009 influenza viruses: role of glycans in viral evolution and vaccine design, *Science translational medicine*. **2**, 24ra21-24ra21.
129. Skehel, J., Stevens, D., Daniels, R., Douglas, A., Knossow, M., Wilson, I. & Wiley, D. (1984) A carbohydrate side chain on hemagglutinins of Hong Kong influenza viruses inhibits recognition by a monoclonal antibody, *Proceedings of the National Academy of Sciences*. **81**, 1779-1783.
130. Matrosovich, M. N., Matrosovich, T. Y., Gray, T., Roberts, N. A. & Klenk, H. D. (2004) Neuraminidase is important for the initiation of influenza virus infection in human airway epithelium, *Journal of virology*. **78**, 12665-12667.
131. Wagner, R., Wolff, T., Herwig, A., Pleschka, S. & Klenk, H.-D. (2000) Interdependence of hemagglutinin glycosylation and neuraminidase as regulators of influenza virus growth: a study by reverse genetics, *Journal of virology*. **74**, 6316-6323.
132. Li, S., Schulman, J., Itamura, S. & Palese, P. (1993) Glycosylation of neuraminidase determines the neurovirulence of influenza A/WSN/33 virus, *Journal of virology*. **67**, 6667-6673.
133. Butler, M. & Reichl, I. U. (2017) Animal Cell Expression Systems, *Advances in Glycobiotechnology*, Springer, Cham, 1-36.
134. York, I. A., Stevens, J. & Alymova, I. V. (2019) Influenza virus N-linked glycosylation and innate immunity, *Bioscience reports*. **39**, BSR20171505.
135. Weis, W., Brown, J., Cusack, S., Paulson, J., Skehel, J. & Wiley, D. (1988) Structure of the influenza virus haemagglutinin complexed with its receptor, sialic acid, *Nature*. **333**, 426-431.
136. Hamilton, B. S., Whittaker, G. R. & Daniel, S. (2012) Influenza virus-mediated membrane fusion: determinants of hemagglutinin fusogenic activity and experimental approaches for assessing virus fusion, *Viruses*. **4**, 1144-1168.
137. Sakai, T., Nishimura, S. I., Naito, T. & Saito, M. (2017) Influenza A virus hemagglutinin and neuraminidase act as novel motile machinery, *Scientific reports*. **7**, 45043.
138. Pinto, L. H. & Lamb, R. A. (2006) The M2 proton channels of influenza A and B viruses, *Journal of Biological Chemistry*. **281**, 8997-9000.
139. Bullough, P. A., Hughson, F. M., Skehel, J. J. & Wiley, D. C. (1994) Structure of influenza haemagglutinin at the pH of membrane fusion, *Nature*. **371**, 37-43.
-

-
140. Bui, M., Whittaker, G. & Helenius, A. (1996) Effect of M1 protein and low pH on nuclear transport of influenza virus ribonucleoproteins, *Journal of virology*. **70**, 8391-8401.
141. Bottcher-Friebertshauser, E., Garten, W., Matrosovich, M. & Klenk, H. D. (2014) The hemagglutinin: a determinant of pathogenicity, *Current topics in microbiology and immunology*. **385**, 3-34.
142. Harrison, S. C. (2015) Viral membrane fusion, *Virology*. **479**, 498-507.
143. White, J. M. & Whittaker, G. R. (2016) Fusion of enveloped viruses in endosomes, *Traffic*. **17**, 593-614.
144. Einfeld, A. J., Neumann, G. & Kawaoka, Y. (2015) At the centre: influenza A virus ribonucleoproteins, *Nature Reviews Microbiology*. **13**, 28-41.
145. Dou, D., Revol, R., Östbye, H., Wang, H. & Daniels, R. (2018) Influenza A virus cell entry, replication, virion assembly and movement, *Frontiers in immunology*. **9**, 1581.
146. Pflug, A., Lukarska, M., Resa-Infante, P., Reich, S. & Cusack, S. (2017) Structural insights into RNA synthesis by the influenza virus transcription-replication machine, *Virus research*. **234**, 103-117.
147. York, A. & Fodor, E. (2013) Biogenesis, assembly, and export of viral messenger ribonucleoproteins in the influenza A virus infected cell, *RNA biology*. **10**, 1274-1282.
148. Dou, D., da Silva, D. V., Nordholm, J., Wang, H. & Daniels, R. (2014) Type II transmembrane domain hydrophobicity dictates the cotranslational dependence for inversion, *Molecular biology of the cell*. **25**, 3363-3374.
149. Daniels, R., Kurowski, B., Johnson, A. E. & Hebert, D. N. (2003) N-linked glycans direct the cotranslational folding pathway of influenza hemagglutinin, *Molecular cell*. **11**, 79-90.
150. Wang, N., Glidden, E. J., Murphy, S. R., Pearse, B. R. & Hebert, D. N. (2008) The cotranslational maturation program for the type II membrane glycoprotein influenza neuraminidase, *Journal of Biological Chemistry*. **283**, 33826-33837.
151. Klenk, H.-D., Rott, R., Orlich, M. & Blödorn, J. (1975) Activation of influenza A viruses by trypsin treatment, *Virology*. **68**, 426-439.
152. Trombetta, C. M., Kistner, O., Montomoli, E., Viviani, S. & Marchi, S. (2022) Influenza Viruses and Vaccines: The Role of Vaccine Effectiveness Studies for Evaluation of the Benefits of Influenza Vaccines, *Vaccines*. **10**, 714.
153. Houser, K. & Subbarao, K. (2015) Influenza vaccines: challenges and solutions, *Cell host & microbe*. **17**, 295-300.
154. Schild, G., Oxford, J., De Jong, J. & Webster, R. (1983) Evidence for host-cell selection of influenza virus antigenic variants, *Nature*. **303**, 706-709.
155. Centers for Disease Control and Prevention, N. C. f. I. a. R. D. N. (2019) Influenza vaccines — United States, 2019–20 influenza season in
156. (2018) Quality Guidelines : ICH (2014) in
157. (2017) EMA - Guideline on Influenza Vaccines - Quality module (2017) in
158. Scorza, F. B. & Pardi, N. (2018) New kids on the block: RNA-based influenza virus vaccines, *Vaccines*. **6**, 20.
159. Bartley, J. M., Cadar, A. N. & Martin, D. E. (2021) Better, faster, stronger: mRNA vaccines show promise for influenza vaccination in older adults, *Immunological Investigations*. **50**, 810-820.
160. Pralow, A., Hoffmann, M., Nguyen-Khuong, T., Pioch, M., Hennig, R., Genzel, Y., Rapp, E. & Reichl, U. (2021) Comprehensive N-glycosylation analysis of the influenza A virus proteins HA and NA from adherent and suspension MDCK cells, *The FEBS journal*.
161. Das, S. R., Puigbò, P., Hensley, S. E., Hurt, D. E., Bennink, J. R. & Yewdell, J. W. (2010) Glycosylation focuses sequence variation in the influenza A virus H1 hemagglutinin globular domain, *PLoS pathogens*. **6**, e1001211.
-

162. Hebert, D. N., Zhang, J.-X., Chen, W., Foellmer, B. & Helenius, A. (1997) The number and location of glycans on influenza hemagglutinin determine folding and association with calnexin and calreticulin, *The Journal of cell biology*. **139**, 613-623.
163. Hensley, S. E., Das, S. R., Bailey, A. L., Schmidt, L. M., Hickman, H. D., Jayaraman, A., Viswanathan, K., Raman, R., Sasisekharan, R. & Bennink, J. R. (2009) Hemagglutinin receptor binding avidity drives influenza A virus antigenic drift, *Science*. **326**, 734-736.
164. Wanzeck, K., Boyd, K. L. & McCullers, J. A. (2011) Glycan shielding of the influenza virus hemagglutinin contributes to immunopathology in mice, *American journal of respiratory and critical care medicine*. **183**, 767-773.
165. Alymova, I. V., York, I. A., Air, G. M., Cipollo, J. F., Gulati, S., Baranovich, T., Kumar, A., Zeng, H., Gansebom, S. & McCullers, J. A. (2016) Glycosylation changes in the globular head of H3N2 influenza hemagglutinin modulate receptor binding without affecting virus virulence, *Scientific reports*. **6**, 1-15.
166. Caton, A. J., Brownlee, G. G., Yewdell, J. W. & Gerhard, W. (1982) The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype), *Cell*. **31**, 417-27.
167. Temoltzin-Palacios, F. & Thomas, D. B. (1994) Modulation of immunodominant sites in influenza hemagglutinin compromise antigenic variation and select receptor-binding variant viruses, *J Exp Med*. **179**, 1719-1724.
168. Das, S. R., Hensley, S. E., David, A., Schmidt, L., Gibbs, J. S., Puigbò, P., Ince, W. L., Bennink, J. R. & Yewdell, J. W. (2011) Fitness costs limit influenza A virus hemagglutinin glycosylation as an immune evasion strategy, *Proceedings of the National Academy of Sciences*. **108**, E1417-E1422.
169. Tate, M. D., Job, E. R., Deng, Y. M., Gunalan, V., Maurer-Stroh, S. & Reading, P. C. (2014) Playing hide and seek: how glycosylation of the influenza virus hemagglutinin can modulate the immune response to infection, *Viruses*. **6**, 1294-316.
170. Kosik, I., Ince, W. L., Gentles, L. E., Oler, A. J., Kosikova, M., Angel, M., Magadán, J. G., Xie, H., Brooke, C. B. & Yewdell, J. W. (2018) Influenza A virus hemagglutinin glycosylation compensates for antibody escape fitness costs, *PLoS pathogens*. **14**, e1006796.
171. White, M. R., Boland, P., Tecele, T., Gantz, D., Sorenson, G., Tornoe, I., Holmskov, U., McDonald, B., Crouch, E. C. & Hartshorn, K. L. (2010) Enhancement of antiviral activity of collectin trimers through cross-linking and mutagenesis of the carbohydrate recognition domain, *Journal of innate immunity*. **2**, 267-279.
172. Crouch, E., Nikolaidis, N., McCormack, F. X., McDonald, B., Allen, K., Rynkiewicz, M. J., Cafarella, T. M., White, M., Lewnard, K. & Leymarie, N. (2011) Mutagenesis of surfactant protein D informed by evolution and x-ray crystallography enhances defenses against influenza A virus in vivo, *Journal of Biological Chemistry*. **286**, 40681-40692.
173. Tseng, Y. C., Wu, C. Y., Liu, M. L., Chen, T. H., Chiang, W. L., Yu, Y. H., Jan, J. T., Lin, K. I., Wong, C. H. & Ma, C. (2019) Egg-based influenza split virus vaccine with monoglycosylation induces cross-strain protection against influenza virus infections, *Proceedings of the National Academy of Sciences of the United States of America*. **116**, 4200-4205.
174. Wang, C. C., Chen, J. R., Tseng, Y. C., Hsu, C. H., Hung, Y. F., Chen, S. W., Chen, C. M., Khoo, K. H., Cheng, T. J., Cheng, Y. S., Jan, J. T., Wu, C. Y., Ma, C. & Wong, C. H. (2009) Glycans on influenza hemagglutinin affect receptor binding and immune response, *Proceedings of the National Academy of Sciences of the United States of America*. **106**, 18137-42.
175. Lee, J., Papparoditis, P., Horton, A. P., Fruhwirth, A., McDaniel, J. R., Jung, J., Boutz, D. R., Hussein, D. A., Tanno, Y., Pappas, L., Ippolito, G. C., Corti, D., Lanzavecchia, A. & Georgiou, G. (2019) Persistent Antibody Clonotypes Dominate the Serum Response to Influenza over Multiple Years and Repeated Vaccinations, *Cell Host Microbe*. **25**, 367-376 e5.

176. Lin, S. C., Jan, J. T., Dionne, B., Butler, M., Huang, M. H., Wu, C. Y., Wong, C. H. & Wu, S. C. (2013) Different immunity elicited by recombinant H5N1 hemagglutinin proteins containing pauci-mannose, high-mannose, or complex type *N*-glycans, *PLoS One*. **8**, e66719.
177. Vigerust, D. J., Ulett, K. B., Boyd, K. L., Madsen, J., Hawgood, S. & McCullers, J. A. (2007) N-linked glycosylation attenuates H3N2 influenza viruses, *Journal of virology*. **81**, 8593-600.
178. Tate, M. D., Brooks, A. G. & Reading, P. C. (2011) Specific sites of N-linked glycosylation on the hemagglutinin of H1N1 subtype influenza A virus determine sensitivity to inhibitors of the innate immune system and virulence in mice, *Journal of immunology*. **187**, 1884-94.
179. Tate, M. D., Job, E. R., Brooks, A. G. & Reading, P. C. (2011) Glycosylation of the hemagglutinin modulates the sensitivity of H3N2 influenza viruses to innate proteins in airway secretions and virulence in mice, *Virology*. **413**, 84-92.
180. Abdel-Motal, U. M., Guay, H. M., Wigglesworth, K., Welsh, R. M. & Galili, U. (2007) Immunogenicity of influenza virus vaccine is increased by anti-gal-mediated targeting to antigen-presenting cells, *Journal of virology*. **81**, 9131-41.
181. An, Y., Parsons, L. M., Jankowska, E., Melnyk, D., Joshi, M. & Cipollo, J. F. (2019) *N*-Glycosylation of Seasonal Influenza Vaccine Hemagglutinins: Implication for potency testing and immune processing, *Journal of virology*. **93**, e01693-18.
182. Khatri, K., Klein, J. A., White, M. R., Grant, O. C., Leymarie, N., Woods, R. J., Hartshorn, K. L. & Zaia, J. (2016) Integrated omics and computational glycobiology reveal structural basis for Influenza A virus glycan microheterogeneity and host interactions, *Molecular & cellular proteomics : MCP*.
183. She, Y.-M., Farnsworth, A., Li, X. & Cyr, T. D. (2017) Topological *N*-glycosylation and site-specific *N*-glycan sulfation of influenza proteins in the highly expressed H1N1 candidate vaccines, *Scientific reports*. **7**, 10232.
184. Harvey, D. J. (2018) Mass spectrometric analysis of glycosylated viral proteins, *Expert review of proteomics*. **15**, 391-412.
185. Sharma, V. K., Sharma, I. & Glick, J. (2018) The expanding role of mass spectrometry in the field of vaccine development, *Mass spectrometry reviews*.
186. Hämmerling, F., Pieler, M. M., Hennig, R., Serve, A., Rapp, E., Wolff, M. W., Reichl, U. & Hubbuch, J. (2017) Influence of the production system on the surface properties of influenza A virus particles, *Engineering in Life Sciences*. **17**, 1071-1077.
187. Hussain, S., Miller, J., Harvey, D., Gu, Y., Rosenthal, P., Zitzmann, N. & McCauley, J. (2014) Strain-specific antiviral activity of iminosugars against human influenza A viruses, *Journal of Antimicrobial Chemotherapy*. **70**, 136-152.
188. She, Y.-M., Tam, R. Y., Li, X., Rosu-Myles, M. & Sauvé, S. (2020) Resolving Isomeric Structures of Native Glycans by Nanoflow Porous Graphitized Carbon Chromatography–Mass Spectrometry, *Analytical chemistry*. **92**, 14038-14046.
189. Zhu, B., Shen, J., Zhao, T., Jiang, H., Ma, T., Zhang, J., Dang, L., Gao, N., Hu, Y. & Shi, Y. (2019) Intact glycopeptide analysis of influenza A/H1N1/09 neuraminidase revealing the effects of host and glycosite location on site-specific glycan structures, *Proteomics*. **19**, 1800202.
190. Chen, W., Zhong, Y., Su, R., Qi, H., Deng, W., Sun, Y., Ma, T., Wang, X., Yu, H. & Wang, X. (2017) *N*-glycan profiles in H9N2 avian influenza viruses from chicken eggs and human embryonic lung fibroblast cells, *Journal of virological methods*. **249**, 10-20.
191. Butler, M. & Spearman, M. (2014) The choice of mammalian cell host and possibilities for glycosylation engineering, *Current opinion in biotechnology*. **30**, 107-12.

192. Nguyen-Khuong, T., Pralow, A., Reichl, U. & Rapp, E. (2018) Improvement of electrospray stability in negative ion mode for nano-PGC-LC-MS glycoanalysis via post-column make-up flow, *Glycoconjugate journal*. **35**, 499-509.
193. Kluge, S., Benndorf, D., Genzel, Y., Scharfenberg, K., Rapp, E. & Reichl, U. (2015) Monitoring changes in proteome during stepwise adaptation of a MDCK cell line from adherence to growth in suspension, *Vaccine*. **33**, 4269-80.
194. Serve, A., Pieler, M. M., Benndorf, D., Rapp, E., Wolff, M. W. & Reichl, U. (2015) Comparison of Influenza Virus Particle Purification Using Magnetic Sulfated Cellulose Particles with an Established Centrifugation Method for Analytics, *Analytical chemistry*.
195. Hennig, R., Rapp, E., Kottler, R., Cajic, S., Borowiak, M. & Reichl, U. (2015) N-Glycosylation Fingerprinting of Viral Glycoproteins by xCGE-LIF, *Methods in molecular biology*. **1331**, 123-43.
196. Jensen, P. H., Kolarich, D. & Packer, N. H. (2010) Mucin-type O-glycosylation - putting the pieces together, *Febs Journal*. **277**, 81-94.
197. Selman, M. H., Hemayatkar, M., Deelder, A. M. & Wuhrer, M. (2011) Cotton HILIC SPE microtips for microscale purification and enrichment of glycans and glycopeptides, *Analytical chemistry*. **83**, 2492-9.
198. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004) UCSF Chimera—a visualization system for exploratory research and analysis, *Journal of computational chemistry*. **25**, 1605-1612.
199. Garcia-Montoya, I. A., Cendon, T. S., Arevalo-Gallegos, S. & Rascon-Cruz, Q. (2012) Lactoferrin a multiple bioactive protein: an overview, *Biochimica et biophysica acta*. **1820**, 226-36.
200. Yu, T., Guo, C., Wang, J., Hao, P., Sui, S., Chen, X., Zhang, R., Wang, P., Yu, G., Zhang, L., Dai, Y. & Li, N. (2011) Comprehensive characterization of the site-specific N-glycosylation of wild-type and recombinant human lactoferrin expressed in the milk of transgenic cloned cattle, *Glycobiology*. **21**, 206-24.
201. Lapko, V. N., Smith, D. L. & Smith, J. B. (2000) Identification of an artifact in the mass spectrometry of proteins derivatized with iodoacetamide, *Journal of mass spectrometry : JMS*. **35**, 572-5.
202. Wilm, M. & Mann, M. (1996) Analytical properties of the nanoelectrospray ion source, *Analytical chemistry*. **68**, 1-8.
203. El-Faramawy, A., Siu, K. M. & Thomson, B. A. (2005) Efficiency of nano-electrospray ionization, *Journal of the American Society for Mass Spectrometry*. **16**, 1702-1707.
204. Kawasaki, N., Ohta, M., Hyuga, S., Hyuga, M. & Hayakawa, T. (2000) Application of liquid chromatography/mass spectrometry and liquid chromatography with tandem mass spectrometry to the analysis of the site-specific carbohydrate heterogeneity in erythropoietin, *Analytical Biochemistry*. **285**, 82-91.
205. Kawasaki, N., Itoh, S., Ohta, M. & Hayakawa, T. (2003) Microanalysis of N-linked oligosaccharides in a glycoprotein by capillary liquid chromatography/mass spectrometry and liquid chromatography/tandem mass spectrometry, *Anal Biochem*. **316**, 15-22.
206. Karlsson, N. G., Wilson, N. L., Wirth, H. J., Dawes, P., Joshi, H. & Packer, N. H. (2004) Negative ion graphitised carbon nano-liquid chromatography/mass spectrometry increases sensitivity for glycoprotein oligosaccharide analysis, *Rapid communications in mass spectrometry : RCM*. **18**, 2282-92.
207. Hershberger, L. W., Callis, J. B. & Christian, G. D. (1979) Sub-Microliter Flow-through Cuvette for Fluorescence Monitoring of High-Performance Liquid-Chromatographic Effluents, *Analytical chemistry*. **51**, 1444-1446.

208. Bruins, A. P., Covey, T. R. & Henion, J. D. (1987) Ion Spray Interface for Combined Liquid Chromatography/Atmospheric Pressure Ionization Mass-Spectrometry, *Analytical chemistry*. **59**, 2642-2646.
209. Straub, R. F. & Voyksner, R. D. (1993) Negative-Ion Formation in Electrospray Mass-Spectrometry, *Journal of the American Society for Mass Spectrometry*. **4**, 578-587.
210. Kohler, M. & Leary, J. A. (1995) Lc/Ms/Ms of Carbohydrates with Postcolumn Addition of Metal Chlorides Using a Triaxial Electrospray Probe, *Analytical chemistry*. **67**, 3501-3508.
211. Marginean, I., Kronewitter, S. R., Moore, R. J., Slysz, G. W., Monroe, M. E., Anderson, G., Tang, K. Q. & Smith, R. D. (2012) Improving N-Glycan Coverage using HPLC-MS with Electrospray Ionization at Subambient Pressure, *Analytical chemistry*. **84**, 9208-9213.
212. Ikonomou, M. G., Blades, A. T. & Kebarle, P. (1991) Electrospray Mass-Spectrometry of Methanol and Water Solutions Suppression of Electric-Discharge with Sf6 Gas, *Journal of the American Society for Mass Spectrometry*. **2**, 497-505.
213. Kebarle, P. (2000) A brief overview of the present status of the mechanisms involved in electrospray mass spectrometry, *Journal of Mass Spectrometry*. **35**, 804-817.
214. Kurokawa, T., Wuhrer, M., Lochnit, G., Geyer, H., Markl, J. & Geyer, R. (2002) Hemocyanin from the keyhole limpet *Megathura crenulata* (KLH) carries a novel type of N-glycans with Gal (β 1-6) Man-motifs, *European journal of biochemistry*. **269**, 5459-5473.
215. Schmidt, A., Karas, M. & Dulcks, T. (2003) Effect of different solution flow rates on analyte ion signals in nano-ESI MS, or: When does ESI turn into nano-ESI?, *Journal of the American Society for Mass Spectrometry*. **14**, 492-500.
216. Kebarle, P. & Verkerk, U. H. (2009) Electrospray: From Ions in Solution to Ions in the Gas Phase, What We Know Now, *Mass spectrometry reviews*. **28**, 898-917.
217. Yamashita, M. & Fenn, J. B. (1984) Negative-Ion Production with the Electrospray Ion-Source, *J Phys Chem-Us*. **88**, 4671-4675.
218. Yamashita, M. & Fenn, J. B. (1984) Electrospray Ion-Source - Another Variation on the Free-Jet Theme, *J Phys Chem-Us*. **88**, 4451-4459.
219. Staples, G. O., Naimy, H., Yin, H., Kileen, K., Kraiczek, K., Costello, C. E. & Zaia, J. (2010) Improved hydrophilic interaction chromatography LC/MS of heparinoids using a chip with postcolumn makeup flow, *Analytical chemistry*. **82**, 516-22.
220. Ni, W., Bones, J. & Karger, B. L. (2013) In-depth characterization of N-linked oligosaccharides using fluoride-mediated negative ion microfluidic chip LC-MS, *Analytical chemistry*. **85**, 3127-35.
221. Shaw, M. L., Stone, K. L., Colangelo, C. M., Gulcicek, E. E. & Palese, P. (2008) Cellular proteins in influenza virus particles, *PLoS pathogens*. **4**, e1000085.
222. Everest-Dass, A. V., Abrahams, J. L., Kolarich, D., Packer, N. H. & Campbell, M. P. (2013) Structural Feature Ions for Distinguishing N- and O-Linked Glycan Isomers by LC-ESI-IT MS/MS, *Journal of the American Society for Mass Spectrometry*. **24**, 895-906.
223. Domon, B. & Costello, C. E. (1988) A Systematic Nomenclature for Carbohydrate Fragmentations in Fab-Ms Ms Spectra of Glycoconjugates, *Glycoconjugate journal*. **5**, 397-409.
224. Everest-Dass, A. V., Kolarich, D., Campbell, M. P. & Packer, N. H. (2013) Tandem mass spectra of glycan substructures enable the multistage mass spectrometric identification of determinants on oligosaccharides, *Rapid Commun Mass Sp*. **27**, 931-939.
225. Cruz, E., Cain, J., Crossett, B. & Kayser, V. (2018) Site-specific glycosylation profile of influenza A (H1N1) hemagglutinin through tandem mass spectrometry, *Human vaccines & immunotherapeutics*. **14**, 508-517.
226. Liu, Y.-J., Wu, S.-L., Love, K. R. & Hancock, W. S. (2017) Characterization of site-specific glycosylation in influenza A virus hemagglutinin produced by *Spodoptera frugiperda* insect cell line, *Analytical chemistry*. **89**, 11036-11043.

-
227. Parsons, L. M., An, Y., de Vries, R. P., de Haan, C. A. & Cipollo, J. F. (2017) Glycosylation characterization of an influenza H5N7 hemagglutinin series with engineered glycosylation patterns: implications for structure–function relationships, *Journal of proteome research*. **16**, 398-412.
228. Sarkar, A., Bale, S., Behrens, A.-J., Kumar, S., Sharma, S. K., de Val, N., Pallesen, J., Irimia, A., Diwanji, D. C. & Stanfield, R. L. (2018) Structure of a cleavage-independent HIV Env recapitulates the glycoprotein architecture of the native cleaved trimer, *Nature communications*. **9**, 1-14.
229. De La Pena, A. T., Rantalainen, K., Cottrell, C. A., Allen, J. D., Van Gils, M. J., Torres, J. L., Crispin, M., Sanders, R. W. & Ward, A. B. (2019) Similarities and differences between native HIV-1 envelope glycoprotein trimers and stabilized soluble trimer mimetics, *PLoS pathogens*. **15**, e1007920.
230. Behrens, A.-J., Vasiljevic, S., Pritchard, L. K., Harvey, D. J., Andev, R. S., Krumm, S. A., Struwe, W. B., Cupo, A., Kumar, A. & Zitzmann, N. (2016) Composition and antigenic effects of individual glycan sites of a trimeric HIV-1 envelope glycoprotein, *Cell reports*. **14**, 2695-2706.
231. Watanabe, Y., Allen, J. D., Wrapp, D., McLellan, J. S. & Crispin, M. (2020) Site-specific glycan analysis of the SARS-CoV-2 spike, *Science*.
232. Galili, U. (2020) Amplifying immunogenicity of prospective Covid-19 vaccines by glycoengineering the coronavirus glycan-shield to present α -gal epitopes.
233. Cox, M. M. & Hollister, J. R. (2009) FluBlok, a next generation influenza vaccine manufactured in insect cells, *Biologicals*. **37**, 182-9.
234. Cox, M. M., Izikson, R., Post, P. & Dunkle, L. (2015) Safety, efficacy, and immunogenicity of Flublok in the prevention of seasonal influenza in adults, *Ther Adv Vaccines*. **3**, 97-108.
235. Almansa, R., Martínez-Orellana, P., Rico, L., Iglesias, V., Ortega, A., Vidaña, B., Martínez, J., Exposito, A., Montoya, M. & Bermejo-Martin, J. F. (2017) Pulmonary transcriptomic responses indicate a dual role of inflammation in pneumonia development and viral clearance during 2009 pandemic influenza infection, *PeerJ*. **5**, e3915.
236. Stavenhagen, K., Plomp, R. & Wuhrer, M. (2015) Site-Specific Protein N- and O-Glycosylation Analysis by a C18-Porous Graphitized Carbon-Liquid Chromatography-Electrospray Ionization Mass Spectrometry Approach Using Pronase Treated Glycopeptides, *Analytical chemistry*. **87**, 11691-9.
237. Katz, J. M. & Webster, R. G. (1989) Efficacy of inactivated influenza A virus (H3N2) vaccines grown in mammalian cells or embryonated eggs, *The Journal of infectious diseases*. **160**, 191-8.
238. Chen, Z., Glover, M. S. & Li, L. (2018) Recent advances in ion mobility–mass spectrometry for improved structural characterization of glycans and glycoconjugates, *Current opinion in chemical biology*. **42**, 1-8.

APPENDIX

A. Supplementary information

A.1. Improvement of the glycoproteomic toolbox with the discovery of a unique C-terminal cleavage specificity of flavastacin for *N*-glycosylated asparagine

Supplementary Table A.1-1: **Proteome Discoverer results of a nanoRP-LC-ESI-OT-MS²(HCD) measurement after sequential digestion of hLTF with trypsin and flavastacin.** MASCOT search against unspecific in-silico digestion of mammalian taxonomy (UniProt-KB/SwissProt database). Peptide sequences with the N-terminal cleavage of aspartic acid or deamidated asparagine (as indicated from the supplier of flavastacin) are highlighted in red. Adapted from Pralow et al. (2017) [81].

Sequence	# PSM	Modifications	MH+ [Da]	IonScore	ΔM [Da]
AVTLDDGGFIYEAGLAPYK	2		1884.96916	118	-0.005
VTLDDGGFIYEAGLAPYK	2		1813.93694	115	0.000
ADAVTLDDGGFIYEAGLAPYK	8		2071.03203	115	-0.005
DVTVLQNTDGNNEAWAK	3		1988.93157	106	0.001
GEADAMSLDGGVYTAGK	3		1804.80339	105	-0.002
DAMSLDGGVYTAGK	3		1547.70452	96	0.000
PFLNWTGPPEPIEAA	3	N4(Deamidated)	1639.79436	94	-0.006
MSLDGGVYTAGK	2		1361.63091	92	-0.009
LNWTGPPEPIEAAVAR	3	N2(Deamidated)	1721.88347	85	-0.002
AMSLDGGVYTAGK	4		1432.66716	85	-0.011
SLDGGVYTAGK	6		1230.59307	82	-0.007
KGGSFQLNELQGLK	8		1518.81743	82	-0.010
GGSFQLNELQGLK	7		1390.73186	81	-0.001
SDTSLTWNSVK	4	N8(Deamidated)	1238.58220	81	-0.008
KGGSFQLNELQGLK	5	N8(Deamidated)	1519.80046	78	-0.011
QLFGSPSGQK	2		1048.54045	77	-0.002
DGAGDVAFIR	6		1020.50469	77	-0.006
PVAAEVYGTTER	1		1191.59941	76	-0.001
DGGFIYEAGLAPYK	4		1500.73662	73	0.000
GGSFQLNELQGLK	2	N7(Deamidated)	1391.71404	73	-0.003
(L)DGGVYTAGK(C)	4		1030.48296	72	-0.001
IYEAGLAPYK	2		1124.59172	70	-0.007
LRPVAAEVYK	1		1074.58684	70	-0.007
LDGGVYTAGK	4		1143.56816	69	-0.002
TAGWNVPIGTLRPF	3		1528.82671	67	0.000
(L)NWTGPPEPIEAAVAR(F)	3	N1(Deamidated)	1608.80254	67	0.001
THYYAVAVVK	2		1150.62476	63	-0.001
RSDTSLTWNSVK	2		1393.70500	63	-0.002
LRPVAAEVYGTTER	4		1460.78345	63	-0.002
(L)DGGFIYEAGL(A)	1		1041.48467	63	-0.004
(L)DGGFIYEAGLAPY(K)	1		1372.63616	62	-0.006
TAGWNVPIGTLR	8		1284.70586	61	0.000
LNWTGPPEPIEAAVAR	2		1720.90019	60	-0.002
LRPVAAEVY	1		1017.56499	60	-0.008
VPIGTLRPF	2		999.59953	59	0.001
EPIEAAVAR	4		955.51976	59	-0.001
WTGPPEPIEAAVAR	1		1493.77837	59	0.004
YYGTGAFR	9		1097.49968	59	-0.005
SDTSLTWNSVK	2		1237.59917	58	-0.007
EDAIWNLLR	3	N6(Deamidated)	1130.58367	57	0.000
TAGWNVPIGTLRP	1		1381.75627	57	-0.002
GPQYVAGITNLK	1		1260.68547	57	-0.009
TAIQNLR	1		815.47423	56	0.001
VAGITNLK	1	N6(Deamidated)	816.48180	54	-0.001
VAGITNLK	9		815.49779	54	0.001
YVAGITNLK	1		978.56163	53	0.000
VPIGTLRPFL	5		1112.68254	51	0.000
(W)NVPIGTLRPFL(N)	1		1226.72600	51	0.000
VAGITNLKK	3		943.59502	49	0.001
LAVAVVR	2		727.48247	49	-0.000

APPENDIX

Supplementary Table A.1-2: **Proteome Discoverer results of a nanoRP-LC-ESI-OT-MS²(HCD) measurement after digestion of BSA with trypsin. MASCOT search against unspecific in-silico digestion of mammalian taxonomy (UniProt-KB/SwissProt database). Adapted from Pralow et al. (2017) [81].**

Sequence	# PSM	Modifications	MH+ [Da]	IonScore	ΔM [Da]
LGEYGFQNALIVR	642		1479.79180	112	-0.004
KVPQVSTPTLVEVSR	86		1639.93987	110	0.000
KVPQVSTPTLVEVSR	6	Q4(Deamidated)	1640.92436	110	0.003
MPCTEDYLSLILNR	6	M1(Oxidation); C3(Carbamidomethyl)	1740.82976	108	0.000
VPQVSTPTLVEVSR	2	Q3(Deamidated)	1512.82854	104	0.002
MPCTEDYLSLILNR	1	C3(Carbamidomethyl)	1724.83562	96	0.001
TMENFVAFVVDK	7		1399.69560	93	0.003
TMENFVAFVVDK	14	M3(Oxidation)	1415.68877	93	0.002
LGEYGFQNALIVR	12	N8(Deamidated)	1480.78032	92	0.006
VPQVSTPTLVEVSR	9		1511.84307	90	0.000
RHPEYAVSVLLR	44		1439.81119	90	-0.001
DAFLGSFLYEYSR	616		1567.74260	82	0.000
YICDNQDTISSK	3	C3(Carbamidomethyl)	1443.64311	80	0.001
GLVLIAFSQYLQCCPFDEHVK	9	C14(Carbamidomethyl)	2492.27080	77	0.006
DAIPENLPPLTADFAEDK	17		1955.96416	76	0.004
LVNELTEFAK	705		1163.63213	75	0.001
DAIPENLPPLTADFAEDK	2	N6(Deamidated)	1956.95293	75	0.009
HLVDEPQNLIK	652		1305.71573	73	0.000
FYAPPELLYYANK	21		1491.75286	71	0.001
PQVSTPTLVEVSR	1		1412.77715	70	0.003
RHPYFYAPPELLYYANK	5		2045.03081	69	0.003
LKPDNPTLCDEFKADEK	6	C9(Carbamidomethyl)	2019.96952	69	0.000
LFTFHADICTLPDTEK	3	C9(Carbamidomethyl)	1907.92160	68	0.001
KQTALVELLK	13		1142.71428	68	0.000
AFLGSFLYEYSR	3		1452.71758	68	0.002
LGEYGFQNAL	2		1111.54143	66	0.000
FYAPPELLYYANK	1	N11(Deamidated)	1492.74236	66	0.006
DAIPENLPPLTADFAEDKDVCK	3	C21(Carbamidomethyl)	2458.18448	66	0.004
SHCIAEVEK	3	C3(Carbamidomethyl)	1072.51055	60	0.000
LGSFLYEYSR	2		1234.61003	60	0.002
DDSPDLPK	4		886.41557	60	0.000
HPYFYAPPELLYYANK	6		1888.93045	58	0.004
HPEYAVSVLLR	9		1283.71208	57	0.001
VSTPTLVEVSR	2		1187.66374	56	0.001
SQYLQCCPFDEHVK	1	C7(Carbamidomethyl)	1778.82256	54	0.006
KQTALVELLK	2	Q2(Deamidated)	1143.69890	54	0.000
DAIPENLPPLTADFAEDK	1		2070.98711	54	0.000
RPCFSALTPDETVPK	3	C3(Carbamidomethyl)	1880.92327	53	0.005
LKPDNPTLCDEFK	1	C9(Carbamidomethyl)	1576.77104	53	0.003
QATALVELLK	2	Q1(Deamidated)	1015.60143	52	-0.002
HPYFYAPPELLYYAN	1		1760.83354	52	0.002
FKDLGEEHFK	30		1249.61954	52	0.002
QATALVELLK	1		1014.61962	51	0.000
PNTLCDEFK	1	C5(Carbamidomethyl)	1123.50969	51	0.001
RHPEYAVSVL	1		1170.62676	50	0.000
SHKDDSPDLPK	3		1238.60154	48	0.000
LVNELTEFAK	1	N3(Deamidated)	1164.61833	47	0.004
SLHTLFGDELCK	1	C11(Carbamidomethyl)	1419.69524	46	0.002
HPYFYAPPELLYY	1		1575.75444	46	0.003
AEFVEVTK	6		922.48815	46	0.001
KLVDLTK	2		917.56658	45	0.000
LSQKFPK	1		847.50133	42	-0.002
SEIAHR	1		712.37370	41	0.000
LVDLTK	13		789.47087	41	0.000
LYEYIAR	7		927.49590	40	0.001

APPENDIX

Supplementary Table A.1-3: **Proteome Discoverer results of a nanoRP-LC-ESI-OT-MS²(HCD) measurement after digestion of hLTF with trypsin. MASCOT search against unspecific in-silico digestion of mammalian taxonomy (UniProt-KB/SwissProt database). Adapted from Pralow et al. (2017) [81].**

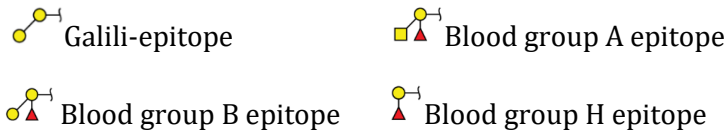
Sequence	# PSM	Modifications	MH+ [Da]	IonScore	ΔM [Da]
ADAVTLDGGFIYEAGLAPYK	1015		2071.03203	142	-0.003
AVTLDGGFIYEAGLAPYK	14		1884.96941	118	-0.005
GEADAMSLDGGVYVTAGK	48	M6(Oxidation)	1820.79863	117	-0.002
IDSGLYLGSYFTAIQNLR	102		2088.08208	116	-0.002
TLDGGFIYEAGLAPYK	6		1714.86785	115	-0.001
DVTVLQNTDGNNEAWAK	10		1988.93132	111	0.000
ESTVFEDLSDEAER	30		1626.71208	105	0.001
YLG PQYVAGITNLK	1126		1536.84294	104	0.000
CLAENAGDVAFVK	2	C1(Carbamidomethyl)	1393.67656	100	-0.002
IDSGLYLGSYFTAIQNLRK	6		2216.16996	100	-0.001
ADAVTLDGGFIYEAGLAPY	10		1942.94328	90	-0.002
LGSYFTAIQNLR	4		1439.76567	87	0.001
KGGSFQLNELQGLK	56		1518.82707	87	0.003
VPPRIDSGLYLGSYFTAIQNLR	6		2537.34977	86	-0.001
LADFALLCLDGK	6	C8(Carbamidomethyl)	1335.69707	85	0.000
YLG PQYVAGITNLK	4	Q5(Deamidated)	1537.82891	85	0.003
SGLYLGSYFTAIQNLR	6		1859.96196	84	-0.003
SVNGKEDAIWNLLR	10	N3(Deamidated)	1615.84221	83	-0.002
AMSLDGGVYVTAGK	10	M2(Oxidation)	1448.67143	83	-0.001
SLDGGVYVTAGK	12		1230.60100	82	-0.001
YLG PQYVAGITNLK	28		1664.93682	81	0.000
KGGSFQLNELQGLK	32	Q6(Deamidated)	1519.80997	81	0.000
DSPIQCIQIAIENR	6	C6(Carbamidomethyl)	1614.78862	77	-0.002
SVNGKEDAIWNLLR	2		1614.86150	77	0.002
GGSFQLNELQGLK	32		1390.73210	77	0.001
PVAAEVYGTERT	7		1191.60002	76	-0.001
RSDSL TWNSVK	8		1393.70696	75	0.000
DGAGDVAFIR	40		1020.51061	74	0.000
GGFYEAGLAPYK	8		1385.71037	74	0.000
GGSFQLNELQGLK	21	N7(Deamidated)	1391.71709	72	0.000
LRPVAAEVYGTERT	484		1460.78545	72	0.001
THYYAVAVVKKG	27		1335.74089	71	-0.001
IYEAGLAPYK	4		1124.59795	70	-0.001
GAGDVAFIR	4		905.48308	66	-0.001
SSQEPYFSYSGAFK	12		1597.71636	65	-0.001
YYGTGAFR	117		1097.50408	64	0.000
TGPPPEIAA VAR	12		1307.69414	64	-0.001
RPVEGYLAVAVVR	6		1428.83183	64	0.000
THYYAVAVVK	76		1150.62493	63	0.000
THYYAVAVVK	10		1278.71782	61	-0.003
SDSL TWNSVK	6		1237.60454	61	-0.001
LKQVLLHQQAK	22		1305.79851	59	-0.001
VPSHAVVAR	38		935.54033	57	0.000
RPVAAEVYGTERT	2		1347.70329	57	0.002
EDAIWNLLR	14		1129.59892	56	0.000
TAIQNLR	4		815.47435	56	0.001
KYLG PQYVAGITNLK	12		1664.93680	56	0.000
DGGVYVTAGK	2		1030.48345	53	-0.001
QLNELQGLK	2		1042.58916	52	0.000
YVAGITNLK	2		978.56017	52	-0.002
FQLFGSPSGQK	307		1195.60991	52	0.000
RKPVTEAR	6		956.56316	51	0.000
LRPVAAEVY	5		1017.57176	50	-0.001
DSAIGFSR	18		852.41979	50	-0.001
YLG PQYVAGITN	6		1295.66326	49	0.000
SFQLNELQGLK	4		1276.68999	49	0.000
PSHAVVAR	11		836.47368	47	0.000
MDKVER	1	M1(Oxidation)	793.38671	47	-0.001
GGVYVTAGK	2		915.45488	47	-0.002
FSYSGAFK	3		906.43413	47	-0.001
THYYAVAVVKKGG	9		1392.76201	46	-0.001
TAGWNVPIGTLR	4		1284.70500	46	-0.001
SAIGFSR	2		737.39324	46	-0.001
SPKFQLFGSPSGQKDLLFKD	4		2239.17917	46	0.003
FFSASCVP GADK	2	C6(Carbamidomethyl)	1285.58806	46	0.000
KSEEEVAAR	6		1018.51573	44	-0.001
GEADAMSLDGGVYVY	5	M6(Oxidation)	1463.59929	44	0.000

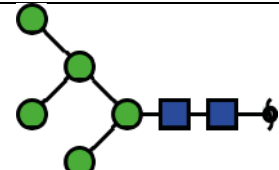
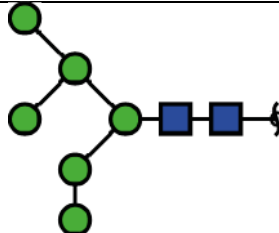
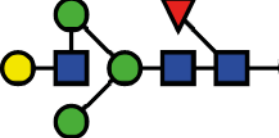
APPENDIX

Sequence	# PSM	Modifications	MH+ [Da]	IonScore	ΔM [Da]
QVLLHQAKFGRN	5		1538.85440	44	-0.001
QVLLHQAK	36		1064.61968	40	0.000
GPQYVAGITNLKK	1		1388.78960	38	0.000
YAVAVVK	1		749.45525	35	0.000

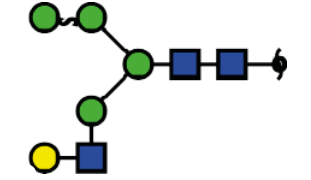
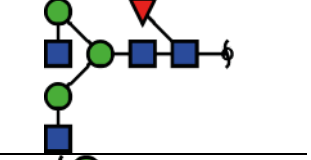
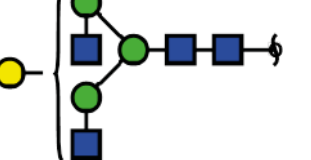
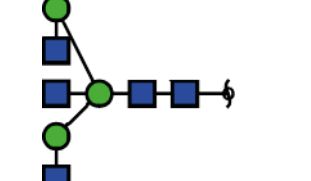
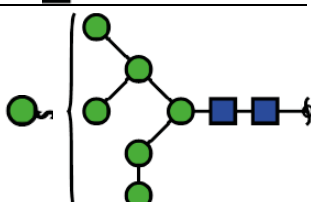
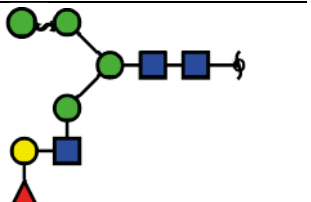
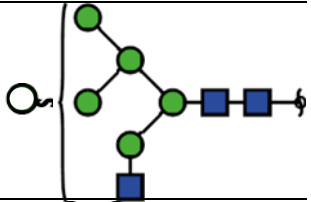
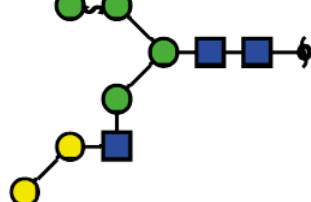
A.2. Comprehensive N-glycosylation analysis of the IAV proteins HA and NA from adherent and suspension MDCK cells

Supplementary Table A.2-1: **Viral N-glycome list of influenza A virus (IAV) produced in MDCK.ADH and MDCK.SUS2 cells.** All N-glycan fine-structures of IAV are listed as illustrations. For each structure, the corresponding glycan-composition, ID, double negatively charged molecular mass ($[M-2H]^{-2}$) and retention time (tR) (of one representative dataset) is given. Furthermore, for all five N-glycan samples ($n=3$ for IAV propagated in MDCK.ADH cells and $n=2$ for IAV propagated in MDCK.SUS2 cells) the number of samples with the detected structure is shown. The ID was established based on the N-glycan type (M = high mannose, C = complex and H = hybrid, X = not identified) and the N-glycan mass (numbered from low to high molecular mass). Symbolic representations of N-glycan structures were drawn with GlycoWorkbench Version 1.1, following the Symbol Nomenclature for Graphical Representations of Glycans [6]. Adapted from Pralow et al. (2021) [160].



N-Glycan Composition	ID	$[M-2H]^{-2}$	tR		MDCK.AD H cell	MDCK.SUS 2 cell	Structure
			[M-2H] ⁻²	(min)			
(Hex) ₂ + (Man) ₃ (GlcNAc) ₂	M01	617.22	32.8	3	2		
(Hex) ₃ + (Man) ₃ (GlcNAc) ₂	M02	698.24	29.8	3	2		
(Hex) ₁ (HexNAc) ₁ (dHex) ₁ + (Man) ₃ (GlcNAc) ₂	C01	710.76	36.8	0	2		

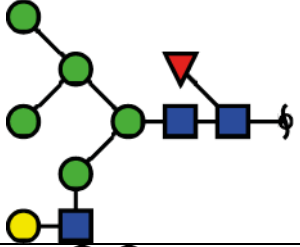
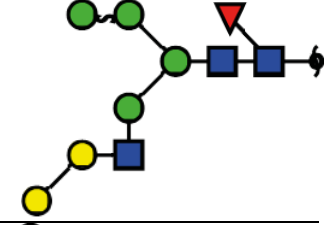
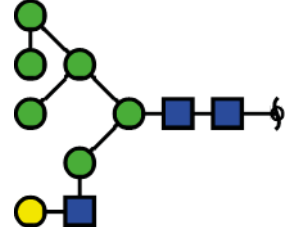
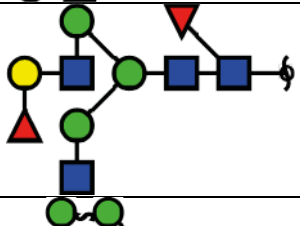
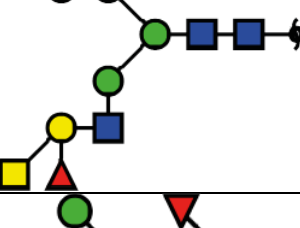
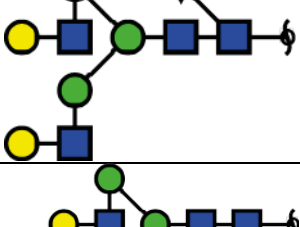
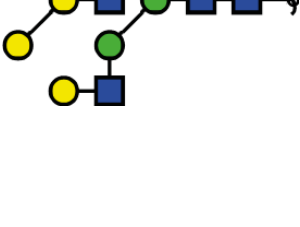
APPENDIX

(Hex) ₂ (HexNAc) ₁ + (Man) ₃ (GlcNAc) ₂	C02	718.76	29.3	3	2	
(HexNAc) ₂ (dHex) ₁ + (Man) ₃ (GlcNAc) ₂	C03	731.22	30.1	0	2	
(Hex) ₁ (HexNAc) ₂ + (Man) ₃ (GlcNAc) ₂	C04	739.27	30.7	0	2	
(HexNAc) ₃ + (Man) ₃ (GlcNAc) ₂	C05	759.78	25.4	3	2	
(Hex) ₄ + (Man) ₃ (GlcNAc) ₂	M03	779.26	29.4	3	2	
(Hex) ₂ (HexNAc) ₁ (dHex) ₁ + (Man) ₃ (GlcNAc) ₂	H01	791.32	31.0	3	2	
(Hex) ₃ (HexNAc) ₁ + (Man) ₃ (GlcNAc) ₂	H02	799.77	30.9	3	2	
	C/H 01	799.78	31.8	3	0	

APPENDIX

(Hex)1 (HexNAc)2 (dHex)1 + (Man)3(GlcNAc)2	C06	812.30	33.2	2	2	
(Hex)2 (HexNAc)2 + (Man)3(GlcNAc)2	C07	820.30	31.6	3	2	
(HexNAc)3 (dHex)1 + (Man)3(GlcNAc)2	C08	832.81	28.7	3	2	
(Hex)1 (HexNAc)3 + (Man)3(GlcNAc)2	C09	840.81	27.0	3	2	
(Hex)5 + (Man)3(GlcNAc)2	M04	860.28	29.4	3	2	
(Hex)2 (HexNAc)1 (dHex)2 + (Man)3(GlcNAc)2	H03	864.81	33.6	3	2	
(Hex)3 (HexNAc)1 (dHex)1 + (Man)3(GlcNAc)2	H04	872.81	32.8	3	2	

APPENDIX

	H05	872.81	33.3	3	2	
	H06	872.81	33.9	3	0	
(Hex) ₄ (HexNAc) ₁ + (Man) ₃ (GlcNAc) ₂	H07	880.81	33.1	3	2	
(Hex) ₁ (HexNAc) ₂ (dHex) ₂ + (Man) ₃ (GlcNAc) ₂	C10	885.32	34.5	0	2	
(Hex) ₂ (HexNAc) ₂ (dHex) ₁ + (Man) ₃ (GlcNAc) ₂	H08	893.32	30.0	0	2	
	C11	893.32	34.0	3	2	
(Hex) ₃ (HexNAc) ₂ + (Man) ₃ (GlcNAc) ₂	C12	901.32	31.0	0	2	

APPENDIX

	C13	901.32	33.6	3	2	
(Hex)1 (HexNAc)3 (dHex)1 + (Man)3(GlcNAc)2	C14	913.84	29.7	3	2	
(Hex)2 (HexNAc)3 + (Man)3(GlcNAc)2	C15	921.83	27.8	3	2	
	C16	921.84	28.7	3	0	
(Hex)6 + (Man)3(GlcNAc)2	M05	941.31	29.6	3	2	
(Hex)2 (HexNAc)2 (dHex)2 + (Man)3(GlcNAc)2	H09	966.31	31.8	0	1	
	C17	966.35	35.2	3	2	
(Hex)3 (HexNAc)2 (dHex)1 + (Man)3(GlcNAc)2	H10	974.34	30.5	1	0	

APPENDIX

	C18	974.35	35.4	3	2	
(Hex)4 (HexNAc)2 + (Man)3(GlcNAc)2	C19	982.3	34.5	0	2	
(Hex)1 (HexNAc)3 (dHex)2 + (Man)3(GlcNAc)2	C20	986.86	32.0	3	0	
(Hex)2 (HexNAc)3 (dHex)1 + (Man)3(GlcNAc)2	C21	994.85	30.1	3	2	
	C22	994.86	31.2	3	2	
	C23	994.87	35.0	3	2	
	C24	994.87	35.5	0	2	2x
(Hex)3 (HexNAc)3 + (Man)3(GlcNAc)2	C25	1002.86	34.3	3	2	

APPENDIX

(Hex)7 + (Man)3(GlcNAc)2	M06	1022.34	30.5	0	2	
(Hex)2 (HexNAc)4 + (Man)3(GlcNAc)2	C26	1023.37	28.5	0	2	
(Hex)2 (HexNAc)2 (dHex)3 + (Man)3(GlcNAc)2	X01	1039.38	37.2	3	0	No sufficient data
(Hex)3 (HexNAc)2 (dHex)2 + (Man)3(GlcNAc)2	C27	1047.37	36.5	1	0	
(Hex)4 (HexNAc)2 (dHex)1 + (Man)3(GlcNAc)2	C28	1055.38	36.9	3	2	
(Hex)2 (HexNAc)3 (dHex)2 + (Man)3(GlcNAc)2	C29	1067.85	32.5	3	2	
	C30	1067.89	33.5	0	2	

APPENDIX

(Hex)3 (HexNAc)3 (dHex)1 + (Man)3(GlcNAc)2	C31	1075.87	36.4	3	2	
	C32	1075.89	31.9	3	0	
(Hex)2 (HexNAc)4 (dHex)1 + (Man)3(GlcNAc)2	C33	1096.40	31.0	3	2	
(Hex)3 (HexNAc)4 + (Man)3(GlcNAc)2	C34	1104.41	29.2	0	2	3x
	C35	1104.41	34.4	0	2	3x
(Hex)2 (HexNAc)3 (dHex)3 + (Man)3(GlcNAc)2	X02	1140.92	36.6	3	0	No sufficient data
(Hex)3 (HexNAc)3 (dHex)2 + (Man)3(GlcNAc)2	C36	1148.87	32.1	3	0	
	C37	1148.92	37.5	3	2	
(Hex)4 (HexNAc)3 (dHex)1 + (Man)3(GlcNAc)2	C38	1156.86	37.6	3	2	
(Hex)2 (HexNAc)4 (dHex)2 + (Man)3(GlcNAc)2	X03	1169.43	33.3	3	2	No sufficient data

APPENDIX

(Hex)3 (HexNAc)4 (dHex)1 + (Man)3(GlcNAc)2	C38	1177.41	31.3	3	2	3x
	C39	1177.42	32.5	3	2	2x
	C40	1177.43	33.7	0	2	2x
	C41	1177.44	36.4	3	2	3x
(Hex)4 (HexNAc)4 + (Man)3(GlcNAc)2	C42	1185.42	34.7	0	2	
(Hex)3 (HexNAc)3 (dHex)3 + (Man)3(GlcNAc)2	C43	1221.95	41.0	3	2	
(Hex)4 (HexNAc)3 (dHex)2 + (Man)3(GlcNAc)2	C44	1229.94	38.7	3	2	
(Hex)5 (HexNAc)3 (dHex)1 + (Man)3(GlcNAc)2	C45	1237.92	38.7	3	2	
(Hex)2 (HexNAc)4 (dHex)3 + (Man)3(GlcNAc)2	X04	1242.46	38.7	3	0	No sufficient data

APPENDIX

(Hex)3 (HexNAc)4 (dHex)2 + (Man)3(GlcNAc)2	C46	1250.46	33.7	3	2	
	C47	1250.46	35.0	3	2	
(Hex)4 (HexNAc)4 (dHex)1 + (Man)3(GlcNAc)2	C48	1258.44	36.4	0	2	
	C49	1258.45	39.2	0	2	
(Hex)5 (HexNAc)4 + (Man)3(GlcNAc)2	X05	1266.44	32.8	3	2	No sufficient data
(Hex)4 (HexNAc)3 (dHex)3 + (Man)3(GlcNAc)2	C50	1302.98	41.8	0	2	
(Hex)5 (HexNAc)3 (dHex)2 + (Man)3(GlcNAc)2	C51	1310.97	38.9	2	0	
(Hex)6 (HexNAc)3 (dHex)1 + (Man)3(GlcNAc)2	X06	1318.97	39.8	3	0	No sufficient data
(Hex)3 (HexNAc)4 (dHex)3 + (Man)3(GlcNAc)2	C52	1323.49	39.0	3	2	

APPENDIX

(Hex) ₄ (HexNAc) ₄ (dHex) ₂ + (Man) ₃ (GlcNAc) ₂	C53	1331.49	35.2	3	2	
(Hex) ₅ (HexNAc) ₄ (dHex) ₁ + (Man) ₃ (GlcNAc) ₂	C54	1339.48	34.0	3	2	
(Hex) ₃ (HexNAc) ₅ (dHex) ₂ + (Man) ₃ (GlcNAc) ₂	C55	1351.99	31.8	1	2	
(Hex) ₄ (HexNAc) ₄ (dHex) ₃ + (Man) ₃ (GlcNAc) ₂	C56	1404.52	40.7	2	2	
(Hex) ₅ (HexNAc) ₄ (dHex) ₂ + (Man) ₃ (GlcNAc) ₂	X07	1412.51	36.0	3	0	No sufficient data
(Hex) ₃ (HexNAc) ₅ (dHex) ₃ + (Man) ₃ (GlcNAc) ₂	C57	1425.03	34.9	0	2	
(Hex) ₄ (HexNAc) ₅ (dHex) ₂ + (Man) ₃ (GlcNAc) ₂	X08	1433.02	35.9	0	2	No sufficient data
(Hex) ₅ (HexNAc) ₅ (dHex) ₁ + (Man) ₃ (GlcNAc) ₂	C58	1441.01	39.8	0	2	
(Hex) ₇ (HexNAc) ₄ (dHex) ₁ + (Man) ₃ (GlcNAc) ₂	X09	1501.54	34.5	3	0	No sufficient data

APPENDIX

Supplementary Table A.2-2: *Site-specific glycopeptide analysis of influenza A virus glycoproteins propagated in MDCK.SUS2 and MDCK.ADH cells using nano-RP-LC MS(/MS)*. All identified N-glycopeptides for each site are listed. For MDCK.ADH (n=3) and MDCK.SUS2 (n=2) cells, the number of samples is specified, in which a respective N-glycopeptide could be detected. N-glycan compositions highlighted in red, could not be identified in the N-glycomic analysis. Adapted from Pralow et al. (2021) [160].

Protein	Site	N-glycan Composition	MDCK.ADH	MDCK.SUS2
			cell	cell
Hemagglutinin	N27/28	(Hex)3 (HexNAc)3 (dHex)1 + (Man)3(GlcNAc)2	2	2
		(Hex)4 (HexNAc)3 (dHex)1 + (Man)3(GlcNAc)2	3	2
		(Hex)4 (HexNAc)4 (dHex)1 + (Man)3(GlcNAc)2	3	2
		(Hex)5 (HexNAc)3 (dHex)1 + (Man)3(GlcNAc)2	3	/
		(Hex)5 (HexNAc)4 (dHex)1 + (Man)3(GlcNAc)2	2	/
		(Hex)3 (HexNAc)3 (dHex)2 + (Man)3(GlcNAc)2	3	2
		(Hex)3 (HexNAc)4 (dHex)2 + (Man)3(GlcNAc)2	3	2
		(Hex)4 (HexNAc)3 (dHex)2 + (Man)3(GlcNAc)2	3	/
		(Hex)4 (HexNAc)4 (dHex)2 + (Man)3(GlcNAc)2	3	/
		(Hex)5 (HexNAc)3 (dHex)2 + (Man)3(GlcNAc)2	3	/
		(Hex)5 (HexNAc)4 (dHex)2 + (Man)3(GlcNAc)2	3	/
		(Hex)6 (HexNAc)4 (dHex)2 + (Man)3(GlcNAc)2	1	/
		(Hex)3 (HexNAc)3 (dHex)3 + (Man)3(GlcNAc)2	2	/
		(Hex)3 (HexNAc)4 (dHex)3 + (Man)3(GlcNAc)2	3	/
		(Hex)4 (HexNAc)3 (dHex)3 + (Man)3(GlcNAc)2	3	/
(Hex)3 (HexNAc)5 (dHex)3 + (Man)3(GlcNAc)2	/	2		
	N40	(Hex)3 (HexNAc)3 (dHex)1 + (Man)3(GlcNAc)2	/	2
	N285	(Hex)7 (HexNAc)4 (dHex)1 + (Man)3(GlcNAc)2	2	/
		(Hex)2 (HexNAc)2 (dHex)1 + (Man)3(GlcNAc)2	/	2
		(Hex)2 (HexNAc)3 (dHex)1 + (Man)3(GlcNAc)2	/	2
		(Hex)3 (HexNAc)2 (dHex)1 + (Man)3(GlcNAc)2	/	2

APPENDIX

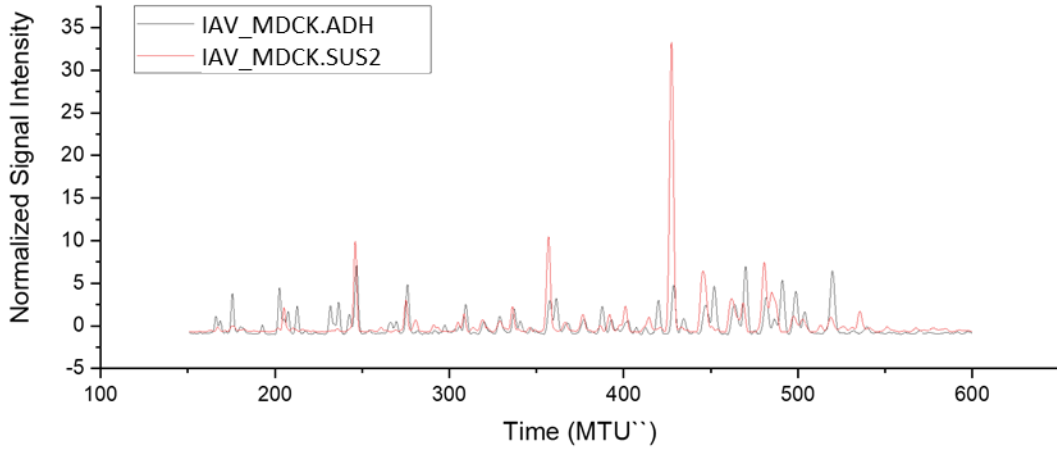
(Hex)3 (HexNAc)3 (dHex)1 + (Man)3(GlcNAc)2	3	2
(Hex)3 (HexNAc)4 (dHex)1 + (Man)3(GlcNAc)2	3	2
(Hex)4 (HexNAc)3 (dHex)1 + (Man)3(GlcNAc)2	3	2
(Hex)4 (HexNAc)4 (dHex)1 + (Man)3(GlcNAc)2	3	2
(Hex)5 (HexNAc)3 (dHex)1 + (Man)3(GlcNAc)2	3	2
(Hex)5 (HexNAc)4 (dHex)1 + (Man)3(GlcNAc)2	3	2
(Hex)5 (HexNAc)5 (dHex)1 + (Man)3(GlcNAc)2	2	/
(Hex)6 (HexNAc)3 (dHex)1 + (Man)3(GlcNAc)2	3	/
(Hex)6 (HexNAc)4 (dHex)1 + (Man)3(GlcNAc)2	3	/
(Hex)2 (HexNAc)2 (dHex)2 + (Man)3(GlcNAc)2	/	2
(Hex)2 (HexNAc)3 (dHex)2 + (Man)3(GlcNAc)2	3	2
(Hex)2 (HexNAc)4 (dHex)2 + (Man)3(GlcNAc)2	3	2
(Hex)3 (HexNAc)3 (dHex)2 + (Man)3(GlcNAc)2	3	2
(Hex)3 (HexNAc)4 (dHex)2 + (Man)3(GlcNAc)2	3	2
(Hex)3 (HexNAc)5 (dHex)2 + (Man)3(GlcNAc)2	/	2
(Hex)4 (HexNAc)3 (dHex)2 + (Man)3(GlcNAc)2	3	2
(Hex)4 (HexNAc)4 (dHex)2 + (Man)3(GlcNAc)2	3	2
(Hex)5 (HexNAc)3 (dHex)2 + (Man)3(GlcNAc)2	3	/
(Hex)5 (HexNAc)4 (dHex)2 + (Man)3(GlcNAc)2	3	/
(Hex)6 (HexNAc)4 (dHex)2 + (Man)3(GlcNAc)2	1	/
(Hex)2 (HexNAc)4 (dHex)3 + (Man)3(GlcNAc)2	3	/
(Hex)3 (HexNAc)3 (dHex)3 + (Man)3(GlcNAc)2	/	2
(Hex)3 (HexNAc)4 (dHex)3 + (Man)3(GlcNAc)2	3	2
(Hex)3 (HexNAc)5 (dHex)3 + (Man)3(GlcNAc)2	/	2
(Hex)4 (HexNAc)3 (dHex)3 + (Man)3(GlcNAc)2	3	/
(Hex)4 (HexNAc)4 (dHex)3 + (Man)3(GlcNAc)2	3	/

APPENDIX

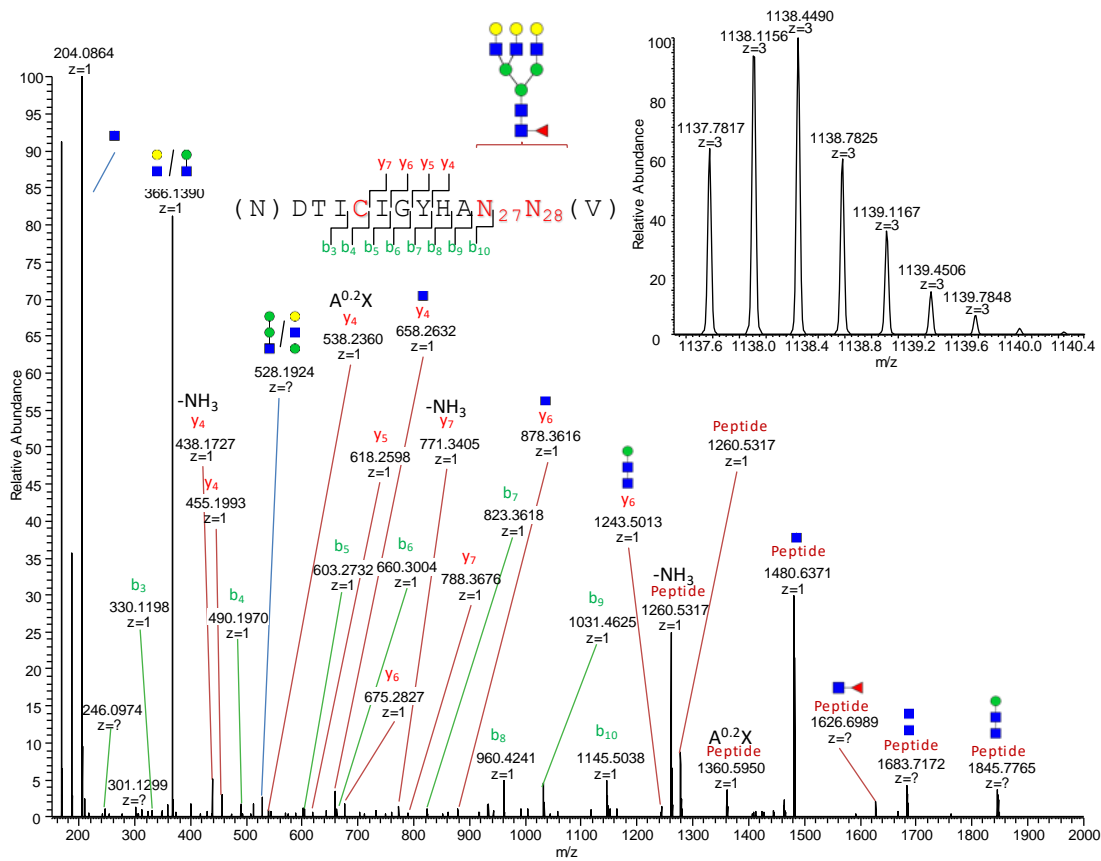
	(Hex)5 (HexNAc)4 + (Man)3(GlcNAc)2	1	/
N303	(Hex)7 (HexNAc)4 (dHex)1 + (Man)3(GlcNAc)2	2	/
	(Hex)1 (HexNAc)3 (dHex)1 + (Man)3(GlcNAc)2	3	2
	(Hex)2 (HexNAc)2 (dHex)1 + (Man)3(GlcNAc)2	1	2
	(Hex)2 (HexNAc)3 (dHex)1 + (Man)3(GlcNAc)2	3	2
	(Hex)2 (HexNAc)4 (dHex)1 + (Man)3(GlcNAc)2	/	2
	(Hex)3 (HexNAc)2 (dHex)1 + (Man)3(GlcNAc)2	/	1
	(Hex)3 (HexNAc)3 (dHex)1 + (Man)3(GlcNAc)2	/	2
	(Hex)3 (HexNAc)4 (dHex)1 + (Man)3(GlcNAc)2	3	1
	(Hex)4 (HexNAc)3 (dHex)1 + (Man)3(GlcNAc)2	3	2
	(Hex)4 (HexNAc)4 (dHex)1 + (Man)3(GlcNAc)2	3	/
	(Hex)5 (HexNAc)3 (dHex)1 + (Man)3(GlcNAc)2	3	2
	(Hex)5 (HexNAc)4 (dHex)1 + (Man)3(GlcNAc)2	3	/
	(Hex)6 (HexNAc)3 (dHex)1 + (Man)3(GlcNAc)2	3	/
	(Hex)2 (HexNAc)2 (dHex)2 + (Man)3(GlcNAc)2	/	2
	(Hex)2 (HexNAc)3 (dHex)2 + (Man)3(GlcNAc)2	3	2
	(Hex)2 (HexNAc)4 (dHex)2 + (Man)3(GlcNAc)2	3	2
	(Hex)3 (HexNAc)3 (dHex)2 + (Man)3(GlcNAc)2	3	2
	(Hex)3 (HexNAc)4 (dHex)2 + (Man)3(GlcNAc)2	3	2
	(Hex)3 (HexNAc)5 (dHex)2 + (Man)3(GlcNAc)2	/	2
	(Hex)4 (HexNAc)3 (dHex)2 + (Man)3(GlcNAc)2	2	2
	(Hex)4 (HexNAc)4 (dHex)2 + (Man)3(GlcNAc)2	3	2
	(Hex)4 (HexNAc)5 (dHex)2 + (Man)3(GlcNAc)2	/	2
	(Hex)5 (HexNAc)3 (dHex)2 + (Man)3(GlcNAc)2	3	/
	(Hex)5 (HexNAc)4 (dHex)2 + (Man)3(GlcNAc)2	3	/
	(Hex)3 (HexNAc)3 (dHex)3 + (Man)3(GlcNAc)2	/	2

APPENDIX

		(Hex)3 (HexNAc)4 (dHex)3 + (Man)3(GlcNAc)2	3	2
		(Hex)3 (HexNAc)5 (dHex)3 + (Man)3(GlcNAc)2	/	2
		(Hex)4 (HexNAc)4 (dHex)3 + (Man)3(GlcNAc)2	3	/
N497		(Hex)3 (HexNAc)3 (dHex)1 + (Man)3(GlcNAc)2	2	/
		(Hex)3 (HexNAc)3 (dHex)2 + (Man)3(GlcNAc)2	3	/
		(Hex)3 (HexNAc)4 (dHex)2 + (Man)3(GlcNAc)2	/	2
Neuraminidase	N73	(Hex)1 (HexNAc)3 (dHex)1 + (Man)3(GlcNAc)2	3	/
		(Hex)2 (HexNAc)1 (dHex)1 + (Man)3(GlcNAc)2	3	2
		(Hex)2 (HexNAc)2 (dHex)1 + (Man)3(GlcNAc)2	3	2
		(Hex)2 (HexNAc)3 (dHex)1 + (Man)3(GlcNAc)2	3	1
		(Hex)3 (HexNAc)2 (dHex)1 + (Man)3(GlcNAc)2	3	2
		(Hex)3 (HexNAc)3 (dHex)1 + (Man)3(GlcNAc)2	/	2
		(Hex)4 (HexNAc)2 (dHex)1 + (Man)3(GlcNAc)2	3	/
		(Hex)1 (HexNAc)3 (dHex)2 + (Man)3(GlcNAc)2	3	2
		(Hex)2 (HexNAc)1 (dHex)2 + (Man)3(GlcNAc)2	3	2
		(Hex)2 (HexNAc)2 (dHex)2 + (Man)3(GlcNAc)2	3	2
		(Hex)2 (HexNAc)3 (dHex)2 + (Man)3(GlcNAc)2	3	2
		(Hex)2 (HexNAc)4 (dHex)2 + (Man)3(GlcNAc)2	1	/
		(Hex)3 (HexNAc)2 (dHex)2 + (Man)3(GlcNAc)2	3	/
		(Hex)3 (HexNAc)3 (dHex)2 + (Man)3(GlcNAc)2	3	2
		(Hex)3 (HexNAc)4 (dHex)2 + (Man)3(GlcNAc)2	3	/
		(Hex)2 (HexNAc)2 (dHex)3 + (Man)3(GlcNAc)2	3	2
		(Hex)2 (HexNAc)3 (dHex)3 + (Man)3(GlcNAc)2	3	2
		(Hex)2 (HexNAc)4 (dHex)3 + (Man)3(GlcNAc)2	/	2
		(Hex)3 (HexNAc)4 (dHex)3 + (Man)3(GlcNAc)2	3	/

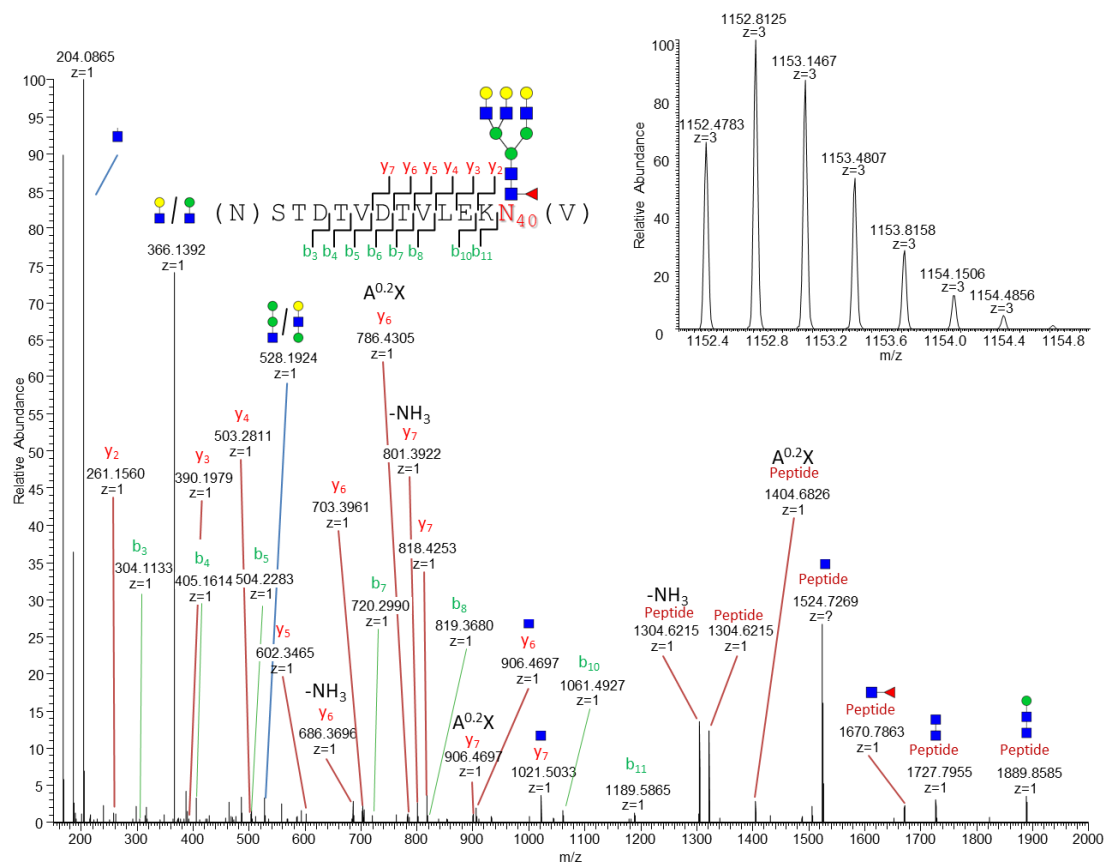


Supplementary Figure A.2-1: **Overlay of xCGE-LIF fingerprints (normalized electropherograms) of APTS labeled N-glycans derived from influenza A virus (IAV) propagated in MDCK.ADH and MDCK.SUS2 cells. Adapted from Pralow et al. (2021) [160].**

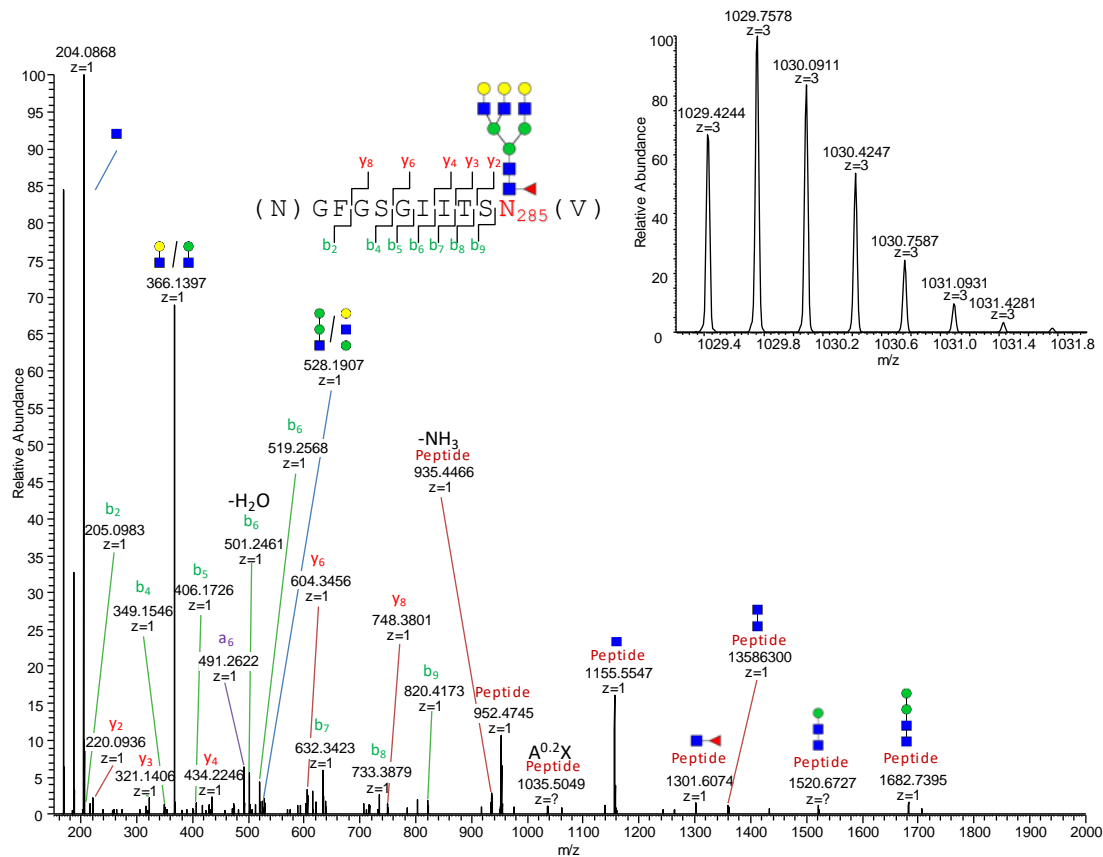


Supplementary Figure A.2-2: **Site-specific N-glycopeptide analysis using nano-RP-LC-MS(/MS). Fragment ion spectrum of a HA N-glycopeptide, derived from a sequential proteolytic digestion using trypsin and flavastacin. Amino acids highlighted in red indicate the carbamidomethylation of cysteine and the N-glycosylation of asparagine. For the validation of the (glyco-)peptide sequence, specific b- (green) and y- (red), as well as B-(oxonium-, blue) and Y-ions (peptide + glycan moiety) are annotated. The isotopic pattern of the precursor is located at the upper right corner. Symbolic representations of N-glycan structures were**

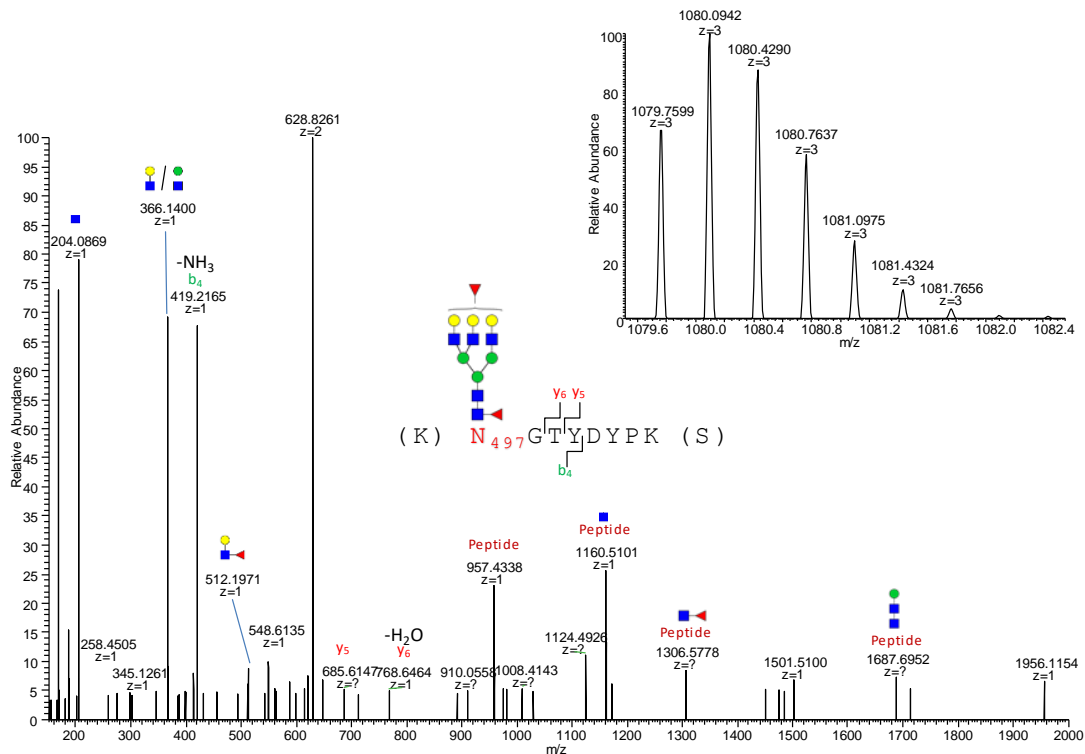
drawn with GlycoWorkbench Version 1.1, following the Symbol Nomenclature for Graphical Representations of Glycans [6]. Adapted from Pralow et al. (2021) [160].



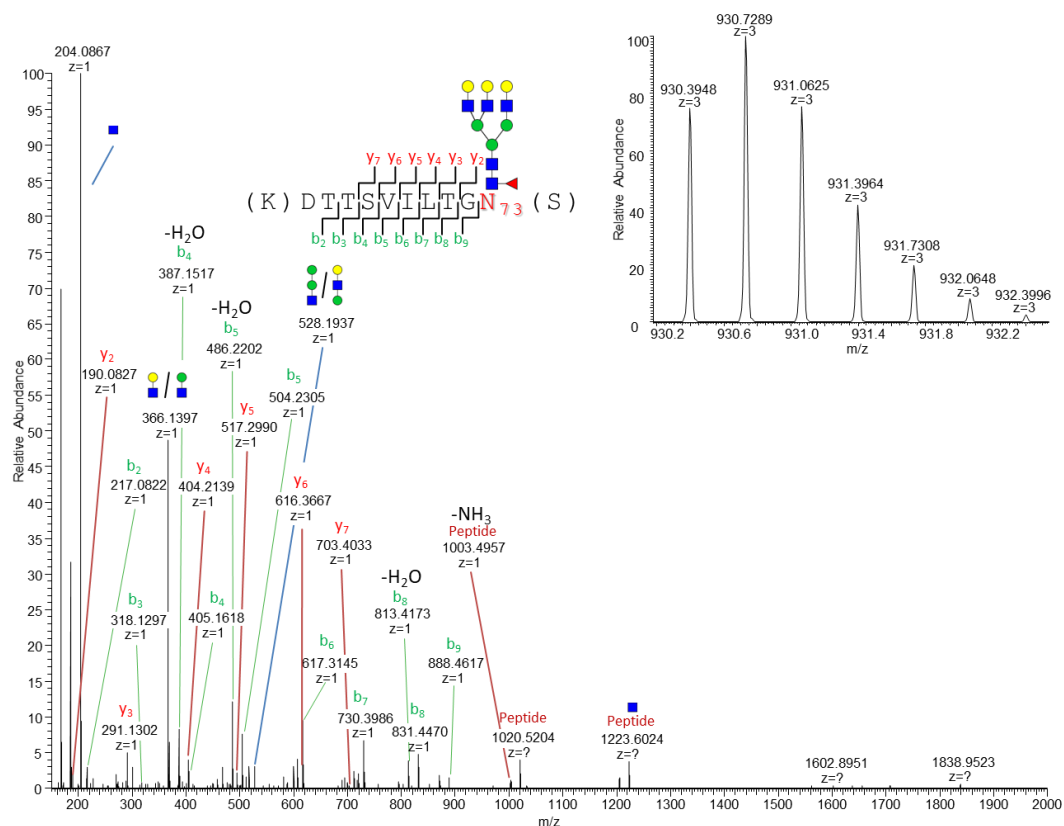
Supplementary Figure A.2-3: **Site-specific N-glycopeptide analysis using nano-RP-LC-MS(/MS)**. Fragment ion spectrum of a HA N-glycopeptide, derived from a sequential proteolytic digestion using trypsin and flavastacin. Amino acids highlighted in red indicate the carbamidomethylation of cysteine and the N-glycosylation of asparagine. For the validation of the (glyco-)peptide sequence, specific b- (green) and y- (red), as well as B-(oxonium-, blue) and Y-ions (peptide + glycan moiety) are annotated. The isotopic pattern of the precursor is located at the upper right corner. Symbolic representations of N-glycan structures were drawn with GlycoWorkbench Version 1.1, following the Symbol Nomenclature for Graphical Representations of Glycans [6]. Adapted from Pralow et al. (2021) [160].



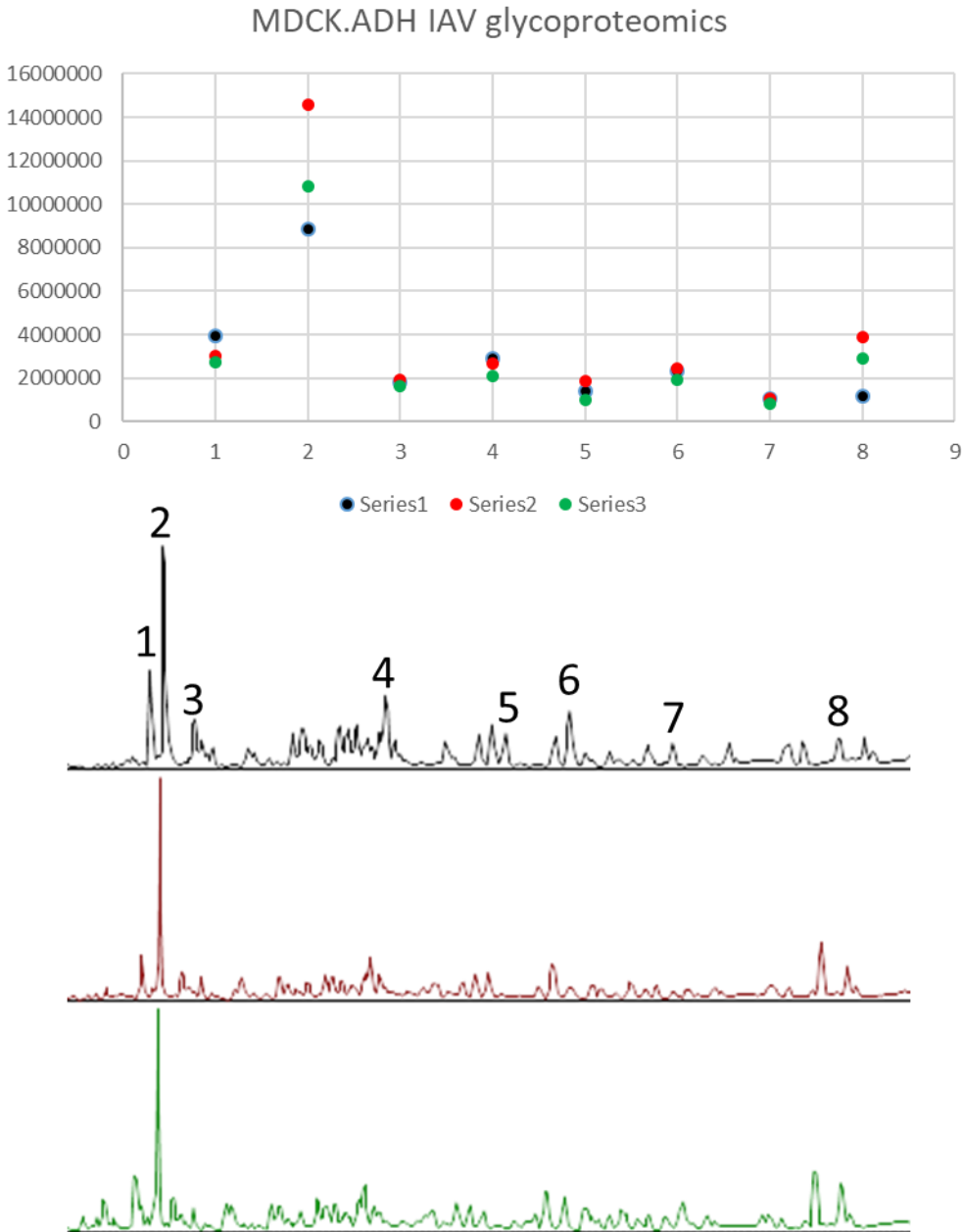
Supplementary Figure A.2-4: **Site-specific N-glycopeptide analysis using nano-RP-LC-MS(/MS)**. Fragment ion spectrum of a HA N-glycopeptide, derived from a sequential proteolytic digestion using trypsin and flavastacin. Amino acids highlighted in red indicate the carbamidomethylation of cysteine and the N-glycosylation of asparagine. For the validation of the (glyco-)peptide sequence, specific b- (green) and y- (red), as well as B-(oxonium-, blue) and Y-ions (peptide + glycan moiety) are annotated. The isotopic pattern of the precursor is located at the upper right corner. Symbolic representations of N-glycan structures were drawn with GlycoWorkbench Version 1.1, following the Symbol Nomenclature for Graphical Representations of Glycans [6]. Adapted from Pralow et al. (2021) [160].



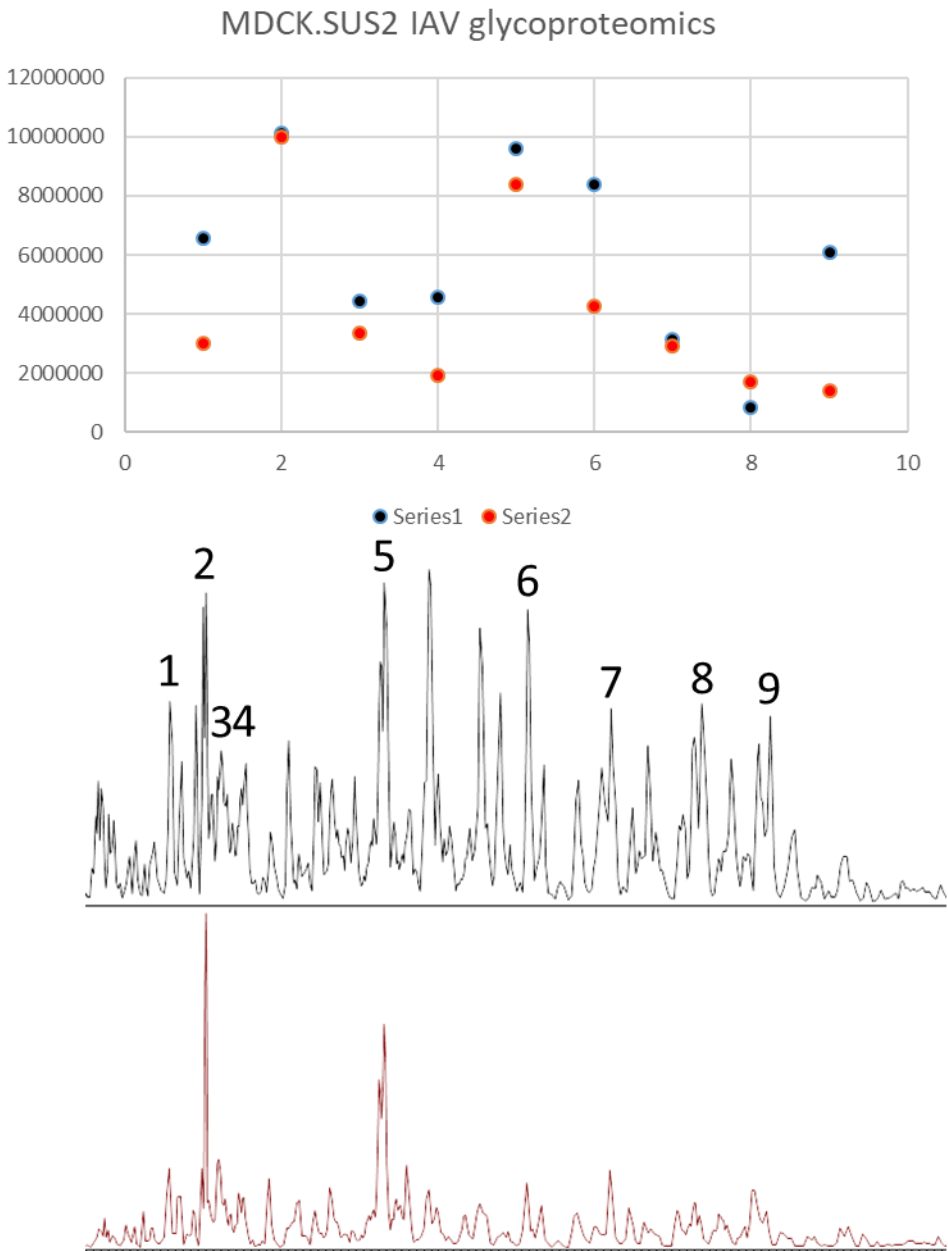
Supplementary Figure A.2-5: **Site-specific N-glycopeptide analysis using nano-RP-LC-MS(/MS)**. Fragment ion spectrum of a HA N-glycopeptide, derived from a proteolytic digestion using trypsin. Amino acids highlighted in red indicate the carbamidomethylation of cysteine and the N-glycosylation of asparagine. For the validation of the (glyco-)peptide sequence, specific b- (green) and y- (red), as well as B-(oxonium-, blue) and Y-ions (peptide + glycan moiety) are annotated. The isotopic pattern of the precursor is located at the upper right corner. Symbolic representations of N-glycan structures were drawn with GlycoWorkbench Version 1.1, following the Symbol Nomenclature for Graphical Representations of Glycans [6]. Adapted from Pralow et al. (2021) [160].



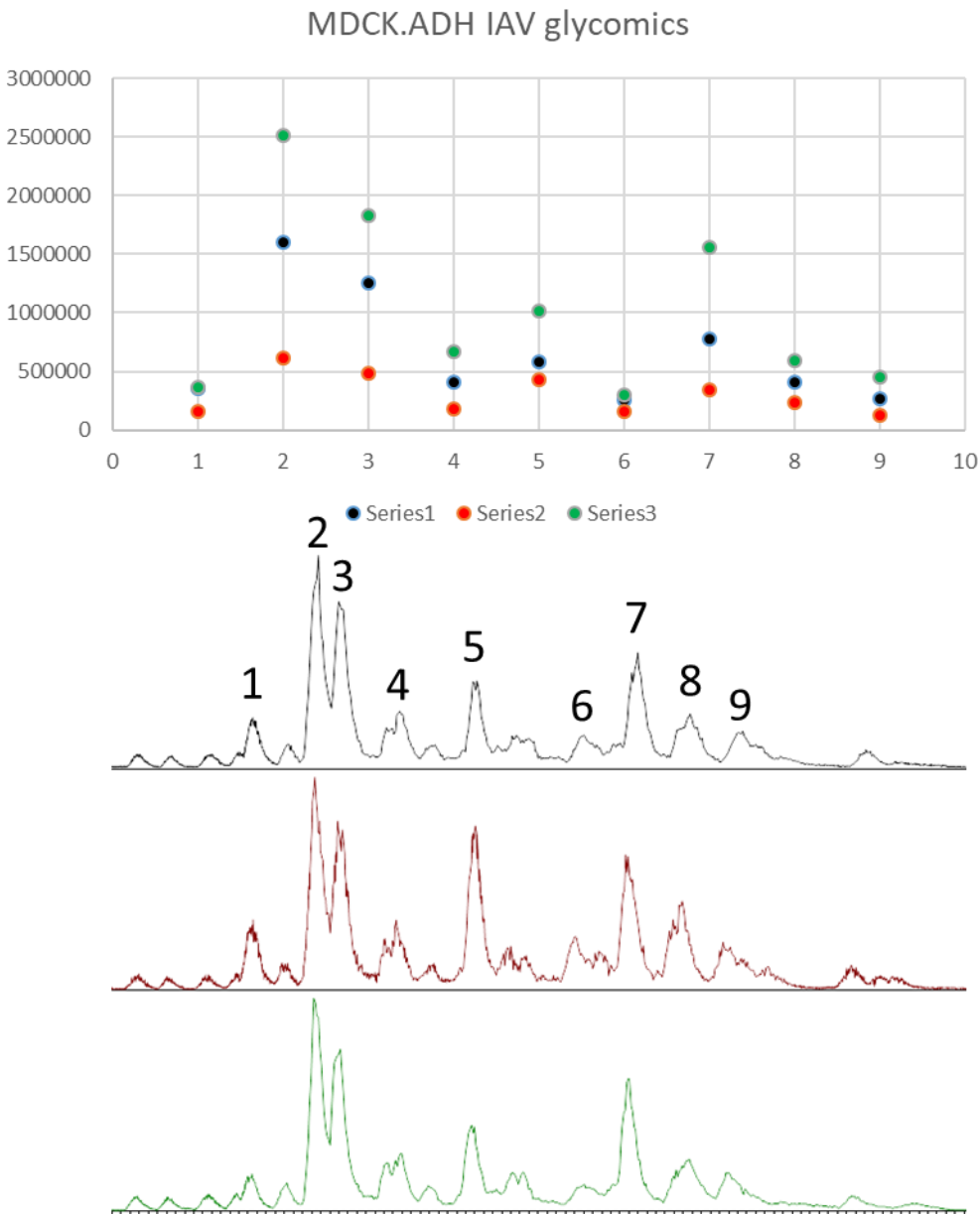
Supplementary Figure A.2-6: **Site-specific N-glycopeptide analysis using nano-RP-LC-MS(/MS)**. Fragment ion spectrum of a NA N-glycopeptide, derived from a sequential proteolytic digestion using trypsin and flavastacin. Amino acids highlighted in red indicate the carbamidomethylation of cysteine and the N-glycosylation of asparagine. For the validation of the (glyco-)peptide sequence, specific b- (green) and y- (red), as well as B-(oxonium-, blue) and Y-ions (peptide + glycan moiety) are annotated. The isotopic pattern of the precursor is located at the upper right corner. Symbolic representations of N-glycan structures were drawn with GlycoWorkbench Version 1.1, following the Symbol Nomenclature for Graphical Representations of Glycans [6]. Adapted from Pralow et al. (2021) [160].



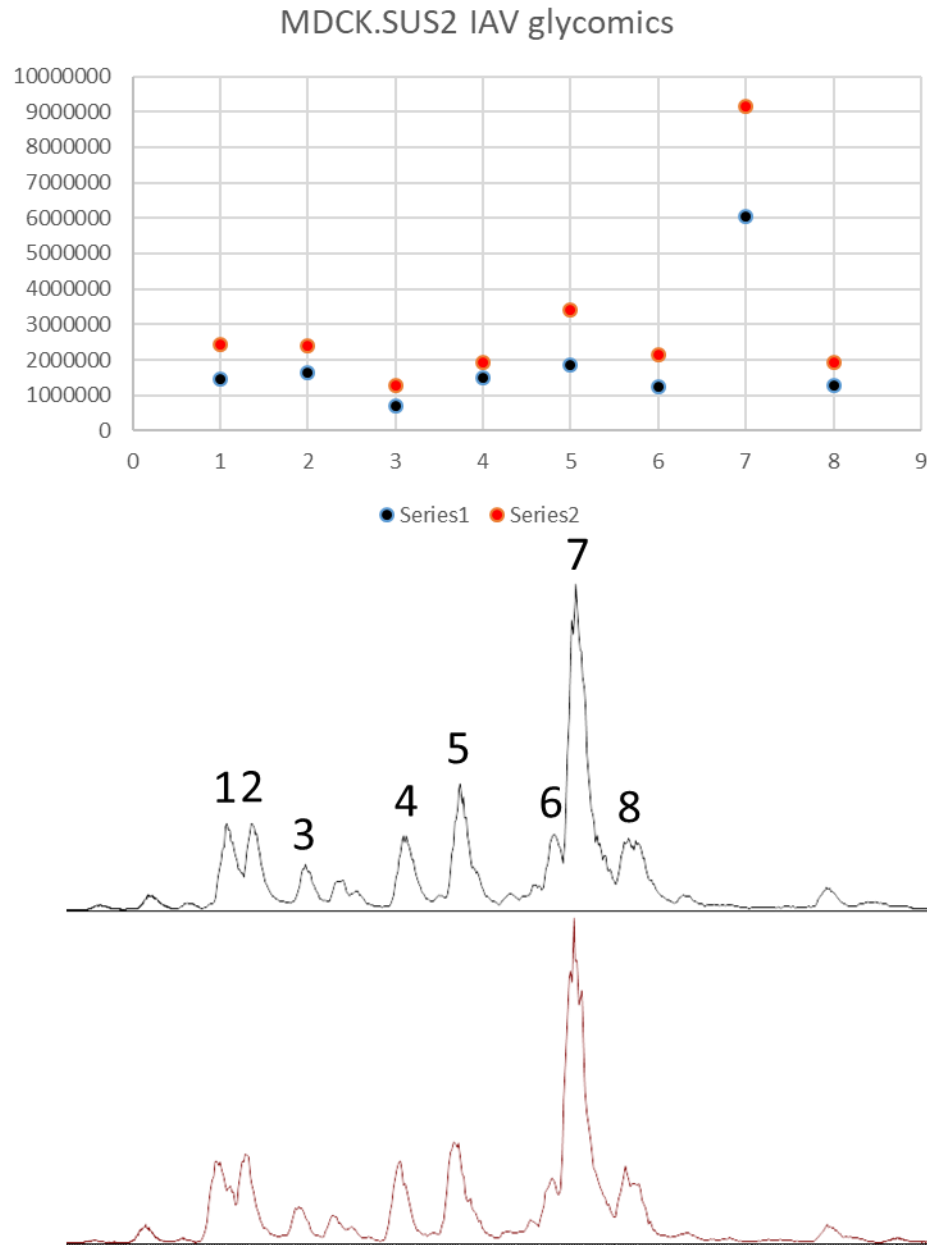
Supplementary Figure A.2-7: **Base peak chromatogram (BPC) of the replicates from the nano-RP-LC-MS(/MS) glycopeptide analysis of IAV propagated in MDCK.ADH.** The upper figure shows the distribution of the relative abundance of selected peaks from the replicates. The lower figure shows the BPC of the replicates and the numbers of the selected peaks. Adapted from Pralow et al. (2021) [160].



Supplementary Figure A.2-8: **Base peak chromatogram (BPC) of the replicates from the nano-RP-LC-MS(/MS) glycopeptide analysis of IAV propagated in MDCK.SUS2.** The relative abundances of respective peaks are illustrated. The upper figure shows the distribution of the relative abundance of selected peaks from the replicates. The lower figure shows the BPC of the replicates and the numbers of the selected peaks. Adapted from Pralow et al. (2021) [160].



Supplementary Figure A.2-9: **Base peak chromatogram (BPC) of the replicates from the nano-PGC-LC-MS(/MS) glycan analysis of IAV propagated in MDCK.ADH.** The relative abundances of respective peaks are illustrated. The upper figure shows the distribution of the relative abundance of selected peaks from the replicates. The lower figure shows the BPC of the replicates and the numbers of the selected peaks. Adapted from Pralow et al. (2021) [160].



Supplementary Figure A.2-10: **Base peak chromatogram (BPC) of the replicates from the nano-PGC-LC-MS(/MS) glycan analysis of IAV propagated in MDCK.SUS2.** The relative abundances of respective peaks are illustrated. The upper figure shows the distribution of the relative abundance of selected peaks from the replicates. The lower figure shows the BPC of the replicates and the numbers of the selected peaks. Adapted from Pralow et al. (2021) [160].

B. List of materials

Supplementary Table B-1: **Standard proteins and enzymes**

Title	Manufacturer	Bestellnummer
Albumin from bovine serum	Sigma-Aldrich (Steinheim, Germany)	A3912-100G
Bovine fetuin	Sigma-Aldrich (Steinheim, Germany)	F2379-100MG
Endoproteinase AspN	New England Biolabs (Frankfurt am Main, Germany)	P8104S
Lactoferrin from human milk	Sigma-Aldrich (Steinheim, Germany)	L4894-5MG
Peptide-N-Glycosidase F from <i>Elizabethkinga miricola</i> Proteomics Grade	Sigma-Aldrich (Steinheim, Germany)	P7367-300UN
Trypsin Sequencing Grade Modified	Promega (Mannheim, Germany)	V5111

APPENDIX

Supplementary Table B-2: Chemicals

Title	Manufacturer	Order number
2,2,2-Trifluoroethanol for synthesis	Merck KGaA (Mannheim, Germany)	8.08259.1000
Acetic acid	Sigma-Aldrich (Steinheim, Germany)	49199-50ML-F
Acetonitril LC-MS Chromasolv	Fluka Analytical Sigma-Aldrich (Steinheim, Germany)	34967-1L
Ammonium acetate	Sigma-Aldrich (Steinheim, Germany)	A-1542
Ammonium Bicarbonat	Fluka Analytical Sigma-Aldrich (Steinheim, Germany)	09830-500G
Calcium chloride dehydrate Molecular biology grade	AppliChem (Darmstadt, Germany)	A4689,0250
Di-Natriumhydrogenphosphat	Merck KGaA (Mannheim, Germany)	1.06585.5000
DL-Dithiothreitol	Sigma-Aldrich (Steinheim, Germany)	D5545-5G
Ethanol 99.6 %	Carl Roth GmbH (Karlsruhe, Germany)	9065.6
<i>Formic acid</i> eluent additive for LC-MS	Fluka Analytical Sigma-Aldrich (Steinheim, Germany)	56302-10X1ML-F
GMEM	Gibco Invitrogen (Carlsbach, USA)	22100-093
HCl	Carl Roth GmbH (Karlsruhe, Germany)	4625.1
Iodacetamide	Sigma-Aldrich (Steinheim, Germany)	I1149-25G
Kaliumchloride	Merck KGaA (Mannheim, Germany)	1.04935.5000
Kaliumdihydrogenphosphat	Merck KGaA (Mannheim, Germany)	1.04873.1000
Methanol LC/MS-grade	Thermo Fisher Scientific (Schwerte, Germany)	10031094
Natriumchloride	Carl Roth GmbH (Karlsruhe, Germany)	P029.3
Natriumhydroxid (NaOH)	Sigma-Aldrich (Steinheim, Germany)	S8045-500G
Sodium borohydride	Sigma-Aldrich (Steinheim, Germany)	71320-100G
<i>Trifluoroacetic acid</i>	Thermo Fisher Scientific (Schwerte, Germany)	28904
Tris-Hydrochlorid für die Molekularbiologie	AppliChem (Darmstadt, Germany)	A3452,1000
Urea, SigmaUltra	Sigma-Aldrich (Steinheim, Germany)	UO631-1KG

APPENDIX

Supplementary Table B-3: Consumables

Title	Manufacturer	Order number
20 µL Sample Loop, Polyetheretherketon (PEEK), WPS- 3000PL	Dionex (Amsterdam, Netherlands)	6820.0018
Acclaim PepMap RSLC nano 75 µm x 15 cm C18, 2 µm, 100 Ångström (Å)	Dionex (Amsterdam, Netherlands)	164534
Acclaim PepMap100 100 µm x 2 mm C18, 5 µm, 100Å	Dionex (Amsterdam, Netherlands)	164564
AG50W-X8 cation-exchange resin	Bio-Rad (München, Germany)	142-1431
C18 StageTip Frit	Thermo Fisher Scientific (Waltham, USA)	SP201
Carbograph	Grace (Worms, Germany)	1769
epT.I.P.S. Reloads 0,1-10 µL	Eppendorf (Hamburg, Germany)	0030073.401
epT.I.P.S. Reloads 1-10 mL	Eppendorf (Hamburg, Germany)	0030073.363
epT.I.P.S. Reloads 2-200 µL	Eppendorf (Hamburg, Germany)	0030073.428
epT.I.P.S. Reloads 50-1000 µL	Eppendorf (Hamburg, Germany)	0030073.460
Hypercarb Kappa, 320 µm x 3 cm	Thermo Fisher Scientific (Waltham, USA)	35005-030315
Hypercarb Kappa, 75 µm x 10 cm	Thermo Fisher Scientific (Waltham, USA)	35003-100065
LC-Pak Polisher	Merck Millipore (Darmstadt, Germany)	LCPAK0001
Millipak 0,22 µm	Merck Millipore (Darmstadt, Germany)	MPGP04001
Nano-bore emitter, 30 µm InD, 1/32" OD, 40 mm length	Thermo Fisher Scientific (Waltham, USA)	ES542
Nanosep 10K OMEGA	Pall Life Science (Dreieich, Germany)	FZ1356
NanoViper capillary (20 µm InD, 1/32" OD, 650 mm length	Thermo Fisher Scientific (Waltham, USA)	6041.5275
NanoViper capillary (75 µm InD, 1/32" OD, 550 mm length	Thermo Fisher Scientific (Waltham, USA)	6041.5760
Progard TL1CL2	Merck Millipore (Darmstadt, Germany)	PROGTLCS1
QuantIT protein assay	Life technologies(Darmstadt, Germany)	Q33210
Safe Lock 0.5 mL	Eppendorf (Hamburg, Germany)	0030120.094
Safe Lock 1.5 mL	Eppendorf (Hamburg, Germany)	0030120.086
Safe Lock 2.0 mL	Eppendorf (Hamburg, Germany)	0030121.023
SilicaTip	New Objective (Cambridge, USA)	FS360-20-10-D-20
T-piece	Thermo Fisher Scientific (Waltham, USA)	SC901

APPENDIX

Supplementary Table B-4: Instruments and Devices

Instrument/ Device	Title	Manufacturer
Centrifuge	Galaxy Mini	VWR (Darmstadt, Germany)
Centrifuge	Heareus Tresco 17 Centrifuge	Thermo Fisher Scientific (Schwerte, Germany)
Centrifuge	Avanti J-20XP	Beckmann Coulter (Brea, USA)
Chromatography System	Ultimate 3000 (Pumpe, Flowmanager, Autosampler)	Dionex (Amsterdam, Netherlands)
Cooling trap	Alpha 2-4L0 plus	Martin Christ Gefriertrocknungsanlagen (Osterode, Germany)
Mass spectrometer	LTD Orbitrap Elite	Thermo Fisher Scientific (Bremen, Germany)
pH-meter	Seven Multi	Mettler Toledo (Hildesheim, Germany)
Pump	Chemistry Hybrid Pump	Vacuubrand (Wertheim, Germany)
Rotating vacuum concentrator	RVC 2-33 CO plus	Martin Christ Gefriertrocknungsanlagen (Osterode, Germany)
Scale	Feinwaage Excellent plus	Mettler Toledo (Hildesheim, Germany)
Scale	Classic plus	Mettler Toledo (Hildesheim, Germany)
Stirrer	RCT basic	IKA-Werke (Staufen, Germany)
Thermomixer	Thermomixer comfort	Eppendorf (Hamburg, Germany)
Ultracentrifuge	Optima TM LE-80K	Beckmann Coulter (Brea, USA)
Vortex mixer	Reax top	Heidolph (Schwabach, Germany)
Water purification system	Elix 100 (No.: ZLX550100)	Merck Millipore (Darmstadt, Germany)
Water purification system	Milli Q Reference A+	Merck Millipore (Darmstadt, Germany)

APPENDIX

Supplementary Table B-5: **Software**

Title	Manufacturer	Version
Byonic and Byologic	ProteinMetrics (Cupertino, USA)	
Cromleon Xpres	Dionex (Amsterdam, Netherlands)	6.8
Excel	Microsoft (Unterschleißheim, Germany)	Professional Plus 2010
ExPASy GlycoMod	Swiss Institute of Bioinformatics (SIB) (Genf, Switzerland)	Online tool
ExPASy-FindPept	Swiss Institute of Bioinformatics (SIB) (Genf, Switzerland)	Online tool
GlycoWorkbench	European Carbohydrates DataBase (Hinxton, Großbritannien)	1.1
glyXtool ^{MS}	In-house developed by Pioch et al. (2018)	Free available
<i>MASCOT</i>	Matrix Science (London, UK)	2.5
MS-Product	University of California, San Francisco (San Francisco, USA)	Online tool
Power Point	Microsoft (Unterschleißheim, Germany)	Professional Plus 2010
Proteome Discoverer	Thermo Fisher Scientific (Waltham, USA)	1.4
UCSF Chimera	University of California, San Francisco	1.10.2
Xcalibur	Thermo Fisher Scientific (Waltham, USA)	2.2

C. Publications

The dissertation includes content that has already been published in or is submitted to international peer-reviewed scientific journals. Subsequently, the published and submitted publications are listed in detail, including the specific contribution of the author of this doctoral thesis. First authorship is indicated with “*”.

Pralow, A.*, Hoffmann, M., Nguyen-Khuong, T., Rapp, E., & Reichl, U. (2017). Improvement of the glycoproteomic toolbox with the discovery of a unique C-terminal cleavage specificity of flavastacin for N-glycosylated asparagine. *Scientific reports*, 7(1), 11419. *(first author)*

Nguyen-Khuong, T.* , **Pralow, A.***, Reichl, U., & Rapp, E. (2018). Improvement of electrospray stability in negative ion mode for nano-PGC-LC-MS glycoanalysis via post-column make-up flow. *Glycoconjugate journal*, 35(6), 499-509. *(first author)*

Pioch, M.* , Hoffmann, M., **Pralow, A.**, Reichl, U., & Rapp, E. (2018). glyXtoolMS: An Open-Source Pipeline for Semiautomated Analysis of Glycopeptide Mass Spectrometry Data. *Analytical chemistry*, 90(20), 11908-11916. *(co-author, helping in the development process, testing the software)*

Hoffmann, M.* , Pioch, M., **Pralow, A.**, Hennig, R., Kottler, R., Reichl, U., & Rapp, E. (2018). The Fine Art of Destruction: A Guide to In-Depth Glycoproteomic Analyses—Exploiting the Diagnostic Potential of Fragment Ions. *Proteomics*, 18(24), 1800282. *(co-author, help conceiving the study and developing the cotton-spin HILIC)*

Gränicher, G.* , Coronel, J., **Pralow, A.**, Marichal-Gallardo, P., Wolff, M., Rapp, E., ... & Reichl, U. (2019). Efficient influenza A virus production in high cell density using the novel porcine suspension cell line PBG. PK2. 1. Vaccine. *(co-author, glycoproteomic analysis of IAV samples)*

Hinneburg, H.*, Pedersen, J. L., Bokil, N. J., **Pralow, A.**, Schirmeister, F., Kawahara, R., ... & Thaysen-Andersen, M. (2020). High-resolution longitudinal N-and O-glycoprofiling of human monocyte-to-macrophage transition. *Glycobiology. (co-author, glycoproteomic and proteomic analysis)*

Pralow, A.*, Cajic, S., Alagesan, K., Kolarich, D., & Rapp, E. (2020) State-of-the-Art Glycomics Technologies in Glycobiotechnology. *Advances in Glycobiotechnology, Springer, Cham, 379-411. (first author)*

Pralow, A.*, Hoffmann, M., Nguyen-Khuong, T., Pioch, M., Hennig, R., Genzel, Y., ... & Reichl, U. (2021). Comprehensive N-glycosylation analysis of the influenza A virus proteins HA and NA from adherent and suspension MDCK cells. *The FEBS Journal. (first author)*

Pralow, A.*, Nikolay, A., Leon, A., Genzel, Y., Rapp, E., & Reichl, U. (2021). Site-specific N-glycosylation analysis of animal cell culture-derived Zika virus proteins. *Scientific reports, 11(1), 1-7. (first author)*

D. Attended conferences

D.1. Talks

Pralow, A., Nguyen-Khuong, T., Pioch, M., Hennig, R., Hoffmann, M., Rapp, E., & Reichl, U. (2017). Structural and site-specific N-glycan analysis of influenza A/PR8/34 virus propagated in adherent and suspension MDCK cells. In 28th Joint Glycobiology Meeting 2017.

D.2. Posters

Pralow, A., Hoffmann, M., Kottler, R., Reichl, U., & Rapp, E. (2014). Optimized Workflow for Sample Preparation and Site-specific Glycosylation Analysis of Human Milk Lactoferrin. In 25th Joint Glycobiology Meeting.

Pralow, A., Nguyen-Khuong, T., Pioch, M., Hennig, R., Hoffmann, M., Rapp, E., & Reichl, U. (2017). Comprehensive structural and site-specific N-glycan analysis of influenza A/PR8/34 virus propagated in adherent and suspension MDCK cells. In GlycoBioTec Symposium 2017.

Pralow, A., Hoffmann, M., Pioch, M., Hennig, R., Nguyen-Khuong, T., Genzel, Y., ... & Reichl, U. (2019). Comprehensive N-glycosylation analysis of influenza A virus antigens derived from different production cell lines reveal immunogenic glycan epitopes. In GlycoBioTec 2019.

E. Supervised projects

E.1. Students

During the time this dissertation was accomplished, no student projects were supervised by the doctoral candidate.

E.2. Other projects

- Supervision of the “Schülerpraktikum”
- Organization of the “Long Night of Science” for the BPE group