N-glycolylneuraminic acid binding of avian H7 influenza A viruses

Cindy M. Spruit[1], Xueyong Zhu[2], Frederik Broszeit[1], Alvin X. Han[3], Roosmarijn van der Woude[1], Kim M. Bouwman[1], Michel M. T. Luu[1], Colin A. Russell[3], Ian A. Wilson[2,4], Geert-Jan Boons[1,5], and Robert P. de Vries[1*]

[1] Department of Chemical Biology & Drug Discovery, Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Utrecht, The Netherlands

[2] Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA, United States of America

[3] Department of Medical Microbiology, Amsterdam University Medical Center, Amsterdam, The Netherlands

[4] Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, CA, United States of America

[5] Complex Carbohydrate Research Center, University of Georgia, Athens, GA, United States of America

* Corresponding author

E-mail: r.vries@uu.nl

Abstract

Influenza A viruses initiate infection by binding to glycans with terminal sialic acids present on the cell surface. Hosts of influenza A viruses variably express two major forms of sialic acid, N-acetylneuraminic acid (NeuAc) and N-glycolylneuraminic acid (NeuGc). NeuGc is produced in the majority of mammals including horses, pigs, and mice, but is absent in humans, ferrets, and birds. Intriguingly, the only known naturally occurring influenza A viruses that exclusively bind NeuGc are the extinct highly pathogenic equine H7N7 viruses. We determined the crystal structure of a representative equine H7 hemagglutinin (HA) in complex with its NeuGc ligand and observed a high similarity in the receptor-binding domain with an avian H7 HA. To determine the molecular basis for NeuAc and NeuGc specificity, we performed systematic mutational analyses, based on the structural insights, on two distant avian H7 HAs. We found that mutation A135E is key for binding α2,3-linked NeuGc but does not abolish NeuAc binding. Interestingly, additional mutations S128T, I130V, or a combination of T189A and K193R, converted from NeuAc to NeuGc specificity as determined by glycan microarrays. However, specific binding to NeuGc-terminal glycans on our glycan array did not always correspond with full NeuGc specificity on chicken and equine erythrocytes and tracheal epithelium sections. Phylogenetic analysis of avian and equine H7 HAs that investigated the amino acids at positions 128, 130, 135, 189, and 193 reveals a clear distinction between equine and avian residues. The highest variability in amino acids (four different residues) is observed at key position 135, of which only the equine glutamic acid leads to binding of NeuGc. The results demonstrate that avian H7 viruses, although genetically distinct from equine H7 viruses, can bind NeuGc after the introduction of two to three mutations, providing insights into the adaptation of H7 viruses to NeuGc receptors.

Author summary

Influenza A viruses cause millions of cases of severe illness and deaths annually. To initiate infection and replicate, the virus first needs to bind to a structure on the cell surface, like a key fitting in a lock. For influenza A virus, these 'keys' (receptors) on the cell surface are chains of sugar molecules (glycans). The terminal sugar on these glycans is often either N-acetylneuraminic acid (NeuAc) or N-glycolylneuraminic acid (NeuGc). Most influenza A viruses bind NeuAc, but a small minority binds NeuGc. NeuGc is present in species like horses, pigs, and mice, but not in humans, ferrets, and birds. Therefore, NeuGc binding could be a determinant of an Influenza A virus species barrier. Here, we investigated the molecular determinants of NeuGc specificity and the origin of viruses that bind NeuGc.

Introduction

Influenza A viruses can infect a broad range of animals, including mammalian and avian species. Infection is initiated when the hemagglutinin (HA) on the outside of a virus particle binds to glycans with terminal sialic acid on the cell surface. The vast majority of Influenza A viruses use a glycan with a terminal N-acetylneuraminic acid (NeuAc) as their receptor, although some strains use N-glycolylneuraminic acid (NeuGc) instead. Sialic acids are bound in the receptor-binding site (RBS) of the HA, consisting of conserved residues (Y98, W153, H183, and Y195) and structural features (130-, 150-, and 220-loop and 190-helix) [1]. Amino acid mutations in or near the RBS can change HA binding specificity, as shown extensively for HAs binding to either α2,3-linked or α2,6-linked NeuAc [2-6].

The ability of viruses to bind either α2,3-linked or α2,6-linked sialic acids is a host determinant. Binding to either NeuAc or NeuGc could likewise affect the host range. NeuGc is only present in species that express an active form of the enzyme cytidine monophosphate CMP-N-acetyl neuraminic acid hydroxylase (CMAH), which facilitates the hydroxylation of NeuAc to convert it to NeuGc. The gene encoding CMAH, mainly expressed in mammalian species, has been partially or completely lost at several distinct events during evolution [7], causing NeuGc to be absent in, among others, humans, ferrets, European dogs, and avian species [8-10].

In species that generate NeuGc, its percentage of the total sialic acid content varies. For instance, pig trachea contains an equal amount of NeuAc and NeuGc, while 90% of the sialic acids on equine trachea and erythrocytes is NeuGc [11-14]. The high NeuGc content in horses may explain why equine H7N7 viruses are the only known

influenza A viruses that bind α2,3-linked NeuGc [15,16]. Highly pathogenic equine H7 viruses have not been isolated since 1978 and are, therefore, thought to be extinct [17,18]. Currently, horses are mainly infected by circulating H3N8 influenza viruses [19] and bind NeuAc instead. Unlike equine H7 strains, avian and human H7 viruses bind NeuAc [16]. It is still unclear where equine H7 viruses originated from and what the molecular determinants of NeuGc specificity are.

Here, we investigated receptor binding specificities of avian H7 HAs to identify the origin of equine H7 viruses. Inspired by the crystal structure of the equine H7 HA in complex with its ligand NeuGc, we performed targeted mutagenesis of avian H7 viruses. Several combinations of mutations were found that enabled avian H7 viruses to bind NeuGc. Our results demonstrate a phenotypical relationship between avian and equine H7 viruses despite their substantial genetic distance.

Results

Crystal structure of an equine H7 HA in complex with receptor analog 3′-GcLN and its similarity to an avian H7 HA

We previously reported the crystal structure of the HA of A/Equine/New York/49/73 H7N7 (H7eq) without a ligand (PDB: 6N5A [15]). To understand the structural basis for NeuGc specificity of H7eq, we determined the crystal structure of H7eq in complex with its natural ligand 3′-GcLN (NeuGcα2-3Galβ1-4GlcNAc) at 2.05 Å resolution (Fig 1A and S1 Table). The electron density for the 3′-GcLN ligand could be fitted well for all three monosaccharides (Fig 1B). H7eq binds 3′-GcLN mainly through NeuGc-1, but the interactions extend over the 220-loop with hydrogen bonds between Gal-2 and the main-chain carbonyl oxygen of G225 and between GlcNAc-3 and the side chain of Q222 (Fig 1A).

The RBS structures of H7eq and A/Turkey/Italy/214845/02 H7N3 (H7tu) (PDB: 4SBI, [20] appear to be very similar (Fig 1C-D), although the turkey strain binds NeuAc instead of NeuGc and was isolated almost 30 years after the equine strain. Nevertheless, 85% of HA1 residues are identical and the amino acid sequences around the RBS differ at 13 positions (Fig 1E and S1). The NeuGc-Gal bond of 3′-GcLN in the H7eq complex adopts a *cis* conformation, which is consistent with our previous findings in the structure of 3′-GcLN in complex with the A/Vietnam/1203/2004 H5N1 Y161A mutant that shifts receptor specificity from NeuAc to NeuGc [15]. On the contrary, the NeuAc-Gal bond in the avian receptor analog 3′-SLN (NeuAcα2-3Galβ1-4GlcNAc) in complex with H7tu adopts a *trans* NeuAc-Gal bond (Fig 1C-D).

In the H7eq 3′-GcLN structure, the 1-hydroxyl group of NeuGc-1 forms a hydrogen bond with the main-chain nitrogen of E135 and the E135 side chain forms a salt bridge with R144 (Fig 1A). The amino acid at position 193 is known to be an important determinant of receptor specificity [21-24]. In the crystal structure of H7eq and 3′-GcLN, R193 forms a hydrogen bond with NeuGc-1 (Fig 1A). In comparison, K193 in H7tu, with a shorter side-chain, is not in hydrogen bond distance with the NeuAc-1 of 3′-SLN (Fig 1C).

Despite the similarities in RBS structures, we found that H7tu binds solely to α2,3-linked NeuAc on the glycan array (Fig 2B), whereas H7eq exclusively binds to α2,3-linked NeuGc [15]. To decipher which residues determine NeuGc and NeuAc receptor specificity, targeted mutagenesis was performed on H7tu by replacing residues in the RBS with H7eq-like amino acids.

Amino acid 135 is essential for binding N-glycolylneuraminic acid

To identify which amino acids are critical for NeuGc binding, we mutated H7tu towards H7eq at eight different locations (128, 130, 135, 144, 159+160, 189, 193, and 219). Using the previously published glycan microarray containing glycans with terminal NeuAc or NeuGc (Fig 2A, [15]), we assessed the binding specificities of the mutants.

In the 130-loop, mutations S128T and I130V did not induce clear changes in the NeuAc/NeuGc specificity (Fig 2C-D). The amino acid at position 135 of H7 HAs is naturally diverse and has been associated with the adaptation of viruses between avian species and humans during a zoonotic outbreak of an H7N9 virus [25]. We observed that mutating position 135 (A135E) caused H7tu to acquire binding to NeuGc while maintaining binding to NeuAc (Fig 2E). Residue 143 has previously been

suggested to be relevant for NeuGc recognition in H3 viruses [26]. In H7eq, R144 forms salt bridges with the 130-loop residue E135, but mutation G144R alone in H7tu did not change HA binding specificity (Fig 2F). In the 150-loop, a highly conserved tyrosine is present at position 161 in all HA subtypes except H7, H10, H12, H15, H17, and H18 [22,27,28]. Previously, it was demonstrated that a Y161A mutation changes the binding properties of an H5 HA from NeuAc to NeuGc [15,27]. However, introducing Y161A in several other HA subtypes (H1, H2, and H4) did not change binding specificities (S2 Fig). We made mutations A159G and A160V simultaneously in H7tu, but unlike the Y161A mutation in H5, we did not observe NeuGc binding with this double mutation (Fig 2G). In the 190-helix, residue 189 is next to E190, which makes hydrogen bonds to the ligand, in both H7eq and H7tu (Fig 1A and 1C). Mutation T189A in H7tu did not change receptor specificity when introduced on its own (Fig 2H). Residue 193 is important for ligand recognition (Fig 1A) [21-24]. Introducing K193R in H7tu seems to abolish all binding to the glycan array (Fig 2I), even when illuminating the glycan array with higher laser powers. Despite residue 219 being very close to the 220-loop, mutation A219P did not change binding properties of H7tu from NeuAc to NeuGc (Fig 2J). In summary, while most mutations performed on H7tu did not affect binding specificity, the introduction of mutation K193R abolished glycan-binding and A135E seems to be key for binding NeuGc while maintaining binding to NeuAc.

Various combinations of mutations switch from NeuAc to NeuGc binding

Starting from the key mutation A135E, we continued mutagenesis by adding mutations at previously stated positions (Fig 3A). Mutating more amino acids in the 130-loop, at position 128 (S128T) or 130 (I130V), appears to abolish NeuAc binding while maintaining binding to NeuGc on the glycan microarray. The combination of mutation

A135E and mutation G144R, A159G+A160V, T189A, or A219P did not change binding specificity compared to mutation A135E solely since both NeuAc and NeuGc were still bound. Whereas almost all binding was abolished when introducing mutation K193R by itself (Fig 2I), adding mutation A135E restored binding to both NeuAc and NeuGc.

Since mutations A135E and K193R both affected the receptor binding properties, we further combined these two mutations with mutations that did not change binding specificity so far (Fig 3B). We found that adding mutation G144R or A159G+A160V abolished binding to the array. Adding the T189A mutation in the A135E+K193R background, switched H7tu to binding mainly NeuGc. The addition of mutation A219P did not affect the binding specificity since both NeuAc and NeuGc were still bound almost equally.

In short, we were able to modify H7tu for binding NeuGc specifically on the glycan microarray by combinations of mutations A135E+S128T, A135E+I130V, or A135E+T189A+K193R. These results show that residues in the 130-loop or 190-helix modify the specificity towards NeuGc.

No binding specificity to avian or equine erythrocytes and tracheal epithelium observed for avian H7 mutants that bind NeuGc in the glycan microarray

A glycan microarray, as used in this study, is a sophisticated tool to investigate the binding of proteins to synthetic glycans of which we know the exact structure. However, not all glycans can be present on the array and, therefore, it is necessary to investigate the binding specificities of HAs to host cells and tissues. Therefore, we performed a

hemagglutination assay with avian and equine erythrocytes and tissue staining on tracheal epithelium, which is the natural location of infection, of the same species.

While only binding NeuAc on the glycan array, the wild-type (WT) H7tu agglutinates both chicken erythrocytes, which contain only NeuAc [7,8], and horse erythrocytes, which contain mainly NeuGc [11,13,14] (Fig 4A). Therefore, a loss of binding to chicken erythrocytes would indicate a loss of NeuAc-binding. However, both types of erythrocytes were still bound by HAs with combinations of all investigated mutants (A135E, A135E+S128T, A135E+I130V, and A135E+T189A+K193R, Fig 4A) and therefore no conclusions concerning NeuGc specificity could be drawn from the hemagglutination assay. Similarly, WT and all mutants of H7tu bound both horse and chicken tracheal tissue (Fig 4C). The results demonstrate that there are some differences in NeuGc specificity between the glycan array, hemagglutination assay, and tissue staining.

NeuGc binding specificity can also be achieved in another avian H7 strain

To investigate whether the mutations that were found to switch H7tu, a virus from the Eurasian lineage, towards NeuGc binding are universal among H7 strains, we analyzed the HA of another avian strain from the North American lineage (A/Chicken/Jalisco/12283/2012 H7N3). Alignment of the HA sequences showed that the two strains differ at four amino acid positions (158, 188, 208, and 214) in the otherwise very similar RBS (Fig 5A). In glycan array analysis, similar to WT H7tu, the WT HA of A/Chicken/Jalisco/12283/12 H7N3 binds NeuAc (Fig 5B). Solely introducing mutation A135E acquires NeuGc binding and already seems to abolish some binding to NeuAc. Furthermore, NeuGc binding specificity on the glycan array was achieved

by combining mutation A135E with mutations I130V or T189A+K193R. A combination of A135E and S128T resulted in a loss of glycan binding on the array.

The combinations of mutations A135E, A135E+S128T, and A135E+I130V did not change the binding specificity of the HA in either the hemagglutination assay using chicken or horse erythrocytes (Fig 5C) or on tissue slides containing equine and avian trachea (Fig 5D), similar to observations in H7tu. The combination of mutations A135E+T189A+K193R did not change the binding specificity in the hemagglutination assay either, but binding to both chicken and horse tracheal tissue was lost. Nevertheless, based on the glycan array analysis, we conclude that distant avian H7 HAs from different lineages can acquire NeuGc binding through identical amino acid changes.

Equine and avian H7 strains are evolutionarily distant

The fact that avian H7 HA can be mutated towards binding NeuGc suggests that equine and avian H7 strains are phenotypically related. To investigate the genetic relationship between H7 strains, we reconstructed a maximum likelihood (ML) phylogenetic tree using HA sequences of equine H7 strains and the closest related Eurasian avian H7 strains (Fig 6A-E and S3). All equine strains cluster under a single monophyletic clade. Strains A/FPV/Dutch/1927 H7N7 and A/Fowl/Weybridge/1934 H7N7 appear to be the closest related avian strains to the equine viruses. We investigated the natural variation in amino acids at positions for which binding specificity changed (128, 130, 135, 189, and 193).

For each selected amino acid position, we annotated the ML tree based on the variation in residues (Fig 6A-E). The predicted most recent common ancestor (MRCA) at all positions contains avian-like amino acids. Of the five investigated amino acid positions, the highest variability in amino acids is present at key position 135, although we observe a clear distinction between a glutamic acid in the equine strains and a variation of alanine, valine, and threonine in the avian strains (Fig 6A). At position 128, there is an obvious distinction between the threonine in equine strains and mainly serine in the avian strains (Fig 6B). Again, a clear difference is observed at position 130 between avian strains (isoleucine) and equine strains (valine). Surprisingly, a closely related avian strain, A/Turkey/England/1963 H7N3, also contains a valine at position 130, just like the equine strains (Fig 6C). At position 189, all but one of the equine strains contain an alanine, while there is a mixture of mainly alanine and threonine in the avian strains (Fig 6D). At position 193, nearly all equine strains contain an arginine, whereas most avian strains, except for a small clade of viruses from chickens in Pakistan, contain a lysine (Fig 6E). We conclude that there is a clear distinction between the amino acids in the equine and avian strains, with the highest variability in residues being present at position 135, which we investigated further using targeted mutagenesis.

Four different amino acids (glutamic acid, alanine, valine, and threonine) are naturally present at position 135 of avian and equine H7 viruses (Fig 6A). When the alanine is encoded by either GCG or GCA, changing a single base pair will change the amino acid to glutamic acid, valine, or threonine. We introduced all of these residues at position 135 of H7tu to investigate whether acquirement of NeuGc binding is specific

268 for the glutamic acid. This was indeed the case, as the introduction of a threonine or

269 a valine at position 135 did not promote binding to NeuGc (Fig 6F).

Discussion

To elucidate the molecular determinants for NeuGc binding, we determined the crystal structure of the HA of A/Equine/New York/49/73 H7N7 in complex with its receptor analog 3′-GcLN (NeuGcα2-3Galβ1-4GlcNAc). The overall RBS structures of H7eq and A/Turkey/Italy/214845/02 H7N3 were shown to be similar. To examine the critical amino acids for NeuGc binding, we performed mutational analysis on two distant avian H7 HAs that bind NeuAc. Previously, we have demonstrated that HAs can bind either NeuAc or NeuGc [15]. Here, we demonstrate that HAs can bind both NeuAc and NeuGc, by the introduction of A135E in H7tu. The combination of mutations A135E+T189A+K193R, A135E+S128T, and A135E+I130V switched two avian HAs to bind NeuGc with some residual binding to α2,3-linked NeuAc, as determined on the glycan microarray.

We previously studied the NeuAc-specific HA of A/Vietnam/1203/2004 H5N1 (H5VN) and its Y161A mutant that is specific for NeuGc, which showed complete specificity on the glycan microarray, in the hemagglutination assay [15], and on tracheal epithelium tissue (Fig 4C). Similarly, we found that WT H1, H2, and H4 HAs bind chicken, but not horse, erythrocytes in the hemagglutination assay (S2 Fig). Contrarily, we observed that WT avian H7 HAs bound both chicken (NeuAc [7,8]) and horse (mainly NeuGc [11,13,14]) erythrocytes and tracheal tissue while binding specifically to NeuAc on the glycan microarray. This distinguishes these H7 viruses from other subtypes of influenza A. Possibly, the presence of NeuAc on horse erythrocytes and tracheal tissue, although estimated to be less than 10% of the total sialic acids [11,13,14], is sufficient to be bound by the WT avian H7 HAs. Additionally, the residual NeuAc-binding of the mutant avian H7 HAs may explain the binding to chicken

295 erythrocytes and tissue. Another important point is that not all compounds that are
296 present in the host are represented on the glycan microarray. Therefore, missing
297 glycans on the array may explain the binding of avian H7 HAs to horse and chicken
298 erythrocytes and tissue.
299
300 Besides the glycans currently present on the array (Fig 2A), glycans can be further
301 elongated, either symmetrically or asymmetrically, be tri- or tetra-antennary, and
302 contain one or multiple sialic acids. The addition of fucose at different positions on
303 LacNAc structures give rise to different Lewis antigens and additional sulfate or O-
304 acetyl groups can be present, adding another layer of complexity. Fucosylated (Lewis
305 X) and sulfated glycans are present in the human lung [29,30] and sulfated glycans
306 have been observed on porcine lungs [31]. For equine and avian species, the glycans
307 in the respiratory tract have not been studied in detail yet. One study describes the
308 presence of sialyl Lewis X structures in the respiratory tract of chickens [32]. Little is
309 known about the glycans present on erythrocytes of different species, apart from two
310 studies that describe that very few glycans with fucoses are present on chicken and
311 mouse erythrocytes [33,34]. Influenza A viruses of different subtypes (H1, H3, H4, H5,
312 H6, H7, H9, H13, H14) are known that (specifically) bind, or do not bind at all, to
313 fucosylated and sulfated glycans [35-42]. Most relevant, avian, human, and seal H7
314 HAs also prefer to bind sulfated sialyl Lewis X structures [16,36,39]. In conclusion, fine
315 receptor binding specificity regarding fucosylation and sulfation is observed in many
316 different influenza A viruses and it could well be that the viruses used in this study also
317 have additional receptor binding preferences, besides the distinction between NeuAc
318 and NeuGc.
319

320 It has been suggested that recognition of NeuGc by Influenza A viruses is essential
321 for viral replication in horses [11]. The most prevalent influenza viruses among horses
322 currently and in the past are H3N8 and H7N7 viruses [19]. While equine H7 viruses
323 have been shown to prefer binding to NeuGc [15,16], equine H3 viruses bind to NeuAc
324 [16]. H3N8 viruses from horses often infect dogs [19,43], which are not able to make
325 glycans containing NeuGc due to lack of a functional CMAH. Thus, it appears to be
326 advantageous for H3 viruses to maintain NeuAc specificity while circulating in horses.
327
328 Equine and avian H7 strains are estimated, by phylogenetic analysis, to have
329 separated in the mid to late 1800s [44]. Nevertheless, we demonstrated that avian H7
330 HAs, although genetically distinct from equine H7 viruses, are able to bind NeuGc after
331 the introduction of two to three mutations, suggesting that these viruses are
332 phenotypically related. The mutations converting to NeuGc binding were shown to be
333 universal in avian H7 influenza A viruses from both the Eurasian and Northern
334 American lineage. It would be interesting to investigate whether these same mutations
335 induce NeuGc binding in other subtypes of group 2 influenza A viruses as well.
336

Material and Methods

Expression, crystallization, and structural determination of the equine H7 HA in complex with receptor analog 3′-GcLN

The HA ectodomain of A/equine/NY/49/73/H7N7 (GenBank ID LC414434) was cloned and expressed as described previously [15]. Briefly, cDNA corresponding to residues 11 to 327 of HA1 and 1 to 179 of HA2 (H3 numbering) was cloned into a pFastbac vector. The HA was expressed in Hi5 insect cells as described [45], after which it was purified, the trimerization domain and His_6-tag were removed, and the HA was concentrated to 6 mg/ml.

Crystals of the H7eq HA were obtained at 20°C using the vapor diffusion sitting drop method against a reservoir solution containing 32% (w/v) polyethylene glycol 400 and 0.1 M CAPS at pH 10. The complex of HA protein with 3′-GcLN was obtained by soaking HA crystals in a reservoir that contained 3′-GcLN to a final concentration of 10 mM for 1 hour at 20°C. The crystals were flash cooled in liquid nitrogen, without additional cryoprotectant, before data collection at the Advanced Photon Source (APS) (S1 Table). Data integration and scaling were performed using HKL2000 [46]. The molecular replacement method using Phaser [47] was used to solve the H7eq complex structure, for which an apo H7eq HA structure (PDB: 6N5A) was used as the search model. REFMAC5 [48] was used for structure refinement and modeling was done with COOT [49]. The final refinement statistics are outlined in S1 Table.

Expression and purification of HA for binding studies

HA encoding cDNAs of A/Turkey/Italy/214845/02 H7N3 [20], synthesized by GenScript, and A/Chicken/Jalisco/12283/12 H7N3 (a kind gift from Florian Krammer, Mt Sinai Medical School) were cloned into the pCD5 expression vector as described

previously [50,51]. The pCD5 expression vector was adapted to clone the HA-encoding cDNAs in frame with DNA sequences coding for a secretion signal sequence, the Twin-Strep (WSHPQFEKGGGSGGGSWSHPQFEK); IBA, Germany), a GCN4 trimerization domain (RMKQIEDKIEEIESKQKKIENEIARIKK), and a superfolder GFP [52] or mOrange2 [53]. Mutations in HAs were generated by site-directed mutagenesis. The HAs were purified from cell culture supernatants after expression in HEK293S GnTI(-) cells as described previously [50]. In short, transfection was performed using the pCD5 expression vectors and polyethyleneimine I. The transfection mixtures were replaced at 6 h post-transfection by 293 SFM II expression medium (Gibco), supplemented with sodium bicarbonate (3.7 g/L), Primatone RL-UF (3.0 g/L), glucose (2.0 g/L), glutaMAX (Gibco), valproic acid (0.4 g/L) and DMSO (1.5%). At 5 to 6 days after transfection, tissue culture supernatants were collected and Strep-Tactin sepharose beads (IBA, Germany) were used to purify the HA proteins according to the manufacturer's instructions.

Glycan microarray binding of HA

The glycan microarray as earlier presented [15] was utilized. HAs were pre-complexed with mouse anti-streptag-HRP and goat anti-mouse-Alexa555 antibodies in a 4:2:1 molar ratio respectively in 50 µL PBS with 0.1% Tween-20. The mixture was incubated on ice for 15 min and afterward incubated on the surface of the array for 90 min in a humidified chamber. Then, slides were rinsed successively with PBS-T (0.1% Tween-20), PBS, and deionized water. The arrays were dried by centrifugation and immediately scanned as described previously [15]. Processing of the six replicates was performed by removing the highest and lowest replicate and subsequently calculating the mean value and standard deviation over the four remaining replicates.

388 **Hemagglutination assay**

389 Hemagglutination assays were performed with pre-complexed HAs, as described for
390 the glycan microarray, on 1.0% erythrocytes as previously described [50] with a
391 starting concentration of 10 µg/ml of HA. Erythrocytes were provided by the
392 Department of Equine Sciences and the Department of Farm Animal Health of the
393 Faculty of Veterinary Medicine, Utrecht University, the Netherlands. The blood was
394 taken from adult animals that are in the educational program of the Faculty of
395 Veterinary Medicine.

396

397 **Protein histochemistry**

398 Sections of formalin-fixed, paraffin-embedded chicken (*Gallus gallus domesticus*) and
399 equine (*Equus ferus caballus*) trachea were obtained from the Department of
400 Veterinary Pathobiology, Faculty of Veterinary Medicine, Utrecht University, the
401 Netherlands. Protein histochemistry was performed as previously described [54,55].
402 In short, tissue sections of 4 µm were deparaffinized and rehydrated, after which
403 antigens were retrieved by heating the slides in 10 mM sodium citrate (pH 6.0) for 10
404 min. Endogenous peroxidase was inactivated using 1% hydrogen peroxide in MeOH
405 for 30 min. Tissues were blocked overnight at 4°C using 3% BSA (w/v) in PBS with
406 0.1% Tween-20 and subsequently stained using 5 µg/ml pre-complexed HAs, as
407 previously described for the glycan microarray. After washing in PBS, binding was
408 visualized using 3-amino-9-ethylcarbazole (AEC) (Sigma-Aldrich, Germany) and
409 slides were counterstained using hematoxylin.

410

411 **Phylogenetic trees**

412 All available, high-quality HA nucleotide sequences (i.e. sequence length is >90% of
413 full-length HA gene segment and has <1% of ambiguous base) of avian H7Nx and

414 equine H7N7 influenza viruses dated between 1905 and 2005 from the NCBI Genbank
415 database were downloaded (N=944). The maximum-likelihood phylogenetic tree was
416 reconstructed using IQ-TREE [56] using the optimal nucleotide substitution model (i.e.
417 GTR+F+R3) based on the Bayesian Information Criterion as determined by
418 ModelFinder [57]. Ancestral sequences were reconstructed using treetime [58].

Data deposition

The atomic coordinates and structure factors of the HA of A/equine/NY/49/73/H7N7 in complex with 3′-GcLN are being deposited in the Protein Data Bank (PDB) under accession codes xxxx.

Acknowledgments

We would like to thank the Department of Equine Sciences and the Department of Farm Animal Health of Utrecht University for supplying erythrocytes. We thank Andrea Gröne and Hélène Verheije from the Department of Veterinary Pathobiology of Utrecht University for providing paraffin-embedded tissues.

Funding

R.P.dV. is a recipient of an ERC Starting Grant from the European Commission (802780) and a Beijerinck Premium of the Royal Dutch Academy of Sciences. C.A.R. and A.X.H. are supported by an ERC Consolidator Grant from the European Commission (818353). Synthesis and microarray analysis were funded by a grant from the Netherlands Organization for Scientific Research (NWO TOPPUNT 718.015.003) to G.-J.B.. This work was funded in part by the Bill and Melinda Gates Foundation (OPP1170236) to I.A.W. X-ray data were collected at the beamline 23ID-D 9GM/CA CAT). The use of the APS was supported by the U.S. Department of Energy (DOE), Basic Energy Sciences, Office of Science, under contract DE-AC02-06CH11357.

References

1. Skehel JJ, Wiley DC (2000) Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. Annu Rev Biochem 69: 531-569.
2. Long JS, Mistry B, Haslam SM, Barclay WS (2019) Host and viral determinants of influenza A virus species specificity. Nat Rev Microbiol 17: 67-81.
3. Ji Y, White YJ, Hadden JA, Grant OC, Woods RJ (2017) New insights into influenza A specificity: an evolution of paradigms. Curr Opin Struct Biol 44: 219-231.
4. de Vries RP, Zhu X, McBride R, Rigter A, Hanson A, et al. (2014) Hemagglutinin receptor specificity and structural analyses of respiratory droplet-transmissible H5N1 viruses. J Virol 88: 768-773.
5. de Vries RP, de Vries E, Moore KS, Rigter A, Rottier PJ, et al. (2011) Only two residues are responsible for the dramatic difference in receptor binding between swine and new pandemic H1 hemagglutinin. J Biol Chem 286: 5868-5875.
6. de Vries RP, Tzarum N, Peng W, Thompson AJ, Ambepitiya Wickramasinghe IN, et al. (2017) A single mutation in Taiwanese H6N1 influenza hemagglutinin switches binding to human-type receptors. EMBO Mol Med 9: 1314-1325.
7. Peri S, Kulkarni A, Feyertag F, Berninsone PM, Alvarez-Ponce D (2018) Phylogenetic distribution of CMP-Neu5Ac hydroxylase (CMAH), the enzyme synthetizing the proinflammatory human xenoantigen Neu5Gc. Genome Biol Evol 10: 207-219.
8. Schauer R, Srinivasan GV, Coddeville B, Zanetta JP, Guerardel Y (2009) Low incidence of N-glycolylneuraminic acid in birds and reptiles and its absence in the platypus. Carbohydr Res 344: 1494-1500.
9. Ng PS, Bohm R, Hartley-Tassell LE, Steen JA, Wang H, et al. (2014) Ferrets exclusively synthesize Neu5Ac and express naturally humanized influenza A virus receptors. Nat Commun 5: 5750.
10. Yasue S, Handa S, Miyagawa S, Inoue J, Hasegawa A, et al. (1978) Difference in form of sialic acid in red blood cell glycolipids of different breeds of dogs. J Biochem 83: 1101-1107.
11. Suzuki Y, Ito T, Suzuki T, Holland RE, Jr., Chambers TM, et al. (2000) Sialic acid species as a determinant of the host range of influenza A viruses. J Virol 74: 11825-11831.
12. Suzuki T, Horiike G, Yamazaki Y, Kawabe K, Masuda H, et al. (1997) Swine influenza virus strains recognize sialylsugar chains containing the molecular species of sialic acid predominantly present in the swine tracheal epithelium. FEBS Lett 404: 192-196.
13. Suzuki Y, Matsunaga M, Matsumoto M (1985) N-acetylneuraminyllactosylceramide, GM3-NeuAc, a new influenza A virus receptor which mediates the adsorption-fusion process of viral infection. Binding specificity of influenza virus A/Aichi/2/68 (H3N2) to membrane-associated GM3 with different molecular species of sialic acid. J Biol Chem 260: 1362-1365.
14. Barnard KN, Alford-Lawrence BK, Buchholz DW, Wasik BR, LaClair JR, et al. (2020) Modified sialic acids on mucus and erythrocytes inhibit influenza A virus hemagglutinin and neuraminidase functions. J Virol 94.
15. Broszeit F, Tzarum N, Zhu X, Nemanichvili N, Eggink D, et al. (2019) N-glycolylneuraminic acid as a receptor for influenza A viruses. Cell Rep 27: 3284-3294 e3286.

16. Gambaryan AS, Matrosovich TY, Philipp J, Munster VJ, Fouchier RA, et al. (2012) Receptor-binding profiles of H7 subtype influenza viruses in different host species. J Virol 86: 4370-4379.
17. Webster RG, Bean WJ, Gorman OT, Chambers TM, Kawaoka Y (1992) Evolution and ecology of influenza A viruses. Microbiol Rev 56: 152-179.
18. Webster RG (1993) Are equine 1 influenza viruses still present in horses? Equine Vet J 25: 537-538.
19. Murcia PR, Wood JL, Holmes EC (2011) Genome-scale evolution and phylodynamics of equine H3N8 influenza A virus. J Virol 85: 5312-5322.
20. Russell RJ, Gamblin SJ, Haire LF, Stevens DJ, Xiao B, et al. (2004) H1 and H7 influenza haemagglutinin structures extend a structural classification of haemagglutinin subtypes. Virology 325: 287-296.
21. de Vries RP, Peng W, Grant OC, Thompson AJ, Zhu X, et al. (2017) Three mutations switch H7N9 influenza to human-type receptor specificity. PLoS Pathog 13: e1006390.
22. Tzarum N, de Vries RP, Peng W, Thompson AJ, Bouwman KM, et al. (2017) The 150-loop restricts the host specificity of human H10N8 influenza virus. Cell Rep 19: 235-245.
23. Peng W, Bouwman KM, McBride R, Grant OC, Woods RJ, et al. (2018) Enhanced human-type receptor binding by ferret-transmissible H5N1 with a K193T mutation. J Virol 92: e02016-02017.
24. Medeiros R, Naffakh N, Manuguerra JC, van der Werf S (2004) Binding of the hemagglutinin from human or equine influenza H3 viruses to the receptor is altered by substitutions at residue 193. Arch Virol 149: 1663-1671.
25. Watanabe T, Kiso M, Fukuyama S, Nakajima N, Imai M, et al. (2013) Characterization of H7N9 influenza A viruses isolated from humans. Nature 501: 551-555.
26. Masuda H, Suzuki T, Sugiyama Y, Horiike G, Murakami K, et al. (1999) Substitution of amino acid residue in influenza A virus hemagglutinin affects recognition of sialyl-oligosaccharides containing N-glycolylneuraminic acid. FEBS Lett 464: 71-74.
27. Wang M, Tscherne DM, McCullough C, Caffrey M, Garcia-Sastre A, et al. (2012) Residue Y161 of influenza virus hemagglutinin is involved in viral recognition of sialylated complexes from different hosts. J Virol 86: 4455-4462.
28. Wen F, Li L, Zhao N, Chiang MJ, Xie H, et al. (2018) A Y161F hemagglutinin substitution increases thermostability and improves yields of 2009 H1N1 influenza A virus in cells. J Virol 92.
29. Jia N, Byrd-Leotis L, Matsumoto Y, Gao C, Wein AN, et al. (2020) The human lung glycome reveals novel glycan ligands for influenza A virus. Sci Rep 10: 5320.
30. Sriwilaijaroen N, Nakakita SI, Kondo S, Yagi H, Kato K, et al. (2018) N-glycan structures of human alveoli provide insight into influenza A virus infection and pathogenesis. FEBS J 285: 1611-1634.
31. Byrd-Leotis L, Liu R, Bradley KC, Lasanajak Y, Cummings SF, et al. (2014) Shotgun glycomics of pig lung identifies natural endogenous receptors for influenza viruses. Proc Natl Acad Sci U S A 111: E2241-2250.
32. Hiono T, Okamatsu M, Nishihara S, Takase-Yoden S, Sakoda Y, et al. (2014) A chicken influenza virus recognizes fucosylated alpha2,3 sialoglycan receptors on the epithelial cells lining upper respiratory tracts of chickens. Virology 456-457: 131-138.

538 33. Hua S, Jeong HN, Dimapasoc LM, Kang I, Han C, et al. (2013) Isomer-specific
539 LC/MS and LC/MS/MS profiling of the mouse serum N-glycome revealing a
540 number of novel sialylated N-glycans. Anal Chem 85: 4636-4643.
541 34. Aich U, Beckley N, Shriver Z, Raman R, Viswanathan K, et al. (2011) Glycomics-
542 based analysis of chicken red blood cells provides insight into the selectivity of
543 the viral agglutination assay. FEBS J 278: 1699-1712.
544 35. Gambaryan A, Tuzikov A, Pazynina G, Bovin N, Balish A, et al. (2006) Evolution
545 of the receptor binding phenotype of influenza A (H5) viruses. Virology 344:
546 432-438.
547 36. Gambaryan AS, Tuzikov AB, Pazynina GV, Desheva JA, Bovin NV, et al. (2008)
548 6-sulfo sialyl Lewis X is the common receptor determinant recognized by H5,
549 H6, H7 and H9 influenza viruses of terrestrial poultry. Virol J 5: 85.
550 37. Bateman AC, Karamanska R, Busch MG, Dell A, Olsen CW, et al. (2010) Glycan
551 analysis and influenza A virus infection of primary swine respiratory epithelial
552 cells: the importance of NeuAc{alpha}2-6 glycans. J Biol Chem 285: 34016-
553 34026.
554 38. Stevens J, Chen LM, Carney PJ, Garten R, Foust A, et al. (2010) Receptor
555 specificity of influenza A H3N2 viruses isolated in mammalian cells and
556 embryonated chicken eggs. J Virol 84: 8287-8299.
557 39. Gambaryan A, Yamnikova S, Lvov D, Tuzikov A, Chinarev A, et al. (2005)
558 Receptor specificity of influenza viruses from birds and mammals: new data on
559 involvement of the inner fragments of the carbohydrate chain. Virology 334:
560 276-283.
561 40. Stevens J, Blixt O, Glaser L, Taubenberger JK, Palese P, et al. (2006) Glycan
562 microarray analysis of the hemagglutinins from modern and pandemic influenza
563 viruses reveals different receptor specificities. J Mol Biol 355: 1143-1155.
564 41. Hiono T, Okamatsu M, Igarashi M, McBride R, de Vries RP, et al. (2016) Amino
565 acid residues at positions 222 and 227 of the hemagglutinin together with the
566 neuraminidase determine binding of H5 avian influenza viruses to sialyl Lewis
567 X. Arch Virol 161: 307-316.
568 42. Wen F, Blackmon S, Olivier AK, Li L, Guan M, et al. (2018) Mutation W222L at the
569 receptor binding site of hemagglutinin could facilitate viral adaption from equine
570 influenza A(H3N8) virus to dogs. J Virol 92.
571 43. Collins PJ, Vachieri SG, Haire LF, Ogrodowicz RW, Martin SR, et al. (2014) Recent
572 evolution of equine influenza and the origin of canine influenza. Proc Natl Acad
573 Sci U S A 111: 11175-11180.
574 44. Worobey M, Han GZ, Rambaut A (2014) A synchronized global sweep of the
575 internal genes of modern avian influenza virus. Nature 508: 254-257.
576 45. Stevens J, Blixt O, Paulson JC, Wilson IA (2006) Glycan microarray technologies:
577 tools to survey host specificity of influenza viruses. Nat Rev Microbiol 4: 857-
578 864.
579 46. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in
580 oscillation mode. Methods Enzymol 276: 307-326.
581 47. McCoy AJ, Grosse-Kunstleve RW, Storoni LC, Read RJ (2005) Likelihood-
582 enhanced fast translation functions. Acta Crystallogr D Biol Crystallogr 61: 458-
583 464.
584 48. Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, et al. (2011)
585 REFMAC5 for the refinement of macromolecular crystal structures. Acta
586 Crystallogr D Biol Crystallogr 67: 355-367.

587 49. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics.
588 Acta Crystallogr D Biol Crystallogr 60: 2126-2132.
589 50. de Vries RP, de Vries E, Bosch BJ, de Groot RJ, Rottier PJ, et al. (2010) The
590 influenza A virus hemagglutinin glycosylation state affects receptor-binding
591 specificity. Virology 403: 17-25.
592 51. Zeng Q, Langereis MA, van Vliet AL, Huizinga EG, de Groot RJ (2008) Structure
593 of coronavirus hemagglutinin-esterase offers insight into corona and influenza
594 virus evolution. Proc Natl Acad Sci U S A 105: 9065-9069.
595 52. Nemanichvili N, Tomris I, Turner HL, McBride R, Grant OC, et al. (2019)
596 Fluorescent trimeric hemagglutinins reveal multivalent receptor binding
597 properties. J Mol Biol 431: 842-856.
598 53. Shaner NC, Lin MZ, McKeown MR, Steinbach PA, Hazelwood KL, et al. (2008)
599 Improving the photostability of bright monomeric orange and red fluorescent
600 proteins. Nat Methods 5: 545-551.
601 54. Bouwman KM, Parsons LM, Berends AJ, de Vries RP, Cipollo JF, et al. (2020)
602 Three amino acid changes in avian coronavirus spike protein allow binding to
603 kidney tissue. J Virol 94.
604 55. Wickramasinghe IN, de Vries RP, Grone A, de Haan CA, Verheije MH (2011)
605 Binding of avian coronavirus spike proteins to host factors reflects virus tropism
606 and pathogenicity. J Virol 85: 8903-8912.
607 56. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, et al. (2020)
608 IQ-TREE 2: New models and efficient methods for phylogenetic inference in
609 the genomic era. Mol Biol Evol.
610 57. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS (2017)
611 ModelFinder: fast model selection for accurate phylogenetic estimates. Nat
612 Methods 14: 587-589.
613 58. Sagulenko P, Puller V, Neher RA (2018) TreeTime: Maximum-likelihood
614 phylodynamic analysis. Virus Evol 4: vex042.
615 59. Chen VB, Arendall WB, 3rd, Headd JJ, Keedy DA, Immormino RM, et al. (2010)
616 MolProbity: all-atom structure validation for macromolecular crystallography.
617 Acta Crystallogr D Biol Crystallogr 66: 12-21.
618

619

Figures

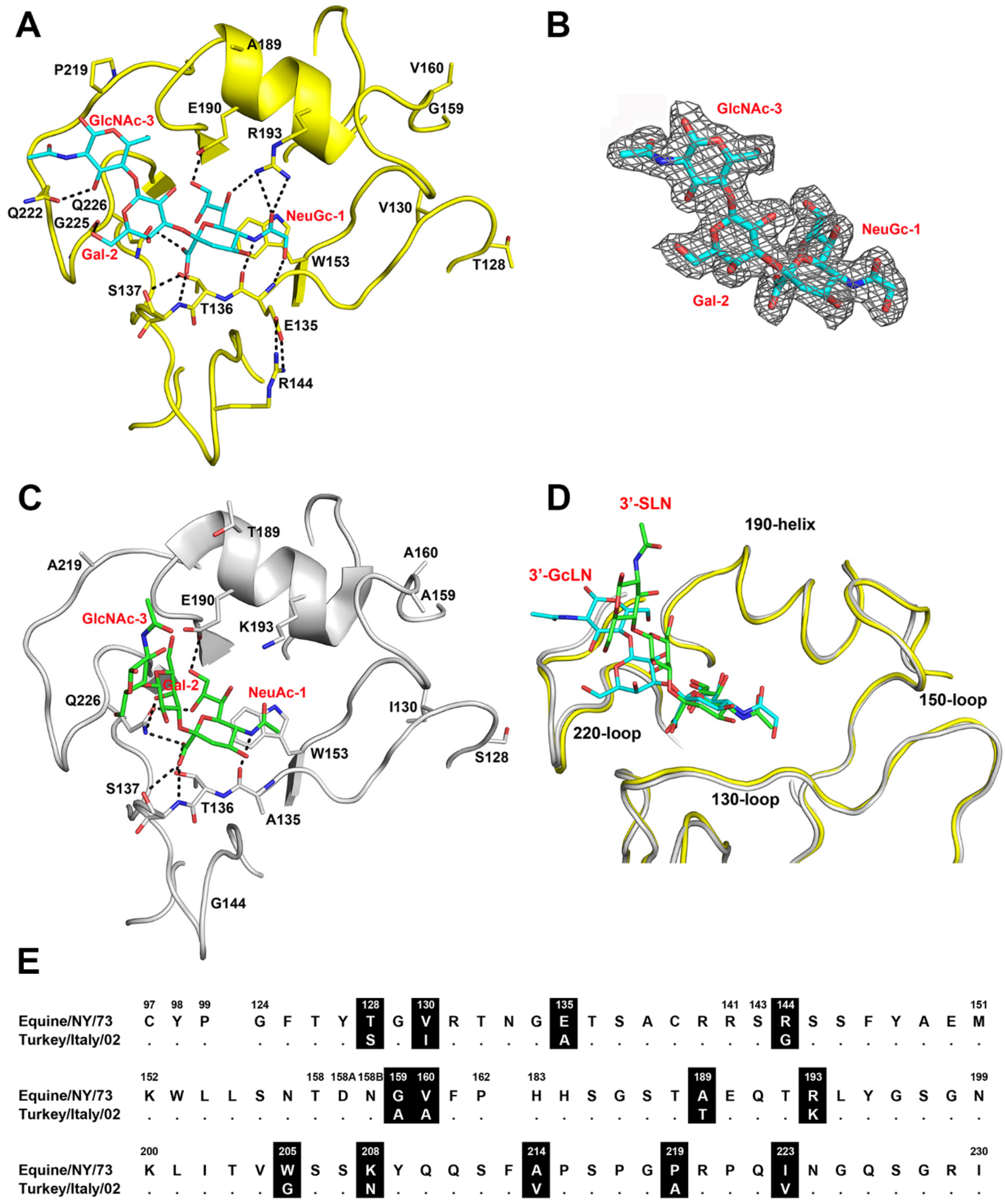
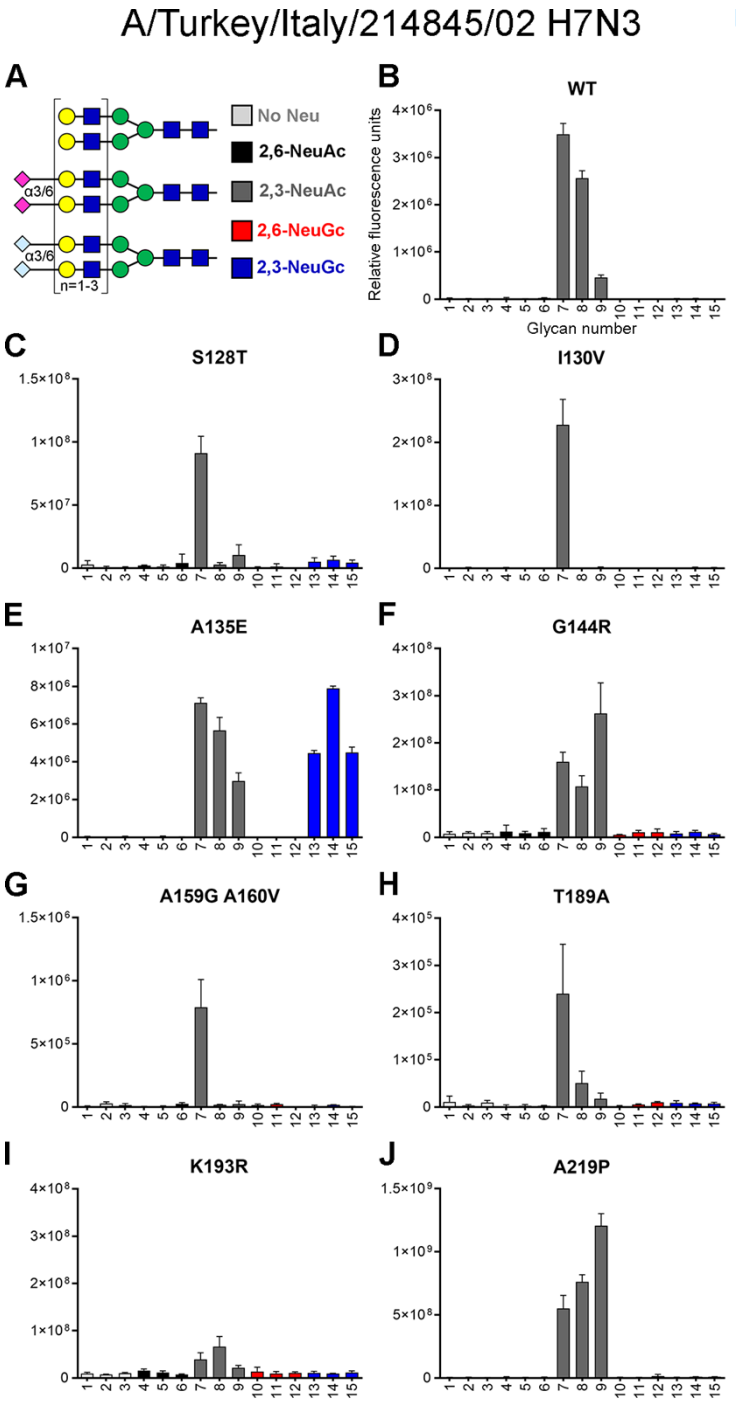


Fig 1. Comparison of the HA of A/Equine/New York/43/73 H7N7 and A/Turkey/Italy/214845/03 H7N3.

(**A**) RBS structure of H7eq (yellow) in complex with 3′-GcLN (NeuGcα2-3Galβ1-4GlcNAc, cyan). (**B**) Electron density 2Fo-Fc map at 1σ level for receptor analog 3′-

626 GcLN. (**C**) RBS structure of H7tu (grey) in complex with 3′-SLN (NeuAcα2-3Galβ1-
627 4GlcNAc, green) (PDB: 4BSI). (**D**) Superimposition of the RBS structures of H7eq and
628 H7tu and their ligands 3′-GcLN and 3′-SLN. The coloring scheme is following A and
629 C. (**E**) Alignment of the RBS residues of H7eq and H7tu, with amino acid positions (H3
630 numbering) indicated above the alignment, non-conserved residues highlighted in
631 black, and dots indicating identical amino acids.
632

633



634

635 **Fig 2. Evaluation of the binding specificities of single mutants of the HA of**

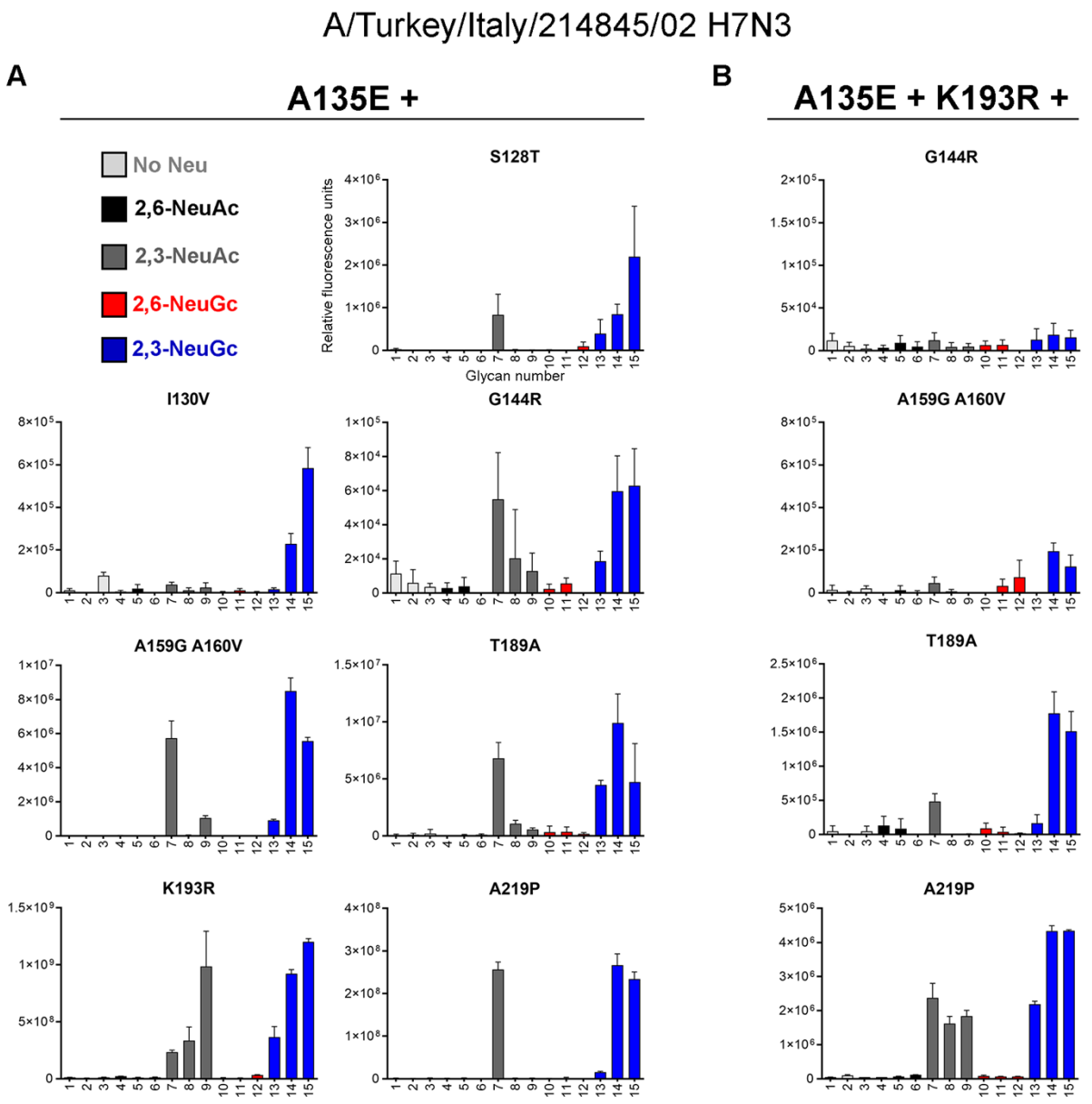
636 **A/Turkey/Italy/214845/03 H7N3.**

637 (**A**) Synthetic glycans printed on the microarray (n=6), either without sialic acid

638 (structures 1-3, light gray), with α2,6-linked NeuAc (4-6, black), α2,3-linked NeuAc (7-9,

639 dark gray), α2,6-linked NeuGc (10-12, red) or α2,3-linked NeuGc (13-15, blue).

Structures 1, 4, 7, 10, and 13 contain one LacNAc repeat, while structures 2, 5, 8, 11, and 14 have two repeats and structures 3, 6, 9, 12, and 15 contain three repeats [15]. The glycan microarray, representative for two independent assays, was used to determine the receptor specificity of (**B**) H7tu wild-type (WT), (**C**) mutant S128T, (**D**) I130V, (**E**) A135E, (**F**) G144R, (**G**) A159G+A160V, (**H**) T189A, (**I**) K193R, and (**J**) A219P.



646

647 **Fig 3. Evaluation of the binding specificities of double and triple mutants of the**

648 **HA of A/Turkey/Italy/214845/03 H7N3.**

649 The glycan microarray as described in Fig 2A was used, containing glycans with

650 terminal NeuAc or NeuGc or without sialic acid. Representative binding specificities

651 for two independent assays for (**A**) mutant HAs containing mutation A135E and an

652 additional mutation (S128T, I130V, G144R, A159G+A160V, T189A, K193, or A219P)

653 and (**B**) mutant HAs containing mutations A135E, K193R, and an additional mutation

654 (G144R, A159G+A160V, T189A, or A219P) are shown.

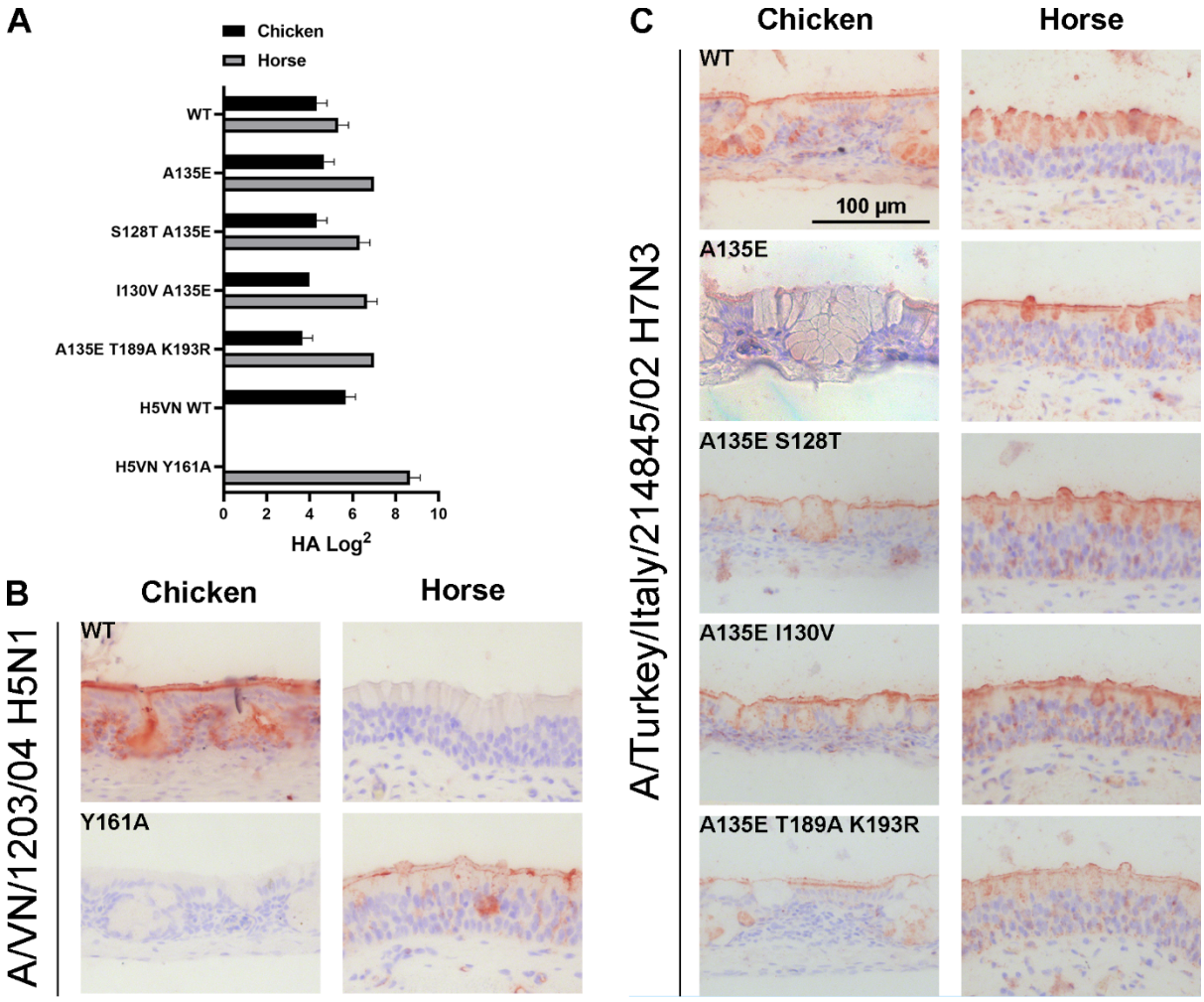


Fig 4. Binding specificities of (mutant) HA of A/Turkey/Italy/214845/03 H7N3 to chicken and horse erythrocytes and tracheal epithelium.

(**A**) A hemagglutination assay (n=3, mean and SD shown) with chicken and horse erythrocytes was performed using H7tu WT and mutant HAs (A135E, A135E+S128T, A135E+I130V, and A135E+T189A+K193R). AEC staining is used to visualize tissue binding and images are representative for three independent assays. Tissue staining of chicken and horse tracheal epithelium is performed with (**B**) WT and Y161A mutant HA of A/Vietnam/1203/2004 H5N1 as a positive and negative control and (**C**) H7tu WT or mutant HAs as in panel A.

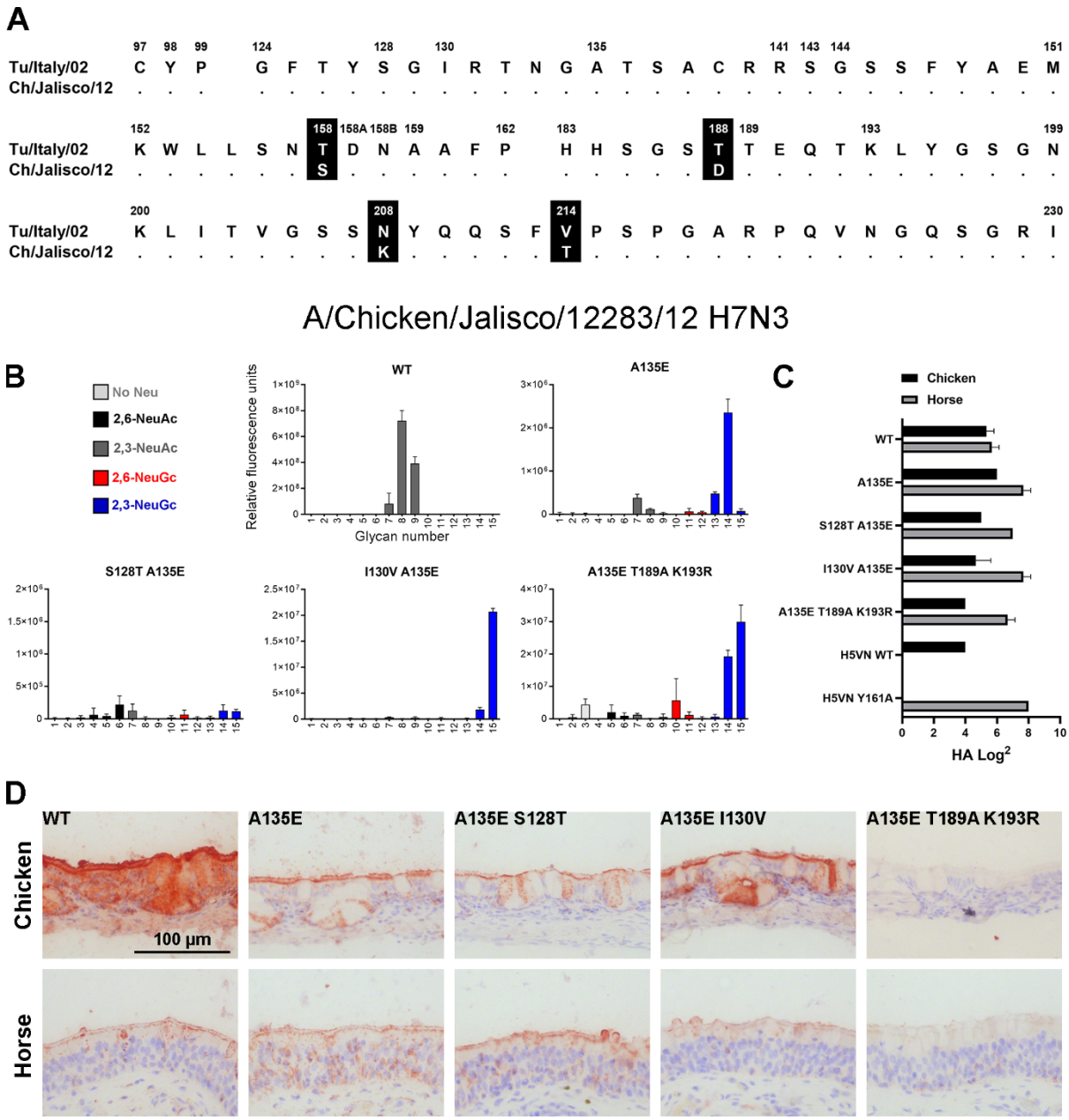
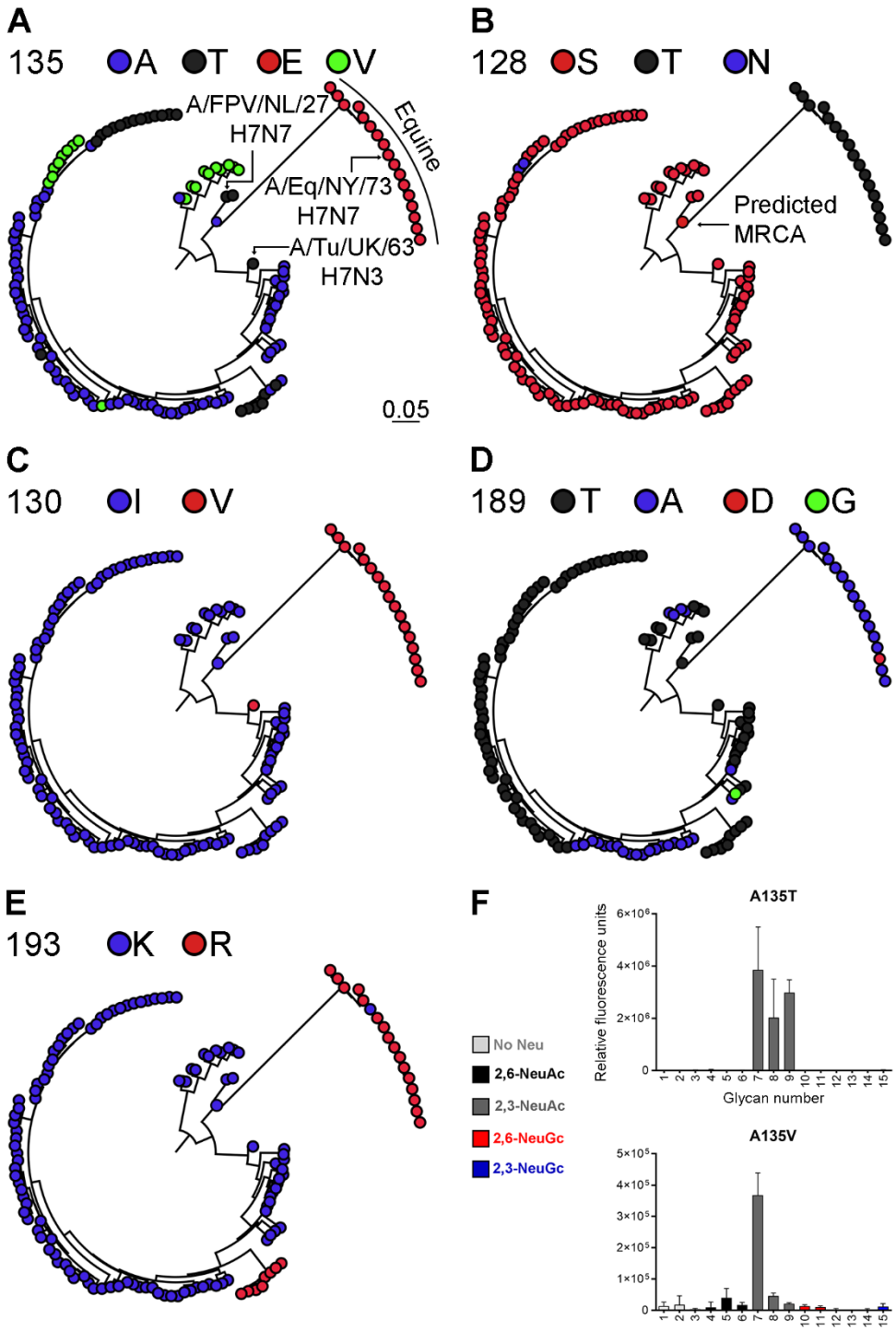


Fig 5. Evaluation of the binding specificities of (mutant) HA of A/Chicken/Jalisco/12283/12 H7N3.

(**A**) Alignment of the RBS of the HAs of A/Turkey/Italy/214845/03 H7N3 and A/Chicken/Jalisco/12283/12 H7N3, with amino acid positions indicated above the alignment and dots indicating identical amino acids. (**B**) The binding specificities of WT HA of A/Chicken/Jalisco/12283/12 H7N3 and mutant A135E, A135E+S128T, A135E+I130V, and A135E+T189A+K193R were evaluated on the glycan microarray as described in Fig 2A. (**C**) Binding specificities of WT and mutant HAs were

674 furthermore tested in a hemagglutination assay on chicken and horse erythrocytes
675 (n=3, mean and SD shown). (**D**) Binding of the WT and mutant HAs to chicken and
676 horse tracheal epithelium (controls shown in Fig 4B) is visualized using AEC staining.
677



678

679 **Fig 6. Phylogenetic tree of equine and avian H7 HA sequences and evaluation**
680 **of the binding specificities of mutants of the HA of A/Turkey/Italy/214845/03**
681 **H7N3 at amino acid position 135.**

682 Phylogenetic trees of equine and avian H7 influenza A strains from the Eurasian
683 lineage were reconstructed. The equine H7 strains cluster as a single monophyletic

684 clade. The avian strains that are closest related to the equine strains
685 (A/FPV/Dutch/1927 H7N7 and A/Turkey/England/1963 H7N3) are indicated, as well
686 as A/Equine/New York/43/73 H7N7. The annotated phylogenetic tree with all strain
687 names is displayed in S3 Fig. The variation in amino acids at positions (**A**) 135 (alanine,
688 threonine, glutamic acid, valine), (**B**) 128 (serine, threonine, asparagine), (**C**) 130
689 (isoleucine, valine), (**D**) 189 (threonine, alanine, aspartic acid, glycine) and (**E**) 193
690 (lysine, arginine) is shown. For all positions, the amino acid of the predicted most
691 recent common ancestor (MRCA) is shown. (**F**) Representative binding specificities
692 on the glycan microarray (as in Fig 2A) for H7tu mutants A135T and A135V are shown.

Supplementary figures

S1 Table. Data collection and refinement statistics of H7eq in complex with 3′-GcLN

Data Collection	
X-ray source	APS 23ID-D
Space group	P3
Unit cell (Å)	$a = b = 112.8$, $c = 130.2$
Resolution (Å)[a]	48.87-2.05 (2.09-2.05)
Unique reflections	116,095 (5,442)
Redundancy[a]	9.4 (5.2)
Average I/σ(I)[a]	16.1 (1.0)
Completeness[a]	99.5 (93.2)
R_{sym}[a,b]	0.10 (0.77)
R_{pim}[a,b]	0.03 (0.34)
$CC_{1/2}$[a]	0.99 (0.70)
No. molecules per ASU[c]	3
Refinement	
Resolution (Å)[a]	48.87-2.05
Reflections in refinement	110,267
Refined residues	1,455
Refined waters	688
R_{cryst}[d]	0.224
R_{free}[e]	0.251
B-values (Å^2)	
Protein	70
RBS subdomain (Residues 117-265 of HA1) of chain A, C, E	33, 40, 64
Ligand of chain A, C, E	30, 53, 100
Waters	48
Wilson *B*-values (Å^2)	28
Ramachandran values (%)[f]	96.2, 0
r.m.s.d. bond (Å)	0.005
r.m.s.d. angle (deg.)	1.36
PDB code	XXXX

[a] Parentheses denote outer-shell statistics.

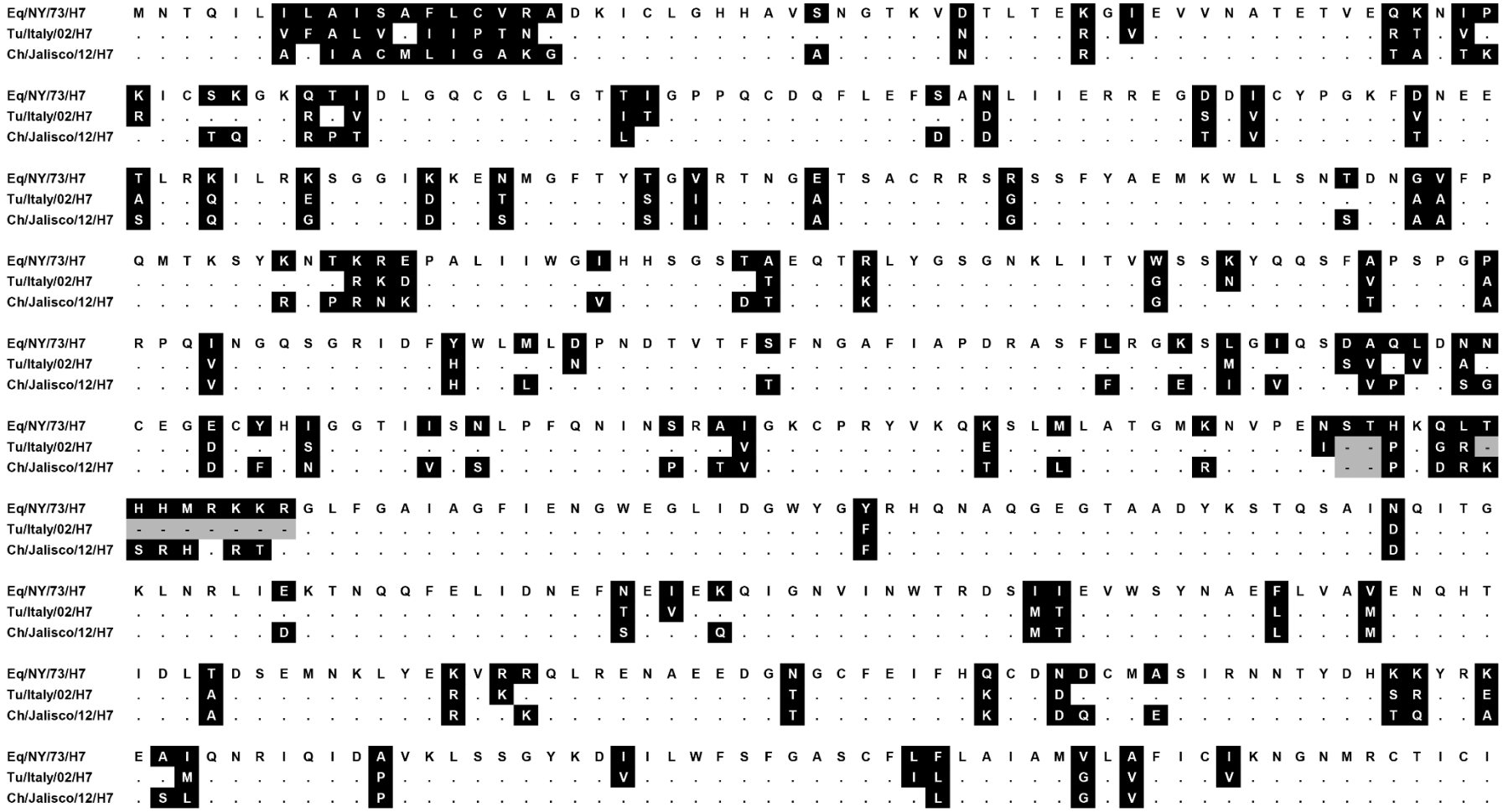
[b] $R_{sym} = \sum_{hkl}\sum_i |I_{hkl,i} - \langle I_{hkl}\rangle| / \sum_{hkl}\sum_i I_{hkl,i}$ and $R_{pim} = \sum_{hkl}[1/(N-1)]^{1/2}\sum_i |I_{hkl,i} - \langle I_{hkl}\rangle| / \sum_{hkl}\sum_i I_{hkl,i}$, where $I_{hkl,i}$ is the scaled intensity of the ith measurement of reflection h, k, l, $\langle I_{hkl}\rangle$ is the average intensity for that reflection, and N is the redundancy. $R_{pim} = \Sigma_{hkl}\,(1/(n-1))^{1/2}\,\Sigma_i\,|\,I_{hkl,i} - \langle I_{hkl}\rangle\,|\,/\,\Sigma_{hkl}\,\Sigma_i\,I_{hkl,i}$, where n is the redundancy

[c] No. molecules for complexes refers to number of HA protomers per asymmetric unit (ASU).

[d] $R_{cryst} = \sum_{hkl} |F_o - F_c| / \sum_{hkl} |F_o|$, where F_o and F_c are the observed and calculated structure factors.

[e] R_{free} was calculated as for R_{cryst}, but on 5% of data excluded before refinement.

[f] The values are the percentage of residues in the favored and outliers regions analyzed by MolProbity [59].

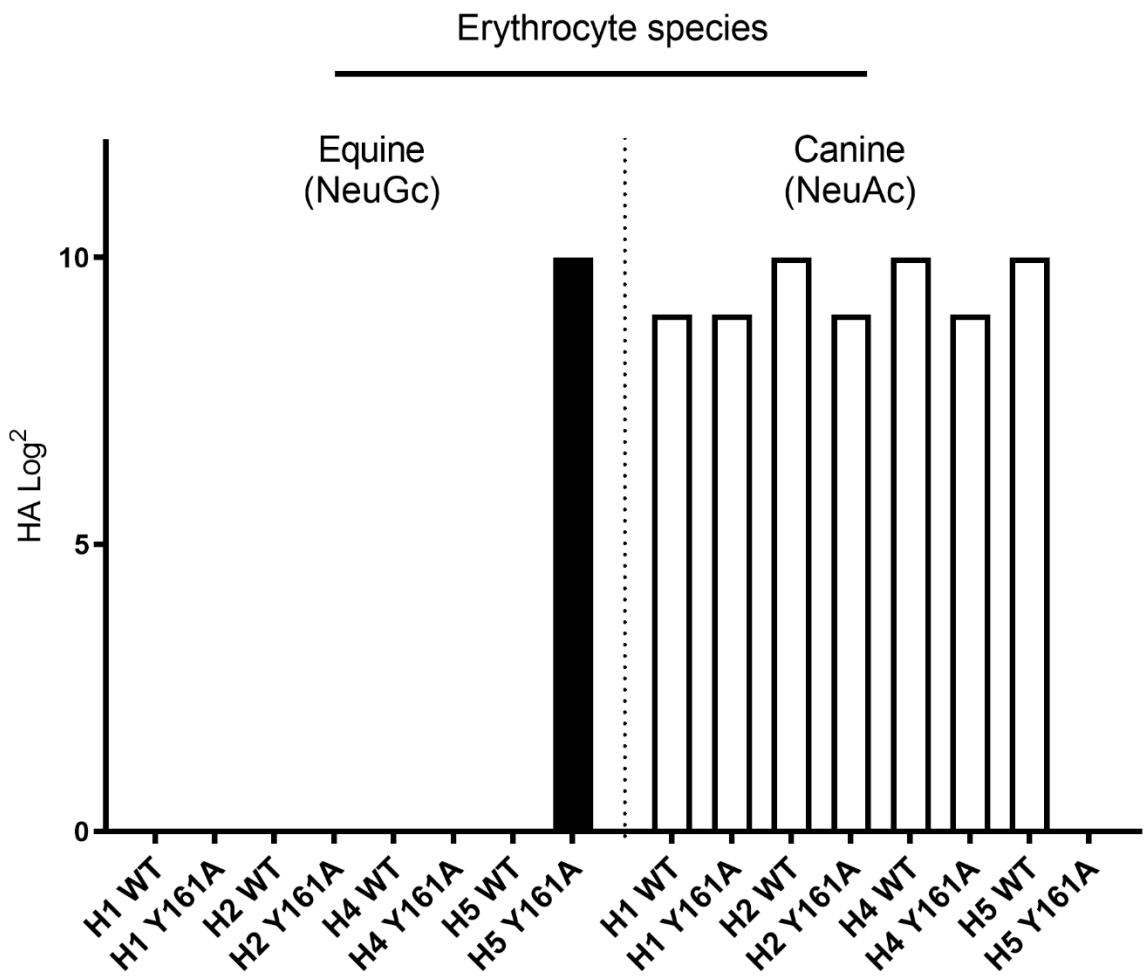


705

706 **S1 Fig. Amino acid alignment of the HAs of A/Equine/New York/49/73 H7N7, A/Turkey/Italy/214845/02 H7N3, and**

707 **A/Chicken/Jalisco/12283/12 H7N3.**

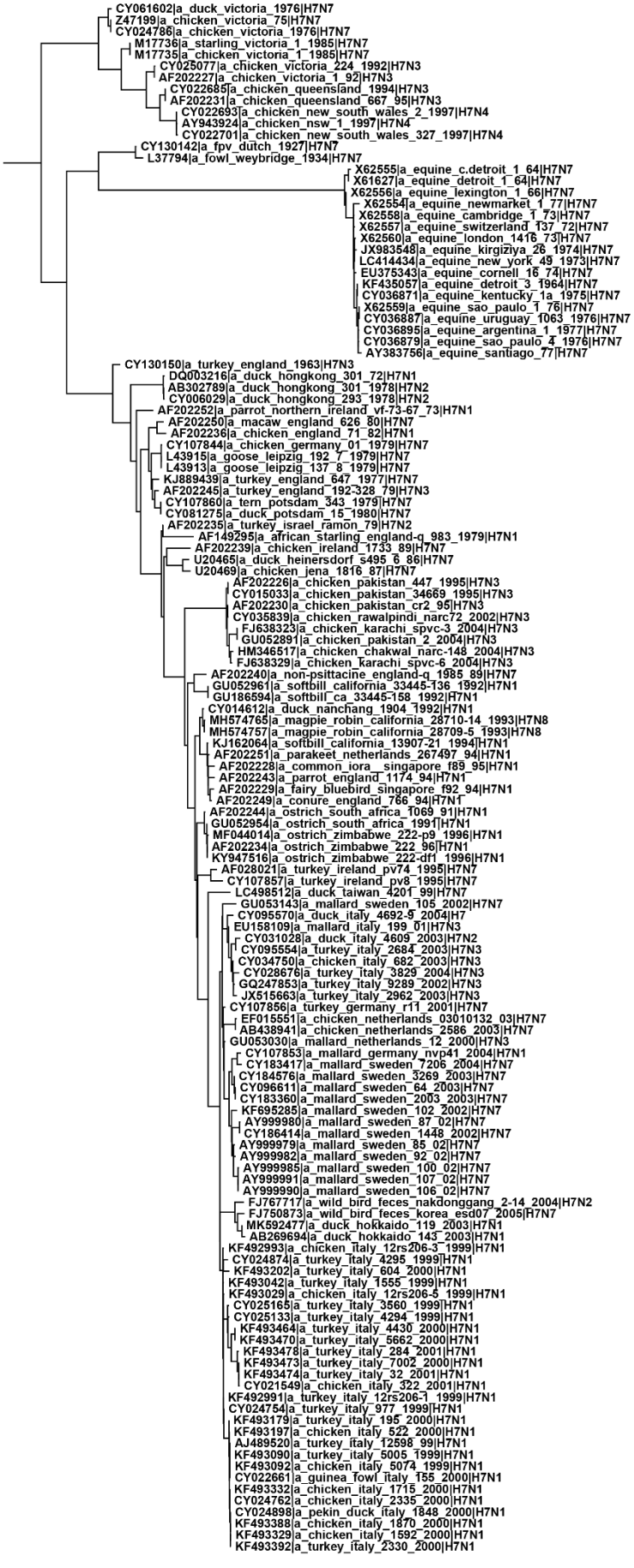
708 Amino acid positions are indicated, dots indicate identical amino acids, and gray squares indicate deletions.



709

710 **S2 Fig. Hemagglutination assay with equine and canine erythrocytes.**

711 Equine erythrocytes contain 90% NeuGc and canine erythrocytes do not contain
712 NeuGc. HAs of wild-type and Y161A mutants of A/Duck/Hokkaido/111/2009 H1N5,
713 A/Duck/Hokkaido/95/2001 H2N2, A/Duck/Hokkaido/138/2007 H4N6, and
714 A/Vietnam/1203/2004 H5N1 were investigated.



716

S3 Fig. Complete annotated phylogenetic tree of equine and Eurasian avian H7 influenza A strains.

The compact trees that show the variation in amino acids at positions 135, 128, 130, 189, and 193 without strain names are shown in Fig 6A-E.