Published in final edited form as:
Cell. 2013 June 20; 153(7): 1486–1493. doi:10.1016/j.cell.2013.05.034.

Glycan-Receptor Binding of the Influenza A Virus H7N9 Hemagglutinin

Kannan Tharakaraman[*], **Akila Jayaraman**[*], **Rahul Raman**, **Karthik Viswanathan**, **Nathan W. Stebbins**, **David Johnson**, **Zachary Shriver**, **V. Sasisekharan**, and **Ram Sasisekharan**
Department of Biological Engineering, Koch Institute of Integrative Cancer Research, Infectious Diseases Inter-disciplinary Research Group, Singapore-MIT Alliance for Research & Technology Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge MA 02139.

SUMMARY

The advent of H7N9, in early 2013, is of concern for a number of reasons, including its capability to infect humans, the etiology of infection is unclear, and that, broadly the human population does not have pre-existing immunity to the H7 subtype. Earlier sequence analyses of H7N9 hemagglutinin (HA) point to amino acid changes that predicted human receptor binding and impinge on the antigenic characteristics of the HA. Herein we report that the H7N9 HA shows limited binding to human receptors; however, should a single amino acid mutation occur, this would result in structural changes within the receptor binding site that allow for extensive binding to human receptors present in upper respiratory tract. Furthermore, a subset of the H7N9 HA sequences demarcating coevolving amino acids appear to be in the antigenic regions of H7, which in turn could impact effectiveness of the current WHO recommended pre-pandemic H7 vaccines.

INTRODUCTION

In late March 2013, a H7N9 influenza A virus subtype was identified that was found to infect humans, causing rapidly progressing severe lower respiratory tract infection and mortality (Gao et al., 2013). As of early May, 131 reported cases have been confirmed, with most cases in Mainland China. Detailed sequencing and analysis of the human isolates of H7N9 have offered insights into their potential origin and factors that govern the virus's virulence and pathogenicity (Li et al., 2013). The transmission and pathogenicity associated with this virus is unanticipated for several reasons. First, transmission of H7 viruses from birds to mammals has been only reported rarely (Kwon et al., 2011). Additionally, until the advent of H7N9, human infections with N9 subtype viruses have not been previously reported. Finally, human infections with other H7 viruses (primarily H7N2, H7N3, and H7N7), even with high pathogenicity viruses containing a polybasic cleavage site in HA, have primarily resulted in conjunctivitis or uncomplicated illness, with few exceptions (Fouchier et al., 2004; Hirst et al., 2004).

Corresponding author: Ram Sasisekharan, 77 Massachusetts Avenue, 76-461, Cambridge, MA 02139, U.S.A. 617-258-9494 (telephone); rams@mit.edu.
[*]Contributed equally

Analysis of newly arising H7N9 strains, including A/Shanghai/1/2013, A/Shanghai/2/2013, and A/Anhui/1/2013 indicates that H7N9 is a reassorted virus incorporating envelope genes from at least two H7 strains (hemagglutinin, HA, from an H7N3 strain and neuraminidase, NA from an avian-adapted H7N9 strain) with the internal genes from at least two H9N2 avian-adapted influenza strains (Li et al., 2013). Further analysis of the H7N9 gene segments have shown the occurrence of signature amino acids associated with adaptation to human host and virulence (summarized in **Table 2**, ref (Gao et al., 2013)). H7N9 strains exhibit hallmark mutations that are thought to correlate with increased virulence and potentially transmission in animal models such as mice and ferrets, including the E627K mutation in PB2 (which is known to play a key role in human-to-human respiratory droplet transmission (Van Hoeven et al., 2009)). Furthermore, genetic analysis indicates the presence of mutations, in at least some strains, within the M2 ion channel and NA that confer drug resistance to the adamantanes and oseltamivir, respectively.

In contrast to the above analysis, several of the hallmark features found in highly pathogenic influenza strains (*eg.*, H5N1) including the N66S mutation in PB1-F2, the aforementioned polybasic sequence in the linker between HA1 and HA2, and the PDZ binding motif in C-terminal of NS1 are absent in the H7N9 human isolates analyzed to date. Within this context and given the somewhat confounding genetic and epidemiological evidence of the relative human-adaptation of H7N9, one of the most important outstanding questions is the “status” of the HA protein for these isolates. Characterization of the HA protein is important given its role in virulence (Pappas et al., 2008) and virus neutralization to preexisting antibodies through antigenic memory (Hensley et al., 2009). Finally, the receptor binding properties of HA, governing a given virus’s tissue and organismic tropism is one of the key factors that critically govern aerosol transmissibility, including human-to-human transmission (Van Hoeven et al., 2009).

Previous studies have demonstrated that one of the key properties governing human adaptation of influenza A virus is a “switch” in the glycan receptor binding specificities of viral HA (Skehel and Wiley, 2000). Therefore, together with hallmark mutations in other genes, such as PB2, describing mutations in HA that lead to such a “switch” become important to for surveillance purposes. From the standpoint of human tissue tropism, the HA from human-adapted viruses including pandemic strains show extensive binding to the apical surface of human upper respiratory tissues (such as trachea) and also show characteristic binding characteristic binding to mucin secreting non-ciliated goblet cells on the apical surface and to submucosal glands in ferret respiratory tract (Matrosovich et al., 2004; Nicholls et al., 2007; Srinivasan et al., 2008). Through lectin staining it has ben demonstrated previously that these regions in human tracheal sections predominantly display diverse glycan receptors terminated by $\alpha 2 \rightarrow 6$ sialic acid linkage (human receptors). (Srinivasan et al., 2008; Viswanathan et al., 2010) We have previously demonstrated with H1, H2 and H3 subtypes that this binding property of human-adapted viruses is one of the key factors that correlate with their ability to efficiently transmit via respiratory droplets in ferrets–a well-established animal model to measure the potential for airborne human-to-human transmission (Jayaraman et al., 2011; Maines et al., 2009).

In the case of H7N9, based on the presence of a leucine residue in the 226 position (H3 numbering) of the HA, earlier studies have predicted that this HA would have strong binding to human receptors (Gao et al., 2013; Liu et al., 2013). These recent studies note that HA-glycan receptor interaction is a critical property of the virus and that experimental characterization of this property for the H7N9 subtype is important.

Herein, we report the molecular and structural features of the glycan receptor-binding site (RBS) of H7N9 HA and the experimental characterization of its binding to physiological

glycan receptors in the human respiratory tract. Contrary to the predicted strong binding to human receptors, H7N9 HA shows limited binding to these receptors in the human upper respiratory tract when compared to human receptor-binding of other human-adapted HAs. The experimentally observed limited binding to human receptors by H7N9 HA is consistent with the analysis of its RBS structural features. The structural and sequence analyses further point to a single Gly228→Ser amino acid change which modifies the network of inter-residue contacts in the RBS for more optimal contacts with both avian and human receptor. Consequently, introducing this amino acid change in H7N9 HA RBS resulted in a mutant HA that extensively bound to the apical surface of human tracheal tissue sections in a fashion similar to that of other human-adapted HAs. Our findings therefore, provide new insights into the physiological glycan receptor tropism of the current H7N9 HA. We also report an increase in human receptor binding should a single Gly228→Ser amino acid change occur in this HA in the context of monitoring the evolution of this emerging subtype, especially as it continues to circulate in humans. Finally, we report amino acid substitutions in the H7N9 HA sequences that distinguish the evolution (including antigenic sites) of this recently emerged subtype from past H7 isolates, which in turn has implications for vaccine development strategies.

RESULTS

Given that H7N9 is a newly emerged subtype, there is a limited set of HA sequences available for this subtype to do a comprehensive analysis of sequence evolution of its RBS. Therefore, in order to understand glycan receptor binding properties of H7N9 HA, we chose a representative human isolate A/Anhui/1/2013 (Anh13). Since there is no crystal structure available, we constructed a homology-based structural model of Anh13 HA and compared its RBS with H3 HA (its phylogenetically closest human-adapted HA Supplemental Figure 1). Anh13 was chosen since it shares substantial sequence identity with many other reported strains of H7N9, including A/Shanghai/2/2013 and A/Hangzhou/1/2013. Additionally, as noted previously (Gao et al., 2013), while there are changes in the HA sequence of Anh13 compared to the HA sequence of other reported H7N9 strains, including A/Shanghai/2/2013, the HA of Anh13 contains a key Q226L mutation which has been reported to be important for altered receptor specificity for group 2 viruses, including H3 and H7. As such, Anh13 (and related viruses) are more likely than viruses such as A/Shanghai/2/2013 to be of concern from the standpoint of altered receptor specificity and hence human transmissibility.

Earlier, we defined a framework that incorporated descriptors of the structural topology of the human glycan receptor as well as inter-amino acid interaction networks within the RBS to define the molecular features of the RBS of H5 HA for high avidity/specificity binding to human glycan receptors (Chandrasekaran et al., 2008; Soundararajan et al., 2011). We therefore compared the molecular features of the Anh13 RBS with those of H3 HA (from A/Aichi/1/68 or Aich68, a strain from the 1967-68 pandemic), which was recently co-crystallized with both avian and human receptors (Lin et al., 2012). Also importantly, since Aich68 represents a human-adapted virus, comparison of its HA to that of Anh13 provide an important benchmark to address the question of whether the HA from Anh13 shares structural characteristics with HAs of human-adapted viruses and if not, which structural characteristics are missing.

Structural analysis of both H7 and H3 HAs indicate that the structural topology of the human glycan receptor bound in the RBS of HA is such that it has a clearly defined base region consisting of the terminal Neu5Acα2→6Galβ1→ motif and an extension region consisting of at least a disaccharide →4GlcNAcβ1→3Galβ1→. On the other hand, the topology of the avian receptor bound within the HA RBS is such that majority of the contacts with residues within the RBS involve the terminal Neu5Acα2→3Galβ1→ motif in

the base region (Figure 1). The 130- (residues 131-138), 140- (residues 140-145) and 220- (residues 219-228) loops in the RBS make critical contacts with the disaccharide motif in the base region and thus play a key role in dictating the avian or human receptor binding preference (residue position numbering is based on Aichi68 HA-glycan co-crystal structure PDB ID:2YPG). Additionally, residues within the 190-helix (residues 190-196) and a part of the 150-loop (residues 156-160), if properly positioned, make critical contacts with the extension region of the human receptor. Taken together, these 5 loops and 1 helix in the RBS that make contact with the base and potentially with the extension region of the human receptor and their network of interactions with spatially proximal residues in the RBS constitute a complete set of molecular features that should be analyzed to understand the receptor binding preference of an HA. Comparison of these features between Anh13 and Aichi68 HA show many similarities as well as some important differences (Table 1). First, many of the residues in the 130-loop, 140-loop and 220- loop are similar between the HA of Anh13 and Aichi68. Based on this level of similarity, we find that the network of inter-residue contacts involving these residues is also similar (Figure 1A and 1C). The key structural differences noted in this analysis include part of the 190-helix and the 150-loop required for contact with the extension region of the human receptor. In the Aich68 HA, the residues interacting with the extension region include Q189, S193, K156 and S159 (Figure 1C), whereas in H7 HA, these residues include K193, T158, D159 and N160 (Figure 1B). Furthermore, in the case of the H7 HA, R131 is positioned to make an additional contact with the extension region (Figure 1B). In addition to the differences in the amino acids at these positions, and also of note, are differences between Aich68 and Anh13 in the inter-residue interaction network governed by the residue at position 228. This position is a Ser in Aichi68 but a Gly in Anh13. The S228 position in H3 is critical for the inter-amino acid network involving S186, T187 and E190, which positions E190 to make critical contacts with the sialic acid of both avian and human receptors (Figure 1C). On the other hand G228 in H7 HA does not possess this inter-amino acid network and therefore the network of inter-residue contacts involving E190 in H7 is different from that of H3 HA and instead includes 193 (and its network) and 189 (Figure 1A).

Extending this analysis, the E190 in H7 HA is positioned to make additional contacts with the extension region instead of the critical contact with the sialic acid in the base region (Figure 1B). Therefore, Anh13 H7 HA lacks at least two critical contacts involving the 190 and 226 HA positions with the Neu5Acα2→3Galβ1→ terminal motif as observed in avian-adapted HAs. In the case of human receptor contacts, Anh13 has lower contacts with the base region owing to the absence of the S228 residue. Our structural analysis pointed to the H7N9 HA having a substantially lower binding to avian receptors than typical avian-adapted HAs as well as lower binding to human receptors than the human-adapted H3 HAs. Further, the RBS of H7N9 HA is such that a single G228→S amino acid change would modify the inter-amino acid network in the RBS to position the E190 and S228 residues for optimal contacts with both avian and human receptors.

To validate the structural analyses, we analyzed binding of Anh13 to tissue sections, representative of human respiratory tract, which display the physiologically relevant receptors for influenza A viruses (Figure S2A). Anh13 stained the apical surface and submucosal region of the human trachea, which express glycans known to be receptors of human adapted viruses (Jayaraman et al., 2012; Shinya et al., 2006). However, the intensity and the extent of tracheal apical surface staining by Anh13 was substantially lower than what is typically observed for human-adapted HAs (Figure S2A). Furthermore, Anh13 HA also showed minimal binding to deep lung alveolar section – a region that is extensively stained by avian-adapted HAs (Figure S2B). In contrast, the G228→S mutant HA showed a dramatic and significant increase in the extent and intensity of staining to apical surface of the tracheal section including extensive staining of non-ciliated goblet cells in a fashion

similar to that of other human-adapted HAs and *Sambucus nigra* agglutinin I (SNA I) (Figure 2 and Figure S2C). Interestingly the G228→S mutation also substantially increased its binding to the alveolar sections (Figure S2B). These results are consistent with the structural analyses of the RBS features of H7N9 HA.

Finally, in the context of H7N9 evolution, two mutations 174S and 226L appear to be unique to the H7N9 HA sequences. As noted above, the residue at 226 is a critical determinant of the receptor-binding specificity of H7 HA, with human viruses favoring L/I, avian viruses favoring Q. These two positions are also part of a larger cluster of coevolving positions (122A, 174S, 186V, 202V, 226L), all within the 50-230 HA region, which demarcates the virus from its previous H7 ancestors (Methods). The above observations indicate that the H7N9 HA has evolved to be distinct from its predecessors. Previous H7 strains carrying single mutations (from the coevolving cluster) are predominantly from the Eurasian lineage suggesting that viruses from this lineage have higher potential to generate variants when compared to the American lineage (Supplemental Figure 3). Nucleotide analyses of the RBS-proximal region of HA (residues 50-230 of HA1) of Eurasian sequences from 1902 to 2013 show strong diversifying (positive) selection at 156 (Methods). The same position has been shown to be under selection pressure in H1 subtype as well (Li et al., 2011).

Previously, we defined a quantitative metric to compare the antigenicity of two HAs. Briefly, the metric, called antigenic intactness (AI), is directly proportional to the fraction of residues conserved in the immunodominant antigenic sites between two HAs. In our previous work, we showed good agreement between AI values and antigenic relatedness metric computed from ferret antisera hemagglutinin inhibition (HI) cross-reactivity data (Tharakaraman et al., 2013), indicating that AI values could be applied to predict vaccine-induced cross-reactive antibody responses. Critically, strains that are antigenically related had AI > 80% whereas strains that are not related to each other had AI < 80% (Tharakaraman et al., 2013). Keeping 80% as the cutoff, the AI values between the recent WHO-recommended H7 vaccines strains (A/Canada/rv444/2004 (H7N3), A/mallard/Netherlands/12/2000 (H7N3) and A/New York/107/2003 (H7N2)) and the A/Anhui/1/2013 H7N9 HA were computed. The H7N9 HA has AI values of 67%, 89% and 70% with A/Canada/rv444/2004 (H7N3), A/mallard/Netherlands/12/2000 (H7N3) and A/New York/107/2003 (H7N2), respectively, suggesting that only /mallard/Netherlands/12/2000 (H7N3) may be effective as a vaccine component. However, amino acid differences between A/Anhui/1/2013 (H7N9) and A/mallard/Netherlands/12/2000 (H7N3) at key antigenic sites (122 in site A, 188 and 189 in site B) might limit this mallard strain to be effective.

In the context of the Q228S mutation and given that H7N9 is a subtype that has just emerged in the human population, it is difficult to predict if this mutation alter antibody response to this region in this subtype. However, it is generally known that the 220 loop, which includes the 228 position, is an antigenic site for some HA subtypes (224 and 225 are part of antigenic site Ca in H1 subtype; 222 is part of antigenic site I-C in H2 subtype; 220 is part of antigenic site D in H3, a subtype that is phylogenetically closest to H7); thus, antibody response targeting this region in H7N9 cannot be excluded. As such, it will be important to complete serological analysis to confirm the vaccine implications of our AI analyses.

DISCUSSION

The emergence of a H7N9 influenza A virus subtype poses significant global health concern given that it has led to severe infection and mortality in humans. Although preliminary genetic analysis of this subtype has led to predictions about its human host adaptation based on hallmark genetic signatures including strong binding to human receptors; this virus has

not yet resulted in a widespread infection in humans resulting from aerosol human-to-human transmission. Earlier studies have highlighted the importance of determining glycan-receptor binding property of H7N9 HA, since this property is one of the many key factors that govern human adaptation of the virus (Gao et al., 2013; Liu et al., 2013). Therefore, in this study we experimentally characterized glycan receptor-binding properties of H7N9 HA to assess the human adaptation of this HA. To the best of our knowledge this is the first report on the glycan binding data of the H7N9 HA.

Our structural analyses and binding of H7N9 HA to human respiratory tissues demonstrate that it possesses a distinct binding tropism when compared to either an avian-adapted or a human-adapted HA. Our findings shed new light on the distinct tropism of this H7N9 HA that is contrary to the expected strong human receptor-binding preference predicted from earlier sequence analyses (Gao et al., 2013; Liu et al., 2013). This distinct tropism would likely impinge on the aerosol transmissibility of the H7N9 viruses in ferrets when compared to the efficient transmission observed in the past pandemic viruses. The limited human- and avian-receptor binding of the H7N9 HA raises a question as to whether this currently circulating subtype is an intermediate in the adaptation to the human host. Answering this question at this point in time is limited by the availability of sequence information for this recently emerged subtype. Nevertheless, our structural and experimental analyses point to an important role for G228→S amino acid change in the RBS, should it emerge in the current H7N9 HA, in substantially increasing its binding to human receptors in the human respiratory tract. Of note, within the context of this study, is the increased binding of the mutant HA to the non-cilicated goblet cells in the human trachea. This is significant as binding to goblet cells is one of the hallmarks of human adapted HAs (Jayaraman et al., 2012; Matrosovich et al., 2004; Srinivasan et al., 2008). Even in cultures of differentiated human airway epithelial cells, the human influenza A viruses are found to predominantly infect non-cilicated cells as compared to avian influenza A viruses which target the cilicated cells (Matrosovich et al., 2004)

Previously, several groups including us have the studied receptor specificity of influenza hemagglutinin using ‘prototypic’ glycan arrays containing limited sets of glycans capped with α2-3 or α2-6 linked sialic acid. We find that in most cases with H1, H2 and H3 HA, the observed binding on glycan array correlated to observed binding of these HAs on physiological glycans from respiratory tract tissue sections. However, as noted in our previous study of HA from H7N2 A/Netherlands/219/2003, we observed that introduction of 226L and 228S resulted in significant staining of the apical region of tracheal tissues despite having modest binding to α2-6 sialyalted glycans on the glycan array (Srinivasan et al., 2013). In the present study, with Anh13 we again observed that the presence of 226L and 228S on the HA results in significant binding to the apical and submucosal regions of the trachea that was not captured by the affinity of the HA to limited set of glycans in the array (data not shown). This result warrants an investigation into binding to a more diverse set of α2-6 sialylated glycans might be warranted for H7 HAs and furthermore highlights the need for caution for interpreting prototypic glycan array data during surveillance of H7N9 viral evolution.

The significance of the change in residue 156 in receptor binding or other HA function is unclear although the neighboring 158 glycosylation is known to have an influence on human receptor binding (Stevens et al., 2008). Significantly, a subset of the H7N9 demarcating coevolving positions (122, 186 and 202) appear to be in the antigenic regions of H7, which could have implications on the effectiveness of the current WHO recommended pre-pandemic H7 vaccines. Furthermore, the reported poor immunogenicity of vaccine candidates based on H7N9 is a key challenge for potential vaccine strategy. The fact that the wild type virus binds poorly to human receptor supports the notion of poor uptake by human

cells to engender an appropriate human immune response. Mutant forms of H7N9 HA, such as G228S with higher specificity to human receptors, can potentially have important applications for the generation of appropriate vaccine countermeasures. Of course, a critical component of vaccine assessment is the use of serological assays to investigate cross-reactivity of heterologous strains. In this context, we have previously demonstrated the correlation between AI score with cross-neutralization responses based on WHO data (Tharakaraman et al., 2013). Taken together, the mutations on the antigenic regions of H7N9 HA together with the results of the AI analysis could impinge on H7 vaccine development.

In summary, our study reports on the experimental glycan-binding properties of the H7N9 HA and also reports on the effects of a single amino acid G228→S amino acid change in dramatically increasing glycan receptor binding of this HA. In light of the continued circulation of H7N9 in human subtypes our study facilitates monitoring the evolution of H7N9 including the acquisition of amino acid changes such as G228→S that would make it closer to human adaptation. Our study warrants further investigation of introducing this single amino acid change to H7N9 virus in the context of other genetic changes characteristic of human-adapted viruses (such as K627 in PB2) and studying its transmission properties in ferrets. This analysis therefore sets the stage for future *in vitro* studies in human respiratory tract cell cultures and *in vivo* studies in ferret and mice to investigate the replication potential, virulence, pathogenicity and respiratory droplet transmission using these recombinant wild-type and mutant HA H7N9 viruses. Taken together, these findings have important implications for surveillance of H7N9 mutations in clinical settings, as well as for vaccine development efforts.

EXPERIMENTAL PROCEDURES

Cloning, baculovirus synthesis, and mammalian expression and purification of HA

Anh13 wild-type and G228→S mutant HA sequences were codon-optimized for mammalian expression, synthesized (DNA2.0, Menlo Park, CA) and sub-cloned into modified pcDNA3.3 vector for expression under CMV promoter. Recombinant expression of HA was carried out in HEK 293-F FreeStyle suspension cells (Invitrogen, Carlsbad, CA) cultured in 293-F FreeStyle Expression Medium (Invitrogen, Carlsbad, CA) maintained at 37 °C, 80% humidity and 8% CO_2. Cells were transfected with Poly-ethylene-imine Max (PEI-MAX, PolySciences, Warrington, PA) with the HA plasmid and were harvested seven days post-infection. The supernatant was collected by centrifugation, filtered through a 0.45 μm filter system (Nalgene, Rochester, NY) and supplemented with 1:1000 diluted protease inhibitor cocktail (Calbiochem filtration and supplemented with 1:1000 diluted protease inhibitor cocktail (EMD Millipore, MA). HA was purified from the supernatant using His-trap columns (GE Healthcare) on an AKTA Purifier FPLC system. Eluting fractions containing HA were pooled, concentrated and buffer exchanged into 1X PBS pH 7.4 using 100K MWCO spin columns (Millipore, Billerica, MA). The purified protein was quantified using BCA method (Pierce, Rockford, IL).

Homology modeling of HA

A structural model of Anh13 HA was built using the MODELLER homology modeling software. The crystal structure of A/Netherlands/219/2003 (Neth03) HA (PDB: 4DJ6) was used as a template to build the model. The structural model of Anh13 bound to avian receptor was constructed by superimposing the HA1 from co-crystal structure of Neth03-avian receptor complex (PDB ID: 4DJ7) on Anh13 HA1. The structural model of Anh13 in complex with human receptor was constructed by superimposing the HA1 from co-crystal structure of Aichi68-human receptor complex (PDB ID: 2YPG) with HA1 of Anh13. The

final models were subject to energy minimization (500 steps conjugate gradient + 500 steps steepest descent) with potentials assigned using AMBER force field.

Coevolution, phylogeny and selection analyses of H7 HA sequences

A total of 625 non-redundant full-length H7 HA sequences were downloaded from GISAID. To further eliminate redundancy, the sequences were grouped according to subtype, host, year and country and a representative sequence was chosen from each group. This led to a total of 231 HA sequences. Coevolving groups of amino acids were predicted using CAPS online server for protein coevolution (http://bioinf.gen.tcd.ie/caps/). The results indicate functionally or structurally linked regions that are subjected to strong selective constraints. A phylogeny tree was constructed from the 231 HA sequences using the Neighbor Joining method found in MEGA 5.1 software (http://www.megasoftware.net/). Protein coding nucleotide sequences were extracted for the 114 Eurasian HA sequences and the region encoding residues 50-230 of HA1 was employed for finding invididual codons under diversifying/positive selection. Positively selected sites were predicted using DataMonkey (http://www.datamonkey.org/), which uses a normalized dN-dS > 0 at p-value < 0.1 threshold to detect positive selection.

Binding of HA to human tissue sections

The human tracheal epithelia has been extensively benchmarked as a tissue section representative of human upper respiratory tract that is a predominant physiological target site for human adapted influenza A viruses (Imai et al., 2012; Jayaraman et al., 2012; Mansfield, 2007; Shinya et al., 2006). The apical surface and submucosal regions of the human trachea have been shown to predominantly display human receptors (Chandrasekaran et al., 2008; Jayaraman et al., 2012; Shinya et al., 2006). On the other hand, human alveolar tissue sections representative of deep lung region have been shown to predominantly express avian receptors and are typically stained by HA from avian-adapted influenza A viruses (Chandrasekaran et al., 2008; van Riel et al., 2007). Paraffinized human tracheal and alveolar (US BioChain) tissue sections were deparaffinized, rehydrated and incubated with 1% BSA in PBS for 30 minutes to prevent non-specific binding. HA was pre-complexed with primary antibody (mouse anti 6X His tag, Abcam) and secondary antibody (Alexa fluor 488 goat anti mouse, Invitrogen) in a molar ratio of 4:2:1, respectively, for 20 minutes on ice. The tissue binding was performed over two different HA concentrations (40 μg/ml and 20 μg/ml) by diluting the pre-complexed stock HA in 1% BSA-PBS. Tissue sections were then incubated with the HA-antibody complexes for 3 hours at RT. The tissue sections were counterstained by propidium iodide (Invitrogen; 1100 in TBST). The tissue sections were mounted and then viewed under a confocal microscope (Zeiss LSM 700 laser scanning confocal microscopy). Sialic-acid specific binding of HAs to tissue sections was confirmed by loss of staining after pre-treatment with Sialidase A (from *Arthrobacter ureafaciens*, Prozyme), This enzyme has been demonstrated to cleave the terminal Neu5Ac from both Neu5Acα2→3Gal and Neu5Acα2→6Gal motifs. In the case of sialidase pretreatment, tissue sections were incubated with 0.2 units of Sialidase A for 3 hours at 37°C prior to incubation with the proteins. Pre-treatment of human tissue sections with Sialidase A resulted in complete loss of HA staining.

Capturing Network inter-amino acid contacts for RBS residues (RBSN)

The coordinates of Neth03 H7 HA – avian receptor (PDB ID: 4DJ7) and Aichi68 H3 HA–human receptor complexes (PDB ID: 2YPG) were uploaded into the PDBePISA server (http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html) to determine key residues in the HA RBS that make contact with the corresponding glycan receptor (interface cut-off of 30% was used). For these residues, their environment was defined using a distance threshold of 7 Å and the contacts including putative hydrogen bonds (including water-bridged ones),

disulfide bonds, pi-bonds, polar interactions, salt bridges, and Van der Waals interactions (non-hydrogen) occurring between pairs of residues within this threshold distance was computed as described previously (Soundararajan et al., 2011). These data were assembled into an array of eight atomic interaction matrices. A weighted sum of the eight atomic interaction matrices were then computed to produce a single matrix that accounts for the strength of atomic interaction between residue pairs within the RBS, using weights derived from relative atomic interaction energies (Soundararajan et al., 2011). The inter-residue interaction network calculated in this fashion generates a matrix that describes all the contacts made by critical RBS residues with spatial proximal neighboring residues in their environment. Each element *i, j* is the sum of the path scores of all paths between residues *i* and *j*. The degree of networking score for each residue was computed by summing across the rows of the matrix, which was meant to correspond to the extent of "networking" for each residue. The degree of networking score was normalized (RBSN score) with the maximum score for each protein so that the scores varied from 0 (absence of any network) to 1 (most networked).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was funded in part by National Institutes of Health (R37 GM057073-13) and the National Research Foundation supported Interdisciplinary Research group in Infectious Diseases of SMART (Singapore MIT alliance for Research and Technology).

References

Chandrasekaran A, Srinivasan A, Raman R, Viswanathan K, Raguram S, Tumpey TM, Sasisekharan V, Sasisekharan R. Glycan topology determines human adaptation of avian H5N1 virus hemagglutinin. Nat Biotechnol. 2008; 26:107–113. [PubMed: 18176555]

Fouchier RA, Schneeberger PM, Rozendaal FW, Broekman JM, Kemink SA, Munster V, Kuiken T, Rimmelzwaan GF, Schutten M, Van Doornum GJ, et al. Avian influenza A virus (H7N7) associated with human conjunctivitis and a fatal case of acute respiratory distress syndrome. Proc Natl Acad Sci U S A. 2004; 101:1356–1361. [PubMed: 14745020]

Gao R, Cao B, Hu Y, Feng Z, Wang D, Hu W, Chen J, Jie Z, Qiu H, Xu K, et al. Human Infection with a Novel Avian-Origin Influenza A (H7N9) Virus. The New England journal of medicine. 2013

Hensley SE, Das SR, Bailey AL, Schmidt LM, Hickman HD, Jayaraman A, Viswanathan K, Raman R, Sasisekharan R, Bennink JR, et al. Hemagglutinin receptor binding avidity drives influenza A virus antigenic drift. Science. 2009; 326:734–736. [PubMed: 19900932]

Hirst M, Astell CR, Griffith M, Coughlin SM, Moksa M, Zeng T, Smailus DE, Holt RA, Jones S, Marra MA, et al. Novel avian influenza H7N3 strain outbreak, British Columbia. Emerging infectious diseases. 2004; 10:2192–2195. [PubMed: 15663859]

Jayaraman A, Chandrasekharan A, Viswanathan K, Raman R, Fox JG, Sasisekharan R. Decoding the distribution of glycan receptors for human-adapted influenza A viruses in ferret respiratory tract. PLoS ONE. 2012 In Press.

Jayaraman A, Pappas C, Raman R, Belser JA, Viswanathan K, Shriver Z, Tumpey TM, Sasisekharan R. A single base-pair change in 2009 H1N1 hemagglutinin increases human receptor affinity and leads to efficient airborne viral transmission in ferrets. PloS one. 2011; 6:e17616. [PubMed: 21407805]

Kwon TY, Lee SS, Kim CY, Shin JY, Sunwoo SY, Lyoo YS. Genetic characterization of H7N2 influenza virus isolated from pigs. Vet Microbiol. 2011; 153:393–397. [PubMed: 21741185]

Li Q, Zhou L, Zhou M, Chen Z, Li F, Wu H, Xiang N, Chen E, Tang F, Wang D, et al. Preliminary Report: Epidemiology of the Avian Influenza A (H7N9) Outbreak in China. The New England journal of medicine. 2013

Li W, Shi W, Qiao H, Ho SY, Luo A, Zhang Y, Zhu C. Positive selection on hemagglutinin and neuraminidase genes of H1N1 influenza viruses. Virology journal. 2011; 8:183. [PubMed: 21507270]

Lin YP, Xiong X, Wharton SA, Martin SR, Coombs PJ, Vachieri SG, Christodoulou E, Walker PA, Liu J, Skehel JJ, et al. Evolution of the receptor binding properties of the influenza A(H3N2) hemagglutinin. Proceedings of the National Academy of Sciences of the United States of America. 2012; 109:21474–21479. [PubMed: 23236176]

Liu D, Shi W, Shi Y, Wang D, Xiao H, Li W, Bi Y, Wu Y, Li X, Yan J, et al. Origin and diversity of novel avian influenza A H7N9 viruses causing human infection: phylogenetic, structural, and coalescent analyses. Lancet. 2013 doi: 10.1016/S0140-6736(13)60938-1.

Maines TR, Jayaraman A, Belser JA, Wadford DA, Pappas C, Zeng H, Gustin KM, Pearce MB, Viswanathan K, Shriver ZH, et al. Transmission and pathogenesis of swine-origin 2009 A(H1N1) influenza viruses in ferrets and mice. Science. 2009; 325:484–487. [PubMed: 19574347]

Matrosovich MN, Matrosovich TY, Gray T, Roberts NA, Klenk HD. Human and avian influenza viruses target different cell types in cultures of human airway epithelium. Proc Natl Acad Sci U S A. 2004; 101:4620–4624. [PubMed: 15070767]

Nicholls JM, Bourne AJ, Chen H, Guan Y, Peiris JS. Sialic acid receptor detection in the human respiratory tract: evidence for widespread distribution of potential binding sites for human and avian influenza viruses. Respir Res. 2007; 8:73. [PubMed: 17961210]

Pappas C, Aguilar PV, Basler CF, Solorzano A, Zeng H, Perrone LA, Palese P, Garcia-Sastre A, Katz JM, Tumpey TM. Single gene reassortants identify a critical role for PB1, HA, and NA in the high virulence of the 1918 pandemic influenza virus. Proc Natl Acad Sci U S A. 2008; 105:3064–3069. [PubMed: 18287069]

Shinya K, Ebina M, Yamada S, Ono M, Kasai N, Kawaoka Y. Avian flu: influenza virus receptors in the human airway. Nature. 2006; 440:435–436. [PubMed: 16554799]

Skehel JJ, Wiley DC. Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. Annu Rev Biochem. 2000; 69:531–569. [PubMed: 10966468]

Soundararajan V, Zheng S, Patel N, Warnock K, Raman R, Wilson IA, Raguram S, Sasisekharan V, Sasisekharan R. Networks link antigenic and receptor-binding sites of influenza hemagglutinin: Mechanistic insight into fitter strain propagation. Sci Rep. 2011; 1

Srinivasan A, Viswanathan K, Raman R, Chandrasekaran A, Raguram S, Tumpey TM, Sasisekharan V, Sasisekharan R. Quantitative biochemical rationale for differences in transmissibility of 1918 pandemic influenza A viruses. Proc Natl Acad Sci U S A. 2008; 105:2800–2805. [PubMed: 18287068]

Srinivasan K, Raman R, Jayaraman A, Viswanathan K, Sasisekharan R. Quantitative description of glycan-receptor binding of influenza a virus h7 hemagglutinin. PloS one. 2013; 8:e49597. [PubMed: 23437033]

Stevens J, Blixt O, Chen LM, Donis RO, Paulson JC, Wilson IA. Recent avian H5N1 viruses exhibit increased propensity for acquiring human receptor specificity. J Mol Biol. 2008; 381:1382–1394. [PubMed: 18672252]

Tharakaraman K, Raman R, Sasisekharan V, Viswanathan K, Stebbins NW, Sasisekharan R. Antigenically intact hemagglutinin in circulating avian and swine influenza viruses and potential for H3N2 pandemic. Scientific Reports. 2013 In Press.

Van Hoeven N, Pappas C, Belser JA, Maines TR, Zeng H, Garcia-Sastre A, Sasisekharan R, Katz JM, Tumpey TM. Human HA and polymerase subunit PB2 proteins confer transmission of an avian influenza virus through the air. Proc Natl Acad Sci U S A. 2009; 106:3366–3371. [PubMed: 19211790]

Viswanathan K, Koh X, Chandrasekaran A, Pappas C, Raman R, Srinivasan A, Shriver Z, Tumpey TM, Sasisekharan R. Determinants of Glycan Receptor Specificity of H2N2 Influenza A Virus Hemagglutinin. PLoS One. 2010; 5:e13768. [PubMed: 21060797]

HIGHLIGHTS

- The hemagglutinin of H7N9 virus does not efficiently bind human receptors
- A single residue change in receptor binding site increases binding to human receptors
- Mutations on hemagglutinin may reduce the effectiveness of current H7 vaccines

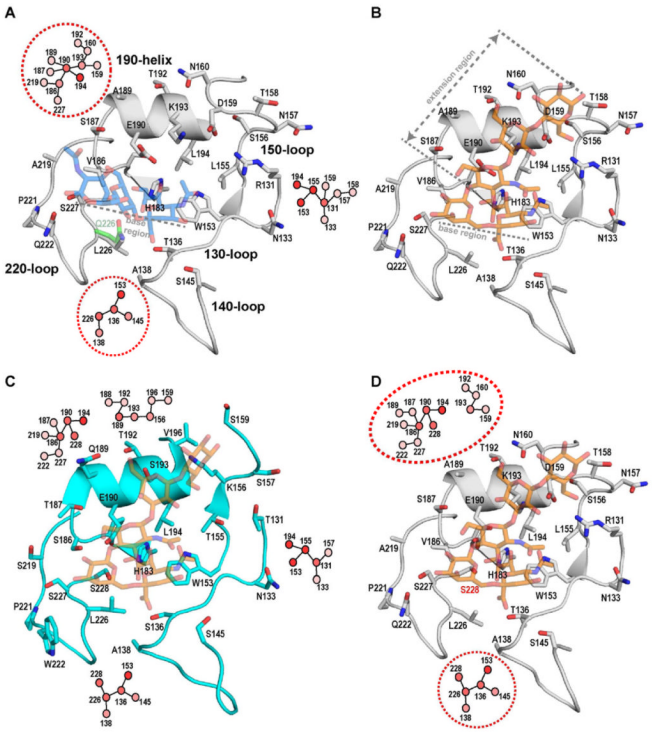

Figure 1. Structural HA-glycan receptor complexes

A, shows structural model of Anh13 H7 HA RBS in complex with avian receptor. The avian receptor is shown as a stick at 40% transparency with carbon atom colored in blue. The loop and helix regions used to define molecular features of the HA RBS are shown. The side chains of the key amino acids in these regions that make contact with the glycan receptor are shown. The side chain of Q at 226 position observed in all other H7 HAs (prior to H7N9 outbreak) is shown as stick with 40% transparency (carbon atom colored green). The networks of inter-residue interaction contacts are shown as two-dimensional maps near the corresponding residue positions. The map comprises of circular nodes representing the amino acid positions (which are labeled above the nodes) and colored according to the degree of inter-residue contact (lightest shade of red indicating lowest contact and darkest shade of red indicating highest contact). **B**, shows the structural model of Anh13 H7 HA RBS in complex with human receptor (shown as a stick with 40% transparency and carbon atom colored in orange). **C**, shows Aichi68 H3 HA RBS – human receptor complex as observed in the X-ray co-crystal structure (PDB ID: 2YPG). The inter-residue interaction network of the key RBS residues is shown similar to what was shown in **A**. See also Figure S1. Note the similarities in the network of residues in 130-, 140- and 220-loop between the H3 and H7 HA (expanded in Table 1). On the other hand the network involving residues in the 190 helix are quite different and this difference is brought about by the amino acid differences and also the Gly in H7 HA versus Ser in H3 HA in the 228 position. **D**, shows the structural model of the G228→S mutant of Anh13 H7 HA RBS in complex with human receptor. Note that network involving residues in 190 helix in the mutant is more similar to that observed in H3 HA than the wild-type. The inter-residue contacts networks that are different between the mutant and wild-type HA are shown in red dotted circle.

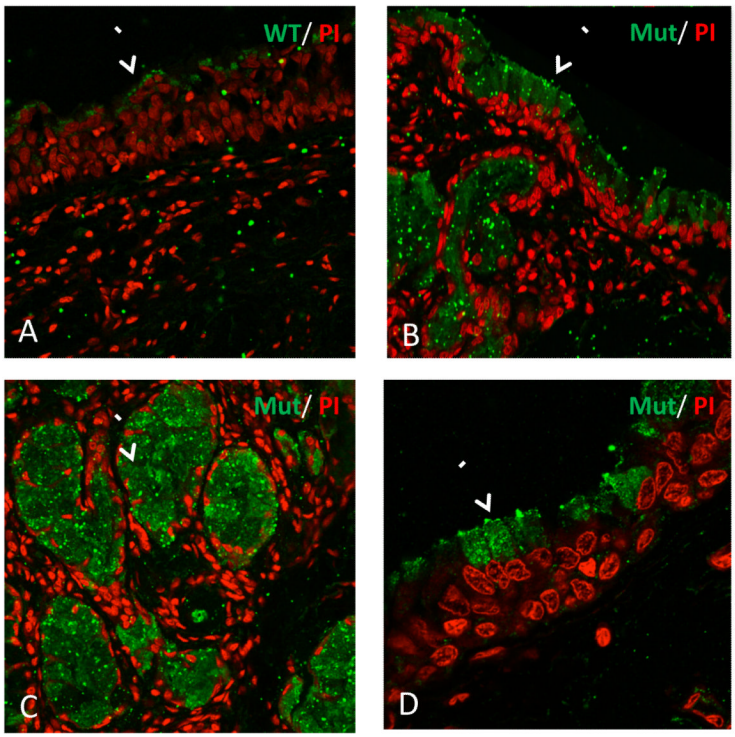

Figure 2. Staining of human trachea with wild-type and G228S A/Anhui/1/13 hemagglutinin

Human Paraffinized tissue sections were stained with recombinant HAs expressed and purified from 293 F cells. Specific staining by recombinant HA (in green), is also demarked by white arrows. The recombinant HAs were precomplexed with primary anti-His and Alexa fluor 488 tagged (green) secondary antibodies (for multivalent presentation) before adding to the tissue sections. The wild-type protein did not stain the trachea (A) as intensely as the G228S mutant HA (B). The G228S HA showed intense staining of the apical surface of the trachea (marked by white arrow). One key feature of the G228S protein is the staining of the submucosal gland (C) and the goblet cells (D) in the human trachea. The staining to goblet cells is similar to staining by other human-adapted influenza A virus HA (**Figure 3**). The tissue was counterstained with propidium iodide (PI) shown in red. Images A, B and C were captured at 25X magnification and Image D was captured at 63X magnification. See also Figure S2.

Table 1

Amino acid differences in the RBS of Aichi68 H3 an Anh13 H7 HAs

The key amino acids in the 130-, 140-, 150-, 220-loop and 190-helix that bind to the base and extension region of the glycan receptors in Aichi68 and Anh13 are shown. The 150 loop has amino acids (positions 153 and 155) that are involved in contacts with the sialic acid in the base region and those that are involved in binding to extension region. See also Figure S3.

	Base (Neu5Acα2->3/6Galβ1-)														**Extension (-4GlcNAcβ1-3Galβ1-4-)**														
HA	**130-loop**				**140-loop**	**220-loop**							**150-loop**		**150-loop**					**190-Helix**									
	131	133	136	138	145	219	221	222	225	226	227	228	153	155	156	157	158	159	160	186	187	188	189	190	192	193	194	196	
Aichi68	T	N	S	A	S	S	P	W	G	L	S	S	W	T	K	S	G	S	T	S	T	N	Q	E	T	S	L	V	
Anh13	R	N	T	A	S	A	P	Q	G	L	S	G	W	L	S	N	T	D	N	V	S	T	A	E	T	K	L	G	