AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Virology®

VIRUS-CELL INTERACTIONS

Influenza Virus Hemagglutinins H2, H5, H6, and H11 Are Not Targets of Pulmonary Surfactant Protein D: *N*-Glycan Subtypes in Host-Pathogen Interactions

Lisa M. Parsons,[a] Yanming An,[b] Li Qi,[c] Mitchell R. White,[d] Roosmarijn van der Woude,[e] Kevan L. Hartshorn,[d] Jeffery K. Taubenberger,[c] Robert P. de Vries,[e] John F. Cipollo[b]

[a]Food and Drug Administration, Center for Biologics Evaluation and Research, Division of Bacterial, Parasitic and Allergenic Products, Silver Spring, Maryland, USA
[b]Food and Drug Administration, Center for Drug Evaluation and Research, DBRRII, Silver Spring, Maryland, USA
[c]Viral Pathogenesis and Evolution Section, Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases, NIH, Bethesda, Maryland, USA
[d]Department of Medicine, Boston University School of Medicine, Boston, Massachusetts, USA
[e]Department of Chemical Biology & Drug Discovery, Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Utrecht, The Netherlands

ABSTRACT Seasonal influenza carrying key hemagglutinin (HA) head region glycosylation sites can be removed from the lung by pulmonary surfactant protein D (SP-D). Little is known about HA head glycosylation of low-pathogenicity avian influenza virus (LPAIV) subtypes. These can pose a pandemic threat through reassortment and emergence in human populations. Since the presence of head region high-mannose glycosites dictates SP-D activity, the ability to predict these glycosite glycan subtypes may be of value. Here, we investigate the activities of two recombinant human SP-D forms against representative LPAIV strains, including H2N1, H5N1, H6N1, H11N9, an avian H3N8, and a human seasonal H3N2 subtype. Using mass spectrometry, we determined the glycan subclasses and heterogeneities at each head glycosylation site. Sequence alignment and molecular structure analysis of the HAs were performed for LPAIV strains in comparison to seasonal H3N2 and avian H3N8. Intramolecular contacts were determined between the protein backbone and glycosite glycan based on available three-dimensional structure data. We found that glycosite "N165" (H3 numbering) is occupied by high-mannose glycans in H3 HA but by complex glycans in all LPAIV HAs. SP-D was not active on LPAIV but was on H3 HAs. Since SP-D affinity for influenza HA depends on the presence of high-mannose glycan on the head region, our data demonstrate that SP-D may not protect against virus containing these HA subtypes. Our results also demonstrate that glycan subtype can be predicted at some glycosites based on sequence comparisons and three-dimensional structural analysis.

IMPORTANCE Low-pathogenicity avian influenza virus (LPAIV) subtypes can reassort with circulating human strains and pandemic viruses can emerge in human populations, as was seen in the 1957 pandemic, in which an H2 virus reassorted with the circulating H1N1 to create a novel H2N2 genotype. Lung surfactant protein D (SP-D), a key factor in first-line innate immunity defense, removes influenza type A virus (IAV) through interaction with hemagglutinin (HA) head region high-mannose glycan(s). While it is known that both H1 and H3 HAs have one or more key high-mannose glycosites in the head region, little is known about similar glycosylation of LPAIV strains H2N1, H5N1, H6N1, or H11N9, which may pose future health risks. Here, we demonstrate that the hemagglutinins of LPAIV strains do not have the required high-mannose glycans and do not interact with SP-D, and that sequence analysis can predict glycan subtype, thus predicting the presence or absence of this virulence marker.

Citation Parsons LM, An Y, Qi L, White MR, van der Woude R, Hartshorn KL, Taubenberger JK, de Vries RP, Cipollo JF. 2020. Influenza virus hemagglutinins H2, H5, H6, and H11 are not targets of pulmonary surfactant protein D: *N*-glycan subtypes in host-pathogen interactions. J Virol 94:e01951-19. https://doi.org/10.1128/JVI.01951-19.

Editor Stacey Schultz-Cherry, St. Jude Children's Research Hospital

Address correspondence to John F. Cipollo, john.cipollo@fda.hhs.gov.

Received 19 November 2019
Accepted 1 December 2019
Accepted manuscript posted online 11 December 2019
Published 14 February 2020

KEYWORDS glycomics, avian viruses, collectin, glycosylation, hemagglutinin, mass spectrometry, pandemic, vaccine, virulence marker

Surfactant D (SP-D) is a key innate immune system collectin that is detected in a range of tissues, with high concentrations in the lung. This secreted soluble protein binds to high-mannose glycans (1) and other glycoconjugates present on host pathogens (2) and serves as an important component of the initial barrier to infection by influenza virus (3). Human SP-D has been crystalized in complex with the canonical *N*-glycan $Man_9GlcNAc_2$ (4). This glycan is minimally processed and arises from the $Glc_3Man_9GlcNAc_2$ that is transferred to all *N*-glycoproteins immediately after passing through the rough endoplasmic reticulum (ER) membranes. X-ray crystal structures show that the Manα1,2Manα1,2Manα1,3 arm of the canonical mannan, which is minimally processed by endoplasmic reticulum (ER) and not at all by Golgi enzymes, binds the SP-D ligand binding site (5, 6).

Site N165 in H3 appears to be important for SP-D interaction. Hartshorn et al. reported that *N*-glycan is absent at position 165 in H3N2-subtype HAs in SP-D collectin-resistant strains (7). Removal of site N165 from H3 by genetic modification or by its glycosylation with the high-mannose-specific endoglycosidase H diminishes binding of SP-D (8–10). X-ray crystal studies by Crouch et al. demonstrated that the high-mannose glycan at N165 binds in the receptor site of SP-D through a calcium-coordinated interaction between mannose, R343, R349, and E333, and additional direct contact between C2 oxygen of the lower-arm alpha1,3-linked mannose and A325 (6).

We have shown previously that site N165 on H3 influenza type A viruses (IAVs) is populated with high-mannose glycans (11, 12). Adding sites N133 and N246 increased sensitivity to SP-D as measured by SP-D hemagglutinin inhibition (HAI) assay (7). Modeling studies showed that these sites are within distances required to interact with SP-D (7). We and others have confirmed that N133, N165, and N246 are populated with high-mannose glycans in H3 HAs (12, 13), and our recent study shows that site N144 can be partially occupied by high-mannose glycan (11, 14).

In vivo experiments suggest that interaction with SP-D reduces virulence. Addition of head glycosites such as N246 decreases viral infection in mouse models (11, 15). Likewise, lack of SP-D activity has been shown to contribute to virulence. Previous analysis of the glycosylation patterns of pandemic and pandemic-like hemagglutinin (HA) proteins from representative viruses of the 1957 H2N2 pandemic (A/Singapore/1/1957), the 1968 H3N2 pandemic (A/Aichi/1/1968), and a close relative of the pandemic 1918 influenza virus (A/Swine/Iowa/1931) have an absence of high-mannose glycosites on the HA head (7, 14, 16). Over time, genetic drift of seasonal strains has led to an increase in glycosylation sites on the head of H3 HAs (14). Addition of historically important glycosylation sites to the head region of the Hong Kong 1968 H3 HA backbone diminished virulence and increased lung clearance in mice (12). When SP-D deficient mice were used, virulence was increased, strongly supporting the notion that SP-D is a key factor for lung clearance. Cytotoxicity in the lungs was also correlated with SP-D presence or absence.

Hartshorn et al. reported that H1N1 follows a similar increased sensitivity to SP-D as the virus gains HA head glycosylation sites (7). In this case, head region glycosylation sites N142/144 and N172 were found to increase SP-D sensitivity. Loss of these sites increases virulence in mice (17). Recently, in A/New Caledonia/20/1999 H1N1, these sites were identified as primarily high-mannose glycosites (18). In H1N1, sites 144, 172, and 177 have been identified as pro-SP-D activity, and site 104 is important in tethering to SP-D (7). Consistent with SP-D removal of highly mannosylated influenza from the lung, H1N1 virus with highly mannosylated HA led to smaller amounts of lung cytopathology (17). Our recent publication (11) demonstrated that N104 (N87 after signal sequence removal in the cited manuscript) is populated with about 50% high-mannose glycoforms when the virus is produced in eggs. This is also consistent with reports by

She et al. (15) and Cruz et al. (18) that H1N1 site N104 is partially populated with high-mannose glycans and N177 is nearly exclusively high-mannose glycans.

LPAIV strains are highly similar in sequence and structure to the H1 and H3 HA types but are not predicted to be heavily glycosylated in the head. This situation may lead to low SP-D activity against these strains. We previously analyzed the site-specific glycosylation patterns of avian H5N1 (19) and H5N7 HAs (20). These viruses are recognized for potential pandemic threat (21–23), and glycosylation at the globular head region was limited. Interestingly, among all H5 hemagglutinins examined, each contained complex glycans exclusively at the site closest to N165 in H3. Consistent with the observation that SP-D binds HA head region high-mannose glycans and not the complex ones, Hillaire et al. showed that porcine SP-D neutralized H1 and H3 IAV but failed to neutralize three H5 IAVs from different clades (24).

To better understand why H3 carries high-mannose glycans and H5 displays complex glycans, we analyzed X-ray crystal structures at the N165 region of H3 and H5 HAs (20). Site N165 in H3 is present at the beginning of a β-strand, close to the top of the HA trimer. The high-mannose core of the H3 N165 glycan forms extensive contacts with the 220 loop, especially with W222, which was present until the year 2003. Subsequent to 2003, arginine replaced Trp(W) at 222 with similar packing to the glycan. In H5 HAs, the glycosylation site is shifted forward on the β-sheet and forms few contacts with the peptide backbone (20). We hypothesized that the extensive contacts to the core glycan in H3 makes the glycan less accessible to ER and Golgi processing, resulting in the observed high-mannose glycan subtype. In comparison, the H5 HAs had fewer peptide backbone contacts, leading to more exposure to ER and Golgi glycan processing enzymes and resulting in the observed complex glycan subclass at that site. The "N165" site in H5 aligns sequentially with glycosylation sites in H2, H6, and H11 HAs, suggesting that they may also be complex and lack interaction with SP-D.

IAVs are an extremely divergent group of RNA viruses that infect a variety of warm-blooded animals, including birds, horses, pigs, and humans. According to the Influenza Research Database (www.fludb.org), 16 out of 18 of the known subtypes occur in birds, lending credence to the hypothesis that the natural reservoir of influenza A viruses is the wild bird population. IAV in wild birds exists as a vast pool of genomic segments that form transient genome constellations (25, 26). Since IAV contains a segmented RNA genome, mixed infection can lead to genetic reassortment, in which the virus may acquire the ability to switch hosts and cause novel, sometimes pandemic, outbreaks in new species. Pandemic IAV containing gene segments derived from avian influenza viruses caused the last four influenza pandemics in humans (25, 26). Based on these data, it is inevitable that there will be future outbreaks due to reassortment and passage of avian IAV to humans. A better understanding of the potential virulence of different IAV subtypes in humans is needed. In this report, we identify the glycoforms on the globular head region of H2N1, H3N8, H5N1, H6N1, H11N9, and seasonal H3N2 strains. MDCK cells were used exclusively to control for cell substrate-specific glycosylation effects. The preference of SP-D for high-mannose forms over complex forms was demonstrated by testing the same HA decorated with different glycoforms. Relative binding preferences to SP-D were examined. Molecular analysis of existing X-ray crystal structures, in combination with glycan compositions at key head sites, revealed structural components and related sequences that may aid in prediction of glycan subtype and thus SP-D activity on infection with these and similar strains.

RESULTS

SP-D binding characteristics of avian and human IAV HAs. We tested the relative binding preferences of H2N1, H3N8, H3N2, H5N1, H6N1, and H11N9 for recombinant human SP-D (rhSP-D [dodecamer]) and a lower-affinity truncated version of SP-D composed of trimers of just the neck and carbohydrate recognition domains (CRDs) of the molecule and containing the D325A and R343V substitutions as previously reported (hnSP-D) (27). The hnSP-D construct was engineered to allow some degree of IAV-inhibiting activity in a trimeric construct, since neck and CRD trimers of wild type SP-D

TABLE 1 Hemagglutinin inhibition assay using turkey erythrocytes

Virus	Subtype	SP-D concn (μg) by HAI assay	
		rhSP-D (dodecamer)	hnSP-D[a]
A/green-winged teal/Ohio/175/1986	H2N1	>8	>8
A/pintail/339/1987	H3N8	0.5 (partial inhibition)	>8
A/NY/470/2004	H3N2 (human)	0.5	8
A/mallard/Maryland/802/2007	H5N1	>8	>8
A/mallard/Ohio/249/1998	H6N1	>8	>8
A/green-winged teal/Ohio/340/1987	H11N9	>8	>8

[a]hnCRD SP-D D325A R343V.

essentially lack antiviral activity. Both turkey red blood cells (RBCs) and human RBCs were used in the hemagglutinin inhibition-based assay (see Materials and Methods). Results are shown in Tables 1 and 2, respectively. For both turkey and human RBCs, high SP-D concentrations (above 8 μg/ml) did not inhibit hemagglutination by H2N1, H5N1, H6N1, and H11N9 strains. High-affinity rhSP-D inhibited hemagglutination of both types of RBCs by H3N2 and H3N8 but in differing amounts. Hemagglutination of turkey RBCs by the human seasonal H3N2 virus was inhibited at 0.5 μg, and avian H3N8 was partially inhibited by rhSP-D at the same amount. Only 0.125 μg of hrSP-D was needed to inhibit hemagglutination of human RBCs by H3N2, but 1 μg was required for partial inhibition of H3N8. Hemagglutination inhibition in the lower-affinity hnSP-D assays was only observed for the H3N2 strain, for which 8 μg protected turkey RBCs and 4 μg protected human RBCs.

Plaque reduction test for neutralization of viruses by SP-D. Plaque reduction neutralization (PRN) assays (Fig. 1) showed that 0.125 μg/ml of SP-D protein was sufficient for the human A/NY470/2004 seasonal H3N2 virus to reduce plaque formation by 100% in MDCK cells (red). The concentration of SP-D needed to reduce plaque formation in H3N2 was significantly different compared to that for all other viruses for both the 0.125 μg/ml and 0.5 μg/ml. A two-way analysis of variance (ANOVA) test showed that with SP-D at 0.125 μg/ml and 0.5 μg/ml concentrations, the human H3N2 virus showed a significant plaque number reduction compared to that of all other avian influenza viruses ($P < 0.037$). Between the rest of the avian strains tested, no significant differences were observed at those SP-D concentrations. At 1 μg/ml SP-D, the H3N2 and avian H3N8 and H2N1 strains showed a significant reduction of plaque number compared to those of other avian strains ($P < 0.05$). The human H3N2 strain, however, caused more reduction than did avian H3N8 and H2N1 ($P < 0.05$). Incubation with the highest concentration of SP-D tested (2 μg/ml) caused a significant (85 to 100%) plaque number reduction in all viruses tested ($P > 0.05$), as shown in Fig. 1. Taken together, these data support the notion that low SP-D binding activity of avian HA proteins may contribute to disease severity and that the head regions of these HAs may lack high-mannose glycans.

Glycomics analysis of avian and pandemic strain HAs. We performed liquid chromatography mass spectrometry (LC/MS^E) analysis of HAs from LPAIV H2N1, H5N1, H6N1, H11N9, and H3N8 strains, as well as a human H3N2 IAV, focusing on the HA head region where interactions with SP-D occur. The workflow is shown in Fig. 2. The H2N1,

TABLE 2 Hemagglutinin inhibition assay using human type O red blood cell

Virus	Subtype	SP-D concn (μg) by HAI assay	
		rhSP-D (dodecamer)	hnSP-D[a]
A/green-winged teal/Ohio/175/1986	H2N1	>8	>8
A/pintail/339/1987	H3N8	1 (partial inhibition)	>8
A/NY/470/2004	H3N2 (human)	0.125	4
A/mallard/Maryland/802/2007	H5N1	>8	>8
A/mallard/Ohio/249/1998	H6N1	>8	>8
A/green-winged teal/Ohio/340/1987	H11N9	>8	>8

[a]hnCRD SP-D D325A R343V.

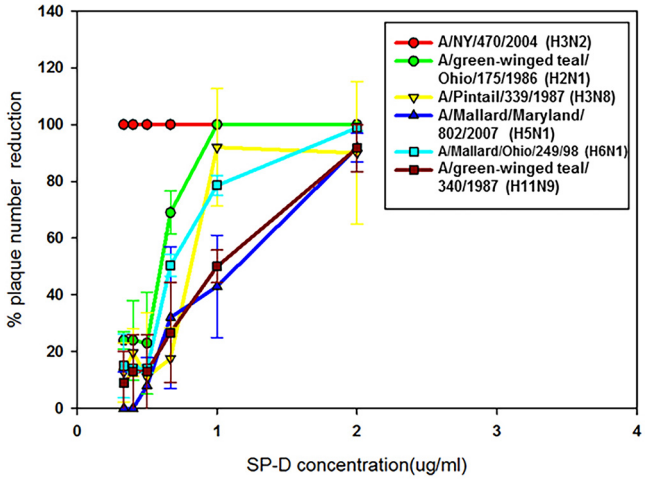

FIG 1 SP-D neutralization and plaque reduction assay. Human H3N2 virus A/NY/470/200 and avian viruses A/green-winged teal/Ohio/175/1986(H2N1), A/pintail/339/1987(H3N8), A/mallard/Maryland/802/2007(H5N1), A/mallard/Ohio/249/1998(H6N1), and A/green-winged teal/Ohio/340/1987(H11N9) (see Table 1) were tested for their ability to bind to SP-D using a SP-D neutralization and plaque reduction assay as described in the Materials and Methods section. Data are presented as mean ± standard deviation (SD). Two-way analysis of variance (ANOVA) was applied for statistics. The difference in the mean among different viruses had a significance of $P \leq 0.001$. The difference in the mean among different SP-D concentrations had a significance of $P \leq 0.001$.

H5N1, H6N1, and H11N9 HAs only have a single glycosylation site on the head region, and this site aligns with that of H3 strain glycosite N165. Here, we refer to these analogous potentially pandemic avian HA glycosites as "N165." The glycan composition of this site and others across the different HA subtypes of various LPAIV strains at this site are shown in Tables S1 to S6. Example MS^E spectra from avian and seasonal strains are shown in Fig. 3. Without exception, every LPAIV HA contained exclusively, or nearly exclusively, complex glycans at "N165." In comparison, the avian H3N8 contained exclusively high-mannose glycans at N165, in line with what has been reported

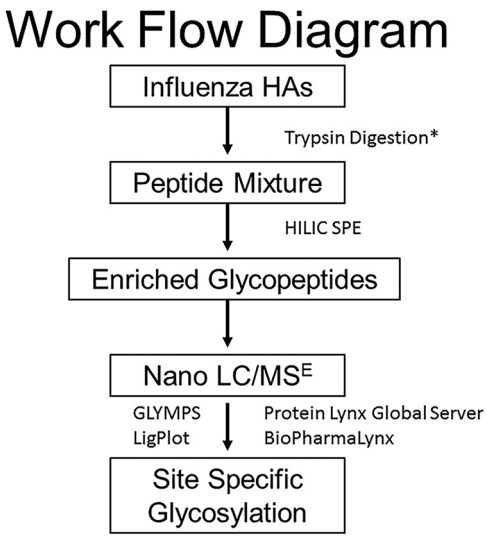

FIG 2 Workflow diagram. Denatured viruses were processed using standard trypsinization after alkylation. The H11N9 was processed by sequential trypsin and chymotrypsin digestion. Peptide mixtures were enriched for glycopeptide analysis using Amide-80 hydrophilic interaction chromatography (HILIC) solid-phase extraction (SPE). Enriched peptides were analyzed by LC/MS^E, and spectral data were processed using Protein Lynx Global Server, BiopharmaLynx, and GLYMPS, the last of which is an in-house-produced software program for analysis of glycopeptides. LigPlot was used to plot the three-dimensional monomer representations shown in Fig. 3.

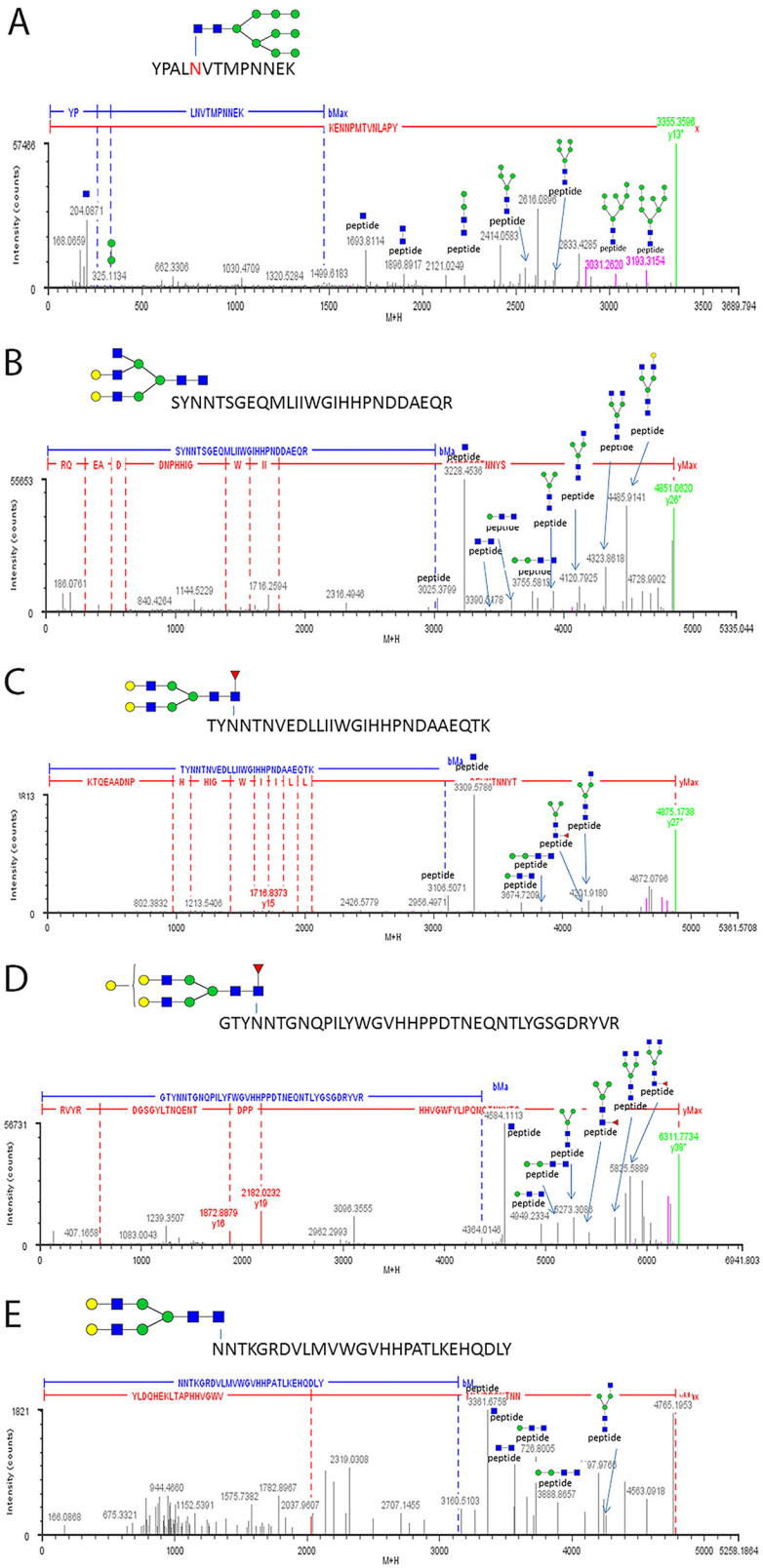

FIG 3 Montage of representative mass spectra of glycopeptides containing "N165" for each HA type. Site "N165" contains high-mannose glycans in HA of H3 and complex glycans in HA of LPAIV strains. The spectra are as follows: A, H3N2 A/NY/470/2004 HA tryptic glycopeptides 161 to 173; B, H2N1 A/green-winged teal/Ohio/175/1986 HA tryptic glycopeptides 162 to 187; C, H5N1 A/mallard/Maryland/802/2007 HA tryptic glycopeptides 163 to 189; D, H6N1 A/mallard/Ohio/249/1998HA tryptic glycopeptides 163 to 200; E, H11N1 A/green-winged teal/Ohio/175/1986 HA chymotryptic glycopeptides 165 to 191.

previously in human H3 IAV (see Table S5). As expected, and in addition to N165, the human H3N2 virus contained high-mannose glycans at the head region glycosylation sites N122, N133, N144, and N246 (see Table S6), likely facilitating the observed greater binding preference seen in SP-D assays. Glycan compositions detected at other HA sites are also shown in Tables S1 to S6. As previously described, the tendency is for increasingly complex glycans moving from the head to the bottom of the stem in HA. H3N2 had complex glycans at head sites N63 and N126. Stem sites, where glycan was detected, were complex and were not predicted to interact with SP-D, as this region is also not in the region demonstrated to interact with SP-D. The H3N8 strain contained predominantly complex glycans at stem sites with a small minority of high-mannose glycans at stem sites N38 and N285. H2N1 HA-detected stem glycans were essentially all complex. In summary, avian strains H2N1, H5N1, H6N1, and H11N9 all contained complex glycans at glycosylation site "N165," while H3N2 seasonal and the avian H3N8 strains contained high-mannose glycans at this site. While the seasonal H3N2 human HA contained several high-mannose sites at the head region of HA, the avian strains had only one site, N165, on the HA head. LPAIV strains also did not have a significant amount of high-mannose glycans in the stem region. These results matched well with expected mannosylation predicted by the SP-D assays.

Structural analysis of the N165 glycosylation site. Previously, we reported that compared to H3, H5 HA "N165" is shifted four amino acids toward the C terminus. Drawing from X-ray crystal structures of 3.0-Å or better resolution, our analysis revealed 20 or more intramolecular contacts of the inner chitobiose of the N165 high-mannose glycan with the intersubunit 220 loop in H3. These contacts were primarily with W222 until 2003, when an arginine replaced tryptophan at position 222 (20). The arrangement is shown in Fig. 4A and is described further below.

The consensus sequences of H2, H3, H5, H6, and H11 HAs were aligned (see Materials and Methods for details). Regions containing glycosite "N165" and the 220 loop are shown in Fig. 4. Compared to the consensus H3 HA sequence, the N165-aligned glycosylation site of consensus H2, H5, H6, and H11 sequences is shifted four amino acids toward the C terminus, suggesting a similar orientation to that observed in H5 HA in our recent study (20). We analyzed available X-ray crystal structure data of H2, H5, and H6 where structures with 3.0-Å or better resolution were available. No such structure was available for H11 HA. The H3 HA has up to 39 close contacts between the N165 glycan $GlcNAc_2$ portion of the core and W222, with additional contacts with S219 and P221. The sugar pair forms a planar arrangement with the W residue. The arrangement of the core $GlcNAc_2$ moiety extending from N165 of H3 HA with the intersubunit 220 loop contacts is shown in Fig. 4A. Contacts are listed in Table 3.

As shown in Fig. 4B, H2 HA (A/ck/New York/91; PDB code 2WR0) does not have W222 and does not form close contacts with the 220 loop. Contacts are observed between the inner aglycone GlcNAc of the core chitobiose unit of "N165" with V241 and W239 on the same subunit. There are 17 van der Waals contacts and one H bond observed (see Table 3). However, unlike the H3 HAs, the arrangement of the glycan and W is nonplanar. The glycan faces out and away from the 220 loop and is extended 4 amino acids away from the HA surface compared to that of H3. The contacts observed in the H2 HA X-ray crystal structures (PDB codes 3QQB, 3KU3, and 4W8N) were also examined in the "N165" and 220 loop regions. Contacts are listed in Table 3. In all three cases, no contacts are made to the neighboring subunit or to the 220 loop. All contacts are to residues on the same subunit, including a tryptophan in the H2 HAs with additional contacts with T171 (PDB codes 3QQB and 3KU3) or V237 and T166 (PDB code 4W8N). The 4W8N structure is most similar to that of the A/green-winged teal/Ohio/175/1986 (H2N1) HA analyzed here by LC/MS^E. In all cases, the arrangement between the inner core chitobiose and the tryptophan is nonplanar. As seen in PDB structure 2WR0 shown in Fig. 4B, "N165" and the attached glycan are extended out and away from the 220 loop by 4 amino acids compared to those of H3 HA. This situation may allow the H2 "N165" glycosylation site and nascent glycan chain to be more accessible

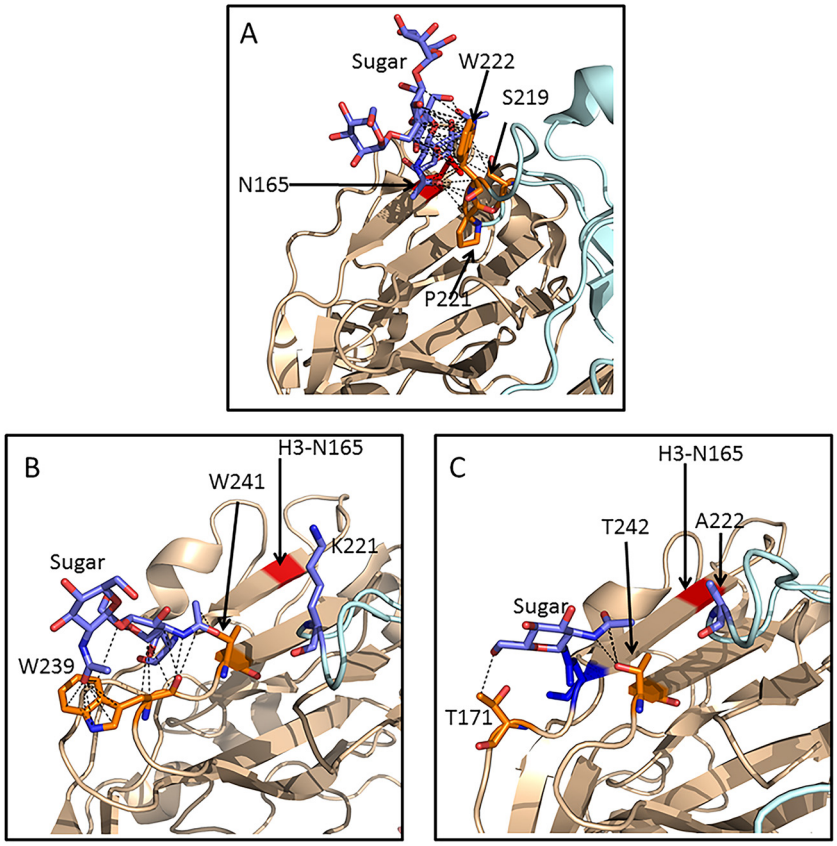

FIG 4 Ribbon representations of H3, H2, and H6 hemagglutinins. Van der Waals contacts between the inner core chitobiose (blue), as seen in crystal structures, and neighboring amino acids are shown as black dashed lines. Separate HA subunits are in wheat and pale cyan. (A) H2 HA from A/ck/New York/91 (PDB code 2WR0). Contacts within the same subunit to W239 and V241 are shown. The red stripe indicates the position of glycosylation in H3 hemagglutinin. K221 is in place of W222 in the H3 protein sequence. (B) H6 HA from A/Taiwan/2013 (PDB code 4XKD). Van der Waals contacts of inner core chitobiose with intrasubunit T171 and T242 shown. As in panel A, red stripe indicates where the H3 N165 glycosylation site resides. A222 on the neighboring subunit is in place of W222 as seen in H3.

to ER and Golgi glycan processing enzymes, allowing processing to occur and giving rise to complex glycans, which is consistent with the observation here that the site is populated exclusively by complex glycans rather than by the comparatively unprocessed high-mannose glycans seen in the H3 HA case.

TABLE 3 Intramolecular and intermolecular contacts of *N*-glycan core chitobiose in H3, H2, and H6 HA crystal structures[c]

PDB code	HA type	Residue(s) within 2.5–4.25 Å	Max no. of contacts[b]	No. of H bonds
4FNK	H3	W222,[a] S219,[a] P221,[a] V223,[a] R224[a]	39	2
1JSM	H5	P217,[a] N236, N237, A238	16	1
2WR0	H2	W239, V2441	17	1
3QQB	H2	W240, T171	9	0
3KU3	H2	W240, T171	9	0
4W8N	H2	W235, V237, T166	12	1
5BNY	H6	T168, T239	6	0
5T08	H6	T242	5	1
5BR0	H6	T240	3	0
4XKD	H6	T171, T242	5	1

[a]Contacts are intersubunit.
[b]Number of contacts varies by subunit. The maximum number found is shown for each PDB file.
[c]All structures were analyzed at 3.0-Å or better resolution.

```
H2       --GSNYPIAKRSYNNTSGEQMLIIWGVHHPNDDAEQRTLYQNVGTYVSVGTSTLNKRSIP     225
H3       --NFKYPALNVTMPNNEQFDKLYIWGVHHPGTDKDQIFLYAQASGRITVSTKRSQQTVIP     231
H5       --NNAYPTIKRSYNNTNQEDLLVLWGIHHPNDAAEQTKLYQNPTTYISVGTSTLNQRLVP     227
H6       -KSAAYPVIKGTYNNTGNQPILYFWGVHHPPDTNEQNTLYGSGDRYVRMGTESMNFAKSP     228
H11      --SGTYPVIRRTFNNTKGRDVLMVWGVHHPATLKEHQDLYKKDSSYVAVGSESYNRRFTP     227
         * .  **  . :  *.     * .**:***    ::  ** .    : :.:.  :    *

H2       EIATRPKVNGQGGRMEFSWTLLDTWDVINFESTGNLIAPEYGFKISKR--GSSGIMKTEK     283
H3       NIGSRPRVRGIPSRISIYWTIVKPGDILLINSTGNLIAPRGYFKIRS---GKSSIMRSDA     288
H5       KIATRSKVNGQSGRMEFFWTILKPNDAINFESNGNFIAPEYAYKIVKK--GDSTIMKSEL     285
H6       EIAARPAVNGQRGRIDYYWSVLKPGETLNVESNGNLIAPWYAYKFVSTN-NKGAVFKSNL     287
H11      EISTRPKVNGQAGRMTFYWTIVKPGEAITFESNGAFLAPRYAFELVSL--GNGKLFRSDL     285
         :*.:*  *.*  .*:   *:::.  : : .:*.* ::**   ::: .  *... :::::
```

FIG 5 Sequence alignment of H2, H3, H5, H6, and H11 consensus sequences. Peptide regions in the region of glycosite "N165" are shown. Compared to H3 consensus, pandemic HA glycosites are 4 residues shifted toward the C terminus. The glycosylation site in question is highlighted in red. Only H3 is out of alignment. The 220 loop is underlined. Arginine (shown as R222 here) replaced tryptophan at position 222 in 2003 and is the predominant residue in the consensus sequence for H3. The tryptophan, which makes contacts with the sugar in H2, is highlighted and shown here as W234. Consensus sequences for each set of HAs in the Influenza Research Database (www.fludb.org) were compared to one another using Clustal Omega (51). Numbering is based upon mature HA sequences.

As stated earlier, the H6 HA consensus sequence reveals the "N165" glycosylation site to also be 4 amino acids closer to the C terminus than that of H3 HA (Fig. 4B). Figure 4C shows the N165 and 220 loop regions of H6 Taiwan 2013 HA (PDB structure 4XKD). In this sequence, W222 has been replaced with an alanine, and there are no contacts between the GlcNAc linked to the glycosylation site and the 220 loop. Rather, the sugar forms up to 6 total close contacts with T171 and T242 on the same subunit. Similar cases are seen in PDB structures 5BNY, 5T08, and 5BRO, as shown in Table 3. The A/mallard/Ohio/249/1998 (H6N1) HA studied here by LC/MS^E is most similar to 5BNY. In all cases, this arrangement projects N165 out and away from the peptide backbone similarly to the orientation seen in the H2 examples. The low number of contacts likely leads to added rotational freedom. As seen in the glycomics results, this site is exclusively occupied by complex glycans. The structural analysis of the available X-ray crystal data supports the notion that the extended orientation, coupled with rotational freedom, leads to increased exposure to ER and Golgi glycan processing machinery compared to that of H3. Like the H2 HA examples, this situation likely leads to exclusively complex glycans at this site in the H6 HAs.

Currently, there is no H11 X-ray crystal structure, However, as seen in Fig. 5, the consensus H11 sequence is similar to those of H2, H5, and H6 in that the N165 glycosylation site is shifted by 4 amino acid residues toward the C terminus. Also, there is no tryptophan or arginine present in the 220 loop like that seen in the H3 HA sequences. Based on the sequence, it can be predicted that "N165" is more accessible to ER and Golgi processing enzymes, leading to a higher probability of complex glycan presence at this site as observed here by mass spectrometry.

The locations of glycosylation sites were mapped to each HA monomer protein backbone as modeled from available X-ray crystal data or, for H11, as modeled by I-TASSER (see Fig. 7). Those where high-mannose glycans were detected are indicated. These figures display the placements of the "N165" glycosite and others in each HA on the globular head. H5 is not shown but is available in Parsons et al. (20).

SP-D binding characteristics of glycans in recombinant LPAIV HAs. To test if placing high-mannose glycan at the "N165" of LPAIV HA would restore SP-D activity, thus supporting the notion that complex glycans at this site prevent this interaction, H6 HA from A/Taiwan/2/13/H6N1 was expressed in HEK293 cells with and without kifunensine (KIF), a potent mannosidase inhibitor that renders all glycans on expressed proteins to have exclusively high-mannose glycans. HA without KIF (−KIF) was treated with *Vibrio cholerae* neuraminidase (VCNA) to remove any sialic acid so that it more closely mimicked native HA, which is devoid of sialic acid due to viral neuraminidase

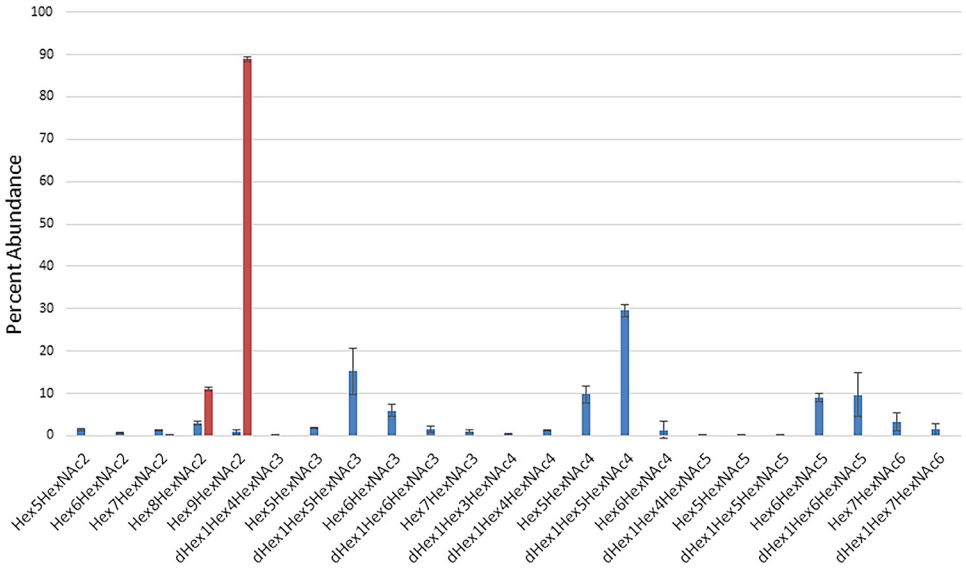

FIG 6 Percent abundances of PNGase F-released, permethylated glycans from VCNA-treated HA (blue) and KIF+ HA (orange). Abundances and standard deviations from positive-mode matrix-assisted laser desorption ionization–time of flight mass spectrometry (MALDI-TOF) spectra collected on three spots, each using 2,5-dihydroxybenzoic acid (DHB) as the matrix. Note that KIF-treated HA contains only high-mannose glycans ($Hex_8HexNAc_2$ and $Hex_9HexNAc_2$ are $Man_8GlcNAc_2$ and $Man_9GlcNAc_2$). VCNA sialidase was used to remove sialic acids, as HA in its natural form is devoid of these due to virus neuraminidase activity.

activity. The *N*-glycans from the resultant HAs were permethylated and examined by matrix-assisted laser desorption ionization–time of flight mass spectrometry (MALDI-TOF) MS analysis. Results are shown in Fig. 6. The HA with KIF (+KIF) contained exclusively high-mannose glycans, while HA −KIF contained primarily complex glycans, as expected. The HA glycopeptides were also analyzed by LC/MS^E (data not shown) revealing that indeed +KIF "N165" was populated with high-mannose glycans and −KIF "N165" had complex glycans.

The +KIF and −KIF H6 HAs were assayed for interaction with rhSP-D using the HAI assay. The −KIF H6 HA protein containing complex glycans at site "N165" was not inhibited with up to 8 μg rhSP-D. However, the H6 HA +KIF containing high-mannose glycans at "N165" was 100% inhibited at 2 μg rhSP-D and partially inhibited at a concentration of 1 μg rhSP-D (Table 4). The assay was performed using human type O red blood cells. A similar result was obtained using turkey red blood cells (data not shown).

DISCUSSION

Both the human and the avian H3 viruses can be expected to contain high-mannose glycans at *N*-glycosylation site N165 (H3 numbering), as previously reported for other H3 viruses, providing a crucial glycan for interaction with SP-D as referenced previously (6, 10). Along with N165, the seasonal H3N2 is predicted to have a number of additional glycosylation sites on the head region of HA. These are located at N63, N123, N126, N133, N144, and N246 (Fig. 7 and Table S6). Additional high-mannose glycans located in the head region have been linked to increased SP-D activity (12, 28). We tested the

TABLE 4 Hemagglutinin inhibition assay of HEK293 cell-derived H6 HA produced with and without KIF using human type O red blood cells[a]

Protein	Glycan type	SP-D concn (μg) by HAI assay (rhSP-D [dodecamer])
H6 HA (VCNA)	Complex	>8
H6 HA (KIF)	High mannose	2 (100% inhibition); 1 (partial inhibition)

[a]Strain used was A/Taiwan/2/13/H6N1, which has 92% identity and identical glycosylation sites to the A/mallard/Ohio/249/1998 H6N1 strain used throughout this study.

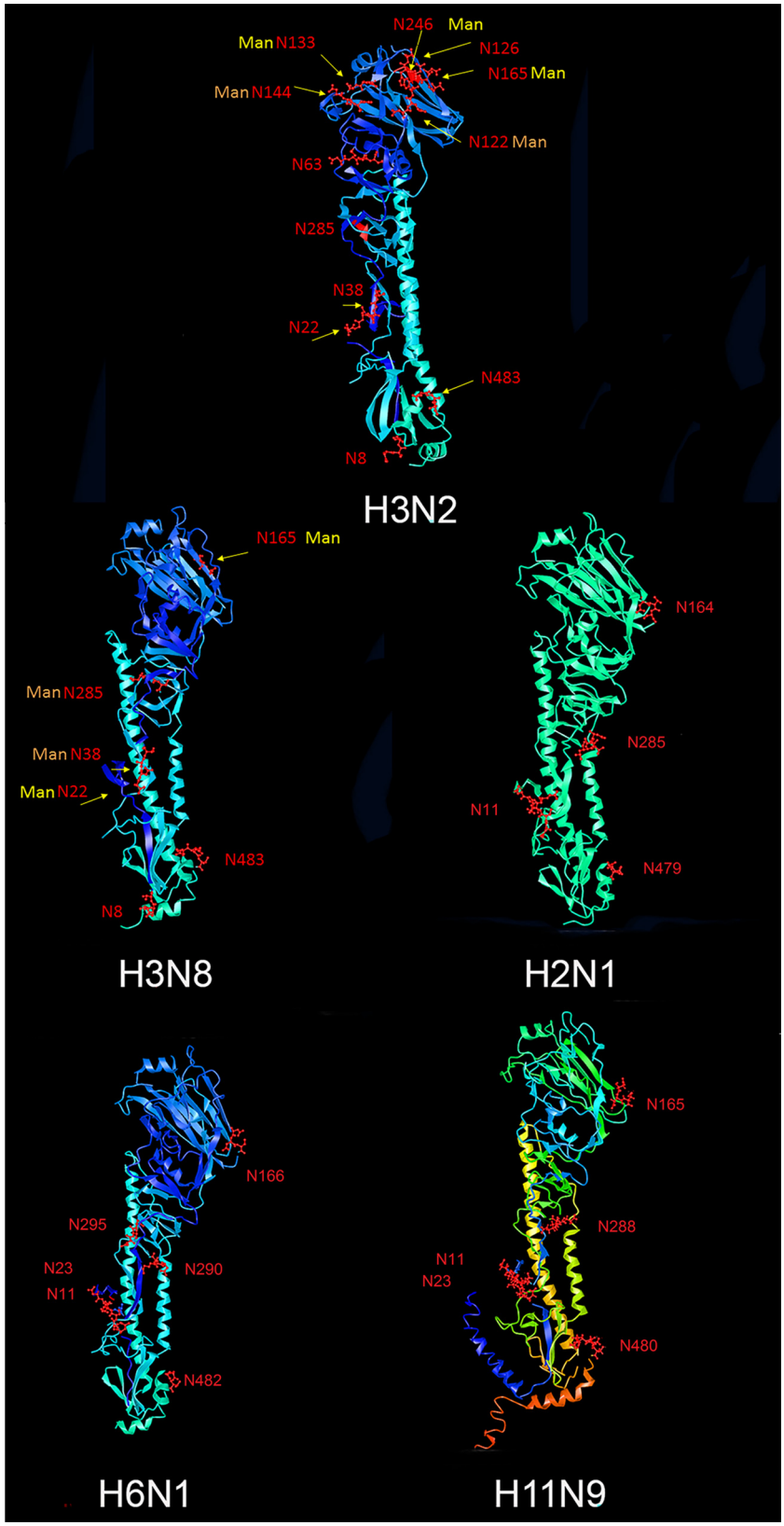

FIG 7 Representative HA monomer images indicating glycosylation site locations. "Man" is indicated in orange if high-mannose glycans were the only subtype found at the site and in yellow if mixed glycans that included high-mannose glycans were found. Glycosites are shown based on the following strains and cited crystal structures: H3N2 A/NY/470/2004 HA glycosylation pattern modeled onto the crystal

(Continued on next page)

relative affinity of LPAIV H2N1, H5N1, H6N1, H11N9, avian strain H3N8, and seasonal strain H3N2 to higher- and lower-affinity recombinant human SP-D strains rhSP-D and hnSP-D. Using a hemagglutinin inhibition assay with either turkey or human red blood cells, we showed that LPAIV viruses were less inhibited by SP-D for RBC hemagglutination than the human seasonal H3N2 or avian H3N8 viruses. H3N8, with only one head glycosite (high mannose at N165), was intermediate in hemagglutination inhibition compared to the H3N2 seasonal strain, supporting the notion that N165 is responsible for HA interaction with SP-D in minimally glycosylated strains. The presence of additional high-mannose sites is consistent with the stronger binding of SP-D to H3N2 observed here. Although H2N1, H5N1, H6N1, and H11N9 LPAIV HAs all possess a glycosylation site at "N165," poor SP-D binding characteristics were observed for these avian strains, suggesting that some other feature was necessary for SP-D binding.

Mass spectrometry analysis confirmed that HAs from the LPAIV strains were all glycosylated but exhibited complex glycans at glycosylation site "N165" rather than the high-mannose glycans seen in avian H3N8 and seasonal H3N2 viruses. The seasonal strain also had exclusively high-mannose glycans at head sites N133 and N246 and some high-mannose glycans at N122 and N144. Moreover, we demonstrated that replacement of complex glycans with high-mannose glycans in H6 LPAIV HA restored interaction with SP-D, indicating that it is the complex glycan at "N165" that is responsible for the poor interaction of SP-D with HAs of these LPAIV strains. These results suggest that the H2N1, H5N1, H6N1, and H11N9 avian strains would likely be SP-D resistant in the host, which could contribute to disease severity if, indeed, human-adapted forms or reassortant(s) emerged in the human population.

Previously, we examined the glycosylation patterns of other avian H5N1 and H5N7 strains produced in different cell lines, including egg and HEK293. In all cases, the H5 N165 site was exclusively or almost exclusively occupied by complex glycans, supporting the notion that the protein conformation, and not the cell substrate, is the primary determinant of glycosylation subtype at this site. The present study supports this notion since all strains were produced in MDCK cells and had the expected glycan subtype at "N165," thus controlling for cell substrate glycosylation effects. This strongly suggests that there are structural elements that can predict glycan subtype in different IAV HAs.

All LPAIV HAs sequence alignments had the "N165" glycosite forward in the β-sheet by 4 amino acids compared to the HA of H3 viruses, placing attached glycans away from the intersubunit 220 loop, as previously reported for H5 HAs (20). Although the H2 HA glycosite forms 10 intramolecular contacts with an intrasubunit W239 in the X-ray structure, the contacts are not enough to prevent it from further processing by the Golgi enzymes. H6 HA is similarly arranged, but lacks W239, although the inner core of the glycan forms contacts with T171 and T242. Again, this may make it more available for processing by ER and Golgi glycan-modifying enzymes, leading to complex compositions. The H11 HA was similarly arranged based on sequence alignment, but unfortunately no crystal or nuclear magnetic resonance (NMR) structural data were available for interrogation. Nonetheless, the detected compositions at the site and coalignment with other avian strains strongly suggest that the positional data approximated by sequence alignment is predictive of glycan subtype at the N165-analogous glycosite. As only complex glycans were detected at "N165," this prediction is borne out. Overall, the alignment of all LPAIV strains studied predicted glycan subtype. Even in the presence of a high number of contacts, the position of the C-terminal shifted

FIG 7 Legend (Continued)
structure of 2005 human H3N2 HA (PBD code 2YP7), H3N8 avian A/pintail/339/1987 HA glycosylation pattern modeled onto the crystal structure of 2005 human H3N2 HA (PBD code 2YP7), H2N1 A/green-winged teal/Ohio/175/1986 HA glycosylation pattern modeled onto the crystal structure of 1957 avian Jena H2 HA (PBD code 2WRF), A/mallard/Ohio/249/1998 H6N1 HA glycosylation pattern modeled onto the crystal structure of H6N1 A/Taiwan/2/2013 HA (PBD code 5BRO), and A/green-winged teal/Ohio/340/1987 H11 HA glycosylation pattern modeled onto an I-TASSER folded model.

"N165" site on the β-sheet and away from the 220 loop predicts increased likelihood of complex glycan presence at this glycosite. Therefore, pandemic strains can be predicted to either contain or lack the crucial high-mannose glycans at this key site based on their similarity to H3 or H5, and thus their susceptibility to SP-D can be proposed. As discussed in the introduction, there is a preponderance of evidence in the literature that supports the notion that high-mannose glycosylation at the head of HAs is required for SP-D interaction and removal of influenza virus from the lung (7, 8, 14, 29, 30), which, in turn, modulates virulence (17, 31), demonstrating that the surfactant is crucial for effective immune response.

How dangerous are the LPAIV strains lacking high-mannose glycans on the head, which are thus not susceptible to SP-D interaction? H5N1 avian viruses, which lack head high-mannose glycans (19, 20) have been associated with severe lung immunopathology (32). Highly pathogenic avian influenza (HPAI) H5Nx viruses have been reported in U.S. domestic poultry (http://www.cdc.gov/flu/avianflu/h5). Similar viruses have infected humans in other countries and caused serious illness and death. H5N1 has a reported human case fatality rate of approximately 60% and has resulted in numerous hospitalizations, deaths, and significant economic loss (33). The H5N1 viruses in central China are undergoing antigenic drift that may be associated with human immune pressure adaptation (34), possibly leading to receptor changes toward the human sialylα2,6 ligands and further suggesting that H5 virus may leap to humans.

The 1957 A/H2N2 pandemic IAV caused an estimated 2 million deaths globally. The H2 subtype continues to infect avian species and swine and, therefore, remains a pandemic threat to humans (35). The pandemic was caused by an H2N2 influenza A virus generated by reassortment of the previously circulating human H1N1 virus and an avian H2N2 virus that contributed the polymerase basic 1 (PB1), hemagglutinin (HA), and neuraminidase (NA) genes to the pandemic strain (36). Jones et al. (37) provided a risk assessment for reintroduction of H2N2 viruses in humans based on the currently circulating strains in wild birds. While the assessment concluded that the current risk is low, the lack of sustained global surveillance in birds and other zoonotic hosts limits the power of risk assessment. Clearly, the H2N2 viruses have made the leap in the past and, if this were to happen again, the new strains would be introduced into a naive population regarding those under 60 years of age.

Wang et al. tested 287 strains of H6 subtype viruses isolated in southern China between 2008 and 2011 and found that 87 recognized human glycan receptors. Some of these transmitted efficiently to mammalian hosts, including mice and guinea pigs (38, 39). LPAIV H6N1 has transmitted to humans (40). Sequence identity analysis of the Taiwan strain indicated that it reassorted from A/chicken/Taiwan/0101/2012(H5N2) (PB2 and PA genes), A/chicken/Taiwan/A1997/2012(H5N2) (M gene), A/chicken/Taiwan/ch1006/04(H6N1) (HA gene), and A/chicken/Taiwan/TC135/2010(H6N1) (PB1, nucleoprotein [NP], NA, and nonstructural protein [NS] genes). Currently, the H6N1 viruses are considered low-pathogenicity strains. However, given that at least one confirmed human infection has occurred, these strains should be monitored.

H11 subtype viruses have been reported in wild waterfowl in regions of South America, Kazakhstan, China, Hong Kong, and Japan, and these have included HA, NA, M, NS, and polymerase acidic protein (PA) genes (41). Reassortants between wild birds, water fowl, and domestic poultry have been detected (42, 43). Recently, during the surveillance of avian influenza viruses in the Dongxi Lake wetland of Hubei, China, an H11N9 virus was isolated from the bean goose (41). The HA belonged to a North American lineage. The virus was able to replicate in mouse lung without prior adaptation. While H11 viruses have not been detected in human or swine to date to our knowledge, they do contribute to reassortment along the East Asian-Australian flyway. Surveillance of duck hunters has revealed serologic evidence of H11 exposure (44). Further surveillance of these strains is warranted.

Here, we have described the glycosylation patterns of H2, H5, H6, and H11 HAs and compared them to low- and high-glycosylation H3 virus strains. Glycan subtype at key glycosylation sites dictated lung surfactant SP-D activity. It was revealed that these

potential pandemic avian IAV strains are not susceptible to lung SP-D activity due to predominantly complex glycans present at key sites that interact with this collectin, a primary first-line defense factor located in human lung tissue. Sequence alignment and three-dimensional structural analysis allows prediction of glycan subtypes at these key regions. This work demonstrates that sequence analysis, with and without three-dimensional structure analysis, can be used to predict glycan subtype and therefore innate immune defense activity by key collectins and lectins of the innate immune system. These assessments may be useful predictors of pandemic virulence.

MATERIALS AND METHODS

Viral strains. The low-pathogenicity avian influenza virus (LPAIV) strains A/green-winged teal/Ohio/175/1986 (H2N1), A/pintail/339/1987(H3N8), A/mallard/Maryland/802/2007(H5N1), A/mallard/Ohio/249/1998(H6N1), and A/green-winged teal/Ohio/340/1987(H11N9) were used (25). All viruses were grown for one passage in MDCK cells. A human H3N2 virus, A/NY/470/2004, was used as a positive control.

Chemicals and reagents. HyperSep C_{18} and porous graphitic carbon (PGC) cartridges were purchased from Thermo Fisher Scientific, Inc. (Waltham, MA). TSKgel Amide-80 particles were purchased from Tosoh Bioscience LLC (Montgomeryville, PA). Sequencing-grade modified trypsin, immobilized trypsin, and Glu-C were purchased from Promega Corp. (Madison, WI). *N*-glycosidase A was purchased from Roche Diagnostics Corporation (Indianapolis, IN). PNGase F was purchased from New England BioLabs, Inc. (Ipswich, MA). Water (^{18}O, 97%) was purchased from Cambridge Isotope Laboratories, Inc. (Tewksbury, MA). Iodomethane, dimethyl sulfoxide (DMSO), sodium hydroxide beads, and other chemicals were purchased from Sigma-Aldrich (St. Louis, MO, USA); solvents were of high-performance liquid chromatography (HPLC) grade or higher. All other reagents were American Chemical Society (ACS) grade or higher.

Surfactant protein D hemagglutinin inhibition (HAI) assay. Recombinant human rhSP-D protein and hnSP-D (D325A and R343V) was prepared as previously described (8, 30). SP-D protein (at 8 μg/ml) was serially diluted in D-phosphate-buffered saline (D-PBS) (containing Ca^{2+} and Mg^{2+} [PBS++]; Quality Biological, Inc., Gaithersburg, MD) in 96-well plates with V-shaped bottoms (Costar). After adding an equal volume (25 μl) of virus at a final concentration of 4 HA units/well, the virus and SP-D mixture was incubated for 15 min at room temperature, followed by addition of 50 μl human type O RBCs (lot no. 23364; Innovative Research) at the ratio of 100 μl RBC cell pellet (1,500 rpm, 5 min, no brake) per 15 ml PBS++, or 0.5% turkey blood cells (lot no. 17F28804; Lampire Biological Laboratories). The plate was incubated at room temperature for 45 min, and hemagglutination inhibition (HAI) was observed. Inhibition was recorded if the sample appeared as a red dot sunken to the bottom of the plate. If no dot appeared at the highest protein concentration used, then the value was expressed as greater than the maximal protein concentration tested.

The recombinant H6 HAs used in HAI assay were prepared as follows. The HA proteins were expressed in HEK293 cells, and VCNA and KIF digests were performed as described previously (45). Before performing the HA assay, the proteins were treated with anti-strep-tag mouse antibody and goat anti-mouse antibody for 30 min in a 4:2:1 ratio.

Plaque reduction neutralization assay. MDCK cells (MDCK.2 [ATCC CRL-2936]) were grown to confluence in 24-well plates. SP-D protein (4 μg/ml) was sequentially diluted 2-fold in D-PBS containing Ca^{2+} and Mg^{2+} (up to 0.25 μg/ml), and each dilution was mixed with an equal volume of virus corresponding to 20 to 30 plaques per well. Final SP-D concentrations ranged from 2 μg/ml to 0.125 μg/ml. Following preincubation at room temperature for 45 to 60 min, the mixed samples (100 μl total) were layered onto the MDCK cells. Each sample was tested in triplicate wells. A virus control without SP-D protein was included. After incubation at 37°C for 1 h, a plaque assay with staining after 48 h was performed.

Microwave-assisted protein trypsin digestion. The viruses and proteins were dialyzed against PBS (pH 7.0), concentrated to ~1μg/μl and denatured at 100°C for 10 min. Dithiothreitol (DTT) was added to reduce the disulfide bonds at a final concentration of 5 mM for 45 min at 60°C. The samples were chilled to room temperature, and iodoacetamide (IAA) was at a concentration of 15 mM added to alkylate the reduced cysteine residues for 45 min in the dark at room temperature. Trypsin was added (enzyme: protein, 1:50 [wt/wt]) and the samples were put into a Discover microwave apparatus (CEM Corp., Matthews, NC) for digestion. The incubation time was 15 min with a fixed microwave irradiation power at 50 W and a temperature of 45°C. After digestion, the samples were boiled for 10 min to inactivate trypsin, and the solutions were vacuum dried for downstream analysis.

Chymotrypsin digestion. The H11N9 viruses were reduced and alkylated as described previously and digested with chymotrypsin. Chymotrypsin was added to the virus solution to a final protease: protein ratio of 1:50 and incubated at 25°C overnight. After digestion, the samples were boiled for 10 min and vacuum dried for downstream analysis.

Glycopeptide enrichment by hydrophilic interaction chromatography. Intact glycopeptides were enriched via solid-phase extraction (SPE) with TSKgel Amide-80 hydrophilic interaction chromatography (HILIC) resin according to the procedure described in our previous report (9). About 200 mg (400 μl of wet resin) of Amide-80 resin was placed into a fritted 1 ml column (Supelco) and washed with 1 ml of 0.1% trifluoroacetic acid (TFA)/water. The column was conditioned with 1 ml of 0.1% TFA/80% acetonitrile (ACN). The peptides were suspended in 0.1% TFA/80% ACN and applied to the column. The hydrophobic species were washed away with 3 ml of 0.1% TFA/80% ACN. The glycopeptides were eluted

with 1 ml of 0.1%TFA/60% ACN and 1 ml of 0.1% TFA/40% ACN. The eluents were mixed and vacuum dried prior to reverse phase LC/MS analysis.

Reverse phase nanoLC/MS^E analysis of glycopeptides. The glycopeptides were reconstituted in 0.1% formic acid in water and injected onto a C_{18} column (HSS T3 nanoAcquity column, 75-μm inside diameter [i.d.] $\times$ 100 mm; 1.8-μm particle; Waters Corporation) for nanoLC/MS^E analysis (11, 12). The solvent system consisted of solvent A (100% water/0.1% formic acid [FA]) and solvent B (100% acetonitrile/0.1% FA). A Waters nanoAcquity ultraperformance liquid chromatography (UPLC) system was used for automatic sample loading and flow control. The gradient was as follows: 3 to 50% of solvent B for 30 min, 85% of solvent B for 5 min, and 3% of solvent B for 25 min. The effluent was introduced into a Waters Synapt G2 high-definition mass spectrometry (HDMS) system (Waters Corp. Milford, MA) via an uncoated 15-μm-i.d. PicoTip Emitter (New Objective Inc., Woburn, MA). The spray voltage was 3,000 V. The mass spectrometer was operated in the positive polarity mode. The instrument was set to perform MS^E experiments. The low-energy precursor ion scan was performed at 4 eV. The elevated collision energy fragmentation scan was performed in a linear ramp from 15 to 45 eV. The scan time was 0.9 s. An auxiliary pump was used to spray a solution of 200 fmol/min Glu-fibrinopeptide B in 50/50 methanol/water with 1% acetic acid for mass calibration (lock mass channel) at a flow rate of 500 nl/min with sampling every 30 s. The system was tuned for a minimum resolution of 20,000 full width at half maximum (FWHM) and calibrated using Glu-fibrinopeptide fragments.

Data analysis for peptide glycosylation identification. The nanoLC/MS^E data were processed by BiopharmaLynx 1.3 (Waters) and in-house-developed GLYMPS software to identify the site-specific glycosylation of HAs. A glycan database was created in BiopharmaLynx using glycans present in the curated glycan database GlycosuiteDB. The search used trypsin and Glu-C digestion with one missed cleavage on the known sequences of HAs. Cysteine carbamidomethylation was set as a fixed modification. All potential glycan compositions expressed in the egg, MDCK, or insect sf9 cells were documented in the glycan database and added as variable modifications in the search parameters. Identified glycopeptides were confirmed manually by MS^E spectra. The oxonium ions (such as *m/z* 204.1, 366.1, and 528.2) present in MS^E spectra were used to help locate and determine the presence of glycopeptides. Assignment criteria included (i) manual observation of oxonium ions, peptide with neutral loss of glycan fragment, and GlcNAc plus peptide fragment, (ii) BiopharmaLynx identification, (iii) identification of peptide fragments, and (iv) mass accuracy of less than 20 ppm and 30 ppm in parent and collision ion scans, respectively.

The publicly accessible program NetNGlyc 1.0 (https://services.healthtech.dtu.dk/service.php?NetNGlyc-1.0), was used as a first approximation of *N*-glycosylation site occupancy. The program uses nine neural net programs in its prediction algorithm. Each neural net is assigned as a juror, the predictions of which are tallied to produce a qualitative likelihood of site occupancy score, as well as a glycosylation probability score (46). A NetNGlyc probability score of 0.5 or higher indicates a high probability of occupancy.

Structural comparison of N165 glycosylation sites. The three-dimensional structures of HAs with PDB codes 2WR0 (H2), 4XKD (H6), and 4FNK (H3) (47–49) were aligned using PyMOL (www.pymol.org). Hydrogen bonds and van der Waals distances between the proteins and the *N*-acetyl glucosamines in the crystal structures were identified with LigPLot (50) and an in-house Python script, respectively.

Alignment of HA subtype protein sequences. All HAs for each subtype were downloaded from www.fludb.org and aligned using Clustal Omega (51). The consensus sequence for each subtype was then aligned.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, PDF file, 0.2 MB.

ACKNOWLEDGMENTS

J.F.C. is funded through U.S. Food and Drug Administration grant Z01 BJ 02044-12 LBP. K.L.H. is funded through NHLBI grant RO1 HL069031. J.K.T. is funded through NIH 1 ZIA AI000986-12. R.P.D.V. is a recipient of an ERC Starting Grant from the European Commission (grant 802780) and a Beijerinck Premium prize from the Royal Dutch Academy of Sciences.

REFERENCES

1. Ogasawara Y, Voelker DR. 1995. Altered carbohydrate recognition specificity engineered into surfactant protein D reveals different binding mechanisms for phosphatidylinositol and glucosylceramide. J Biol Chem 270:14725–14732. https://doi.org/10.1074/jbc.270.24.14725.
2. Kuroki Y, Gasa S, Ogasawara Y, Shiratori M, Makita A, Akino T. 1992. Binding specificity of lung surfactant protein SP-D for glucosylceramide. Biochem Biophys Res Commun 187:963–969. https://doi.org/10.1016/0006-291x(92)91291-w.
3. Hsieh IN, De Luna X, White MR, Hartshorn KL. 2018. The role and molecular mechanism of action of surfactant protein D in innate host defense against influenza A virus. Front Immunol 9:1368. https://doi.org/10.3389/fimmu.2018.01368.
4. Head JF, Mealy TR, McCormack FX, Seaton BA. 2003. Crystal structure of trimeric carbohydrate recognition and neck domains of surfactant protein A. J Biol Chem 278:43254–43260. https://doi.org/10.1074/jbc.M305628200.

5. Crouch E, Hartshorn K, Horlacher T, McDonald B, Smith K, Cafarella T, Seaton B, Seeberger PH, Head J. 2009. Recognition of mannosylated ligands and influenza A virus by human surfactant protein D: contributions of an extended site and residue 343. Biochemistry 48:3335–3345. https://doi.org/10.1021/bi8022703.
6. Crouch E, Nikolaidis N, McCormack FX, McDonald B, Allen K, Rynkiewicz MJ, Cafarella TM, White M, Lewnard K, Leymarie N, Zaia J, Seaton BA, Hartshorn KL. 2011. Mutagenesis of surfactant protein D informed by evolution and X-ray crystallography enhances defenses against influenza A virus *in vivo*. J Biol Chem 286:40681–40692. https://doi.org/10.1074/jbc.M111.300673.
7. Hartshorn KL, Webby R, White MR, Tecle T, Pan C, Boucher S, Moreland RJ, Crouch EC, Scheule RK. 2008. Role of viral hemagglutinin glycosylation in anti-influenza activities of recombinant surfactant protein D. Respir Res 9:65. https://doi.org/10.1186/1465-9921-9-65.
8. Hartshorn KL, White MR, Voelker DR, Coburn J, Zaner K, Crouch EC. 2000. Mechanism of binding of surfactant protein D to influenza A viruses: importance of binding to haemagglutinin to antiviral activity. Biochem J 351:449–458. https://doi.org/10.1042/0264-6021:3510449.
9. Zhang L, Hartshorn KL, Crouch EC, Ikegami M, Whitsett JA. 2002. Complementation of pulmonary abnormalities in SP-D(-/-) mice with an SP-D/conglutinin fusion protein. J Biol Chem 277:22453–22459. https://doi.org/10.1074/jbc.M201632200.
10. Anders EM, Hartley CA, Jackson DC. 1990. Bovine and mouse serum beta inhibitors of influenza A viruses are mannose-binding lectins. Proc Natl Acad Sci U S A 87:4485–4489. https://doi.org/10.1073/pnas.87.12.4485.
11. 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:e01693-18. https://doi.org/10.1128/JVI.01693-18.
12. An Y, McCullers JA, Alymova I, Parsons LM, Cipollo JF. 2015. Glycosylation analysis of engineered H3N2 influenza A virus hemagglutinins with sequentially added historically relevant glycosylation sites. J Proteome Res 14:3957–3969. https://doi.org/10.1021/acs.jproteome.5b00416.
13. 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 Proteomics 15:1895–1912. https://doi.org/10.1074/mcp.M116.058016.
14. 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. https://doi.org/10.1128/JVI.00769-07.
15. She YM, 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. https://doi.org/10.1038/s41598-017-10714-2.
16. Datema R, Schwarz RT. 1981. Effect of energy depletion on the glycosylation of a viral glycoprotein. J Biol Chem 256:11191–11198.
17. 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. https://doi.org/10.4049/jimmunol.1100295.
18. Cruz E, Cain J, Crossett B, Kayser V. 2018. Site-specific glycosylation profile of influenza A (H1N1) hemagglutinin through tandem mass spectrometry. Hum Vaccin Immunother 14:508–517. https://doi.org/10.1080/21645515.2017.1377871.
19. An Y, Rininger JA, Jarvis DL, Jing X, Ye Z, Aumiller JJ, Eichelberger M, Cipollo JF. 2013. Comparative glycomics analysis of influenza hemagglutinin (H5N1) produced in vaccine relevant cell platforms. J Proteome Res 12:3707–3720. https://doi.org/10.1021/pr400329k.
20. 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. https://doi.org/10.1021/acs.jproteome.6b00175.
21. Lee JH, Pascua PN, Song MS, Baek YH, Kim CJ, Choi HW, Sung MH, Webby RJ, Webster RG, Poo H, Choi YK. 2009. Isolation and genetic characterization of H5N2 influenza viruses from pigs in Korea. J Virol 83:4205–4215. https://doi.org/10.1128/JVI.02403-08.
22. Lee CW, Swayne DE, Linares JA, Senne DA, Suarez DL. 2005. H5N2 avian influenza outbreak in Texas in 2004: the first highly pathogenic strain in the United States in 20 years? J Virol 79:11412–11421. https://doi.org/10.1128/JVI.79.17.11412-11421.2005.
23. Donatelli I, Campitelli L, Di Trani L, Puzelli S, Selli L, Fioretti A, Alexander DJ, Tollis M, Krauss S, Webster RG. 2001. Characterization of H5N2 influenza viruses from Italian poultry. J Gen Virol 82:623–630. https://doi.org/10.1099/0022-1317-82-3-623.
24. Hillaire ML, van Eijk M, van Trierum SE, van Riel D, Saelens X, Romijn RA, Hemrika W, Fouchier RA, Kuiken T, Osterhaus AD, Haagsman HP, Rimmelzwaan GF. 2011. Assessment of the antiviral properties of recombinant porcine SP-D against various influenza A viruses *in vitro*. PLoS One 6:e25005. https://doi.org/10.1371/journal.pone.0025005.
25. Dugan VG, Chen R, Spiro DJ, Sengamalay N, Zaborsky J, Ghedin E, Nolting J, Swayne DE, Runstadler JA, Happ GM, Senne DA, Wang R, Slemons RD, Holmes EC, Taubenberger JK. 2008. The evolutionary genetics and emergence of avian influenza viruses in wild birds. PLoS Pathog 4:e1000076. https://doi.org/10.1371/journal.ppat.1000076.
26. Garten RJ, Davis CT, Russell CA, Shu B, Lindstrom S, Balish A, Sessions WM, Xu X, Skepner E, Deyde V, Okomo-Adhiambo M, Gubareva L, Barnes J, Smith CB, Emery SL, Hillman MJ, Rivailler P, Smagala J, de Graaf M, Burke DF, Fouchier RAM, Pappas C, Alpuche-Aranda CM, López-Gatell H, Olivera H, López I, Myers CA, Faix D, Blair PJ, Yu C, Keene KM, Dotson PD, Boxrud D, Sambol AR, Abid SH, St George K, Bannerman T, Moore AL, Stringer DJ, Blevins P, Demmler-Harrison GJ, Ginsberg M, Kriner P, Waterman S, Smole S, Guevara HF, Belongia EA, Clark PA, Beatrice ST, Donis R, Katz J, Finelli L, Bridges CB, Shaw M, Jernigan DB, Uyeki TM, Smith DJ, Klimov AI, Cox NJ. 2009. Antigenic and genetic characteristics of swine-origin 2009 A(H1N1) influenza viruses circulating in humans. Science 325:197–201. https://doi.org/10.1126/science.1176225.
27. Nikolaidis NM, White MR, Allen K, Tripathi S, Qi L, McDonald B, Taubenberger J, Seaton BA, McCormack FX, Crouch EC, Hartshorn KL. 2014. Mutations flanking the carbohydrate binding site of surfactant protein D confer antiviral activity for pandemic influenza A viruses. Am J Physiol Lung Cell Mol Physiol 306:L1036–44. https://doi.org/10.1152/ajplung.00035.2014.
28. Wanzeck K, Boyd KL, McCullers JA. 2011. Glycan shielding of the influenza virus hemagglutinin contributes to immunopathology in mice. Am J Respir Crit Care Med 183:767–773. https://doi.org/10.1164/rccm.201007-1184OC.
29. Hartshorn KL, Crouch EC, White MR, Eggleton P, Tauber AI, Chang D, Sastry K. 1994. Evidence for a protective role of pulmonary surfactant protein D (SP-D) against influenza A viruses. J Clin Invest 94:311–319. https://doi.org/10.1172/JCI117323.
30. Reading PC, Morey LS, Crouch EC, Anders EM. 1997. Collectin-mediated antiviral host defense of the lung: evidence from influenza virus infection of mice. J Virol 71:8204–8212.
31. Qi L, Kash JC, Dugan VG, Jagger BW, Lau YF, Sheng ZM, Crouch EC, Hartshorn KL, Taubenberger JK. 2011. The ability of pandemic influenza virus hemagglutinins to induce lower respiratory pathology is associated with decreased surfactant protein D binding. Virology 412:426–434. https://doi.org/10.1016/j.virol.2011.01.029.
32. Kash JC, Tumpey TM, Proll SC, Carter V, Perwitasari O, Thomas MJ, Basler CF, Palese P, Taubenberger JK, Garcia-Sastre A, Swayne DE, Katze MG. 2006. Genomic analysis of increased host immune and cell death responses induced by 1918 influenza virus. Nature 443:578–581. https://doi.org/10.1038/nature05181.
33. Poovorawan Y, Pyungporn S, Prachayangprecha S, Makkoch J. 2013. Global alert to avian influenza virus infection: from H5N1 to H7N9. Pathog Glob Health 107:217–223. https://doi.org/10.1179/2047773213Y.0000000103.
34. Zou W, Ke J, Zhu J, Zhou H, Jin M. 2012. The antigenic property of the H5N1 avian influenza viruses isolated in central China. Virol J 9:148. https://doi.org/10.1186/1743-422X-9-148.
35. Joseph U, Linster M, Suzuki Y, Krauss S, Halpin RA, Vijaykrishna D, Fabrizio TP, Bestebroer TM, Maurer-Stroh S, Webby RJ, Wentworth DE, Fouchier RA, Bahl J, Smith GJ, CEIRS H2N2 Working Group. 2015. Adaptation of pandemic H2N2 influenza A viruses in humans. J Virol 89:2442–2447. https://doi.org/10.1128/JVI.02590-14.
36. Schafer JR, Kawaoka Y, Bean WJ, Suss J, Senne D, Webster RG. 1993. Origin of the pandemic 1957 H2 influenza A virus and the persistence of its possible progenitors in the avian reservoir. Virology 194:781–788. https://doi.org/10.1006/viro.1993.1319.
37. Jones JC, Baranovich T, Marathe BM, Danner AF, Seiler JP, Franks J, Govorkova EA, Krauss S, Webster RG. 2014. Risk assessment of H2N2 influenza viruses from the avian reservoir. J Virol 88:1175–1188. https://doi.org/10.1128/JVI.02526-13.
38. Wang G, Deng G, Shi J, Luo W, Zhang G, Zhang Q, Liu L, Jiang Y, Li C,

Sriwilaijaroen N, Hiramatsu H, Suzuki Y, Kawaoka Y, Chen H. 2014. H6 influenza viruses pose a potential threat to human health. J Virol 88: 3953–3964. https://doi.org/10.1128/JVI.03292-13.
39. Gillim-Ross L, Santos C, Chen Z, Aspelund A, Yang CF, Ye D, Jin H, Kemble G, Subbarao K. 2008. Avian influenza H6 viruses productively infect and cause illness in mice and ferrets. J Virol 82:10854–10863. https://doi.org/10.1128/JVI.01206-08.
40. Yuan J, Zhang L, Kan X, Jiang L, Yang J, Guo Z, Ren Q. 2013. Origin and molecular characteristics of a novel 2013 avian influenza A(H6N1) virus causing human infection in Taiwan. Clin Infect Dis 57:1367–1368. https://doi.org/10.1093/cid/cit479.
41. Zhang Y, Zou SM, Li XD, Dong LB, Bo H, Gao RB, Wang DY, Shu YL. 2016. Detection of reassortant avian influenza A (H11N9) virus in environmental samples from live poultry markets in China. Infect Dis Poverty 5:59. https://doi.org/10.1186/s40249-016-0149-2.
42. Alexander DJ. 2000. A review of avian influenza in different bird species. Vet Microbiol 74:3–13. https://doi.org/10.1016/s0378-1135(00)00160-7.
43. Alexander DJ. 2007. An overview of the epidemiology of avian influenza. Vaccine 25:5637–5644. https://doi.org/10.1016/j.vaccine.2006.10.051.
44. Gill JS, Webby R, Gilchrist MJ, Gray GC. 2006. Avian influenza among waterfowl hunters and wildlife professionals. Emerg Infect Dis 12: 1284–1286. https://doi.org/10.3201/eid1708.060492.
45. de Vries RP, Smit CH, de Bruin E, Rigter A, de Vries E, Cornelissen LA, Eggink D, Chung NP, Moore JP, Sanders RW, Hokke CH, Koopmans M, Rottier PJ, de Haan CA. 2012. Glycan-dependent immunogenicity of recombinant soluble trimeric hemagglutinin. J Virol 86:11735–11744. https://doi.org/10.1128/JVI.01084-12.
46. Gupta R, Jung E, Brunak K. 2004. Prediction of N-glycsoylation sites in human proteins.
47. Liu J, Stevens DJ, Haire LF, Walker PA, Coombs PJ, Russell RJ, Gamblin SJ, Skehel JJ. 2009. Structures of receptor complexes formed by hemagglutinins from the Asian Influenza pandemic of 1957. Proc Natl Acad Sci U S A 106:17175–17180. https://doi.org/10.1073/pnas.0906849106.
48. Tzarum N, de Vries RP, Zhu X, Yu W, McBride R, Paulson JC, Wilson IA. 2015. Structure and receptor binding of the hemagglutinin from a human H6N1 influenza virus. Cell Host Microbe 17:369–376. https://doi.org/10.1016/j.chom.2015.02.005.
49. Ekiert DC, Kashyap AK, Steel J, Rubrum A, Bhabha G, Khayat R, Lee JH, Dillon MA, O'Neil RE, Faynboym AM, Horowitz M, Horowitz L, Ward AB, Palese P, Webby R, Lerner RA, Bhatt RR, Wilson IA. 2012. Cross-neutralization of influenza A viruses mediated by a single antibody loop. Nature 489:526–532. https://doi.org/10.1038/nature11414.
50. Wallace AC, Laskowski RA, Thornton JM. 1995. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. Protein Eng 8:127–134. https://doi.org/10.1093/protein/8.2.127.
51. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J, Thompson JD, Higgins DG. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 7:539. https://doi.org/10.1038/msb.2011.75.