

Probing altered receptor specificities of antigenically drifting human H3N2 viruses by chemoenzymatic synthesis, NMR, and modeling

Received: 8 June 2023

Accepted: 25 March 2024

Published online: 06 April 2024

Check for updates

Luca Unione ^{1,2,3} , Augustinus N. A. Ammerlaan¹, Gerlof P. Bosman¹, Elif Uslu ¹, Ruonan Liang¹, Frederik Broszeit¹, Roosmarijn van der Woude¹, Yanyan Liu¹, Shengzhou Ma⁴, Lin Liu⁴, Marcos Gómez-Redondo², Iris A. Bermejo², Pablo Valverde², Tammo Diercks², Ana Ardá ^{2,3}, Robert P. de Vries ¹ & Geert-Jan Boons ^{1,4,5,6}

Prototypic receptors for human influenza viruses are *N*-glycans carrying α 2,6-linked sialosides. Due to immune pressure, A/H3N2 influenza viruses have emerged with altered receptor specificities that bind α 2,6-linked sialosides presented on extended *N*-acetyl-lactosamine (LacNAc) chains. Here, binding modes of such drifted hemagglutinin's (HAs) are examined by chemoenzymatic synthesis of *N*-glycans having ¹³C-labeled monosaccharides at strategic positions. The labeled glycans are employed in 2D STD-¹H by ¹³C-HSQC NMR experiments to pinpoint which monosaccharides of the extended LacNAc chain engage with evolutionarily distinct HAs. The NMR data in combination with computation and mutagenesis demonstrate that mutations distal to the receptor binding domain of recent HAs create an extended binding site that accommodates with the extended LacNAc chain. A fluorine containing sialoside is used as NMR probe to derive relative binding affinities and confirms the contribution of the extended LacNAc chain for binding.

Spill-over of influenza A virus (IAV) from an animal reservoir into the human population has caused four pandemics in the past century¹. These viruses became seasonal strains that evade neutralization induced by prior infections or immunization. Seasonal IAV causes significant disease burden, infecting 9–35 million individuals annually and causing 12–56 thousand deaths².

IAV expresses hemagglutinin (HA) and neuraminidase (NA) as two major envelope glycoproteins. HA binds to sialic acid (Neu5Ac) of glycoconjugates on host cells to initiate infection, whereas NA

facilitates the release of progeny viruses from infected cells by cleaving sialosides³. Human IAVs recognize sialosides α 2,6-linked to galactoside (Gal), which are structures that are abundantly found in the upper respiratory tract of humans. On the other hand, HAs of ancestral avian viruses exhibit a preference for α 2,3-linked sialosides that are found in the duck enteric and chicken upper respiratory tract^{4–8}.

Human IAVs have a remarkable ability to evolve and evade neutralization by antibodies elicited by prior infections or vaccination. This antigenic drift is mainly caused by mutations in the receptor

¹Department of Chemical Biology & Drug Discovery, Utrecht Institute for Pharmaceutical Sciences, Utrecht University, 3584 CG Utrecht, The Netherlands.

²CICbioGUNE, Basque Research & Technology Alliance (BRTA), Bizkaia Technology Park, Building 800, 48160 Derio, Bizkaia, Spain. ³Ikerbasque, Basque Foundation for Science, Euskadi Plaza 5, 48009 Bilbao, Bizkaia, Spain. ⁴Complex Carbohydrate Research Center, University of Georgia, 315 Riverbend Rd, Athens, GA 30602, USA. ⁵Bijvoet Center for Biomolecular Research, Utrecht University, Utrecht, The Netherlands. ⁶Department of Chemistry, University of Georgia, Athens, GA 30602, USA. e-mail: lunione@cicbiogune.es; R.Vries@uu.nl; g.j.p.h.boons@uu.nl; gjboons@ccrc.uga.edu

binding site at the globular head of HA^{8–10}. A/H3N2 viruses, which are the leading cause of severe seasonal influenza illness¹¹ resulted in altered glycan binding properties. These changes were first noticed by a lack of agglutination of fowl red blood cells. Furthermore, these viruses also propagate poorly in common laboratory hosts such as MDCK cells and eggs, indicating they require glycan receptors that are not expressed by these cells or hosts^{12–14}. Glycan microarray and other binding studies have indicated that these viruses have lost the ability to bind to simple α 2,6-linked sialoglycans¹⁵ and instead bind to bi-antennary *N*-glycans that have on both arms extended *N*-acetyl lactosamine (LacNAc) moieties terminating in an α 2,6-linked sialoside. Computational studies indicated a binding mode in which the two sialosides of the bi-antennary *N*-glycan bridge the binding sites of a trimeric HA protein resulting in high avidity of binding¹⁵.

We have constructed a glycan microarray populated with bi-antennary *N*-glycans that more closely resemble structures expressed by human respiratory tissue¹⁶. It includes *N*-glycans having multiple consecutive LacNAc repeating units that can be capped by sialic acid. Glycomic analysis has shown that these structures are abundantly expressed by a human epithelial cell line¹⁶ and respiratory tissues^{17,18} and are expected to be relevant receptors for IAVs. Compound **1** (Fig. 1) was established as the minimal epitope for early A/H3N2 viruses, whereas those that emerged after 2000 (e.g., NL2003) did not bind to **1** and required as minimal epitope a compound such as **2** with one of its arms presenting a tri-LacNAc moiety having a terminal α 2,6-linked sialoside¹⁹. Microarray data combined with sequence alignment, reported crystal structures, and MD simulations indicated an alternative binding model for recent A/H3N2 of the 3 C.2 clade in which mutations remote from the receptor binding domain resulted in a rotation of the side chain of Tyr159, creating an extended binding site¹⁶. Furthermore, a 225 G/D mutation reoriented the poly-LacNAc chain of the compound allowing it to establish interactions with this extended binding site.

In this work, we describe the chemoenzymatic synthesis of *N*-glycan **2** and the ¹³C-labeled analogs **3**, **4**, and **5** that differ in the positions of ¹³C-labeled monosaccharides along the poly-LacNAc chain. The ¹³C-labeled glycans are employed in 2D saturation transfer difference (STD)-¹H,¹³C-HSQC experiments to pinpoint those monosaccharides of the poly-LacNAc chain of **2** that engage with HAs from evolutionarily distinct A/H3N2. The NMR data in combination with

in silico and mutagenesis studies demonstrate that mutations distal to the receptor binding domain of HA create an extended binding site that can directly interact with the extended LacNAc chain of glycans. To investigate whether the affinity of HAs of contemporary A/H3N2 strains for the extended glycans has improved or compensated due to mutations at the receptor binding site, we employ a fluorine-containing sialyl-LacNAc derivative (¹⁹F-SLN, compound **6** in Fig. 1) as a ¹⁹F NMR probe to derive relative binding affinities. This demonstrates that the additional contributions of the LacNAc chain improve binding affinity only for recent A/H3N2 strains.

The preparation of isotopomers **3–5** is critical for the STD NMR experiments and uncover the contributions of each monosaccharide of the tri-LacNAc chain of **2** for binding to various HAs. This type of NMR experiment is a powerful approach to characterize glycans interactions with proteins at the atomic level²⁰. Ligand protons in close proximity to the protein receive saturation resulting in strong STD signals. On the other hand, ligand protons that are more remote or do not engage with the protein exhibit weak or no STD signals. Large glycans suffer, however, from very low chemical shift dispersion of their ¹H NMR signals precluding detailed binding studies by ¹H-STD NMR. This is especially problematic when glycans present repeating units, such as poly-LacNAc chains, as these have almost indistinguishable ¹H chemical shifts. These difficulties can be alleviated by using ¹³C isotope labels introduced at specific positions of an oligosaccharide chain. ¹³C-filtered NMR spectra are then specific for these site-selective probes and free of background signals from unlabeled monosaccharides within the *N*-glycan or from the HA glycoprotein. Such isotope labeling preserves binding characteristics while allowing heteronuclear NMR studies with atomic resolution and substantially increased sensitivity^{21,22}. A challenge that we address in this study is the preparation of high-complex glycan having selectively ¹³C labeled monosaccharides.

Results

Synthesis of asymmetrical ¹³C labeled *N*-glycans

Glycan **2** and the three asymmetric isotopomers **3–5** (Fig. 1) were prepared from glycosyl asparagine **8** that could readily be prepared from a sialoglycopeptide extracted from egg-yolk powder followed by pronase treatment and then acid-mediated hydrolysis of the sialosides (Fig. 2a)²³. UDP-GlcNAc and UDP-Gal having ¹³C-isotopes at carbons 1 to 6 in combination with appropriate glycosyl transferases were employed to introduce ¹³C-labeled GlcNAc and Gal moieties into the target oligosaccharides. To selectively extend the α (1,3)-Man arm of **8** by an oligo-LacNAc chain, we exploited the inherent branch selectivity of β -galactoside α 2,6-sialyltransferase 1 (ST6Gal1), which has a 10-fold higher activity for the α 3-arm²⁴, allowing for the convenient preparation of **9**. Next, the remaining terminal galactoside at the α 6-arm of **9** was temporarily modified by an α 1,2-fucoside by treatment with α 1,2-fucosyltransferase 1 (FuT1) in the presence of GDP-Fuc to give **10**. The α 1,2-fucoside temporarily blocks the α 6-arm from modifications by mammalian glycosyl transferases²⁵, and thus after removal of the α 2,6-sialoside of **10** by acid-mediated hydrolysis, the α 3-arm of the resulting **11** could be selectively elaborated into a tri-LacNAc chain by the sequential use of B3GnT2 and B4GalT1 in combination with UDP-GlcNAc and UDP-Gal, respectively. The use of the ¹³C-labeled sugar donors for the installation of the first or second LacNAc moiety gave access to compounds **15** and **20** having the terminal and central LacNAc moiety labeled by ¹³C-isotopes, respectively (Fig. 2b, c). Target glycans **3** and **4** were obtained by treatment **15** and **20** with ST6Gal1 in the presence of CMP-Neu5Ac to give **16** and **21** that were treated with an α 1,2-fucosidase to provide target compounds **3** and **4**.

To introduce the ¹³C labeled galactoside in the inner LacNAc moiety, symmetrical *N*-glycan **8** was treated with galactosyl hydrolase from *E. coli* that has high preference for the α 3-arm²⁶, to give after purification by HILIC-HPLC, asymmetrical glycan **22** (Fig. 3). The

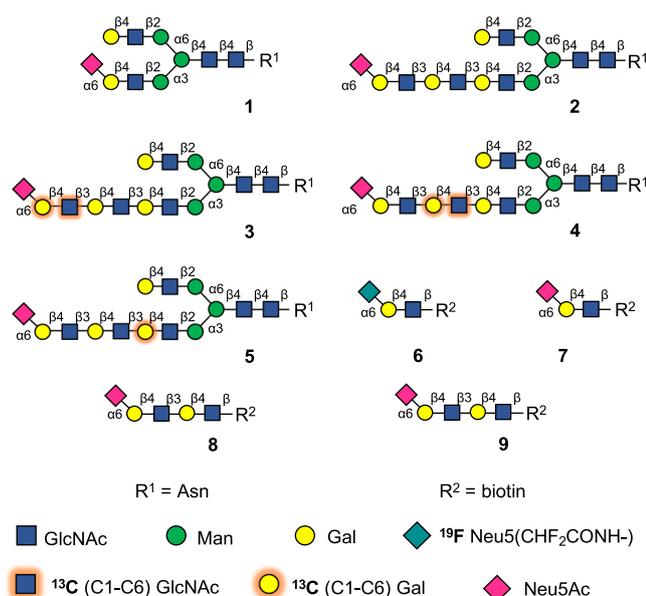


Fig. 1 | Synthesized sialoglycans. Isotopically glycans for NMR studies with evolutionarily distinct HAs.

a. Preparation of common precursor 11

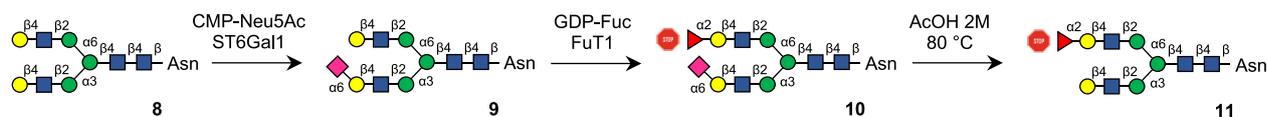
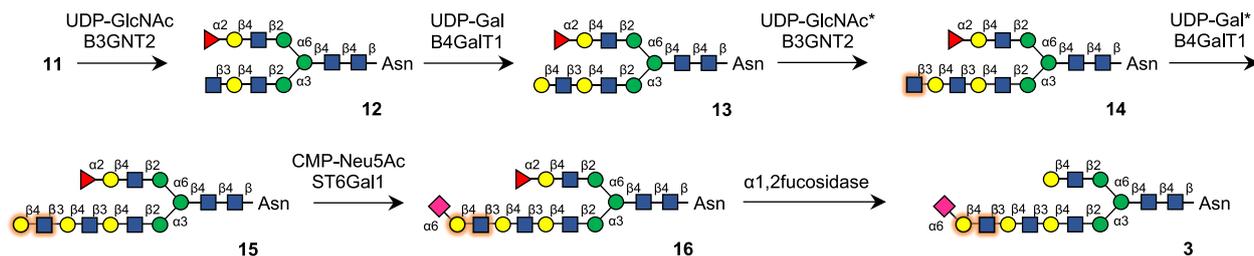
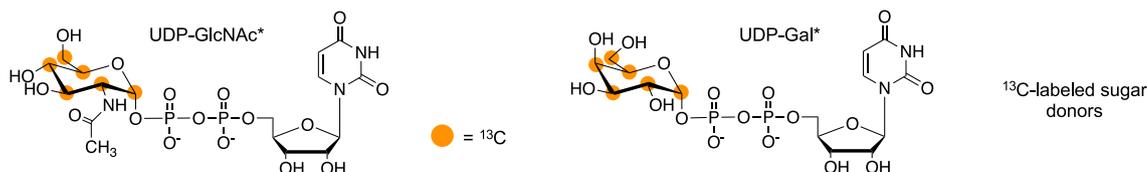
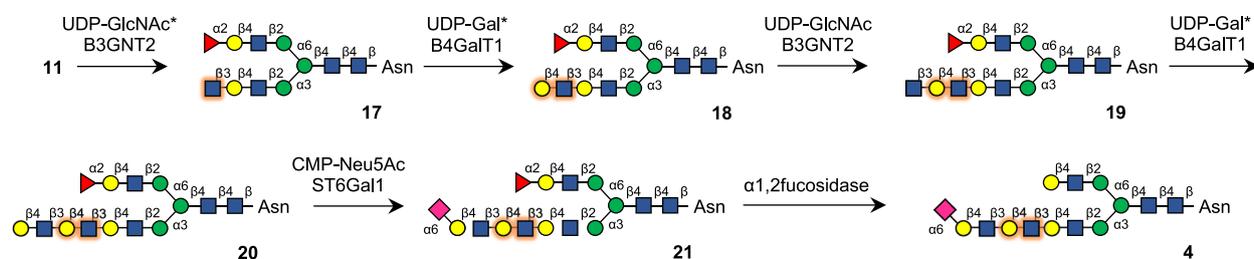
b. Preparation of asymmetrical glycan 3 having terminal ¹³C LacNAc moietyc. Preparation of asymmetrical glycan 4 having a central ¹³C LacNAc moiety

Fig. 2 | Chemoenzymatic synthesis of ¹³C labeled sialoglycans 3 and 4. **a** The $\alpha(1,3)$ -arm of glycan **8**, which was selectively modified as $\alpha(2,6)$ -linked sialoside allowing selectively fucosylation of the $\alpha(1,6)$ -arm which temporarily block its glycosyl accepting properties (stop). After acid hydrolysis of the 2,6-linked sialoside, compound **11** was obtained that allows selective extension of the $\alpha(1,3)$ -arm.

b Glycan **3** was prepared by employing ¹³C labeled UDP-GlcNAc and UDP-Gal during the introduction of the final LacNAc assembly followed by treatment with ST6Gal1 and removal of the fucoside using a 1,2-fucosidase. **c** Glycan **4** was prepared by using ¹³C labeled UDP-GlcNAc and UDP-Gal during the introduction of the central LacNAc moiety.

terminal galactoside at the $\alpha(6)$ -arm of **22** was temporarily blocked as $\alpha(2,3)$ -linked sialoside by treatment with β -galactoside $\alpha(2,3)$ -sialyltransferase 4 (ST3Gal4) in the presence of CMP-Neu5Ac to provide **23**. Next, a ¹³C-labeled galactoside was linked to the terminal GlcNAc moiety of **23** by employing B4GalT1 and ¹³C-labeled UDP-Gal to give **24**. The $\alpha(3)$ -arm of the latter compound was enzymatically extended into a tri-LacNAc structure to give **28** by repeated use of B3GnT2 and B4GalT1. Treatment of **28** with ST6Gal1 in the presence of CMP-Neu5Ac installed an $\alpha(2,6)$ -sialoside thereby providing **29**. Finally, the $\alpha(2,3)$ -linked sialoside at the $\alpha(6)$ -arm of **29** was selectively removed by treatment with $\alpha(2,3)$ -neuraminidase to afford target glycan **5**. The same synthetic strategy was used to synthesize unlabelled **2** from unlabelled UDP-sugar donor.

1D ¹H-STD NMR experiments

HAs from evolutionarily different time points, HK68 (1968), NL91 (1991), and NLO3 (2003), were selected to analyze their binding preferences for unlabelled glycan **2** by 1D ¹H-STD NMR experiments (Fig. 4a). HK68 is the prototypical HA of the 1968 Hong-Kong

pandemic that was adapted from its avian predecessor to bind human-type receptors²⁷. NL91 HA was isolated when these viruses could still hemagglutinate fowl erythrocytes. Later in the antigenic evolution, these viruses lost this property as was observed for in the NLO3 strain.

First, we examined the receptor binding specificities of HK68, NL91, and NLO3 by glycan microarray technology. Thus, a series of *N*-glycans was printed on NHS-activated microarray slides which was probed by several plant lectins for quality control (Supplementary Fig. 4). Next, whole viruses of HK68, NL91, and NLO3 were exposed to the glycan microarray, and binding was determined by using broadly neutralizing antibody CR8020 and an Alexa labeled goat-anti-human antibody (Supplementary Fig. 4). HK68 displays some residual avian-type receptor binding specificity and bound 2,3- as well as 2,6-sialosides. NL91 only bound human-type receptors including compound **1** and **2** that display the 2,6-linked sialoside at a mono- and tri-LacNAc moiety, respectively. The further evolved NLO3 only bound compound **2**. This virus prefers the sialoside at the extended LacNAc moiety to be displayed at the $\alpha(3)$ -mannose antenna as the isomeric compound did

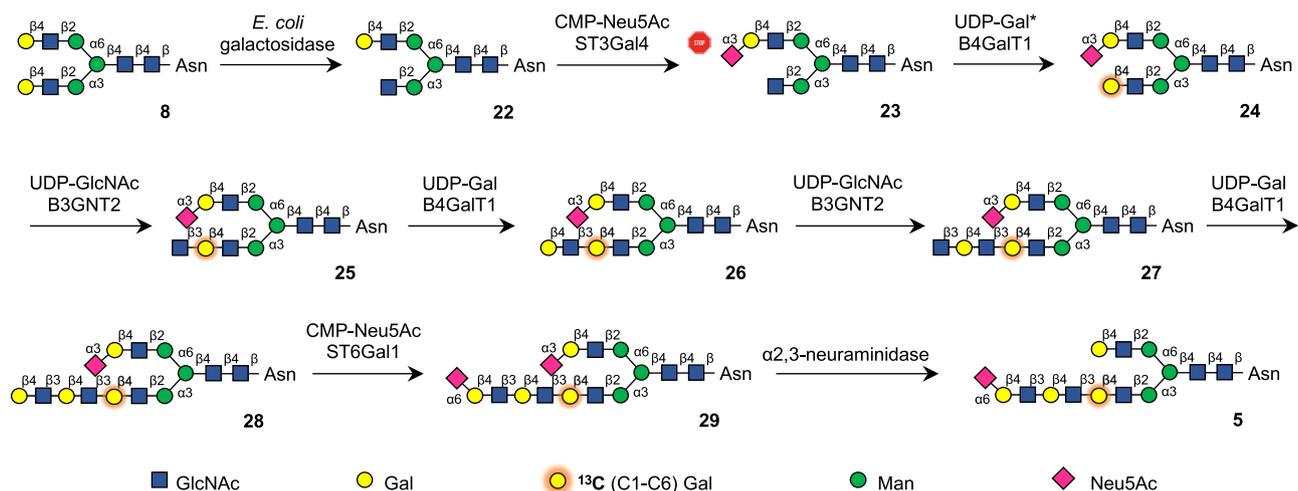


Fig. 3 | Chemoenzymatic synthesis of glycan 5. Selective removal of the galactoside of the α 3-arm allowed the selective installation of the ^{13}C -labeled galactoside at the first residue of the tri-LacNAc moiety to give entry into the synthesis of glycan 5.

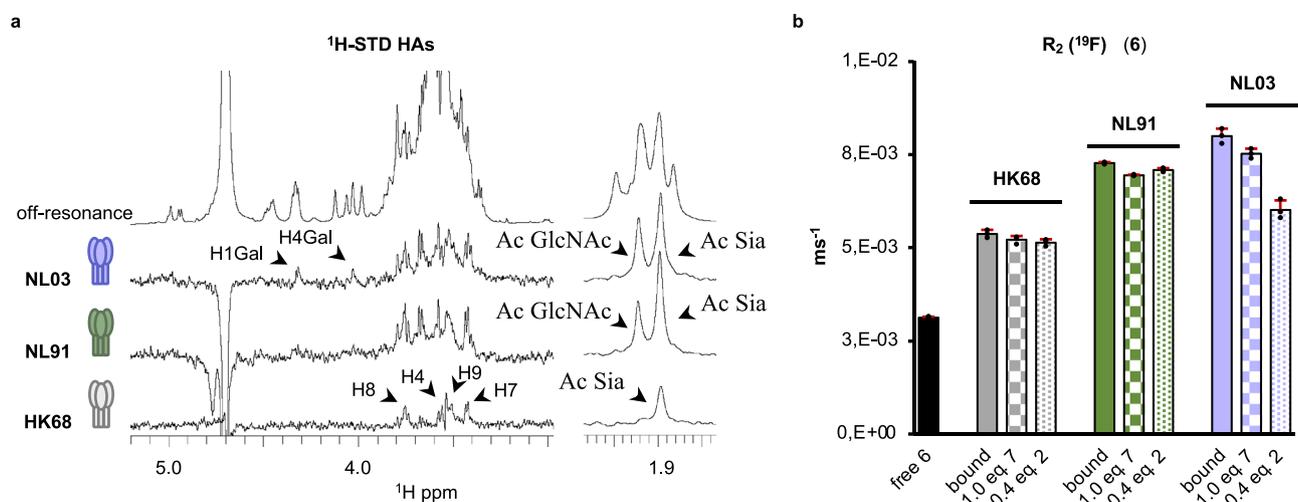


Fig. 4 | NMR binding studies. **a** ^1H -STD NMR. Spectra obtained for the complexes of HK68, NL91, and NL03 HAs with the glycan 2, and inset of the acetyl region. **b** ^{19}F - R_2 NMR. Comparison of transverse ^{19}F relaxation rate, $R_2(^{19}\text{F})$, for the fluorinated probe (6) in the absence or presence of the different HA proteins.

Addition of competitor molecules 7 and 2 causes a decrease in $R_2(^{19}\text{F})$ proportional to their affinities relative to the probe compound 6. The bars represent the average of three replicates \pm SD.

not show binding. Mono- and di-sialylated structures gave similar responsiveness indicating that bidentate binding is not important for recognition. Control compounds lacking sialic acid did not bind.

Next, STD experiments were performed using compound 2 and recombinant HAs from HK68, NL91, and NL03. The STD signals confirmed that all three HAs have functional binding properties. As expected, severe ^1H chemical shift overlap limited ^1H -STD-based analysis that nonetheless revealed clear differences. Thus, HA from HK68 showed STD signals only for the sialic acid moiety, indicating that only this residue contributes to protein binding. The strongest STD signals were observed for H4, H7, H8, H9, and the methyl group of the acetamide moiety of the sialoside. In contrast, the ^1H -STD spectrum of NL91 HA showed several additional STD signals, including the methyl group of an acetamide of a GlcNAc residue. Finally, the ^1H -STD spectrum of NL03 HA was substantially more complex and showed, among others, STD signals above 4 ppm that correspond to the H1 and H4 of galactosides. It was, however, not possible to identify specific monosaccharides due to the chemical shift degeneration of the nuclei of the tri-LacNAc chain. These results indicate that from 1968 to 2003, A/H3N2 HAs evolved to recognize

additional residues of an extended LacNAc chain beyond the terminal α 2,6-sialoside.

Transverse relaxation rate $R_2(^{19}\text{F})$ NMR experiments to establish relative binding affinities

To identify the minimal binding epitope and derive relative binding affinities for the three HAs, we used fluorine-containing Neu5DFA(α 2,6)LacNAc (^{19}F -SLN, compound 6, Fig. 1) as probe (see supplementary section 2 for details of the synthesis of 6). The strategic incorporation of a non-endogenous fluorine atom (^{19}F) enables ligand-protein interaction studies by molecule-specific ^{19}F NMR experiments monitoring $R_2(^{19}\text{F})$ -relaxation rates that do not suffer from overlap or dispersion problems as in ^1H -based NMR experiments described above^{28,29}. First, the overall transverse relaxation rate was measured, $R_2(^{19}\text{F})$, of ^{19}F -SLN in the absence of HA. Addition of the HAs to this fluorinated probe caused a large increase in $R_2(^{19}\text{F})$ relative to the free form indicating ^{19}F -SLN binding (Fig. 4b, filled bars). $R_2(^{19}\text{F})$ increases for NL91 and NL03 were substantially larger than for HK68. These observations indicate a substantial affinity increase for the ^{19}F -SLN trisaccharide from HK68 to NL91 and slightly increased from NL91 to

NL03 as previously also observed by ELISA assays¹⁶. Together with the results from the 1D ¹H-STD NMR experiments, it is concluded that the higher affinity of NL91 and NL03 HAs for this trisaccharide is due to additional recognition of the Gal and GlcNAc residues which does not occur for HK68.

As the replacement of two hydrogen atoms by fluorine at the acetamide moiety may alter ligand binding properties³⁰, the suitability of compound **6** as a probe was examined by a competition experiment using unlabeled sialyl LacNAc **7** (Fig. 1). Thus, 1.0 equivalent of **7** was added to the NMR tube containing the three different HA proteins and ¹⁹F-SLN (**6**). Addition of the non-fluorinated competitor caused a small, but significant decay of R₂(¹⁹F) (*ca.* 5%) that is consistent with similar and low affinities of compounds **6** and **7** for the three HAs (Fig. 4b, square-pattern filled bars) and confirms they bind at the same site of the HAs. The R₂(¹⁹F) decay was simulated which confirmed the observations (Supplementary Section 4.3).

Next, we examined the relative HA affinities for compound **2** that present an α2,6-linked sialoside on its extended tri-LacNAc chain. Thus, 0.4 equivalent of **2** was added relative to the fluorinated probe to samples containing the three different HA proteins and ¹⁹F-SLN (**6**), (Fig. 4b, dotted-pattern filled bars). For the HAs from HK68 and NL91, the addition of the competing glycan **2** caused a small reduction of R₂(¹⁹F) similar to that induced by trisaccharide **7**, indicating that the two additional LacNAc moieties in **2** do not enhance the binding affinity.

On the other hand, addition of compound **2** to the HA of NL03 and ¹⁹F-SLN (**6**) resulted in a substantially larger reduction of R₂(¹⁹F) with a decay of *ca.* 30%. This value is consistent with a K_D that is at least an order of magnitude smaller than that of **6** (Supplementary Fig. 5a, b). These results indicate that the inner di-LacNAc moiety of **2** substantially contributes to recognition by NL03 by providing additional intermolecular interactions in agreement with the ¹H-STD results.

¹H-STD competition experiments were performed to quantify the contribution of the extended LacNAc moiety to the K_D of NL03. Thus, the intensity of unique ¹H-STD signal of the di-fluoro acetamide moiety of **6** was monitored as function of different concentrations of competitors **7**, **8**, and **9** (Fig. 1). Consistent with the ¹⁹F-NMR experiments, the mono- (**7**) and the di-LacNAc (**8**) containing sialosides bound to the protein with a K_D of ~600 μM, which is similar to that derived for the fluorinated compound **6**. The K_D of the tri-LacNAc containing sialoside **9** was determined to be 30 μM. These results further support that the underlying tri-LacNAc chain of **9** substantially contributes to binding of NL03, resulting in a 10-fold higher affinity compared to the mono- and di-LacNAc containing sialosides (The corresponding ¹H-STD spectra are reported in Supplementary Fig. 5c).

2D ¹H-STD-¹H,¹³C-HSQC experiments

To elucidate specific contributions of each monosaccharide of the poly-LacNAc chain of **2** to HA binding, we recorded 2D ¹H-STD-¹H,¹³C-HSQC NMR experiments for isotopomers **3**, **4**, and **5** in the presence of the three HAs. This variant of the ¹H-STD experiment covers ¹H[¹³C]-STD effects only for protons bound to a ¹³C isotope and, therefore unambiguously demonstrates whether specific ¹³C labeled monosaccharides contribute to recognition or not. ¹H-STD-¹H,¹³C-HSQC experiments with isotopomer **3** showed clear differences in recognition between HK68, NL91, and NL03 HA (Fig. 5a). For HK68, no ¹H[¹³C]-STD signals were observed indicating that the terminal LacNAc unit is not part of the recognized epitope. In contrast, ¹H[¹³C]-STD signals were observed with moderate and strong intensities for NL91 HA and NL03 HA, respectively, indicating that these HAs evolved to recognize sialic acid linked to a LacNAc moiety.

Similar experiments with isotopomer **4**, in which the central LacNAc unit is ¹³C-labeled uncovered further differences in recognition by the three HAs (Fig. 5b and Supplementary Fig. 7). As expected, HK68

HA did not produce any ¹H[¹³C]-STD signal while the HA from NL91 and NL03 clearly invoked signals. The ¹H[¹³C] STD signals are derived from both Gal and GlcNAc and were stronger for NL03 HA, where the sum of all STD signal was almost twice as those observed for NL91 HA, indicating that the central LacNAc moiety is more intimately bound by NL03. Finally, ¹H-STD-¹H,¹³C-HSQC experiments with isotopomer **5** (Fig. 5c and Supplementary Fig. 8), which has a ¹³C-labeled Gal farthest away from the sialoside, did not show any ¹H[¹³C]-STD signals for HK68 HA. Thus, this most ancient viral HA recognizes the glycan exclusively through the terminal α2,6-linked sialoside. NL91 HA produced weak ¹H[¹³C]-STD signals while NL03 HA gave rise to the strongest STD signals, where the sum of ¹H[¹³C]-STD intensities was, again, almost twice as those observed for NL91. The most intense ¹H[¹³C]-STD signals from **5** in complex with NL03 HA corresponded to the H3, H4 and H5 protons of the ¹³C labeled Gal residue (Supplementary Fig. 9). In galactose-configured monosaccharides, these protons point in the same direction, forming a so-called alpha-face of the pyranoside rings, and often participate in CH-π interactions with aromatic amino acids receptor proteins^{31,32}.

The 2D ¹H-STD-¹H,¹³C-HSQC experiment in combination with residue-specific ¹³C labeling of glycan **2** made it possible to disentangle the contributions of individual monosaccharides of the extended poly-LacNAc chain for recognition by three different HAs. Increasing ¹H[¹³C]-STD signal intensities were observed during evolution of HA indicating more intimate contacts with the poly-LacNAc chain. Thus, for the original pandemic A/H3N2 human virus (HK68) the poly-LacNAc moiety does not contribute to binding, whereas, for later strains, site-specific mutations close to its glycan binding site enable further interactions with the extended glycan chain. For the evolutionarily early NL91 HA, these interactions are mostly limited to the sialic acid-linked galactoside while the more contemporary NL03 HA recognizes almost the entire poly-LacNAc chain.

Modeling of glycan-HA complexes

To uncover the binding modes of the three HAs, we combined the results from the NMR experiments with in silico modeling (Fig. 6). Full atom 1 μs molecular dynamic (MD) simulations in explicit water were performed for asymmetric N-glycan **2** in complex with the three HAs. In agreement with previous X-ray studies³³, the complex of **2** with HK68 HA showed that the recognition of sialic acid is mediated by hydrophobic interactions with amino acids Y98, H183, W153, and L226, as well as by hydrogen bonding interactions with G135-N137, S228, and E190 (Supplementary Fig. 10a). Analysis of the MD trajectory showed that E190 engages the Sia-1 OH9 through hydrogen bonding with an average distance of ~2.9 Å, while L226 makes hydrophobic interactions with the sialic acid-galactose glycosidic linkage. Beyond the sialic acid residue, the MD simulation did not reveal stable intermolecular interactions and throughout the MD trajectory, the poly-LacNAc chain was flexible and mainly solvent-exposed (Supplementary Fig. 11).

The MD simulation for NL91 HA (Supplementary Fig. 10b and 12) revealed that the S145K mutation at the sialic acid binding site enables a new hydrogen bond interaction with the OH4 of Neu5Ac. To probe the importance of this interaction, we created reverse mutations in NL03. Indeed, the K145 is essential for binding and when mutated to serine almost no response was detected when examined for binding on the glycan microarray (Supplementary Fig. 14). On the other hand, the Y137 side chain in NL91 HA is properly oriented to face the sialic acid-galactose glycosidic bond, which is consistent with the STD signals detected for H4 to H6_s of this Gal residue (Fig. 5a) that were not observed for HK68 HA. We also confirmed the effect of this mutation by creating NL03 N137Y which did not significantly alter glycan binding on the array.

Another relevant NL91 HA mutation is Q189R in the 190-helix where the more extended side chain of the R residue is oriented towards solvent. This leaves space for the acetamide group of the

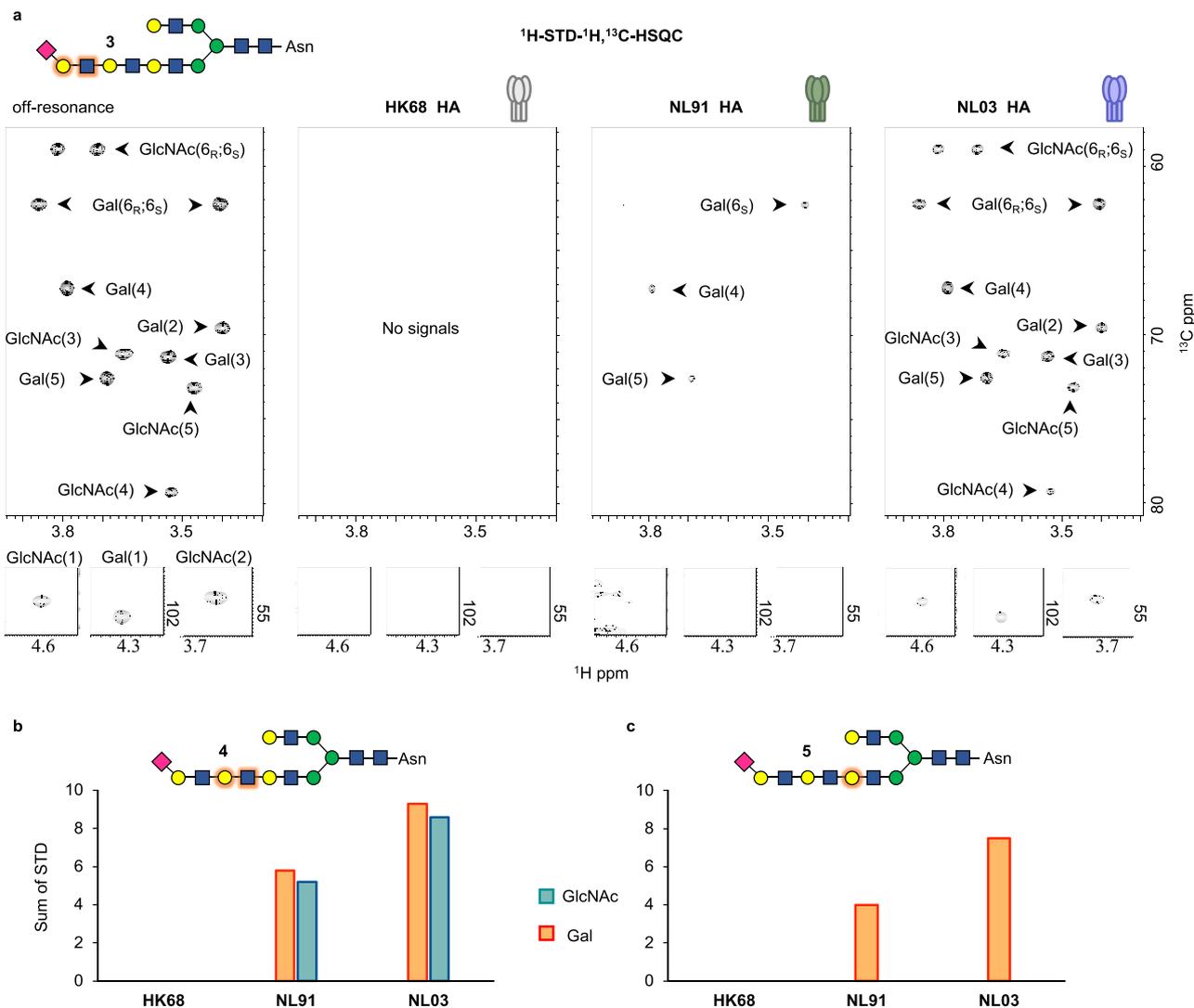


Fig. 5 | 2D $^1\text{H-STD-}^1\text{H,}^{13}\text{C-HSQC}$ NMR experiments probing the interaction of compounds 3 to 5 with different HA proteins. **a Off-resonance reference (left) and associated $^1\text{H-STD-}^1\text{H,}^{13}\text{C-HSQC}$ spectra for the indicated three HA proteins in**

complex with compound **3**. **b**, **c** Sum of all $^1\text{H}^{13}\text{C}$ -STD signals of the ^{13}C labeled Gal (orange) and GlcNAc (cyan) moieties in compounds **4** (**b**) and **5** (**c**) in complex with the three HA proteins.

GlcNAc (GlcNAc-3) to be accommodated into a hydrophobic pocket of the protein, where its methyl group makes van der Waals contacts with L194. These results are consistent with the observed $^1\text{H-STD}$ signal for the acetamide moiety of the GlcNAc residue in NL91 HA (Fig. 4a).

The MD trajectory for NL03 HA in complex with **2** (Supplementary Figs. 10c and 13) showed that the further Y137S mutation restores a hydrogen bond with the carboxylic acid of Neu5Ac as originally observed for HK68³³, while maintaining the K145 hydrogen bond to OH4 of the sialoside, as in NL91. Yet, the concomitant E190D mutation now precludes hydrogen bonding with OH9 of the sialoside due to the shorter D-side chain. The G225D mutation appears crucial to switch the glycan orientation in the receptor binding site consistent with the observed stronger $^1\text{H-STD}$ signal for the terminal Gal in NL03 relative to NL91 HA and with previous X-ray studies^{34,35}. The new hydrogen bonds between D225 side chain and both OH4 and OH3 in Gal forces the $\alpha,6$ -sialyl-Gal glycosidic linkage in a conformation that moves the underlying glycan chain towards the HA 190-helix (Fig. 6d). In this orientation, the hydrophobic interaction between GlcNAc-3 acetamide moiety and L194 is stabilized because S193 can establish an additional hydrogen bond with Gal-4. Further mutations distal to the receptor binding site (RBS) in NL03 HA reorient the Y159 side chain to face the

Gal-6 and establish CH- π interactions (Supplementary Fig. 10c). The MD results are in accordance with the NMR data showing a substantial increase in $^1\text{H-STD}$ signal intensities for H3, H4, and H5 of the Gal-6 residue. Thus, four residues (D225, L194, S193, and Y159) in the binding pocket of NL03 HA accommodate the extended glycan chain and contribute to the strong $^1\text{H}^{13}\text{C}$ -STD signals observed for compounds **3**, **4**, and **5** in the presence of NL03 HA.

Structural studies using additional variants and mutagenesis to confirm the binding models

We expanded our study by including two additional HAs having further mutations in the extended binding site. NL09 HA (2009) has K145N, Y159F, S193F, and D225N mutations that according to our models are expected to impair recognition of glycan **2**. Indeed, the 2D $^1\text{H-STD-}^1\text{H,}^{13}\text{C-HSQC}$ experiment with the three isotopomers **3**, **4**, and **5** showed substantially weaker $^1\text{H}^{13}\text{C}$ -STD signals than observed for NL03 HA (Fig. 7b). MD simulation revealed that N145 cannot engage the sialic acid by hydrogen bonding interaction in the way K145 does for NL03 HA (Supplementary Fig. 10c). Furthermore, the substitution of D by N at residue 225 results in a weaker hydrogen bond acceptor while the S193F mutation precludes hydrogen bond interaction with

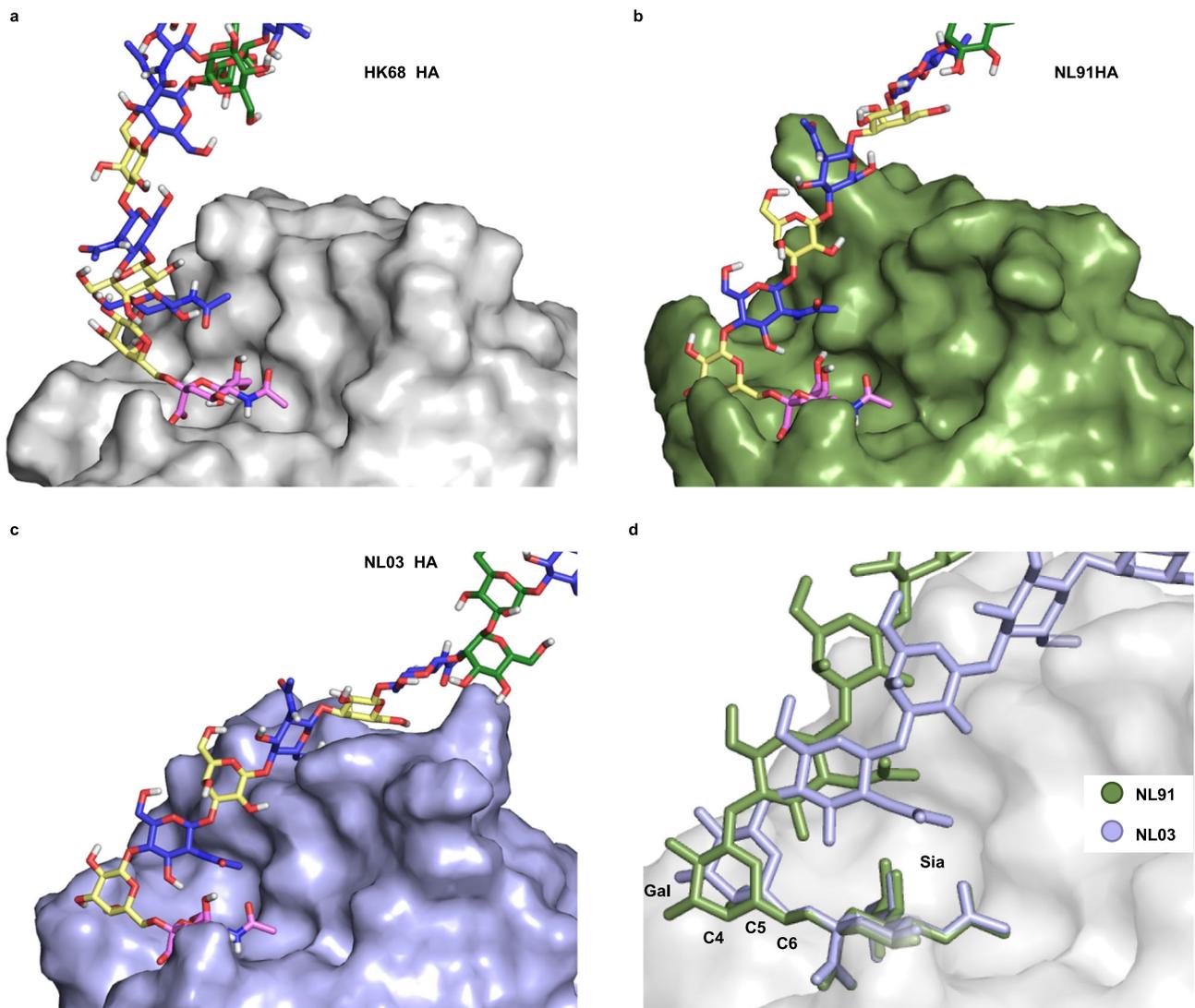


Fig. 6 | Models of *N*-glycan **2 binding by different HA proteins. **a** Binding of **2** by HK68 HA. Only the terminal sialic acid fits into the protein binding pocket while the rest of the poly-LacNAc chain is solvent-exposed. **b** Binding of **2** by NL91 HA. The sialic acid and nearby parts of the α 2,6-linked galactoside establish contacts with the protein. **c** Binding of **2** by NL03 HA. The entire poly-LacNAc chain establishes**

contacts with the protein. **d** Comparison of glycan binding modes for NL91 and NL03 HA. A change in the C5-C6 torsion angle adjacent to the Neu5Ac-Gal glycosidic linkage facilitates a reorientation of the underlying glycan chain for extended interaction with NL03 HA.

the Gal in the central LacNAc unit. Finally, the Y159F mutation produces a weaker CH- π acceptor³².

Mutational reversibility has been observed in HAs³⁴, and for example in Sing16 (2016) residues 225 and 159 reverted back to D and Y, respectively, as in NL03 HA, and residue 193 back to F, as in NL09 HA. MD simulations confirmed that the key hydrogen bond between D225 and Gal-2 has been regained (Supplementary Fig. 10e) while the bulky side chain of F193 provokes steric clashes with the glycan chain as in NL09 HA. Finally, Y159 re-established strong CH- π interactions with Gal-6. Accordingly, NMR experiments showed strong $^1\text{H}[^{13}\text{C}]$ -STD signals for the H3, H4, and H5 in Gal-6 as observed for NL03 HA.

To further assess the key contribution of Y159 for receptor binding, we produced a mutant of the Sing16 HA variant in which this tyrosine moiety is replaced by alanine, thereby abrogating the CH- π interaction. The NMR experiments showed substantial weaker $^1\text{H}[^{13}\text{C}]$ -STD signals for Gal-6 (Fig. 7c, d) indicating much less intimate contacts. These results underpin that several residues cooperate in glycan binding.

Discussion

Human influenza A viruses have a remarkable ability to evolve and evade neutralization by antibodies elicited by prior infections or vaccination^{36–39}. This antigenic drift is mainly caused by mutations of amino acid in the globular head of HA where binding occurs with sialic acid-containing receptors of host cells. As a result, receptor binding modes of IAVs usually co-evolve by orchestrated mutations of several amino acids that allow immune evasion but maintain glycan binding capabilities³³. A/H3N2 viruses, which are the leading cause of severe seasonal influenza illness⁷, exhibit a particularly fast antigenic drift and have evolved to recognize sialosides presented on extended LacNAc moieties. Predicting influenza evolution requires an understanding, at a molecular level, how mutational changes driven by immune pressure shape glycan recognition. X-ray crystallography has provided key understandings of the recognition of sialosides by HAs^{14,33,39,40}. Implementing such an approach for complex glycans such as **2** is, however, very difficult due to glycan flexibility that results in diffuse electron densities complicating refinement of glycan structures within the

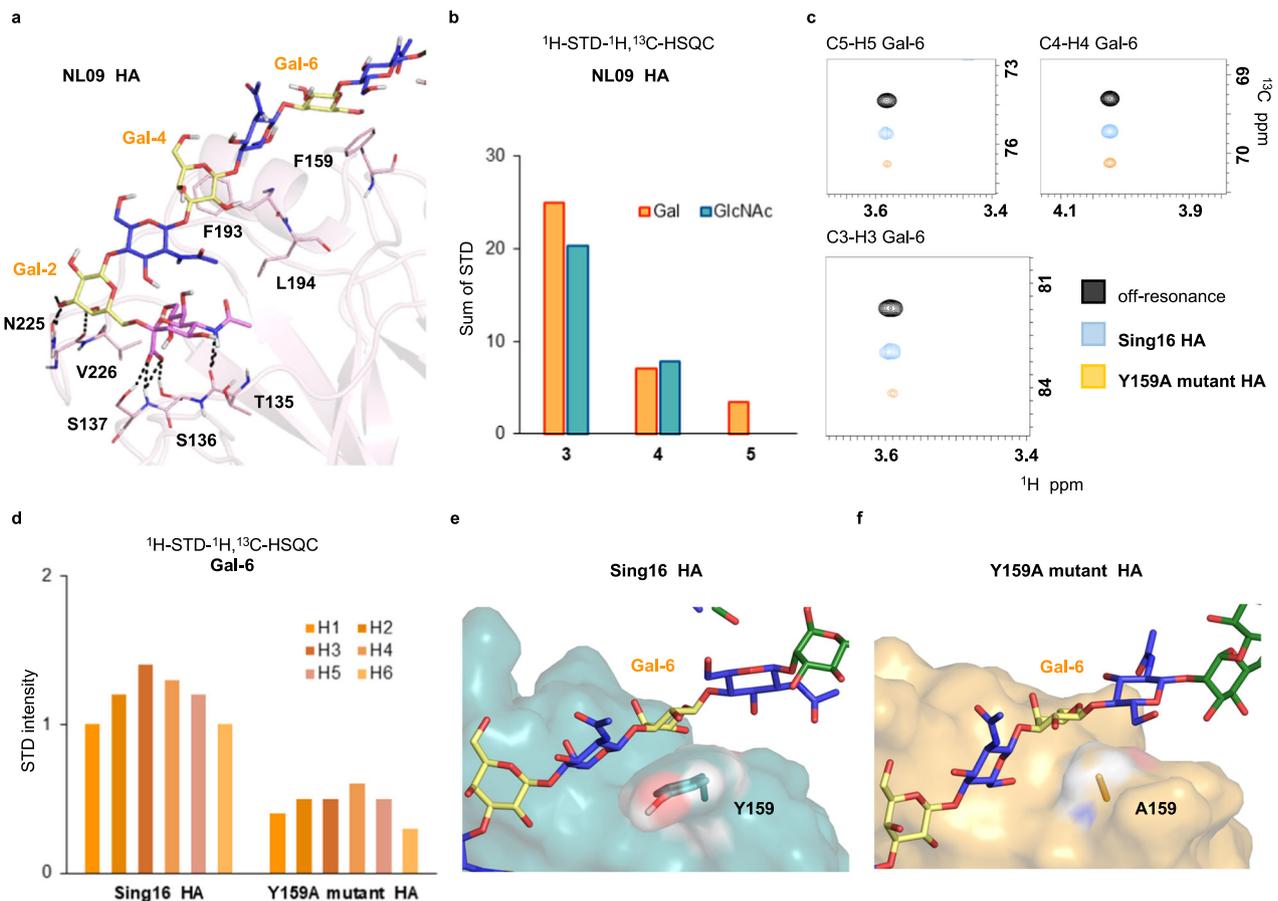


Fig. 7 | Glycan binding by additional HA variants. **a, b** Glycan binding by NL09 HA. MD derived complex structure with **2** (**a**), and $^1\text{H}/^{13}\text{C}$ -STD intensities from the 2D ^1H -STD- ^1H , ^{13}C -HSQC spectrum of **3–5** (**b**). **c** Comparison of $^1\text{H}/^{13}\text{C}$ -STD signals of **5** in the presence of Sing16 HA and its Y159A HA mutant. The STD spectra are offset, in the vertical ^{13}C dimension, to allow a comparison of their intensities. Note the

significantly weaker STD signals for mutant Sing16 Y159A HA relative to the wild type. **d** ^1H -STD intensities for the ^{13}C -labeled galactoside residue in **5**. **e, f** Model of glycan-HA interactions in Sing16 HA (**e**) and its Y159A HA mutant (**f**). The absence of an aromatic moiety in the Y159A mutant precludes the key CH- π interaction with Gal-6.

receptor binding site⁴¹. NMR signals reflect population-weighted averages of component conformers and binding modes under physiological conditions⁴², and is recognized as a versatile technique to study glycans and glycan-receptor interactions with atomic resolution. Various experiments can be considered^{43–47} and among these STD NMR is an attractive solution-based approach that can reveal which protons of a ligand interact with a protein and in combination with computational approaches, can provide models of glycan-protein complexes^{48,49}. The implementation of such an approach for complex glycans is, however, complicated by NMR proton degeneracy. Recently, the binding of a symmetric sialylated-di-LacNAc containing *N*-glycan to an H3 protein was accomplished using an unnatural lanthanide tag to break the NMR signal degeneracy⁵⁰. In this study, we addressed the problem of degeneracy by chemically synthesizing complex *N*-glycans having at specific positions ^{13}C labeled monosaccharides without the need to introduce an artificial probe that may interfere in binding. To obtain the targeted isotopomers, we employed an *N*-glycan isolated from a natural source as starting material, recombinant human glycosyltransferases, and ^{13}C labeled sugar nucleotide donors. Arm-selective extension of one of the arms of a symmetrical *N*-glycan starting material was possible by exploiting inherent branch selectivities of glycosyltransferases and glycosidases combined with modifications that temporarily block one of the arms from further extension⁵¹. The use of the synthetic isotopomers in 2D ^1H -STD- ^1H , ^{13}C -HSQC experiments made it possible to unravel the contribution of individual monosaccharides of an extended poly-LacNAc

chain of an *N*-glycan for binding to evolutionary distinct HAs. In the case of the HA of the original pandemic A/H3N2 (HK68), only the sialoside engages with the protein and no contribution of the poly-LacNAc chain for binding was detected. For later strains, mutations close to the glycan binding site enabled interactions with the LacNAc chain. For the HA of the evolutionarily early NL91, these interactions are mostly limited to the sialic acid-linked galactoside while the more contemporary NLO3 HA recognized the entire tri-LacNAc chain. The latter was confirmed by a fluorine-containing sialyl-LacNAc derivative as NMR probe that provided relative binding affinities and demonstrated additional contributions of the extended LacNAc chain for binding. The importance of an extended LacNAc moiety for binding of NLO3 was also demonstrated by glycan microarray in which only a 2,6-sialoside presented at a tri-LacNAc moiety gave strong responsiveness. Thus, NMR solution-based binding studies as well as a binding assay in which glycans are immobilized on surface demonstrated the importance of the extended LacNAc moiety for binding of A/H3N2 viruses of the 3C2 clade such as NLO3. Previous studies have shown that fowl erythrocytes do not display sialosides on extended LacNAc moieties and accordingly NLO3 cannot agglutinate these cells¹⁶. These viruses also propagate poorly in MDCK cells which also do not express these receptors^{52,53}. Respiratory tissues abundantly display, however, *N*-glycans having extended LacNAc moieties modified by sialosides making these appropriate receptors for IAV infection. It can, however, not be excluded that a high density of low-affinity receptor may also contribute to infection¹⁹.

Combination of the NMR data with molecular modeling revealed how specific residues along the extended binding site of HA synergistically operate in glycan recognition. In particular, a G225D mutation changes the glycan orientation and moves the LacNAc chain closer to the protein surface. A distal mutation to the receptor binding site in NL03 HA reorient the Y159 side chain to face the Gal-6 making it possible to establish CH- π interactions. Mutagenesis experiments confirmed the importance of the Y159 side chain for interacting with the extended LacNAc chain.

IAV have a remarkable ability to evolve and evade neutralization by antibodies elicited by prior infections or vaccinations. Understanding this evolution at a molecular level is important for vaccine design and the development predictive models for the evolution of IAV variants. Here we demonstrate for, the first time, that H3N2 viruses of the 3C.2 clade such as NL03 have evolved to make direct interactions with 2,6-sialoside presented on an extended LacNAc moiety resulting in a much higher affinity in binding.

Methods

Enzymatic synthesis

Materials and general protocols. Glycosyltransferase were expressed as previously described (St6GalI, FutI, St3Gal4, B4GalT1, and B3GnT2)⁵⁴. Galactosidase of *E. coli* was purchased from Sigma–Aldrich [Cat# G5635] α 1,2-Fucosidase [Cat# P0724S] and α 2,3-Neuraminidase [Cat# P0743] were purchased from New England BioLab. Alkaline phosphatase (FastAP) was purchased from Thermo Scientific [Cat# EF0651]. Uridine 5'-diphosphogalactose (UDP-Gal), uridine 5'-diphospho-N-acetyl-glucosamine (UDP-GlcNAc) and cytidine-5'-monophospho-N-acetylneuraminic acid (CMP-Neu5Ac) were obtained from Roche Diagnostics [UDP-Gal: Cat# 07703562103; UDP-GlcNAc: Cat# 06369855103; CMP-Neu5Ac: Cat# 05974003103]. ¹³C labeled Uridine 5'-diphosphogalactose (UDP-Gal) and uridine 5'-diphospho-N-acetyl-glucosamine (UDP-GlcNAc) were obtained from Omicron [UDP-GlcNAc CI-C6: Cat# NTS-009; UDP-Gal CI-C6: Cat# NTS-005]. Reaction mixtures were purified by a size exclusion Biogel (P2) from BioRad in Econo glass columns (0.7 × 30 cm/1.5 × 30 cm/1.5 × 50 cm) coupled to a BioFrac fraction collector (BioRad). Carbohydrate-containing fractions were detected by thin-layer chromatography and an appropriate staining reagent (15 mL AcOH and 3.5 mL p-anisaldehyde in 350 mL EtOH and 50 mL H₂SO₄). Final products were purified by high-performance liquid chromatography (HPLC) using an XBridge HILIC column (10 mm (Ø) × 250 mm (l), 5 μ m particle size) on a semi-preparative liquid chromatography system from Shimadzu (LC-20AT, SIL-20A, CBM-20A, SPD-20A, FRC-10A). The purification was performed using 10 mM NH₄HCO₃ in 10% H₂O in MeCN (buffer B) and 10 mM NH₄HCO₃ in 100% H₂O (buffer A). The progress of the reactions was monitored on a Shimadzu liquid chromatography–mass spectrometry (LC–MS) (system controller: SCL10A-VP; HPLC pumps: LC10AD-VP; injector: SIL10AD-VP) using a ZIC HILIC column (ZeQuant, PEEK coated guard HPLC column, 3.5 μ m particle size, 20 × 2.1 mm). The LC system was attached to a Bruker Daltonics microTOF-Q mass spectrometer.

General procedures for synthesis

Acid-mediated hydrolysis of Neu5Ac. The substrate was dissolved in an aqueous solution of acetic acid (2 M) and kept at 65 °C for 24 h. The solvents were removed under an N₂ flow and the resulting mixture was applied to size exclusion chromatography. Carbohydrate-containing fractions were lyophilized and used without further purification.

Installation of α 2,6-linked Neu5Ac using ST6GalI. Acceptor and CMP-Neu5Ac (1.5 eq) were dissolved in a Tris buffer (50 mM, pH 7.3, 0.1 wt% BSA) to obtain an acceptor concentration of 2 mM. ST6GalI (42 μ g per μ mol acceptor) was added to the mixture and the resulting reaction mixtures were incubated overnight at 37 °C with gentle

shaking. The progress of the reaction was monitored by LC–MS. In case of incomplete conversion, additional ST6GalI (20 μ g per μ mol acceptor) was added and the reaction mixture was incubated at 37 °C for an additional 24 h. After completion of the reaction, the enzyme was removed by spin filtration using a 10 kDa cut-off filter, the filtrate was lyophilized and applied to size exclusion chromatography. Carbohydrate-containing fractions were collected, concentrated by freeze drying, and either used for further modification or further purification by HPLC.

Installation of α 1,2-linked fucoside using FuT1. Acceptor and GDP-Fuc (1.3 eq) were dissolved in a Tris buffer (50 mM, pH 7.3, 0.1 wt% BSA) containing MnCl₂ (10 mM) to a final acceptor concentration of 5 mM. FuT1 (7 μ g per μ mol acceptor) were added and the reaction mixture was incubated at 37 °C for 72 h with gentle shaking. The progress of the reaction was monitored by LC–MS.

Installation of β 1,3-linked GlcNAc using B3GnT2. Acceptor and UDP-GlcNAc (1.5 eq) were dissolved in a HEPES buffer (50 mM, pH 9.6, 0.1 wt% BSA) containing DTT (1 mM) and MnCl₂ (20 mM) to obtain a concentration of 5 mM. B3GnT2 (30 μ g per μ mol acceptor) and CIAP (1 u/ μ L, 1 u per μ mol of added nucleotide) were added to the mixture and then incubated overnight at 37 °C with gentle shaking. The progress of the reaction was monitored by LC–MS. In case of incomplete conversion, additional UDP-GlcNAc (0.5 eq), CIAP (1 u/ μ L, 1 u per μ mol of added nucleotide), and B3GnT2 (15 μ g per μ mol acceptor) were added and the reaction mixture incubated at 37 °C for an additional 24 h. The reaction mixture was lyophilized and applied to size exclusion chromatography. Carbohydrate-containing fractions were lyophilized and used without further purification.

Installation of β 1,4-linked galactoside using B4GalT1. Acceptor and UDP-Gal (1.5 eq) were dissolved in a Tris buffer (50 mM, pH 7.3, 0.1 wt% BSA) containing MnCl₂ (20 mM) to obtain a concentration of 5 mM. B4GalT1 (20 μ g per μ mol acceptor) and CIAP (1 u/ μ L, 1 u per μ mol of added nucleotide) were added and the resulting reaction mixture was incubated for 18 h at 37 °C with gentle shaking. The progress of the reaction was monitored by LC–MS. In case of incomplete conversion, additional UDP-Gal (0.5 eq), CIAP (1 u/ μ L, 1 u per μ mol of added nucleotide), and B4GalT1 (10 μ g per μ mol acceptor) were added and the reaction mixture incubated at 37 °C for a further 24 h. The reaction mixture was lyophilized and applied to size exclusion chromatography. Carbohydrate-containing fractions were lyophilized.

Fucoside hydrolysis. The fucosylated glycan was dissolved into 50 mM sodium acetate buffer at pH 5.5, 5 mM CaCl₂, and treated with the broad specific α 1-2,3,4,6 fucosidase (80 units). The reaction mixture was incubated at 37 °C with gentle shaking. The progress of the reaction was monitored by LC–MS.

Galactoside hydrolysis. The galacto-containing glycan was dissolved in 50 mM TRIS buffer at pH 7.3, 15 mM MgCl₂ and treated with galactosidase (20 units). The reaction mixture was incubated at 37 °C with gentle shaking. The progress of the reaction was monitored by LC–MS.

Installation of α 2,3-linked Neu5Ac using ST3Gal4. Acceptor and CMP-Neu5Ac (1.5 eq) were dissolved in a Tris buffer 50 mM, pH 7.5, 1% v/v BSA (from stock solution 1 mg/mL), and 1% v/v CIAP (from stock solution 20 U/ μ L) to a final acceptor concentration of 5 mM. ST3Gal4 (42 μ g per μ mol acceptor) was added and the reaction mixture was incubated overnight at 37 °C with gentle shaking. The progress of the reaction was monitored by LC–MS.

Enzymatic hydrolysis of α 2,3-linked Neu5Ac. The α 2,3-sialylated glycan was dissolved into 50 mM sodium acetate buffer at pH 5.5, 5 mM

CaCl₂, and treated with the high specific α ,2,3- Neuraminidase S (160,000 U/mg). The reaction mixture was incubated at 37 °C with gentle shaking. The progress of the reaction was monitored by LC–MS.

HA expression

Recombinant trimeric IAV hemagglutinin proteins open reading frames were cloned into the pCD5 expression vector as described previously⁵⁵, in frame with a GCN4 trimerization motif (KQIED-KIEEIESKQKKEIENIARIKK), a superfolder GFP⁵⁶ and the Twin-Strep-tag (WSHPQFEKGGGSGGGSWSHPQFEK); (IBA, Germany). The open reading frames of the HAs of A/Hong-Kong/1/1968 H3 (AFG71887.1), A/Netherlands/816/1991 H3 (EPI_ISL_114608), A/Netherlands/109/2003 H3 (EPI_ISL_113016), A/Netherlands/761/2009 H3 (EPI_ISL_1107270), A/Singapore/INFH-16-0019/2016 H3 (3 C.2a) (QQY97257.1), were synthesized and codon optimized by GenScript.

The trimeric HAs were expressed in HEK293T (CRL-11268) and HEK 293 S GNTI(-) (ATCC CRL-3022) cells with polyethyleneimine I (PEI) in a 1:8 ratio (μ g DNA: μ g PEI) for the HAs as previously described¹⁶. The transfection mix was replaced after 6 h by 293 SFM II suspension medium (Invitrogen, 11686029), supplemented with sodium bicarbonate (3.7 g/L), Primatone RL-UF (3.0 g/L, Kerry, NY, USA), glucose (2.0 g/L), glutaMAX (1%, Gibco), valproic acid (0.4 g/L) and DMSO (1.5%). Culture supernatants were harvested 5 days post-transfection and the HAO was purified with sepharose strep-tactin beads (IBA Life Sciences, Germany) according to the manufacturer's instructions.

NMR sample preparation and analysis

Proteins and ligands were dissolved in a deuterated TRIS-d₁₁ 50 mM buffer pD 7.8 in D₂O. Shigemi NMR tubes with a diameter of 3 mm were used. ¹H-STD and ¹H,¹³C-STD HSQC experiments were acquired with a HA proteins concentration of 5 μ M and a ligand at 300 μ M (ratio 1:60). STD experiments were acquired on a 800 MHz BRUKER AVANCE III spectrometer, equipped with a TCI cryo-probe with z-gradient coil and TopSpin 3.2.7 (BRUKER) software was employed for data acquisition and processing. The temperature was set to 293 K. The 2D STD ¹H,¹³C-HSQC experiment was previously reported²¹. Briefly, the ¹H,¹H-STD module was implemented with saturation by a train of 4 \times 90° PC9_4 shaped pulses (with 1 ms separation) during d₉ = 2 s, applied at the methyl ¹H peak (0.84 ppm) with the non-saturation cut-off at 1.18 ppm (i.e. distance $\Delta\nu$ = 0.34 ppm), resulting in a PC9_4 pulse length of 33 ms at 800 MHz field strength. Protein saturation was alternated with off-resonant irradiation (at -25 ppm) in successive scans. The STD spectrum was then constructed by simple subtraction of both 2D ¹H,¹³C-HSQC spectra. Blank 2D STD-¹H,¹³C-HSQC experiment of the HA glycoproteins alone and glycans 3–5 were acquired as control.

¹⁹F CPMG NMR experiments were acquired on a 600 MHz spectrometer equipped with a Bruker selective ¹⁹F-¹H decoupling (SEF) probe at 298 K, using samples containing the protein at the concentration of 10 μ M, and the F-SLN 6 at 125 μ M. Competitor glycans 7 and 2 were successively added to the sample. The standard CPMG Bruker pulse sequence was modified as described⁵⁷. Twenty-four (24) points were acquired with total echo times from 8 to 5200 ms, with τ = 2 ms. Data were analyzed with the T₁T₂ relaxation module of Topspin 3.5.

In silico studies

MD simulations. When available, starting poses of the HAs were derived from X-ray crystal structures: HK68, A/HK/1/ 1968 H3N2 influenza virus hemagglutinin in complex with 6'-SLN (pdb code 6TZB)³³. NL03, A/Wy/3/03 influenza virus hemagglutinin in complex with 6'-SLN (pdb code 6BKR)³⁴. The starting poses of NL91, NL09, Sing16, and Sing16 (Y159A), were generated by superimposition of the model-derived structure [[http:// 3dfu.cent.uw.edu.pl/index.html](http://3dfu.cent.uw.edu.pl/index.html)]⁵⁸ onto the X-ray crystal structure of the closest variant. The 3D

coordinates of glycan 2 were generated by using the carbohydrate builder GLYCAM-web site [<http://glycam.org>]. The glycosidic torsion angles of the monosaccharides were maintained as observed by X-ray crystallography, while those not resolved were defined according to the lower energy values predicted by the GLYCAM-web modeling tool. The resulting poses were used as starting points for molecular dynamics (MD) simulations. The MD simulations were performed using the Amber16 program⁴ with the protein.ff14SB, the GLYCAM_06j-1, and the water.tip3p force fields parameters^{59,60}. Next, the starting 3D geometries were placed into a 10 Å octahedral box of explicit TIP3P waters, and counter ions were added to maintain electro-neutrality. Two consecutive minimization steps were performed involving (1) only the water molecules and ions and (2) the whole system with a higher number of cycles, using the steepest descent algorithm. The system was subjected to two rapid molecular dynamic simulations (heating and equilibration). The equilibrated structures were the starting points for a final MD simulation at constant temperature (300 K) and pressure (1 atm). 1 μ s molecular dynamic simulations without constraints were recorded, using an NPT ensemble with periodic boundary conditions, a cut-off of 10 Å, and the particle mesh Ewald method. A total of 500,000,000 molecular dynamics steps were run with a time step of 1 fs per step. Coordinates and energy values were recorded every 500,000 steps (500 ps) for a total of 1000 MD models. The detailed analysis of the H-bond and CH- π interactions was performed along the MD trajectory using the cpptraj module included in Amber-Tools 16 package.

Glycan microarray binding of H3N2 viruses, HA proteins, and lectins

Virus isolates (25 μ L) were diluted with PBS-T (PBS + 0.1% Tween, 25 μ L) and applied to the array surface in the presence of oseltamivir (200 nM) in a humidified chamber for 1 h, followed by successive rinsing with PBS-T (PBS + 0.1% Tween), PBS, and deionized water (2 \times) and dried by centrifugation. The virus-bound slide was incubated for 1 h with the CR8020 A/H3N2 influenza hemagglutinin stem-specific antibody (made in-house, 100 μ L, 5 μ g/mL in PBS-T) and washed according to previous washing procedure. A secondary goat-anti-human Alexa 648 antibody (Thermo A-21445, 100 μ L, 2 μ g/mL in PBS-T) was applied, incubated for 1 h in a humidified chamber, and washed again as described above. The control lectins containing a biotin tag were visualized with Streptavidin-AlexaFluor555. Slides were dried by centrifugation after the washing step and scanned immediately. HAs were pre-complexed with human anti-streptag (made in-house) and goat-anti-human Alexa 555 (Thermo A-21433) in a 4:2:1 molar ratio, respectively in 50 μ L PBS-T on ice for 15 min. The mixtures were added to the subarrays for 90 min in a humidified chamber. Wash steps after each incubation (e.g., enzyme treatment, HA, or antibody incubation) involved six successive washes of the whole slides with either twice PBS-T, twice PBS, and twice deionized water. The arrays were dried by centrifugation and immediately scanned as described previously²⁰. Processing of the six replicates was performed by removing the highest and lowest replicates and subsequently calculating the mean value and standard deviation over the four remaining replicates.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The authors declare that the data supporting the findings of this study are available within the paper and its supplementary information files. Source data are provided with this paper. The sequence and 3D structures of HA of HK68 and NL03 were derived from <https://www.rcsb.org/> under the access code 6TZB and 6BKR, respectively. Source data are provided with this paper.

References

1. Taubenberger, J. K. & Morens, D. M. Pandemic influenza including a risk assessment of H5N1. *Rev. Sci. Tech.* **28**, 187–202 (2009).
2. Paget, J. et al. Global mortality associated with seasonal influenza epidemics: new burden estimates and predictors from the GLaMOR project. *J. Glob. Health* **9**, 020421 (2019).
3. Gamblin, S. J. et al. Hemagglutinin structure and activities. *Cold Spring Harb. Perspect. Med.* **11**, a038638 (2021).
4. Long, J. S. et al. Host and viral determinants of influenza A virus species specificity. *Nat. Rev. Microbiol.* **17**, 67–81 (2019).
5. Rogers, G. N. & Paulson, J. C. Receptor determinants of human and animal influenza virus isolates: differences in receptor specificity of the H3 hemagglutinin based on species of origin. *Virology* **127**, 361–373 (1983).
6. Thompson, A. J. & Paulson, J. C. Adaptation of influenza viruses to human airway receptors. *J. Biol. Chem.* **296**, 100017 (2021).
7. Shinya, K. et al. Avian flu: influenza virus receptors in the human airway. *Nature*. **440**, 435–436 (2006).
8. Allen, J. D. & Ross, T. M. H3N2 influenza viruses in humans: viral mechanisms, evolution, and evaluation. *Hum. Vaccin. Immunother.* **14**, 1840–1847 (2018).
9. Rosu, M. E. et al. Substitutions near the HA receptor binding site explain the origin and major antigenic change of the B/Victoria and B/Yamagata lineages. *Proc. Natl Acad. Sci. USA* **119**, e2211616119 (2022).
10. Gouma, S., Weirick, M. & Hensley, S. E. Antigenic assessment of the H3N2 component of the 2019–2020 Northern Hemisphere influenza vaccine. *Nat. Commun.* **11**, 2445 (2020).
11. Jester, B. J., Uyeki, T. M. & Jernigan, D. B. Fifty years of influenza A(H3N2) following the pandemic of 1968. *Am. J. Public Health* **110**, 669–676 (2020).
12. Medeiros, R. et al. Hemagglutinin residues of recent human A(H3N2) influenza viruses that contribute to the inability to agglutinate chicken erythrocytes. *Virology* **289**, 74–85 (2001).
13. Lee, H. K. et al. Comparison of mutation patterns in full-genome A/H3N2 influenza sequences obtained directly from clinical samples and the same samples after a single MDCK passage. *PLoS ONE* **8**, e79252 (2013).
14. Lin, Y. P. et al. Evolution of the receptor binding properties of the influenza A(H3N2) hemagglutinin. *Proc. Natl Acad. Sci. USA* **109**, 21474–21479 (2012).
15. Peng, W. et al. Recent H3N2 viruses have evolved specificity for extended, branched human-type receptors, conferring potential for increased avidity. *Cell Host Microbe* **21**, 23–34 (2017).
16. Broszeit, F. et al. Glycan remodeled erythrocytes facilitate antigenic characterization of recent A/H3N2 influenza viruses. *Nat. Commun.* **12**, 5449 (2021).
17. Chandrasekaran, A. et al. Glycan topology determines human adaptation of avian H5N1 virus hemagglutinin. *Nat. Biotechnol.* **26**, 107–113 (2008).
18. Walther, T. et al. Glycomic analysis of human respiratory tract tissues and correlation with influenza virus infection. *PLoS Pathog.* **9**, e1003223 (2013).
19. Liu, M. et al. Human-type sialic acid receptors contribute to avian influenza A virus binding and entry by hetero-multivalent interactions. *Nat. Commun.* **13**, 4054 (2022).
20. Mayer, M. & Meyer, B. Characterization of ligand binding by saturation transfer difference NMR spectroscopy. *Angew. Chem. Int. Ed.* **38**, 1784–1788 (1999).
21. Moure, M. J. et al. Selective 13 C-labels on repeating glycan oligomers to reveal protein binding epitopes through NMR: polylactosamine binding to galectins. *Angew. Chem. Int. Ed.* **60**, 18777–18782 (2021).
22. Unione, L. et al. The SARS-CoV-2 spike glycoprotein directly binds exogenous sialic acids: a NMR view. *Angew. Chem. Int. Ed.* **61**, e202201432 (2022).
23. Liu, L. et al. Improved isolation and characterization procedure of sialylglycopeptide from egg yolk powder. *Carbohydr. Res.* **452**, 122–128 (2017).
24. Meng, L. et al. Enzymatic basis for N-glycan sialylation: Structure of rat alpha2,6-sialyltransferase (ST6GAL1) reveals conserved and unique features for glycan sialylation. *J. Biol. Chem.* **288**, 34680–34698 (2013).
25. Prudden, A. R. et al. Synthesis of asymmetrical multiantennary human milk oligosaccharides. *Proc. Natl Acad. Sci. USA* **114**, 6954–6959 (2017).
26. Van den Eijnden, D. H., Blanken, W. M. & van Vliet, A. Branch specificity of beta-D-galactosidase from *Escherichia coli*. *Carbohydr. Res.* **151**, 329–335 (1986).
27. Van Poucke, S. et al. Role of substitutions in the hemagglutinin in the emergence of the 1968 pandemic influenza virus. *J. Virol.* **89**, 12211–12216 (2015).
28. Dalvit, C. Ligand- and substrate-based 19F NMR screening: Principles and applications to drug discovery. *Prog. Nucl. Magn. Reson. Spectrosc.* **51**, 243–271 (2007).
29. Fittolani, G. et al. Automated glycan assembly of 19F-labeled glycan probes enables high-throughput NMR studies of protein-glycan interactions. *Angew. Chem. Int. Ed.* **60**, 13302–13309 (2021).
30. Unione, L. et al. Fluoroacetamide moieties as NMR spectroscopy probes for the molecular recognition of GlcNAc-containing sugars: modulation of the CH- π stacking interactions by different fluorination patterns. *Chem. Eur. J.* **3**, 3957–3965 (2017).
31. Hudson, K. L. et al. Carbohydrate-aromatic interactions in proteins. *J. Am. Chem. Soc.* **137**, 15152–15160 (2015).
32. Kiessling, L. L. & Diehl, R. C. CH- π interactions in glycan recognition. *ACS Chem. Biol.* **16**, 1884–1893 (2021).
33. Wu, N. C. et al. Major antigenic site B of human influenza H3N2 viruses has an evolving local fitness landscape. *Nat. Commun.* **11**, 1233 (2020).
34. Wu, N. C. et al. A complex epistatic network limits the mutational reversibility in the influenza hemagglutinin receptor-binding site. *Nat. Commun.* **9**, 1264 (2018).
35. Wu, N. C. et al. A structural explanation for the low effectiveness of the seasonal influenza H3N2 vaccine. *PLoS Pathog.* **13**, e1006682 (2017).
36. Wu, N. C. & Wilson, I. A. Influenza hemagglutinin structures and antibody recognition. *Cold Spring Harb. Perspect. Med.* **10**, a038778 (2020).
37. Zost, S. J. et al. Identification of antibodies targeting the H3N2 hemagglutinin receptor binding site following vaccination of humans. *Cell Rep.* **29**, 4460–4470 (2019).
38. Chambers, B. S. et al. Identification of hemagglutinin residues responsible for H3N2 antigenic drift during the 2014–2015 influenza season. *Cell Rep.* **12**, 1–6 (2015).
39. Lee, P. S. et al. Receptor mimicry by antibody F045-092 facilitates universal binding to the H3 subtype of influenza virus. *Nat. Commun.* **5**, 3614 (2014).
40. Wu, N. C. & Wilson, I. A. Structural biology of influenza hemagglutinin: an amaranthine adventure. *Viruses* **12**, 1053 (2020).
41. Agirre, J. et al. Carbohydrate anomalies in the PDB. *Nat. Chem. Biol.* **11**, 303 (2015).
42. Unione, L., Arda, A., Jimenez-Barbero, J. & Millet, O. NMR of glycoproteins: profiling, structure, conformation and interactions. *Curr. Opin. Struct. Biol.* **68**, 9–17 (2021).
43. Blaum, B. S., Neu, U., Peters, T. & Stehle, T. Spin ballet for sweet encounters: saturation-transfer difference NMR and X-ray

- crystallography complement each other in the elucidation of protein-glycan interactions. *Acta Crystallogr. F Struct. Biol. Commun.* **74**, 451–462 (2018).
44. Meyer, B. & Peters, T. NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors. *Angew. Chem. Int. Ed.* **42**, 864–890 (2003).
 45. Rademacher, C. & Peters, T. Molecular recognition of ligands by native viruses and virus-like particles as studied by NMR experiments. *Top. Curr. Chem.* **273**, 183–202 (2008).
 46. Angulo, J. et al. NMR analysis of carbohydrate-protein interactions. *Methods Enzymol.* **416**, 12–30 (2006).
 47. Gimeno, A., Valverde, P., Ardá, A. & Jiménez-Barbero, J. Glycan structures and their interactions with proteins. A NMR view. *Curr. Opin. Struct. Biol.* **62**, 22–30 (2020).
 48. Peters, T. & Pinto, B. M. Structure and dynamics of oligosaccharides: NMR and modeling studies. *Curr. Opin. Struct. Biol.* **6**, 710–720 (1996).
 49. Angulo, J. et al. Structural biology of glycan recognition. In *Essentials of Glycobiology* 4th edn, Vol 30 (eds Varki, A., Cummings, R. D., Esko, J. D., et al.) (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2022).
 50. Fernández de Toro, B. et al. Avenues to characterize the interactions of extended N-glycans with proteins by NMR spectroscopy: the influenza hemagglutinin case. *Angew. Chem. Int. Ed.* **57**, 15051–15055 (2018).
 51. Liu, L. et al. Streamlining the chemoenzymatic synthesis of complex N-glycans by a stop and go strategy. *Nat. Chem.* **11**, 161–169 (2019).
 52. Kikuchi, C. et al. Glyco-engineered MDCK cells display preferred receptors of H3N2 influenza absent in eggs used for vaccines. *Nat. Commun.* **14**, 6178 (2023).
 53. Spruit, C. M. et al. Contemporary human H3N2 influenza A viruses require a low threshold of suitable glycan receptors for efficient infection. *Glycobiology* **33**, 784–800 (2023).
 54. Moremen, K. W. et al. Expression system for structural and functional studies of human glycosylation enzymes. *Nat. Chem. Biol.* **14**, 156–162 (2018).
 55. de Vries, R. P. et al. The influenza A virus hemagglutinin glycosylation state affects receptor-binding specificity. *Virology* **403**, 17–25 (2010).
 56. Nemanichvili, N. et al. Fluorescent trimeric hemagglutinins reveal multivalent receptor binding properties. *J. Mol. Biol.* **431**, 842–856 (2019).
 57. Urick, A. K. et al. Protein-observed fluorine NMR is a complementary ligand discovery method to 1H CPMG ligand-observed NMR. *ACS Chem. Biol.* **11**, 3154–3164 (2016).
 58. Mazzocco, G. et al. 3DFlu: database of sequence and structural variability of the influenza hemagglutinin at population scale. *Database* **2**, baw130 (2016).
 59. Salomon-Ferrer, R. et al. An overview of the Amber biomolecular simulation package. *WIREs Comput. Mol. Sci.* **3**, 198–210 (2013).
 60. Case, D. A. et al. The Amber biomolecular simulation programs. *J. Comput. Chem.* **26**, 1668–1688 (2005).

Acknowledgements

Research reported in this publication was supported by the Human Frontier Science Program Organization (HFSP) grant LT000747/2018-C (to L.U.), an ERC starting grant (802780) from the European Commission (to R.P.d.V.), and the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under Award Number R01 AI165692 (to G.J.B.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author contributions

L.U., A.N.A.A., G.P.B., E.U., R.L., F.B., R.W., Y.L., S.M., L.L., M.G.R., and I.A.B. performed the experiments. L.U., P.V., T.D., A.A., and R.P.V. analyzed the data and L.U., A.A., and R.P.V. interpreted the results. L.U., R.P.V., and G.J.B. wrote the manuscript. L.U., R.P.V., and G.J.B. supervised the research project. All authors have given approval to the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-47344-y>.

Correspondence and requests for materials should be addressed to Luca Unione, Robert P. de Vries or Geert-Jan Boons.

Peer review information *Nature Communications* thanks the anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024