The FEBS Journal

FEBS PRESS science publishing by scientists

Comprehensive *N*-glycosylation analysis of the influenza A virus proteins HA and NA from adherent and suspension MDCK cells

Alexander Pralow[1], Marcus Hoffmann[1], Terry Nguyen-Khuong[1], Markus Pioch[1], René Hennig[1,2], Yvonne Genzel[1], Erdmann Rapp[1,2] and Udo Reichl[1,3]

1 Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg, Germany
2 glyXera GmbH, Magdeburg, Germany
3 Chair of Bioprocess Engineering, Otto von Guericke University, Magdeburg, Germany

Keywords
glycoproteomics; hemagglutinin; influenza A virus; mass spectrometry; MDCK cells; neuraminidase; *N*-glycosylation; porous graphitized carbon; vaccine production

Correspondence
E. Rapp, Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstrasse 1, 39106 Magdeburg, Germany.
Tel: +493916110314
Email:

(Received 2 November 2020, revised 4 February 2021, accepted 22 February 2021)

doi:10.1111/febs.15787

Glycosylation is considered as a critical quality attribute for the production of recombinant biopharmaceuticals such as hormones, blood clotting factors, or monoclonal antibodies. In contrast, glycan patterns of immunogenic viral proteins, which differ significantly between the various expression systems, are hardly analyzed yet. The influenza A virus (IAV) proteins hemagglutinin (HA) and neuraminidase (NA) have multiple *N*-glycosylation sites, and alteration of *N*-glycan micro- and macroheterogeneity can have strong effects on virulence and immunogenicity. Here, we present a versatile and powerful glycoanalytical workflow that enables a comprehensive *N*-glycosylation analysis of IAV glycoproteins. We challenged our workflow with IAV (A/PR/8/34 H1N1) propagated in two closely related Madin–Darby canine kidney (MDCK) cell lines, namely an adherent MDCK cell line and its corresponding suspension cell line. As expected, *N*-glycan patterns of HA and NA from virus particles produced in both MDCK cell lines were similar. Detailed analysis of the HA *N*-glycan microheterogeneity showed an increasing variability and a higher 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. Almost all *N*-glycan structures were fucosylated. Furthermore, HA and NA *N*-glycan structures were exclusively hybrid- and complex-type structures, to some extent terminated with alpha-linked galactose(s) but also with blood group H type 2 and blood group A epitopes. In contrast to the similarity of the overall glycan pattern, differences in the relative

Abbreviations
ABC, ammonium bicarbonate; ACN, acetonitrile; AspN, endoproteinase AspN; BPC, base peak chromatogram; dHex, deoxyhexose; DTT, DL-dithiothreitol; ECACC, European collection of authenticated cell cultures; EIC, extracted ion chromatogram; FASP, filter-aided sample preparation; FCS, fetal calf serum; FLD, fluorescence detector; Fuc, fucose; Galili, galactose-alpha-(1,3)-galactose; GalNAc, *N*-acetylgalactosamine; Glc, glucose; GlcNAc, *N*-acetylglucosamine; GMEM, Glasgow minimum essential medium; HA, hemagglutinin; HCD, higher-energy collisional dissociation; HEK293, human embryonic kidney 293; Hex, hexose; HexNAc, *N*-acetylhexosamine; HILIC, hydrophilic interaction chromatography; IAA, iodoacetamide; IAV, influenza A virus; ID, identifier; IM, ion mobility; KOH, potassium hydroxide; LacNAc, *N*-acetyllactosamine; LC, liquid chromatography; LC-MS, liquid chromatography coupled to mass spectrometry; M2, matrix-protein 2; MALDI-TOF-MS, matrix-assisted laser desorption ionization–time-of-flight mass spectrometry; Man, mannose; MDBK, Madin–Darby bovine kidney; MDCK, Madin–Darby canine kidney; MDCK.ADH, MDCK—adherent cell line; MDCK.SUS2, MDCK—suspension cell line; Mp, peptide moiety; MS, mass spectrometry; MS/MS, tandem mass spectrometry; NA, neuraminidase; $NaBH_4$, sodium borohydride; NeuAc, *N*-acetylneuraminic acid; PCMF, postcolumn make-up flow; PDB, protein data bank; PGC, porous graphitized carbon; PNGase F, peptide *N*-glycosidase F; rHA, recombinant HA; RP, reversed-phase; RT, retention time; sf, *spodoptera frugiperda*; SP-D, surfactant protein-D; SPE, solid-phase extraction; TCID50, tissue culture infection dose 50; TFA, trifluoroacetic acid; U, units; Vero, African green monkey.

abundance of individual structures were identified. This concerned, in particular, oligomannose-type, alpha-linked galactose, and multiantennary complex-type *N*-glycans.

Introduction

Influenza virus infects over a billion people around the world and accounts for approximately half a million deaths per year [1]. Due to antigenic changes, influenza virus is a leading cause for pandemics and seasonal epidemics [2]. Hemagglutinin (HA) is a homotrimeric surface glycoprotein and plays a key role in terms of pathogen–host interactions. It is the major antigen of influenza A virus (IAV), and each monomer comprises the two domains, HA1 and HA2. The glycosylation of HA is known to affect IAV virulence by modulating virus receptor binding [3], to stimulate the host immune response [4], and to mask antigenic sites [5]. Glycosylation of the homotetrameric surface glycoprotein neuraminidase (NA) influences virus entry [6], release [7], and neurovirulence [8]. The number of potential *N*-glycosylation sites among different influenza viruses is rather variable for HA and more or less conserved for NA [9]. In this study, we used the influenza A/PR/8/34 H1N1 virus, which displays six potential *N*-glycosylation sites at the HA protein (five at the HA1 and one at the HA2 domain) and five at the NA protein.

The majority of influenza virus vaccines is produced in embryonated chicken eggs. However, more and more influenza vaccines are produced in highly defined culture conditions using mammalian cell lines, such as MDCK [10,11] or Vero cells, enabling a more flexible and efficient production. Furthermore, insect cell lines are used to express recombinant hemagglutinin (rHA) [12].

Commercial production of influenza virus in mammalian cell lines has mainly focused on Vero (African Green Monkey) and MDCK (Madin–Darby canine kidney) cell lines [13–15]. Although adherent cells often show higher cell-specific yields, scale-up can be technically very challenging and suspension cells are considered the preferred process option. Besides others, our group has established several adherent and suspension MDCK cell lines [13,16,17], evaluated virus replication dynamics, and optimized production processes in various bioreactor systems (batch, fed-batch, plug flow, and hollow fiber) [16,18,19]. Furthermore, we have conducted comprehensive analytical studies including proteomics [20–22] and capillary gel electrophoresis-based glycoprofiling [23–25]. So far, however, there has not been a detailed investigation on the glycan pattern of the viral envelope proteins HA and NA propagated in our aforementioned MDCK cell lines regarding alterations of *N*-glycan micro- and macroheterogeneity.

Glycosylation is a nontemplate driven (co-)post-translational modification whereby monosaccharides or complex oligosaccharides (most commonly *N*- and *O*-linked glycans) are enzymatically attached to the proteins at distinct glycosylation sites [26]. *N*-glycans are linked to the amino group of asparagine according to a specific consensus sequence (NXS/T(C/V); $X \neq P$). Moreover, *N*-glycans are characterized by a common core-structure $GlcNAc_2Man_3$ [*N*-acetylglucosamine (GlcNAc), mannose (Man)], which can be extended to form complex-, oligomannose-, or hybrid-type *N*-glycan structures [27]. The presence of such *N*-glycan structures will affect the tertiary structure of the IAV antigens HA and NA upon their presentation to the host immune system and therefore might modulate vaccine efficacy.

The advent of powerful glycoanalytical techniques has recently enabled investigators to perform comprehensive glycosylation studies including both, characterization of all glycan compositions present at one glycosylation site (microheterogeneity) and determination of the glycosylation site occupancy (macroheterogeneity) of influenza virus glycoproteins [28–32]. Furthermore, in-depth structural analysis of the influenza virus *N*-glycome can be performed. Multiplexed capillary gel electrophoresis with laser-induced fluorescence detection (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 [,33]. Hussain *et al.* [34] used ion-mobility mass spectrometry to identify *N*-glycans of two H3N2 and one H1N1 IAV strains propagated in adherent MDCK II cells. Recently, She et al. [35] used nanoporous graphitized carbon liquid chromatography coupled to tandem mass spectrometry [nano-PGC-LC-MS(/MS)] to identify sulfated complex-type *N*-glycans derived from IAV propagated in different host cells (including MDCK cells). Harvey et al. [31] reviewed 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, MDBK, HEK293), insect cells (e.g., High Five, express SF+, sf21), tobacco plants, and embryonated chicken eggs. There are a few studies combining the compositional analysis of the IAV glycome and glycoproteome [28,29,36,37] of influenza viruses. 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 fully 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 [23], 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 [38–41]. In strong contrast, compositional analyses only provide information about glycan masses summarizing different amounts of monosaccharides [e.g., hexoses (Hex), *N*-acetylhexosamine (HexNAc), desoxyheoses (dHex)].

In this report, we established a comprehensive MS-based workflow, that is, 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, we used nano-PGC-LC-MS(/MS). 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 [42]), 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 [43–45].

N-glycopeptides were measured on a nanoreversed-phase LC (nano-RP-LC) coupled to MS/MS. Data acquisition was followed by manual and semi-automated analysis of higher-energy collisional dissociation (HCD)-generated *N*-glycopeptide fragment-ion spectra using glyXtool^{MS} [46], an in-house-developed glycopeptide analysis software, and the search engine Byonic [47].

We challenged this workflow 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.

RESULTS

A comprehensive workflow was established for in-depth structural site-specific glycoanalysis of viral glycoproteins using a combination of nano-PGC-LC-MS (/MS) and nano-RP-LC-MS(/MS). To challenge this workflow, we compared the antigen *N*-glycosylation pattern of HA and NA of IAV propagated in two closely related MDCK cell lines regarding differences in the fine structure of the viral *N*-glycome (nano-PGC) and in individual *N*-glycosylation sites (nano-RP).

In Fig. 1, the established workflow is illustrated. Virus harvests are purified via *g*-force step-gradient centrifugation to obtain whole virus particle fractions without contaminating cell debris as described previously [24,25]. In the next step, the virus is inactivated using SDS and the protein concentration is determined. For glycomic analysis (bottom, Fig. 1), the *N*-glycans are 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 are enriched via PGC-solid-phase extraction (SPE) and analyzed using nano-PGC-LC-MS(/MS) equipped with a postcolumn make-up flow (PCMF) [48] followed by manual data analysis. For glycoproteomic analysis (top, Fig. 1), viral proteins are digested using trypsin and flavastacin [49] followed by glycopeptide enrichment using hydrophilic interaction chromatography (HILIC)-SPE [50]. Subsequently, glycopeptides are measured using nano-RP-LC-MS(/MS) [50] and analyzed manually as well as semi-automatically.

Characterization of IAV N-Glycan Fine Structures

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) [43,51,52]. Here, we use nano-PGC-LC-MS(/MS) for the analysis of *N*-glycans derived from

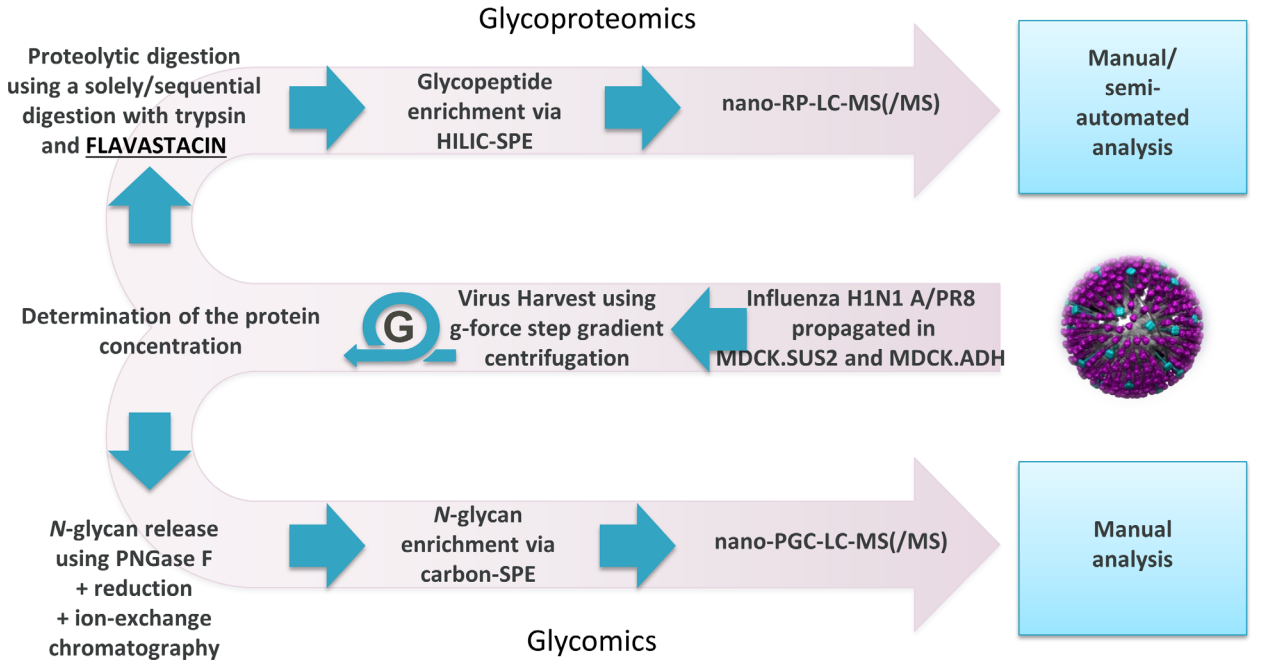


Fig. 1. Workflow—Comprehensive fine-structural and site-specific glycoanalysis of influenza A virus (IAV) surface glycoproteins. Virus harvest of in MDCK cell-derived IAV followed by: *N*-glycomic analysis using nano-PGC-LC-MS(/MS) after *N*-glycan release, reduction of the reducing end, and *N*-glycan enrichment (lower part). *N*-glycoproteomic analysis via nano-RP-LC-MS(/MS) after proteolytic digestion and *N*-glycopeptide enrichment (upper part)

virus glycoproteins to confirm rich data sets derived from chromatographic separation (separation of isobaric glycans) [42] and MS/MS fragmentation (i.e., specific cross-ring fragment ions) [43,44]. Please note, that despite virus particle purification and concentration, detection of a small amount of glycans from host cell glycoproteins is to be expected, due to attachment and/or incorporation of host cell proteins into the viral membrane [53]. Therefore, a complete list of identified proteins after tryptic digestion of IAV samples and g-force step-gradient centrifugation is provided in the supplementary information (Tables S3 and S4).

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 Fig. 2. To facilitate understanding and assignments in figures and tables, *N*-glycan structures are labeled using a unique glycan composition ID in Table S1). The EIC of the *N*-glycan (Hex)2 (HexNAc)3 (dHex)1 + (Man)3(GlcNAc)2 (m/z 994.86^{2-}) 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 fucose (3-arm), terminated with an *N*-acetylgalactosamine (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 (Fig. 2A and B). The fourth structure (C24), however, was only detected for virions produced in MDCK.SUS2 cells (Fig. 2B). Using only compositional or partly structural information obtained from conventional glycomic measurement methods, that is, HILIC with fluorescence detection optional coupled to MS (HILIC-FLD-MS) and matrix-assisted laser desorption ionization—time-of-flight mass spectrometry (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.

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 Fig. 2 as illustrated exemplarily in Fig. 3. In this instance, a combination of different fragment ions found in the MS/MS fragment-

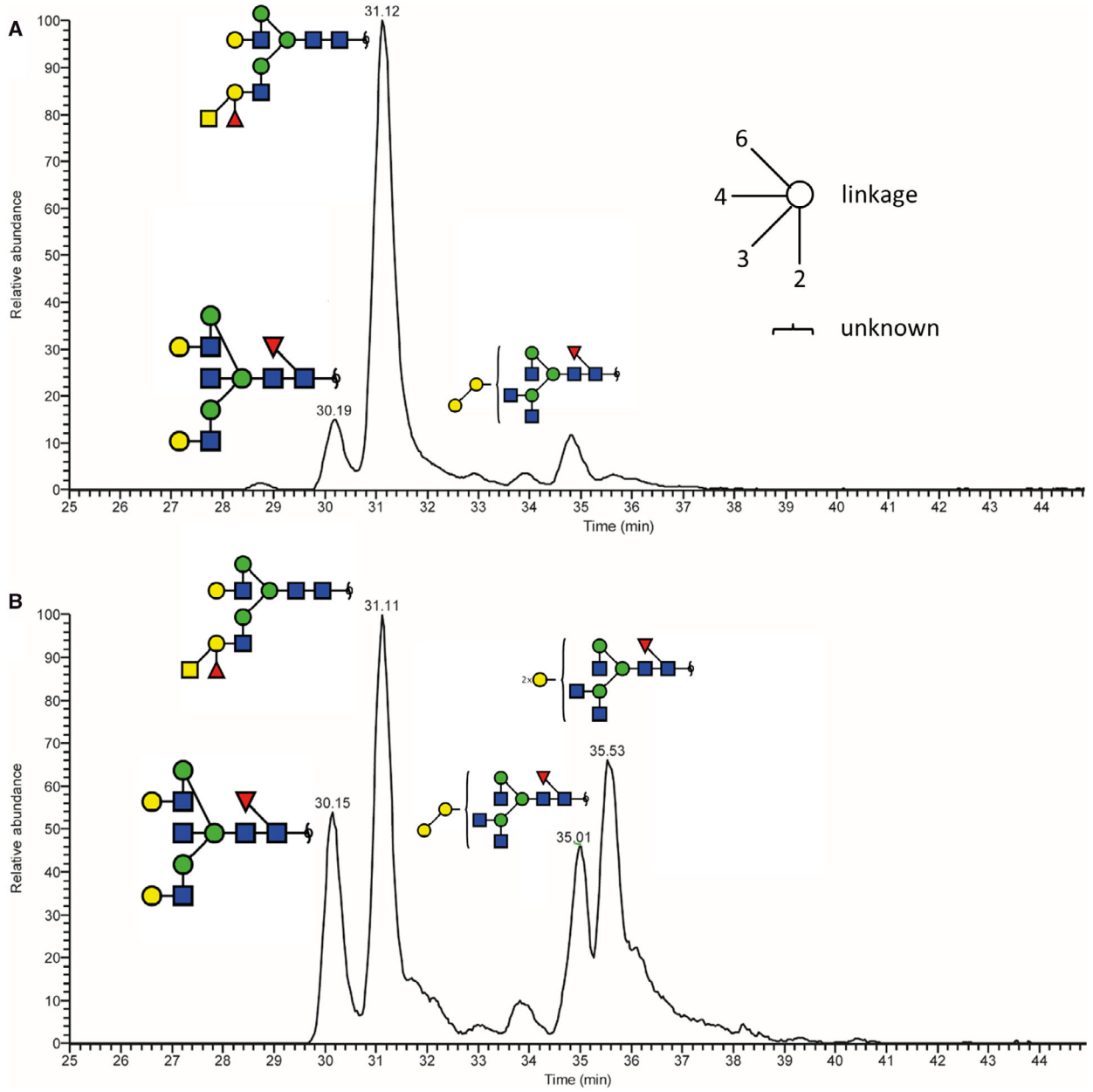
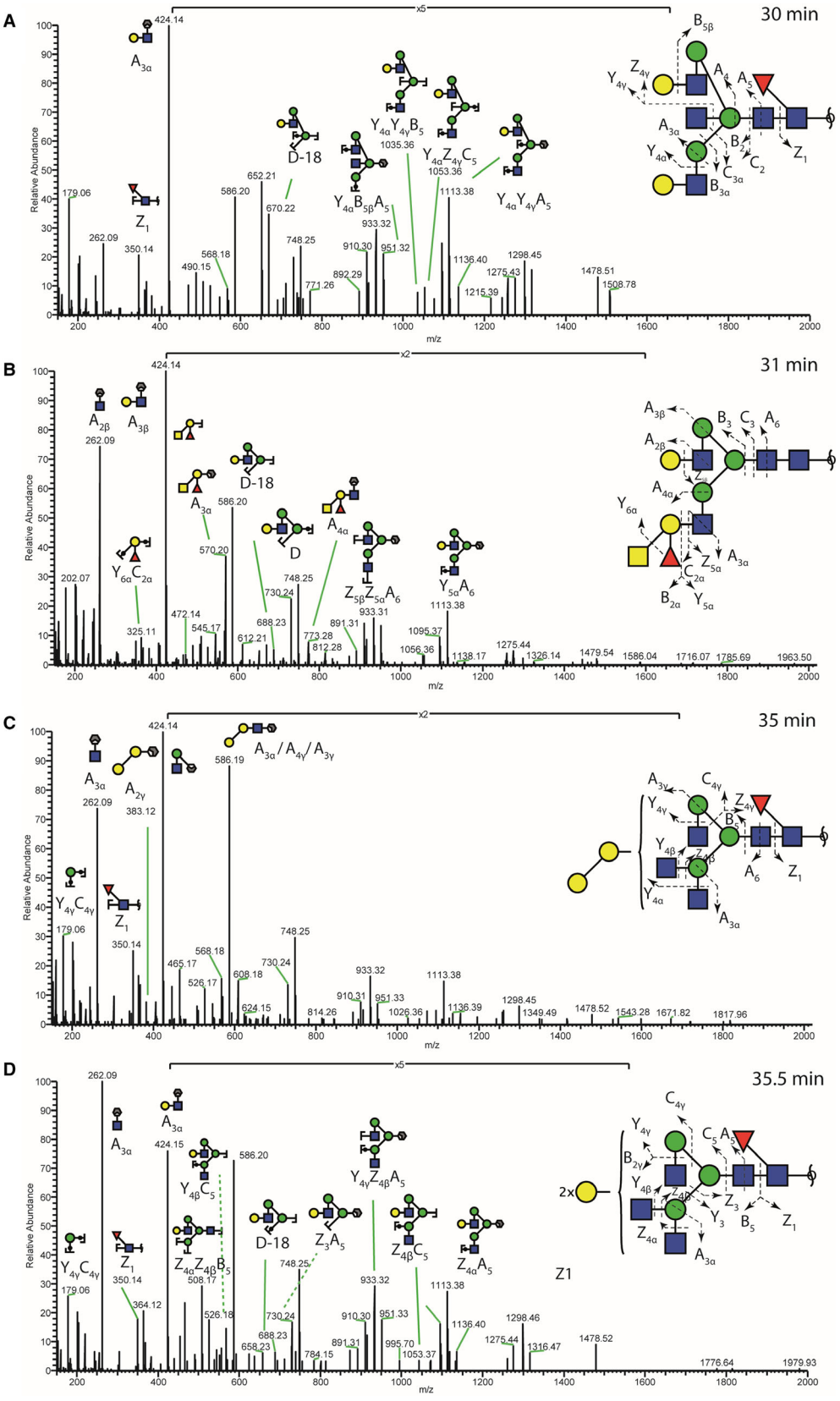


Fig. 2. Separation of *N*-glycan isomers using nano-PGC-LC-MS(/MS). Extracted ion chromatogram (EIC) of the *N*-glycan composition (Hex)2 (HexNAc)3 (dHex)1 + (Man)3(GlcNAc)2 (m/z 994.86^{2-}) derived from influenza A virus (IAV) propagated in (A) MDCK.ADH and (B) MDCK.SUS2 cells. Symbolic representations of *N*-glycan structures were drawn with GlycoWorkbench version 1.1, following the guideline of the Symbol Nomenclature for Graphical Representations of Glycans [84]. Linkage legend is given in the upper right corner.

ion spectra was used to identify, for example, the presence of either core or antenna fucosylation, the presence of galactose-alpha-(1,3)-galactose (Galili)-epitopes, or bisecting structures. For example, the fragment-ion *m*/*z* 350^{1-} can be found in the core-fucosylated *N*-glycan structures depicted in Fig. 3 (A, C, and D), but is missing in the antenna-fucosylated structure depicted in Fig. 3 (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 383^{1-} and 586^{1-} which together indicated the presence of a Galili epitope (Fig. 3C). Furthermore, structural



The FEBS Journal **288** (2021) 4869–4891 © 021 The Authors. The *FEBS Journal* published by John Wiley & Sons Ltd on behalf of Federation of European Biochemical Societies.

Fig. 3. 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 994.86^{2-}). 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 guideline of the Symbol Nomenclature for Graphical Representations of Glycans [84].

differentiation of the *N*-glycan C22 was achieved due to the observation of the specific diagnostic fragment ion for GalNAc-Gal(Fuc)-GlcNac (galactose (Gal), fucose (Fuc)) (m/z 773^{1-}, Fig. 3B). In addition, antenna fucosylation of the 3-arm galactose was identified using the fragment ions m/z 325^{1-}, 570^{1-}, and 773^{1-} (Fig. 3B). In addition, the D-ion can be used to distinguish the arm specificity of the glyco-epitope. For example, the 688^{1-} ion enabled to distinguish the presence of a terminal galactose on the 6-arm of the C22 glycan, as opposed to the 3-arm epitope GalNac-Gal(Fuc)-GlcNac (Fig. 3B).

All identified *N*-glycan structures of the IAV *N*-glycome propagated in MDCK.ADH and MDCK.SUS2 cells (all replicates) are listed in Table S1.

In total, 85 unique *N*-glycan structures from IAV glycoprotein samples were identified based on retention time (RT) and MS/MS fragment-ion spectra analysis. Nine of these structures do 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; that is, 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 409^{1-} and 427^{1-} [43,54,55]). In a lower abundance, also, the glyco-epitope GalNAc-Gal(Fuc)-GlcNAc (blood group A, identified by the diagnostic fragment-ion m/z 773^{1-} [43,54]), was detected in IAV derived from both cell lines. In particular, the latter was found with a high relative abundance (Fig. 4) and frequency (Table S1), 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 341^{1-}, 383^{1-}, and 586^{1-} [54]), with a slightly higher occurrence in MDCK.ADH cells. Glycans eluting at earlier RTs 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 (Fig. 4, RT 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 (Table S1 and Figs S2–S4). Oligomannose-type *N*-glycans were found in all samples, but with a much higher relative abundance in samples derived from MDCK.ADH cells (Fig. 4).

Overall, our 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 (Fig. S1).

Characterization of influenza glycopeptides

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

Figure 5 shows an MS/MS fragment-ion spectrum of an *N*-glycopeptide derived from HA with the precursor mass m/z 1036.4249^{3+} after a sequential digestion using trypsin and flavastacin (as described in Pralow et al. 2017 [49]), respectively. The putative peptide mass m/z 973.4782^{1+} was identified based on the annotated conserved fragment-ion pattern [50]. 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 [56]). The mass of the *N*-glycan moiety was

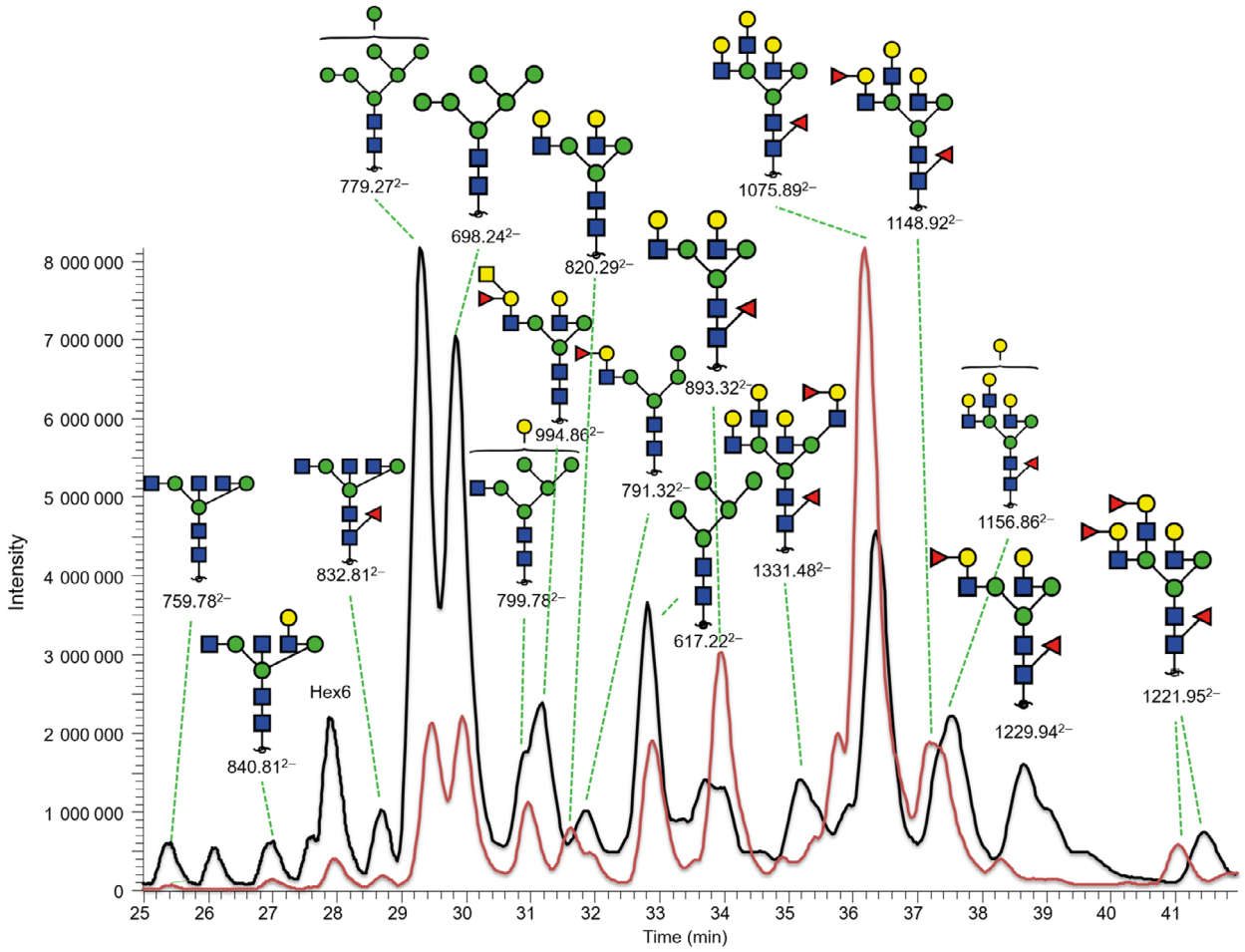


Fig. 4. Comparison of the most abundant *N*-glycan structures. Overlay of the base peak chromatogram (BPC) of the nano-PGC-LC-MS(/MS) measurements of *N*-glycans derived from IAV propagated in MDCK.ADH (black) and MDCK.SUS2 cells (red); retention time (RT); frame 25–42 min. The highest abundant *N*-glycan structure and the corresponding double negatively charged molecular mass of each peak are depicted. Hex6 marks a signal from a polymer. Symbolic representations of *N*-glycan structures were drawn with GlycoWorkbench version 1.1, following the guideline of the Symbol Nomenclature for Graphical Representations of Glycans [83].

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.

In Table 1, we list all identified peptide moieties for each *N*-glycosylation site of HA and NA. 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 1). 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).

In Table S2, 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

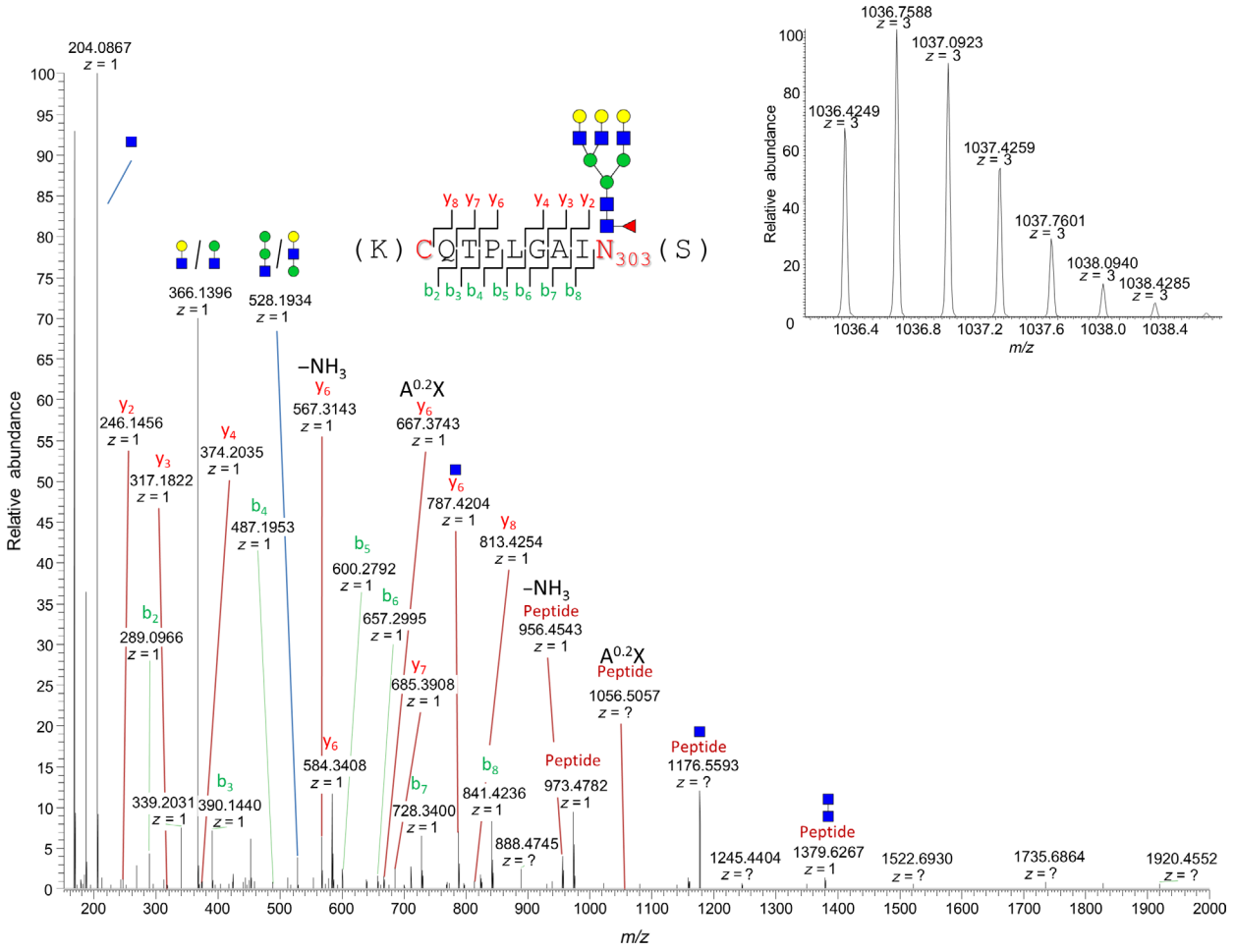


Fig. 5. Site-specific glycopeptide analysis using nano-RP-LC-MS(/MS). Fragment-ion spectrum of a HA *N*-glycopeptide, derived from a sequential proteolytic digestion using trypsin and flavastacin. The *N*-glycopeptide sequence is (K)CQTPLGAIN$_{303}$(S) with the linked *N*-glycan-composition $Hex_3HexNAc_3dHex_1$ + $Man_3GlcNAc_2$ (precursor mass: 1036.4249^{3+}). 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- and y-, as well as B-(oxonium-) and Y-ions 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 guideline of the Symbol Nomenclature for Graphical Representations of Glycans [83].

peak abundancy (Fig. S7) compared to IAV propagated in MDCK.SUS2 (Fig. S8). *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 Fig. 6, we visualize the structure of HA, its potential glycosylation sites, as well as the *N*-glycan compositions detected. 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 (Fig. 6). 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*-glycans 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

Table 1. Identified (glyco)-peptide sequences of the influenza A virus surface glycoproteins hemagglutinin and neuraminidase. For each *N*-glycosylation site, all identified (glyco)-peptide moieties are listed. Furthermore, modified amino acids are underlines and the *N*-glycosylation site is highlighted in bold

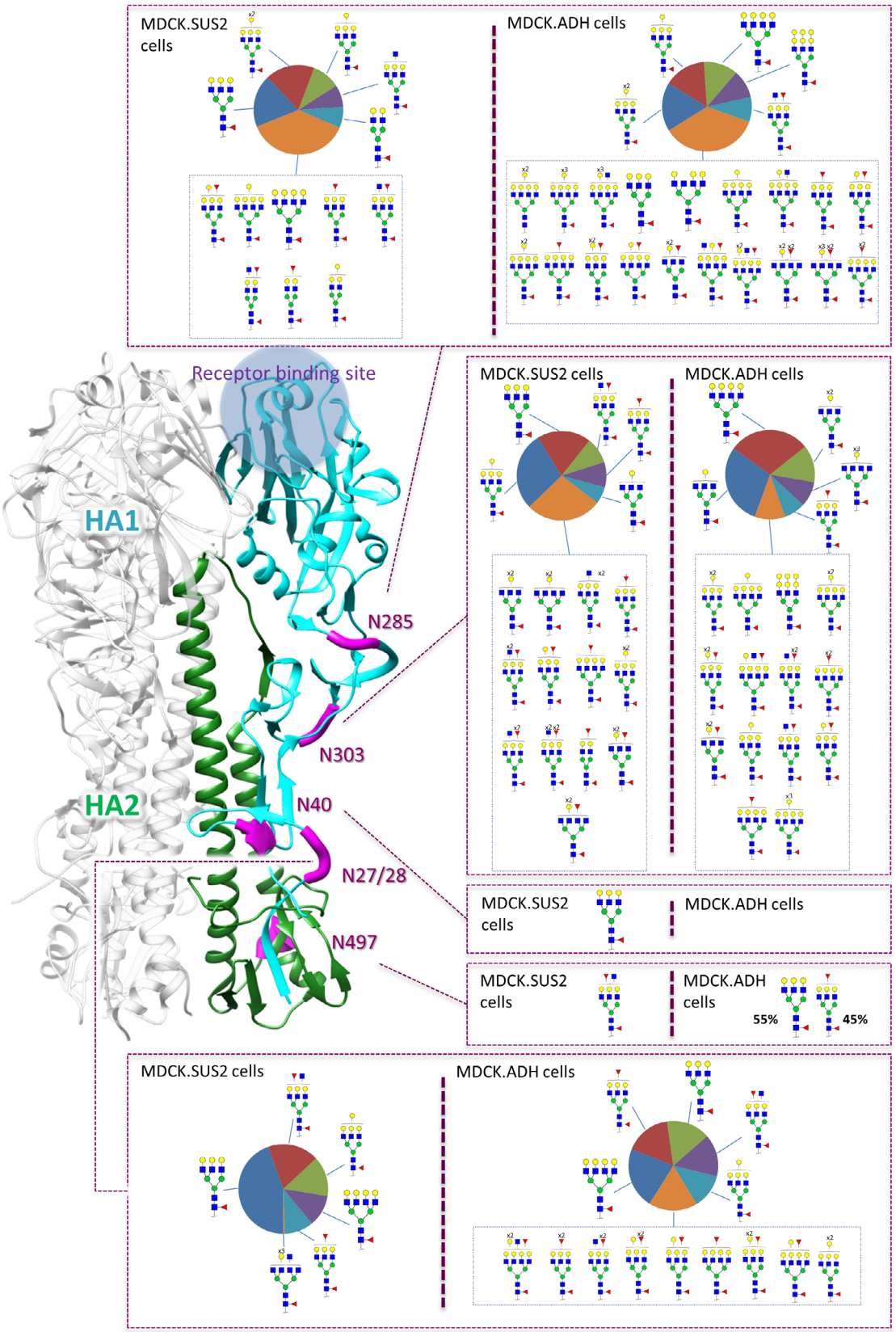
Protein	Site	Peptide Sequence	Modification	Spectrum
Hemagglutinin	N27/28	(A)DTI**C**IGYHA**NN**(S)	(C) Carbamidomethylation	Fig. S2
		(A)DTI**C**IGYHA**NN**(S)	(C) Carbamidomethylation (D) Carbamidomethylation	
	N40	(N)STDTVDTVLEK**N**(V)		Fig. S3
	N285	(R)GFGSGIITS**N**(A)		Fig. S4
		(R)GFGSGIITS**N**(A)	(N-terminal) Carbamidomethylation	
		(G)SGIITS**N**(A)		
	N303	(K)**C**QTPLGAI**N**(S)	(C) Carbamidomethylation	Fig. 5
		(K)**C**QTPLGAI**N**(S)	(C) Carbamidomethylation (Q) NH_3 loss	
		(Q)TPLGAI**N**(S)		
	N497	(R)**N**GTYDYPK(Y)		Fig. S5
Neuraminidase	N73	(K)DTTSVILTG**N**(S)		Fig. S6
		(K)DTTSVILTG**N**(S)	(D) Carbamidomethylation	

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 (Fig. 6), 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 *N*-acetyllactosamine (LacNAc) extensions) dominated. IAV derived from MDCKS.SUS02 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 seemed 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 [33]. In

Fig. 6. Graphical illustration of the trimeric influenza A virus (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 guideline of the Symbol Nomenclature for Graphical Representations of Glycans [83]



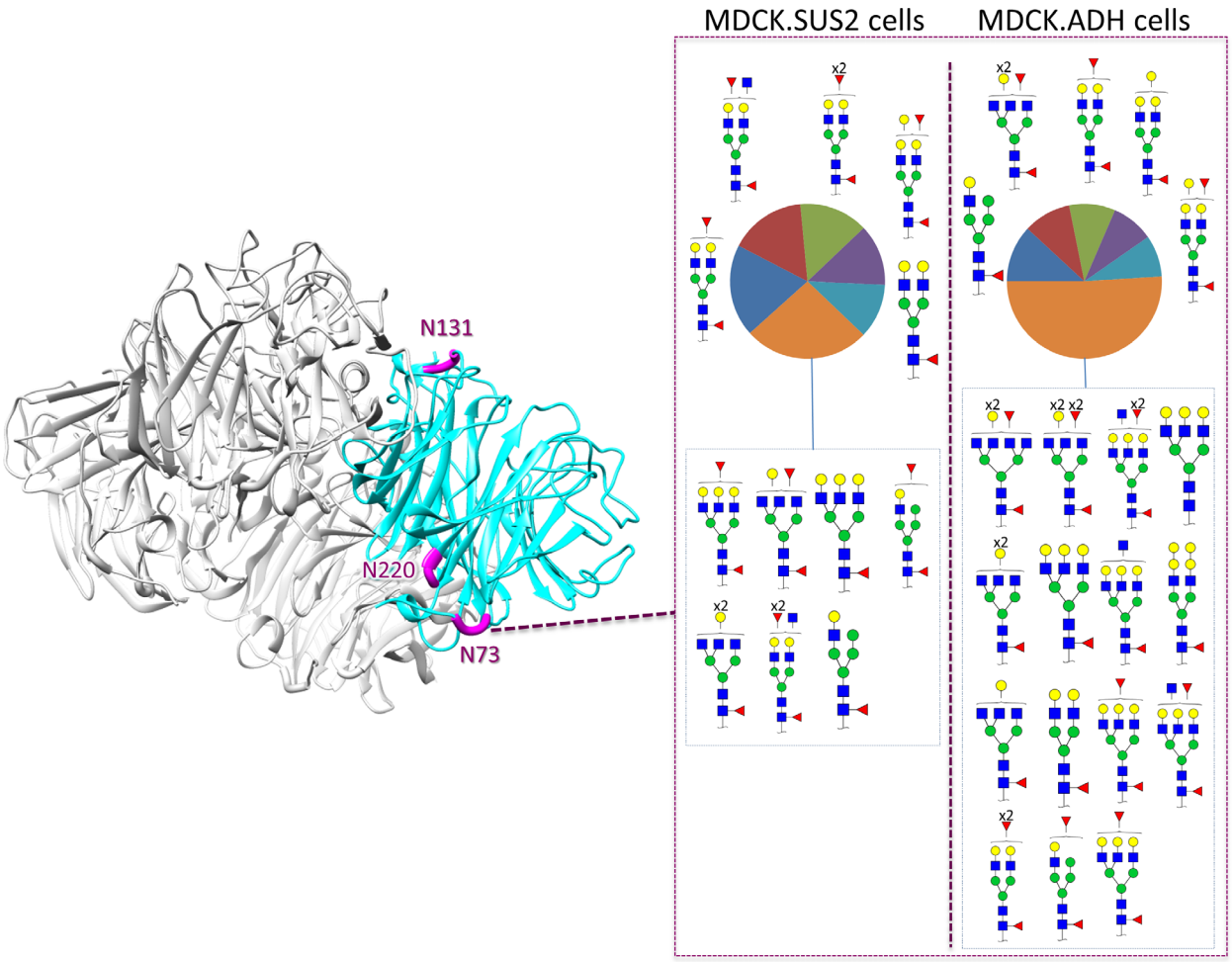


Fig. 7. Graphical illustration of the tetrameric influenza A virus (IAV) antigen NA. The molecular structure of NA was modeled using the Swiss-Model template ID 5hug.1. For model processing and design, the open-source software UCSF Chimera version 1.10.2 was utilized. An NA monomer is highlighted in 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 displayed on the right site. For each site, the five most abundant *N*-glycan compositions are quantified in a pie diagram. The sum of the remaining low abundant *N*-glycans is shown in #6. Symbolic representations of *N*-glycan structures were drawn with GlycoWorkbench version 1.1, following the guideline of the Symbol Nomenclature for Graphical Representations of Glycans [83].

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.

In Fig. 7, 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 Fig. 7 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 abundancy. Comparable to HA, most of the *N*-glycan compositions at NA were antenna- and/or core-fucosylated and to some extent alpha-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.

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 Fig. 6 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 abundancy. When looking up the *N*-glycan list (Table S1), 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 (Fig. 7). As before, in the fine-structural *N*-glycan list (Table S1), 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 (Table S1): 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.

DISCUSSION

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

In this work, we 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 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.

Novel methods used for influenza *N*-glycan analysis

In our previous work, we used only xCGE-LIF for IAV *N*-glycomic analysis [23,24,38,57]. 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 work, 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 [43,54]. Compared to xCGE-LIF-based glycoanalysis, nano-PGC-LC-MS(/MS) has limitations in terms of throughput and reproducibility [42,58] (see, for example, the RT difference of structure C43 in IAV derived from both cell lines, Fig. 4). However, the glycan nano-PGC-LC-MS(/MS) BPCs of the replicates of IAV propagated in MDCK.ADH cells (Fig. S9) as well as MDCK.SUS2 cells (Fig. S10) 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 toward manufacturing of more complex recombinant glycoproteins with multiple glycosylation sites (compared to the less complex monoclonal antibodies with a single glycosylation site) requires not only a glycoanalysis on a global level (e.g., HILIC-FLD-MS, also known as 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 [31]. 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 amino acids 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 study features a sequential digestion workflow using the proteases trypsin and flavastacin as recently described [49,50]. 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 amino acids 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) [28,29,31], the 'glyco'-specific cleavage behavior of flavastacin reduces the amount of redundant peptide moieties and increases 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 is generated. The peptide sequence after tryptic digestion is (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 the 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 [46]. Therefore, in this work, we used glyXtool^{MS} 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 [30,36,59,60], or in combination with MALDI-TOF-MS- [28,37,61], respectively, HILIC-MS-based [29] analysis of released *N*-glycans. In addition, more complex viral glycoproteins like the envelope glycoprotein of HIV [62–64] with up to 30 glycosylation sites and the SARS-CoV-2 spike protein [65] 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 ion-mobility (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 [66] or by analyzing the ratio of specific fragment-ion intensities of glycopeptide MS/MS spectra [50]), additional glycomic analysis is necessary for the fine-structural elucidation of the glycans, as performed in our study. 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 Figure S1). 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, Table S1)) and nonhuman epitopes like the Galili epitope could be structurally elucidated and identified. Without additional fine-structural *N*-glycan analysis, such information would have been missed.

Influenza antigen glycosylation

Comparing glycome-/glycoproteome data among influenza virus-related studies are 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 number and location of potential *N*-glycosylation sites. (II) Besides embryonated chicken eggs (undefined mixture of different cells), a high number of possible host cell systems 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 [41]). (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) contaminant*s* can be found.

To give a short overview of the results of similar studies: An et al. [28] 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. 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 [28]. 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 (≈ 9%) *N*-glycans, but a high amount of highly complex-type *N*-glycans (≈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 [29]. Similar findings were already published earlier [23,38]. She et al. [30] performed a site-specific glycopeptide analysis of two high-yield candidate reassortant vaccines (NIBRG-121xp and NYMC-X181A) propagated in eggs. 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 [35].

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. [28], who refers the partly structural annotation to the known *N*-glycan processing pathways in MDCK cells) 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 our glycomic/glycoproteomic workflow and to obtain further insights into the impact of cell line adaptation to growth in suspension on viral antigen composition, we performed a detailed comparison of the site-specific and structural glycosylation pattern of viral proteins propagated in MDCK.ADH and MDCK.SUS2 cells. 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 high mannose-type *N*-glycans [39]. 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-structural *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 [67]. 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 [68]. 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 nonhuman glyco-epitopes like alpha-galactose 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 [5]. 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.

Increasing interest of cross-reactive IAV vaccines with mono- or nonglycosylation evolves in the scientific community, pointing into the direction of an IAV vaccine, capable of providing protection against different IAV strains [69–71]. 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 [72], or by selecting certain host cell systems to produce specific *N*-glycan patterns [41]. Addition of sites to the head region of A/Hong Kong/1/68 (H3N2) HA resulted in increased sensitivity to surfactant protein (SP)-D *in vitro* and attenuated virulence in mice [73] 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 [74,75]. 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 [76], resulted in sensitivity to SP-D and attenuated virulence in mice [74].

The trend to produce recombinant HA vaccines (e.g., Flublock [77,78]) 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 epitopes) 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 gain importance in the future to better understand antigenicity and immunogenicity of IAV strains expressed in different host systems.

Conclusion

Not only glycosylated biopharmaceuticals such as recombinant proteins but also viral antigens with multiple *N*-glycosylation sites and different types of glycosylation are gaining increasing importance in the pharmaceutical market. Although regulatory authorities do not ask for a glycan analysis of viruses 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 global *N*-glycan analysis (e.g., xCGE-LIF, HILIC-FLD, MALDI-TOF-MS) are optimized for the profiling of monoclonal antibodies, other recombinant proteins (e.g., hormones), and vaccines. Those methods have a high resolution, a very low detection limit. Furthermore, they allow the establishment of high-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.

In this manuscript, we present the use of two orthogonal LC-MS-based methods for the comprehensive structural and site-specific glyco- and glycopeptide analysis of viral antigens. To challenge our approach, IAV propagated in two closely related mammalian cell lines, MDCK.ADH and MDCK.SUS2 cells were used. The workflow established not only allowed to generate a fine-structural *N*-glycan list, but also enable to detect small differences between both expression systems. While the *N*-glycosylation pattern of IAV was similar, it displayed differences in relative abundances of individual *N*-glycan structures. Furthermore, we were able to quantify differences in the glycosylation microheterogeneity of HA and NA and to link this information to the established fine-structural *N*-glycan list. For HA, an increasing complexity and variability of glycosylation at sites closer to the head region (antigen-binding site) could be shown. The identification of alpha-galactosylated *N*-glycan structures in HA as well as NA in both, IAV propagated in MDCK.ADH and MDCK.SUS2 cells, might harbor the potential of modulating antigenicity and immunogenicity of vaccines when using MDCK cells as cultivation system compared to eggs [29,67,79]. To elucidate possible effects of nonhuman glycan epitopes on viral antigens in terms of vaccine efficacy, further structure–function relation as well as animal studies of influenza virus produced in different host cells needs to be performed. Finally, as for manufacturing of other biopharmaceuticals, glycosylation profiling might be requested by regulatory authorities for quality control of viral antigens in the future.

Materials and methods

Enzymes used for proteolytic digestion were trypsin (trypsin sequencing grade modified, V5111) from Promega (Madison, WI, USA), endoproteinase AspN [flavastacin [49]] (AspN, P8104S) from New England Biolabs (Ipswich, MA, USA), and peptide *N*-glycosidase F (PNGase F, P7367) from Sigma-Aldrich (Steinheim, Germany). Potassium hydroxide (KOH, 484016-1KG), dithiothreitol (DTT, D5545-5G), iodoacetamide (IAA, I1149-25G), ammonium bicarbonate (ABC, 09830-500G), sodium borohydride ($NaBH_4$, 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 (A1049), Tris/HCl (A3452), and calcium chloride (A46899) were from AppliChem (Darmstadt, Germany), and trifluoroacetic acid (TFA, 28904) was purchased from Thermo Fisher Scientific (Schwerte, Germany). 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 purification system (18.2 $M\Omega cm^{-1}$ at 25°C, total organic carbon of 3 ppb) from Merck Millipore (Darmstadt, Germany). For LC-MS solvents, water was further purified using the LC-Pak Polisher from Merck Millipore.

Cell lines, cell cultivation, and virus infection

Cell lines, cell cultivation, and virus infection used in this work were adapted from previous work [20,21]. Briefly, MDCK.ADH cells (ECACC, No. 84121903) were cultivated in Glasgow minimum essential medium (GMEM, 22100-093, Gibco Invitrogen, Carlsbad, CA, USA, No. 22100-093), supplemented with 10% (v/v) $FCS_{(aq)}$ (10270-106, Gibco Invitrogen) and 1 % (v/v) $peptone_{(aq)}$ (MC33, Lab M Limited, Heywood, Lancashire, UK). For infections experiments, MDCK.ADH cells were cultivated in triplicates in GMEM supplemented with 5.5 $mmol \cdot L^{-1}$ $glucose_{(aq)}$ (7509, Roth, Karlsruhe, Germany), 47 $mmol \cdot L^{-1}$ $NaHCO_{3(aq)}$ (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 K. Scharfenberg, University of Applied Sciences Emden/Leer, Germany) and cultivated in biological duplicates in chemical defined medium SMIF8 PGd 2x (protein- and peptide-free; Gibco Invitrogen, through contact with K. Scharfenberg) with addition of 23.8 $mmol \cdot L^{-1}$ $NaHCO_{3(aq)}$(8551, Roth), 85.6 $mmol \cdot L^{-1}$ $NaCl_{(aq)}$ (3957, Roth), 0.1% (w/v) Pluronic-F68 (24040032, Gibco Invitrogen), 0.001‰ (w/v) ethanolamine (98%) (398136, Merck, Darmstadt, Germany), 1.6 $mmol \cdot L^{-1}$ L-glutamic acid (G1251, Sigma-Aldrich), 20.3 $mmol \cdot L^{-1}$ D-(+)-glucose (X997, Roth), 4 $mmol \cdot L^{-1}$ glutamine (G3126, Sigma-Aldrich), and 4 $mmol \cdot L^{-1}$ pyruvate (P8574, Sigma-Aldrich) [21]. Cells were infected with human IAV PR/8/34 (H1N1) (no. 3138, Robert Koch Institute (RKI), in-house-generated adherent MDCK-derived viral stock, $TCID_{50}$ titer of 5.17×10^8 infectious virions/mL, HA titer of 2.63 log_{10} HA units/100 µL) with a multiplicity of infection of 0.025 and 2×10^{-6} units trypsin (27250-018, Gibco Invitrogen) per cell. Vented shaker flasks with a volume of 250-mL (MDCK.SUS2) or 850-cm^2 roller

bottles (MDCK.ADH; both VWR, Darmstadt, Germany) were used for all infection experiments. After 72 h postinfection, supernatants containing released virus particles were harvested and subsequently purified.

Virus harvesting, purification, and inactivation

Virus particles were harvested from the cells via *g*-force step-gradient centrifugation as described before [24,25]. This involved pelleting of the cell culture supernatant with three centrifugation steps: (I. 300 ***g***, 4 °C, 20 min; II. 4000 ***g***, 4 °C, 35 min; III. 10 000 ***g***, 4 °C, 45 min). After each step, the supernatant was collected for further centrifugation. After a final ultracentrifugation step (100 000 ***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, Hamburg, Germany).

Protein concentration assay

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

N-glycan analysis: sample preparation and measurement

N-glycan release and clean-up were performed using a modified version of the workflow introduced by Jensen et al. [52]. The entire *N*-glycan analysis workflow was recently described by Nguyen-Khuong et al. [48]. Using a modified filter-aided sample preparation (FASP) approach [80], *N*-glycans were released from 100 µg IAV protein per sample using 1 U PNGase F, followed by the reduction of the reducing end of the *N*-glycans by $NaBH_4$ (converting the α- and β-anomers into alditols). Desalting of the *N*-glycans was performed using a methanolic slurry of AG50W-X8 cation-exchange resin (142-1431, Bio-Rad, München, Germany) packed on top of a 20 µL C18 StageTip frit (SP201, Thermo Fisher Scientific, Waltham, MA, USA). Further clean-up was achieved via PGC-SPE using 5 µL of a 80% ACN_{aq} + 0.1% TFA_{aq} slurry of Carbograph (1769, Grace, Columbia, MD, USA) packed on top of a 20 µL C18 StageTip frit. The SPE was performed threefold using 10 µL of the respective mobile phase in each of the following steps, followed by centrifugation using a table centrifuge (Eppendorf): (I) Therefore, the SPE material was washed with 50% ACN_{aq} + 0.1% TFA_{aq} and (II) equilibrated with 0.1% TFA_{aq}. Afterward, (III) *N*-glycans resuspended in 0.1% TFA_{aq} [10 µL (≈1 $\mu g \cdot \mu L^{-1}$)] were loaded on the SPE material by repetitive pipetting the resulting permeate again on top of the tip, followed by (IV) washing the sample with 0.1% TFA_{aq}. Finally, (V) *N*-glycans were eluted 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, Osterode, Germany) and resuspended in 10 µL 0.1% TFA_{aq}. For nano-PGC-LC-MS(/MS) measurement, *N*-glycans from approximately 500 ng (≈ 1 $\mu g \cdot \mu L^{-1}$) glycoproteins were analyzed on a nano Ultimate 3000 LC system (Thermo Fischer Scientific, Dreieich, Germany) online coupled to an LTQ Orbitrap Elite hybrid mass spectrometer (Thermo Fischer Scientific, Bremen, Germany). To achieve better spray stability, separated *N*-glycans eluting from the LC system were supplemented with a 100% pure ACN PCMF as recently described [48,81]. For details regarding the parameters of the LC-MS setup, see Nguyen-Khuong et al. [48].

Glycopeptide analysis: sample preparation and measurement

Proteolytic digestion of 100 µg IAV glycoproteins per sample was performed using a modified version of the FASP approach described by Wisniewski et al. [50,80]. The samples were proteolytically digested with trypsin, or trypsin followed by flavastacin (sequential digestion) [49]. Afterward, glycopeptide enrichment was performed using a modified version of the HILIC-SPE developed by Selman et al. [50,82]. For nano-RP-LC-MS(/MS) measurement, 500 ng of enriched glycopeptides (≈ 1 $\mu g \cdot \mu L^{-1}$) was analyzed on a nano Ultimate 3000 LC system online coupled to an LTQ Orbitrap Elite hybrid mass spectrometer (Thermo Fischer Scientific). A comprehensive description of the entire glycoproteomic analysis workflow can be found in Pralow et al. and Hoffmann et al. [49,50].

Data analysis

Manual *N*-glycan data analysis

MS datasets obtained from the *N*-glycan analysis of each cell line and their biological replicates (three MDCK.ADH cell and two MDCK.SUS2 cell cultivations) were inspected manually. Specifically, deconvoluted precursor masses were submitted to ExPASy GlycoMod (free online tool: https://web.expasy.org/glycomod/) and UniCarb-DB to identify 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 [45,54]. Identified structures of different *N*-glycan isomers were additionally confirmed based on their retention behavior [42].

Manual/semi-automated glycopeptide data analysis

Manual glycopeptide analysis of a reference data set (first replicate of each cell line) was performed according to

Pralow et al. (2017) and Hoffmann et al. (2018) [49,50]. All acquired MS/MS spectra were manually screened for specific oxonium ions ($[M + H]^+$: HexNAc m/z 204.08, $HexNAc_1Hex_1$ m/z 366.14, $HexNAc_1Hex_1dHex_1$ m/z 512.20, $HexNAc_1Hex_2$ m/z 528.19). Oxonium ions of sialic acid are not supposed to be detected, because of the presence of neuraminidase in IAV containing cell culture supernatants. After the classification of each glycopeptide fragment-ion spectrum, the mass of the peptide moiety (Mp) of each fragment-ion spectrum was determined based on a specific fragment-ion pattern: Peptide mass minus ammonia $[Mp + H^+\text{-}NH_3]^+$, 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) [50]. The putative peptide masses were searched against an *in silico* digestion of HA (Uniprot accession number P03452), NA (Uniprot accession number P03468), and matrixprotein 2 (M2) (Uniprot accession number P06821) using ExPASy FindPept (free online tool https://web.expasy.org/findpept/). Variable modifications were the carbamidomethylation of cysteine and the oxidation of methionine. N-terminal carbamidomethylation was considered 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 into account missed or unspecific cleavage products. The resulting peptide sequences were screened for possible *N*-glycosylation consensus sequences. To confirm the predicted peptide sequence, an *in silico* fragmentation of its peptide sequence 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). Therefore, specific b- and y-ions 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 (first replicate of each sample type), the 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) [46]. The mass error was set to 10 ppm and an unspecific digestion was performed to take into account missed cleavages as well as unexpected cleaved products. 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. The software also allowed 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.

Relative quantification of *N*-glycopeptides

Site-specific relative quantification of *N*-glycopeptides was performed using Byonic and Byologic (ProteinMetrics, Cupertino, CA, USA). 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.

Graphical illustration

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

Acknowledgement

TNK and ER acknowledge support by European Union (EC) under the project 'HTP-GlycoMet' (grant number 324400). MH and ER acknowledge support by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID RA2992/1-1 – Forschungsgruppe FOR 2509. We want so thank Silvana Fischer and Lisa Fichtmüller for their excellent technical assistance.

Conflict of interest.

ER is the founder and CEO of glyXera GmbH, a company, which offers products and services for glycoanalysis and has several patents in the field. UR is coowner; RH is employee of glyXera. All other authors declare no competing interests.

Author contributions

AP, ER, UR, and YG conceived the study, which was supervised by ER, YG, and UR. AP and TNK performed LC-MS measurements, analyzed the data, and prepared figures and tables. RH performed the xCGE-LIF-based *N*-glycomics and analyzed the data. AP, MP, and MH performed the semi-automated *N*-glycopeptide data analysis. AP wrote the original draft with contributions of MH, RH, YG, and ER. All authors reviewed and edited the manuscript.

Data Availability Statement

The datasets generated are available from the corresponding author upon request.

References

1 Lambert LC & Fauci AS (2010) Influenza vaccines for the future. *N Engl J Med* **363**, 2036–2044.
2 Hay AJ, Gregory V, Douglas AR & Lin YP (2001) The evolution of human influenza viruses. *Philos Trans R Soc Lond B Biol Sci* **356**, 1861–1870.
3 de Vries RP, de Vries E, Bosch BJ, de Groot RJ, Rottier PJ & de Haan CA (2010) The influenza A virus hemagglutinin glycosylation state affects receptor-binding specificity. *Virology* **403**, 17–25.
4 Wei C-J, Boyington JC, Dai K, Houser KV, Pearce MB, Kong W-P, Yang Z-Y, Tumpey TM & Nabel GJ (2010) role of glycans in viral evolution and vaccine design. *Sci Transl Med* **2**, 24ra21-24ra21. Cross-neutralization of 1918 and 2009 influenza viruses:.
5 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. *Proc Natl Acad Sci USA* **81**, 1779–1783.
6 Matrosovich MN, Matrosovich TY, Gray T, Roberts NA & Klenk HD (2004) Neuraminidase is important for the initiation of influenza virus infection in human airway epithelium. *J Virol* **78**, 12665–12667.
7 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. *J Virol* **74**, 6316–6323.
8 Li S, Schulman J, Itamura S & Palese P (1993) Glycosylation of neuraminidase determines the neurovirulence of influenza A/WSN/33 virus. *J Virol* **67**, 6667–6673.
9 York IA, Stevens J & Alymova IV (2019) Influenza virus N-linked glycosylation and innate immunity. *Biosci Rep* **39**, BSR20171505.
10 Lamb YN (2019) Cell-based quadrivalent inactivated influenza virus vaccine (Flucelvax® Tetra/Flucelvax Quadrivalent®): a review in the prevention of influenza. *Drugs* **79**, 1337–1348.
11 Manini I, Domnich A, Amicizia D, Rossi S, Pozzi T, Gasparini R, Panatto D & Montomoli E (2015) Flucelvax (Optaflu) for seasonal influenza. *Expert Rev Vaccines* **14**, 789–804.
12 Centers for Disease Control and Prevention, National Center for Immunization and Respiratory Diseases (NCIRD) (2019) Influenza vaccines — United States, 2019–20 influenza season. https://www.cdc.gov/flu/professionals/vaccines.htm
13 Genzel Y, Dietzsch C, Rapp E, Schwarzer J & Reichl U (2010) MDCK and Vero cells for influenza virus vaccine production: a one-to-one comparison up to lab-scale bioreactor cultivation. *Appl Microbiol Biotechnol* **88**, 461–475.
14 U.S. Food and Drug Administration (2020) Influenza Vaccine for the 2020-2021 Season. https://www.fda.gov/vaccines-blood-biologics/lot-release/influenza-vaccine-2020-2021-season
15 Paul-Ehrlich-Institut (2020) Seasonal Influenza Vaccines - Vaccines with a Master Adaptation for 2020/2021. https://www.pei.de/EN/medicinal-products/vaccines-human/influenza-flu/influenza-flu-node.html
16 Genzel Y, Rodig J, Rapp E & Reichl U (2014) Vaccine production: upstream processing with adherent or suspension cell lines. *Methods Mol Biol* **1104**, 371–393.
17 Lohr V, Genzel Y, Behrendt I, Scharfenberg K & Reichl U (2010) A new MDCK suspension line cultivated in a fully defined medium in stirred-tank and wave bioreactor. *Vaccine* **28**, 6256–6264.
18 Tapia F, Genzel Y & Reichl U (2019) Plug flow tubular bioreactor, system containing the same and method for production of virus in. *Google Patents*.
19 Tapia F, Vogel T, Genzel Y, Behrendt I, Hirschel M, Gangemi JD & Reichl U (2014) Production of high-titer human influenza A virus with adherent and suspension MDCK cells cultured in a single-use hollow fiber bioreactor. *Vaccine* **32**, 1003–1011.
20 Serve A, Pieler MM, Benndorf D, Rapp E, Wolff MW & Reichl U (2015) Comparison of influenza virus particle purification using magnetic sulfated cellulose particles with an established centrifugation method for analytics. *Anal Chem* **87**, 10708–10711.
21 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**, 4269–4280.
22 Vester D, Rapp E, Kluge S, Genzel Y & Reichl U (2010) Virus-host cell interactions in vaccine production cell lines infected with different human influenza A virus variants: a proteomic approach. *J Proteomics* **73**, 1656–1669.
23 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**, 4325–4336.
24 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**, 4203–4214.
25 Hennig R, Rapp E, Kottler R, Cajic S, Borowiak M & Reichl U (2015) N-Glycosylation Fingerprinting of Viral Glycoproteins by xCGE-LIF. *Methods Mol Biol* **1331**, 123–143.
26 Varki A (1993) Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology* **3**, 97–130.
27 Moremen KW, Tiemeyer M & Nairn AV (2012) Vertebrate protein glycosylation: diversity, synthesis and function. *Nat Rev Mol Cell Biol* **13**, 448–462.
28 An Y, Parsons LM, Jankowska E, Melnyk D, Joshi M & Cipollo JF (2019) N-Glycosylation of Seasonal Influenza Vaccine Hemagglutinins: Implication for potency testing and immune processing. *J Virol* **93**, e01693–e1718.
29 Khatri K, Klein JA, White MR, Grant OC, Leymarie N, Woods RJ, Hartshorn KL & Zaia J (2016) Integrated omics and computational glycobiology reveal structural basis for Influenza A virus glycan microheterogeneity and host interactions. *Mol Cell Prot* **15**, 1895–1912.
30 She Y-M, Farnsworth A, Li X & Cyr TD (2017) Topological N-glycosylation and site-specific N-glycan sulfation of influenza proteins in the highly expressed H1N1 candidate vaccines. *Sci Rep* **7**, 10232.
31 Harvey DJ (2018) Mass spectrometric analysis of glycosylated viral proteins. *Exp Rev Prot* **15**, 391–412.
32 Sharma VK, Sharma I & Glick J (2018) The expanding role of mass spectrometry in the field of vaccine development. *Mass Spectrom Rev* **39**, 83–104.
33 Hämmerling F, Pieler MM, Hennig R, Serve A, Rapp E, Wolff MW, Reichl U & Hubbuch J (2017) Influence of the production system on the surface properties of influenza A virus particles. *Eng Life Sci* **17**, 1071–1077.
34 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. *J Antimicrob Chemother* **70**, 136–152.
35 She Y-M, Tam RY, Li X, Rosu-Myles M & Sauvé S (2020) Resolving isomeric structures of native glycans by nanoflow porous graphitized carbon chromatography-mass spectrometry. *Anal Chem* **92**, 14038–14046.
36 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.
37 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. *J Virol Methods* **249**, 10–20.
38 Hutter J, Rodig JV, Hoper D, Seeberger PH, Reichl U, Rapp E & Lepenies B (2013) Toward animal cell culture-based influenza vaccine design: viral hemagglutinin N-glycosylation markedly impacts immunogenicity. *Journal Immunol* **190**, 220–230.
39 de Vries RP, Smit CH, de Bruin E, Rigter A, de Vries E, Cornelissen LA, Eggink D, Chung NP, Moore JP & Sanders RW (2012) Glycan-dependent immunogenicity of recombinant soluble trimeric hemagglutinin. *J Virol* 01084–12.
40 Butler M & Spearman M (2014) The choice of mammalian cell host and possibilities for glycosylation engineering. *Curr Opin Biotechnol* **30**, 107–112.
41 Lin SC, Jan JT, Dionne B, Butler M, Huang MH, Wu CY, Wong CH & Wu SC (2013) Different immunity elicited by recombinant H5N1 hemagglutinin proteins containing pauci-mannose, high-mannose, or complex type N-glycans. *PLoS One* **8**, e66719.
42 Pabst M, Bondili JS, 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. *Anal Chem* **79**, 5051–5057.
43 Everest-Dass AV, Abrahams JL, Kolarich D, Packer NH & Campbell MP (2013) Structural feature ions for distinguishing N- and O-linked glycan isomers by LC-ESI-IT MS/MS. *J Am Soc Mass Spectrom* **24**, 895–906.
44 Harvey DJ (2005) Fragmentation of negative ions from carbohydrates: part 3. Fragmentation of hybrid and complex N-linked glycans. *J Am Soc Mass Spect* **16**, 647–659.
45 Harvey DJ, Royle L, Radcliffe CM, Rudd PM & Dwek RA (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.
46 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. *Anal Chem.*

47 Bern M, Kil YJ & Becker C. (2012) Byonic: advanced peptide and protein identification software. *Current Protocols in Bioinformatics* 40, 13.20. 1–13.20. 14.
48 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. *Glycoconj J* **35**, 499–509.
49 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. *Sci Rep* **7**, 11419.
50 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*.
51 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 Mol Biol* **1321**, 427–435.
52 Jensen PH, Karlsson NG, Kolarich D & Packer NH (2012) Structural analysis of N- and O-glycans released from glycoproteins. *Nat Protoc* **7**, 1299–1310.
53 Shaw ML, Stone KL, Colangelo CM, Gulcicek EE & Palese P (2008) Cellular proteins in influenza virus particles. *PLoS Pathog* **4**, e1000085.
54 Everest-Dass AV, Kolarich D, Campbell MP & Packer NH (2013) Tandem mass spectra of glycan substructures enable the multistage mass spectrometric identification of determinants on oligosaccharides. *Rapid Communications in Mass Spectrometry: RCM*. **27**, 931–939.
55 Nguyen-Khuong T, Everest-Dass AV, Kautto L, Zhao ZJ, Willcox MDP & Packer NH (2015) Glycomic characterization of basal tears and changes with diabetes and diabetic retinopathy. *Glycobiology* **25**, 269–283.
56 Domon B & Costello CE (1988) A Systematic Nomenclature for Carbohydrate Fragmentations in Fab-Ms Ms Spectra of Glycoconjugates. *Glycoconj J* **5**, 397–409.
57 Rodig JV, 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. *Biotechnol Bioeng* **110**, 1691–1703.
58 Ashwood C, Pratt B, MacLean BX, Gundry RL & Packer NH (2019) Standardization of PGC-LC-MS-based glycomics for sample specific glycotyping. *The Analyst*.
59 Cruz E, Cain J, Crossett B & Kayser V (2018) Site-specific glycosylation profile of influenza A (H1N1) hemagglutinin through tandem mass spectrometry. *Hum Vacc Immunotherap* **14**, 508–517.
60 Liu Y-J, Wu S-L, Love KR & Hancock WS (2017) Characterization of site-specific glycosylation in influenza A virus hemagglutinin produced by Spodoptera frugiperda insect cell line. *Anal Chem* **89**, 11036–11043.
61 Parsons LM, An Y, de Vries RP, de Haan CA & Cipollo JF (2017) Glycosylation characterization of an influenza H5N7 hemagglutinin series with engineered glycosylation patterns: implications for structure–function relationships. *J Proteome Res* **16**, 398–412.
62 Sarkar A, Bale S, Behrens A-J, Kumar S, Sharma SK, de Val N, Pallesen J, Irimia A, Diwanji DC & Stanfield RL (2018) Structure of a cleavage-independent HIV Env recapitulates the glycoprotein architecture of the native cleaved trimer. *Nat Commun* **9**, 1–14.
63 De La Pena AT, Rantalainen K, Cottrell CA, Allen JD, Van Gils MJ, Torres JL, Crispin M, Sanders RW & Ward AB (2019) Similarities and differences between native HIV-1 envelope glycoprotein trimers and stabilized soluble trimer mimetics. *PLoS Pathog* **15**, e1007920.
64 Behrens A-J, Vasiljevic S, Pritchard LK, Harvey DJ, Andev RS, Krumm SA, Struwe WB, Cupo A, Kumar A & Zitzmann N (2016) Composition and antigenic effects of individual glycan sites of a trimeric HIV-1 envelope glycoprotein. *Cell Rep* **14**, 2695–2706.
65 Watanabe Y, Allen JD, Wrapp D, McLellan JS & Crispin M (2020) Site-specific glycan analysis of the SARS-CoV-2 glycan shield. *BioRxiv*.
66 Hinneburg H, Hofmann J, Struwe WB, Thader A, Altmann F, Varon Silva D, Seeberger PH, Pagel K & Kolarich D (2016) Distinguishing N-acetylneuraminic acid linkage isomers on glycopeptides by ion mobility-mass spectrometry. *Chem Commun (Camb)* **52**, 4381–4384.
67 Abdel-Motal UM, Guay HM, Wigglesworth K, Welsh RM & Galili U (2007) Immunogenicity of influenza virus vaccine is increased by anti-gal-mediated targeting to antigen-presenting cells. *J Virol* **81**, 9131–9141.
68 Galili U (2020) Amplifying immunogenicity of prospective Covid-19 vaccines by glycoengineering the coronavirus glycan-shield to present α-gal epitopes. *Vaccine* **38**, 6487–6499.
69 Tseng YC, Wu CY, Liu ML, Chen TH, Chiang WL, Yu YH, Jan JT, Lin KI, Wong CH & Ma C (2019) Egg-based influenza split virus vaccine with monoglycosylation induces cross-strain protection against influenza virus infections. *Proc Natl Acad Sci USA* **116**, 4200–4205.
70 Wang CC, Chen JR, Tseng YC, Hsu CH, Hung YF, Chen SW, Chen CM, Khoo KH, Cheng TJ, Cheng YS *et al.* (2009) Glycans on influenza hemagglutinin affect receptor binding and immune response. *Proc Natl Acad Sci USA* **106**, 18137–18142.

71 Lee J, Paparoditis P, Horton AP, Fruhwirth A, McDaniel JR, Jung J, Boutz DR, Hussein DA, Tanno Y, Pappas L *et al.* (2019) Persistent Antibody Clonotypes Dominate the Serum Response to Influenza over Multiple Years and Repeated Vaccinations. *Cell Host Microbe* **25**, 367–376.e5.
72 Tate MD, Job ER, Deng YM, Gunalan V, Maurer-Stroh S & Reading PC (2014) Playing hide and seek: how glycosylation of the influenza virus hemagglutinin can modulate the immune response to infection. *Viruses* **6**, 1294–1316.
73 Vigerust DJ, Ulett KB, Boyd KL, Madsen J, Hawgood S & McCullers JA (2007) N-linked glycosylation attenuates H3N2 influenza viruses. *J Virol* **81**, 8593–8600.
74 Tate MD, Brooks AG & Reading PC (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. *J Immunol* **187**, 1884–1894.
75 Tate MD, Job ER, Brooks AG & Reading PC (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.
76 Caton AJ, Brownlee GG, Yewdell JW & Gerhard W (1982) The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype). *Cell* **31**, 417–427.
77 Cox MM & Hollister JR (2009) FluBlok, a next generation influenza vaccine manufactured in insect cells. *Biologicals* **37**, 182–189.
78 Cox MM, 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.
79 Katz JM & Webster RG (1989) Efficacy of inactivated influenza A virus (H3N2) vaccines grown in mammalian cells or embryonated eggs. *J Infect Dis* **160**, 191–198.
80 Wisniewski JR, Zougman A, Nagaraj N & Mann M (2009) Universal sample preparation method for proteome analysis. *Nat Methods* **6**, 359–362.
81 Hinneburg H, Chatterjee S, Schirmeister F, Nguyen-Khuong T, Packer NH, Rapp E & Thaysen-Andersen M (2019) Post-column make-up flow (PCMF) enhances the performance of capillary-flow PGC-LC-MS/MS-based glycomics. *Anal Chem* **91**, 4559–4567.
82 Selman MH, Hemayatkar M, Deelder AM & Wuhrer M (2011) Cotton HILIC SPE microtips for microscale purification and enrichment of glycans and glycopeptides. *Anal Chem* **83**, 2492–2499.
83 Varki A, Cummings RD, Aebi M, Packer NH, Seeberger PH, Esko JD, Stanley P, Hart G, Darvill A, Kinoshita T *et al.* (2015) Symbol Nomenclature for Graphical Representations of Glycans. *Glycobiology* **25**, 1323–1324.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Viral *N*-glycome list of influenza A virus (IAV) produced in MDCK.ADH and MDCK.SUS2 cells.

Table S2. Site-specifc glycopeptide analysis of influenza A virus glycoproteins propagated in MDCK.SUS2 and MDCK.ADH cells using nano-RP-LC MS(/MS).

Table S3. Exemplary proteomic analysis of g-force step-gradient ultracentrifugation purified IAV propagated in MDCK.ADH.

Table S4. Exemplary proteomic analysis of g-force step-gradient ultracentrifugation purified IAV propagated in MDCK.SUS2.

Fig. S1. 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.

Fig. S2. Site-specific *N*-glycopeptide analysis using nano-RP-LC-MS(/MS).

Fig. S3. Site-specific *N*-glycopeptide analysis using nano-RP-LC-MS(/MS).

Fig. S4. Site-specific *N*-glycopeptide analysis using nano-RP-LC-MS(/MS).

Fig. S5. Site-specific *N*-glycopeptide analysis using nano-RP-LC-MS(/MS).

Fig. S6. Site-specific *N*-glycopeptide analysis using nano-RP-LC-MS(/MS).

Fig. S7. BPC of the replicates from the nano-RP-LC-MS(/MS) glycopeptide analysis of IAV propagated in MDCK.ADH.

Fig. S8. BPC of the replicates from the nano-RP-LC-MS(/MS) glycopeptide analysis of IAV propagated in MDCK.SUS2.

Fig. S9. BPC of the replicates from the nano-PGC-LC-MS(/MS) glycan analysis of IAV propagated in MDCK.ADH.

Fig. S10. BPC of the replicates from the nano-PGC-LC-MS(/MS) glycan analysis of IAV propagated in MDCK.SUS2.