12-14-2018

Characterization of mutations in the receptor binding site of influenza A viruses determining virus host, tissue, and cell tropisms using systems biology approaches

Feng Wen

Characterization of mutations in the receptor binding site of influenza A viruses determining virus host, tissue, and cell tropisms using systems biology approaches

By

Feng Wen

A Dissertation
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in Veterinary Medical Sciences
in the College of Veterinary Medicine

Mississippi State, Mississippi

December 2018

Characterization of mutations in the receptor binding site of influenza A viruses determining virus host, tissue, and cell tropisms using systems biology approaches

By

Feng Wen

Approved:

_____

Xiu-Feng (Henry) Wan
(Major Professor)

_____

G. Todd Pharr
(Committee Member)

_____

Alicia K. Olivier
(Committee Member)

_____

Ying Wang
(Committee Member)

_____

Larry Hanson
(Committee Member/Graduate Coordinator)

_____

Kent H. Hoblet
Dean
College of Veterinary Medicine

Name: Feng Wen

Date of Degree: December 14, 2018

Institution: Mississippi State University

Major Field: Veterinary Medical Sciences

Major professor: Dr. Xiu-Feng (Henry) Wan

Title of Study: Characterization of mutations in the receptor binding site of influenza A viruses determining virus host, tissue, and cell tropisms using systems biology approaches

Pages in Study 188

Candidate for Degree of Doctor of Philosophy

Influenza A viruses (IAVs) cause occasional pandemics and seasonal epidemics, thus presenting continuous challenges to public health. Vaccination is the primary strategy for the prevention and control of influenza outbreaks. The antigenicity matched high-yield seed strain is critical for the success of influenza vaccine. Currently, there are several limitations for the influenza vaccine manufacture: 1) the conventional methods for generating such strains are time consuming; 2) egg-based vaccines, the predominant production platform, have several disadvantages including the emergence of viral antigenic variants that can be induced during egg passage; 3) vaccine seed viruses often do not grow efficiently in mammalian cell lines. Previous studies suggested that mutations in the receptor binding site (RBS) that locates at the globular head of the HA1 can change IAVs' binding specificity, antigenicity, and yield and thus RBS would be an potential target for engineering vaccine seed strain. However, systematic analysis of the mutations on RBS affecting those viral phenotypes is lacking. Specifically, this dissertation has following aims: Firstly, we developed a novel method to rapidly generate high-yield candidate vaccine strains by integrating error-prone PCR, site-directed

mutagenesis strategies, and reverse genetics. The error-prone PCR- based reverse genetic system could also be applied to gain-of-function studies for influenza virus and other pathogens; Secondly, in this dissertation, we identified an Y161F mutation in the hemagglutinin (HA) that enhanced the infectivity and thermostability of virus without changing its original antigenic properties which would prompted the development of cell-based vaccines; Thirdly, the molecular mechanisms underlying host adaption of equine-origin influenza A(H3N8) virus from horses to dogs are unknown. This dissertation identified that a substitution of W222L in the HA of the equine-origin A(H3N8) virus facilitated its host adaption to dogs. This mutation increased binding avidity of the virus specifically to sialyl Lewis X motifs, which were found abundantly in the submucosal glands of dog trachea but not in equine trachea. To summary, this dissertation investigated the role of RBS in IAVs biology and expanded the current knowledge toward IAV vaccine strain engineering, IAV host adaption and evolution.

DEDICATION

This dissertation is dedicated to my parents and my wife for their endless love and support.

ACKNOWLEDGEMENTS

First, I have to thank my Ph.D advisor, Dr. Xiu-Feng (Henry) Wan, for his support during the past five years. I am so grateful to Dr. Wan for his scientific advices and many insightful discussions and suggestions in the dissertation. Dr. Wan is one of the smartest people I know. He is my primary source for getting scientific critical thinking in this dissertation. I hope that I could be as energetic, enthusiastic, and knowledgeable as Dr. Wan.

I am also grateful to my dissertation committee members, Dr. Larry Hanson, Dr. Todd Pharr, Dr. Olivier Alicia, and Dr. Ying Wang. I thank them for their insightful discussions and critical comments in my dissertation. I also thank them for their helpful career advices.

I thank all lab members in Dr. Wan's system biology lab. I was fortunate to be part of such a wonderful group during past five years. Many thanks to: Liping long, Hui Wang, Minhui Guan, Liyuan Liu, Kaijun Jiang, Kaitlyn Walters, Elizabeth Bailey, Jillian Harris, Bianca Quade, Alison Stokley, Sherry Blackmon, Karen Nguyen, Lucas Feguson, Tangwangvivat, Hao Chen and Drs. Hailiang Sun, Yulong Zhao, Jianli Xue, Nan Zhao, Yifei Xu, Brigitte Martin, Lei Li, Lei Zhong, Xiaojian Zhang, Yingzhi Lang, Chun-Kai Yang, Lei Han, Kaijian Luo, Junrong Luo, Jing Han, Yong Li, Huabin Cao, Wenjie You for their help during my Ph.D study.

I thank all the collaborators in this dissertation. I am so grateful to Dr. Richard Webby at Department of Infectious Diseases, St. Jude Children's Research Hospital, Memphis, Tennessee for his critical comments in Chapter II and Chapter III. I also thank him for providing the challenging viruses in the mice infection model. I thank Dr. Jim Cooley at Department of Population and Pathobiology, College of Veterinary Medicine, Mississippi State University for his insightful advices and supports on the pathologic examination in Chapter II and Chapter III. I thank Dr. Lei Li and Dr. Peng Wang at Department of Chemistry, Georgia State University for providing the glycan analogs and N-glycan microarray in Chapter III, Chapter IV and Chapter V. I am so grateful to Dr. Lei Li for his suggestions in the glycan microarray experiment. I thank Ding Liu at Department of Chemistry, Georgia State University for the glycan profiling analysis in Chapter IV. I thank Dr. Meng-Jung Chiang and Dr. Hang Xie at Laboratory of Respiratory Viral Diseases, Division of Viral Products, Office of Vaccines Research and Review, Center for Biologics Evaluation and Research, United States Food and Drug Administration (CBER/FDA) for their help in the fortebio binding assay in Chapter II. It was a great honor for me to collaborate with these distinguished researchers in the field of virology, pathology, and glycan biology.

At last, I would like to thank my soul mate and wife. I thank my wife’s support both academically and personally.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Influenza Virus classification

Influenza causes substantial threats to human and animal health globally. Influenza viruses, members of the *Orthomyxoviridae* family, are enveloped viruses with segmented single-stranded negative-sense RNA genomes. Influenza viruses are categorized into four known genera, namely A, B, C, and D, based on the antigenic difference of the nucleoprotein (NP) and matrix protein (MP)(1). Influenza A and B viruses are responsible for pandemic (type A), epidemic and seasonal outbreaks (type A and B) in humans, whereas influenza C and recently discovered influenza D viruses do not cause epidemic or pandemic outbreaks in humans (2).

In addition to humans, influenza A viruses (IAVs) can infect a wide range of animals including birds (i.e. both domestic poultry and wild birds), pigs, horses, dogs, and marine mammals (i.e. seals and whales). IAVs are classified based on the antigenic properties of the surface glycoproteins hemagglutinin (HA) and neuraminidase (NA). To date, there are 18 known HA subtypes (H1-H18) and 11 known NA subtypes (N1-N11) (3). However, only H1N1, H3N2, and H2N2 subtypes have been reported to cause pandemics (H1N1 for the 1918 and 2009 pandemics; H2N2 for the 1957 pandemic; and H3N2 for the 1968 pandemic) or seasonal outbreaks in humans (4).

IAVs' genome structure

The IAVs' genomes consist of eight negative-sense, single-stranded RNA segments with a total size of approximately 13,500 nucleotides (nt). The size of the segments (1-8) varies between 890 nt and 2,341 nt. Segment 1, 4, 5 and 6 of IAV each encodes one protein per segment: polymerase subunit 2 (PB2), hemagglutinin (HA), nucleoprotein (NP), and NA, correspondingly. Segment 2 encodes the polymerase subunit 1 (PB1). In some IAV strains, when the ribosomal scanning passes the main start codon, the segment 2 also encodes PB1-F2, an 87- amino acid protein with apoptotic activity (5), and PB1-N40, an N-terminal 39 amino acid truncated form of PBS with unknown function (6). The regulation of the translation initiation of PB1-F2 and PB1-N40 may be due to the interaction with the ribosome using the pseudoknot at nucleotides 65-126 in segment 2 of IAV (7). Segment 3 encodes the polymerase acidic (PA) subunit and PA-X, proteins made by a ribosomal frame shifting of PA to modulate the virulence of IAV(8). Newly identified N-truncated proteins of PA, PA-N155 and PA-N182, that are translated from the 11^{th} and 13^{th} in-frame AUG of PA were found universally expressed among IAVs (9). Compared to that with truncated PAs, the virus lacking truncated PAs shows an impaired replication efficiency in Madin-Darby canine kidney (MDCK) cells and have lower pathogenicity in mice; however, the detailed function of those PAs in the IAV life cycle remains unknown. The segment 7 of IAV encodes the matrix 1 (M1) protein and the matrix 2 (M2) iron channel through RNA alternative splicing (10). Segment 8 encodes the interferon antagonist protein NS1 and the NS2 through mRNA splicing (11, 12).

IAV virion structures

IAVs have a lipid bilayer envelope and could exhibit a spherically morphology on the order of 80-120 nm in diameter or a filament morphology with a length over 1 μm under electron microscopy (13). The outer layer lipid envelope of an IAV is spiked with HA and NA proteins in a ratio of approximately four to one, and a small number of M2 proteins serving as ion channels (14). M1 proteins play an important role in determining the filamentous or spherical morphology of influenza viruses (15-17). It has been suggested that the increased levels of influenza virus production by spherical strains increase the efficiency of influenza virus transmission (18). The inner layer of the lipid envelope is attached with matrix 1 (M1) protein, which is associated with the nuclear export protein 2 (NEP or NS2) and the ribonucleoprotein (RNP) complex and is responsible for the association of the RNP complex with the inner layer of the envelope. Each IAV virion has eight well-organized ribonucleoprotein (RNP) complexes, each consisting of one of the eight negative sense RNA segments wrapped on the nucleoprotein (NP) and RNA-dependent RNA polymerases, which contains two polymerase basic (PB2 and PB1) and a polymerase acidic (PA) subunits. The nuclear export protein 1 (NS1) and NS2 are encoded from the same RNA through alternative splicing in the infected cells. NS1 functions to inhibit the 3′-end processing of host pre-mRNAs (19) and NS2 functions in maintaining the attachment of the RNP complex with the inner layer of the envelope.

Biology of HA

HA plays an important role in the IAV life cycle. Two key functions of HA are the attachment of virus to host receptors and membrane fusion during virus entry. In

addition, HA is one of the major viral antigens and thus a major target for host neutralizing antibodies induced by infection and vaccination.

HA is a surface glycoprotein that presents as a trimer with three identical monomers spiked on the virus membrane. The precursor protein (HA0) undergoes signal peptide cleavage and *N*-link glycosylation, which result in a monomer with about 549 amino acid residues and a molecular mass of approximately 60 KDa (20). Each HA monomer consists of two polypeptides, HA1 contains the receptor binding domain and the major antigenic sites, and HA2 contains the membrane-proximal stem and fusion peptide, which are cleaved by host cellular proteases. The cleavage of the HA precursor is important for the virus infectivity (20). The seasonal IAV and low pathogenic avian influenza A virus (LPAIV) usually have a mono-basic cleavage sites (e.g. PEKQTR/GLF), and their HA proteins are cleaved by a few cellular trypsin-like proteases, which are usually produced in the intestinal and the respiratory systems. In contrast, the highly pathogenic avian influenza A virus (HPAIV) typically has a stretch of basic amino acids at the cleavage site (e.g. PQRESRRKK/GLF), which can be cleaved by common cellular proteases such as ubiquitous furin and PC6 (21). This is one of the reasons why, compared with the LPAV, the HPAIV has the potential to infect multiple tissues in the body and therefore causes systemic rather than localized infections.

The initial step of IAV infection is the binding of the HA to the sialylated glycan receptors on host cells. The receptor binding site (RBS) located at the hypervariable globular head of the HA is responsible for interaction with the glycan receptors. Each HA monomer contains a RBS which forms a shallow pocket at the top of the globular head and consists of 130-loop, a 150-loop, a 190 helix, and a 220-loop.

Receptors for IAV

Glycans expressed on the surface of mammalian cells play important roles in many biological processes which include serving as recognition binding targets for a variety of viruses, bacteria and parasites (22). Sialic acids (sias) (Figure 1), derivatives of neuraminic acid with a nine-carbon sugar backbone, were first discovered by Gunnar Blix (23) and then found wildly distributed on mammalian tissues in glycoproteins and gangliosides. Sias usually occupy the terminus of glycoproteins and glycolipids as components of oligosaccharide in the mucins. The agglutination of erythrocytes by IAV was first reported when the erythrocytes were agglutinated in the allantoic fluid of the infected chicken embryos (24). The receptors for the influenza viruses were found to be sias when using erythrocytes treated with *Vibrio cholerae* neuraminidase (VCNA) (25). Later, it was found that the linkage of sias to galactose and the types of sias were key determinates for host range of IAVs (26, 27); the avian IAV prefers sias that are linked to galactose in an α2,3 linkage (SAα2,3 Gal) whereas the human IAV prefers sias that are linked to galactose in a α2,6 linkage (SAα2,6 Gal) (27, 28) (also see the section glycosidic linkages).

Compared to wild-type Chinese hamster ovary (CHO) cells, the IAV infection of *N*-linked glycoprotein deficient CHO cells was reduced, suggesting the importance of *N*-glycans for the entry of IAV (29). On the other hand, it was found that the *N*-linked sialosides are not required for the entry of IAV since the IAV can infect CHO cells with immature glycans that cannot be sialylated (30). It was reported that the H3N2 IAV could infect a glycosphingolipid-deficient mouse skin fibroblast mutant cell line, suggesting that gangliosides are not essential for the IAV entry and infection (31). The glycosylated

H1N1 viruses could infect sialic acid-deficient Lec2 Chinese hamster ovary cells that express human C-type lectins, while the PR8 H1N1 strain which bears low levels of mannose showed insufficient infection. Taken together, those studies suggest that IAV entry and attachment can occur in a sialic acid-independent manner and the host cell surface lectins may serve as a target for IAV. Nevertheless, the specific roles of the *N*-glycans and *O*-glycans in the process of virus-cell attachment and entry remain largely unknown.

The following sections will discuss the terminal structures, linkage, and substructures of sias and the roles of the structurally different sias in influenza virus infection.

Neu5Ac and Neu5Gc

N-acetylneuraminic acid (Neu5Ac) and *N*-glycolylneuraminic acid (Neu5Gc) that have modifications at the 5-carbon position are two of the most commonly found types of sias on the mammalian cells. Neu5Ac was shown to be a common receptor for IAV isolates from human, avian, and mammals (32). Neu5Gc has been reported to be present in horses (26), dogs (33), pigs (34), and mice (35), but not in humans (36, 37). The loss of Neu5Gc in humans was due to an inactivating mutation on the CMAH gene encoding the cytidine monophosphate-N-acetylneuraminic acid hydroxylase (CMAH), which converts Neu5Ac to Neu5Gc (38). The 92 bp deletion mutation in exon 6 of the CMAH gene, which causes the loss of Neu5Gc and excess of Neu5Ac, was estimated to to have happened in hominids about 2.9 million years ago (39). The loss of Neu5Gc in humans may help explain the susceptibility or resistance of humans to certain pathogens including influenza viruses (40). Similarly, the ferrets have been shown to lack Neu5Gc due to an

ancient mutation in the CMAH gene shared by several members in the order of carnivora (41). High-performance liquid chromatography, and mass spectrometry analysis of chicken tissues suggested that chickens cannot synthesize Neu5Gc whereas the small amount of Neu5Gc detected in eggs may come from the diet (42). Another report suggested a low level of Neu5Gc detected in both trachea and intestines of chickens and pigeons by high-performance liquid chromatography (43); however, the Neu5Gc glycan could not be detected in chicken trachea by histological immunostaining (26). It is likely that these conflicting observations on the level of Neu5Gc could be affected by the sensitivity of the detection methods used in the experiments.

It has been suggested that Neu5Gc plays an important role in the infection of equine origin H7N7 viruses and duck origin H3N2 viruses (44, 45). Takahashi *et al* (45) showed that expression of Neu5Gc on human epithelial cancer cells reduced the infectivity of IAV with Neu5Gc binding ability, suggesting that Neu5Gc may work as a decoy receptor but not a functional receptor for IAV. Yang *et al* (33) reported that Neu5Gc was detected in canine trachea tissues but not in chicken trachea tissues, suggesting that the Neu5Gc may play an important role in the interspecies transmission of avian-like H3N2 IAV from avian to canine species. Despite of these reports, the functions of Neu5Gc as a receptor for IAV and the effects of Neu5Gc on IAV tissue and host tropisms are not fully understood.

Other modified forms of sias

In addition to Neu5Ac and Neu5Gc, the primary Sia forms can be modified in the Golgi and generate more than 30 variant types of sia, and those modifications include addition of acetyl, sulfate, lactoyl, or methyl groups at different positions (46, 47). For

example, various forms of sias with acetyl groups *O*-linked to carbon 4, 7, 8, or 9 were found to differ in expression among tissues and hosts (48-50).

Influenza C virus was reported to recognize 9-*O*-acetylated sialic acid as a receptor for cell attachment and entry (51). A more recent study showed that 9-*O*-acetylated sialic acid also serves as the receptor for influenza D viruses (52).

Glycosidic linkages

Although IAVs recognize the terminal sias as their receptor for the viral entry, the binding specificity of IAV was found closely related to the linkage that connects the sias with the other structure of the receptor. A number of studies have suggested that human adapted IAV prefers SAα2,6 Gal, whereas the avian IAV prefers SAα2,3 Gal(28, 53, 54).

The discovery of the preference to α-2,3 or α-2,6 linkage was established through the hemagglutination assays of human and animal IAVs of different host origin to enzymatically modified erythrocytes (28). The proportion of the α-2,3 to the α-2,6 linkage of the erythrocytes varies depending on their source species and this determines the phenotypes in the hemagglutination. The chicken erythrocytes contain a roughly an equal number ratio of SAα2,6 Gal to SAα2,3 Gal (55), whereas the turkey and guinea pig erythrocytes contain more SAα2,6Gal than SAα2,3Gal (56, 57). Horse erythrocytes predominately contain SAα2,3 Gal, which makes it an ideal choice for the hemagglutination inhibition (HAI) assay with H5 strains (26, 56).

In addition to the variations in erythrocytes from different hosts, the distribution of α-2,3 and α-2,6 sias also varies among tissues and among hosts. By lectin staining, the human upper respiratory track epithelium cells primarily express α-2,6 sialylated glycans, whereas the lower respiratory track epithelium cells express both α-2,6 and α-

2,3 sialylated glycans (53, 54, 58). In contrast, duck intestines express predominantly α-2,3 sialylated glycans whereas pig respiratory tract express both α2-6 and α2-3 sialylated glycans, and therefore pigs are hypothesized to be a "mixing vessel" for human and avian IAVs (59, 60). However, it was also demonstrated that chicken and quail intestines express both α2-3 sialylated glycans and α2-6 sialylated glycans (61). With the development of glycomics, a wide spectrum of both α2-3 sialylated glycans and α2-6 sialylated glycans were found on human lung, bronchus and nasopharynx by mass spectrometry in a recent study (62). Similarly, the respiratory tissue of ferret was found to heterogeneously expressing both α2-3 sialylated glycans and α2-6 sialylated glycans by lectin staining and mass spectrometric analysis (63). These recent findings suggests that the cell and tissue tropisms of IAV cannot be determined by glycosidic linkage alone.

Substructures affect viral receptor binding

As described above, increasing evidence has shown that the binding specificity of IAVs is more complex than the glycosidic linkages. IAVs may use a wide range of sias to infect host cells. Modifications such as acylation, methylation, sulfation, fucosylation, and phosphorylation can occur at different locations within the backbone structure and thus may alter its biological functions, including the receptor for IAVs (64). In 1997, Gambaryan *et al* (27) reported that the non-egg adapted human H1, H3 subtype IAV and influenza B virus had strong binding affinity to Neu5Acα2-6Galβ1-4GlcNAc (6′SLN) but not Neu5Acα2-6Galβ1-4Glc (6′SLN). Later, Gambaryan *et al* showed that duck IAVs prefer receptors which possess a β, 1–3 linkage rather than a β, 1–4 linkage between Neu5Acα2-3Gal-disaccharide and penultimate *N*-acetylhexosamine residue (GlcNAc)

(65). Furthermore, the fucosylation and sulfation of GlcNAc had negative and low effect, respectively, on its affinity to duck IAVs (65). Similar binding patterns were also observed in H1, H2, H3, H4, and H11 HAs from duck IAVs(44). In contrast, gull IAVs preferentially bind to the α, 2-3 sialic acid receptors with a fucosylated GlcNAc whereas the chicken and mammalian IAVs preferentially bind to the SAα2,3 Gal receptors with a sulfo group at position 6 of GlcNAc via a β, 1–4 linkage (65). Gambaryan *et al* (66) reported that the H5, H6, H7, and H9 subtypes of IAVs from terrestrial poultry differ from ancestral duck viruses by enhanced binding to fucosylated and/or SAα2,3 Gal terminated receptors, suggesting that the 6-sulfo sialyl Lewis X is the common receptor recognized by those IAVs from terrestrial poultry. Those studies suggested the substructures of sialic acid terminated receptors play important roles in the receptor binding specificity of IAVs.

Fucosylation is one of the most common modifications to *N*-glycans, *O*-glycans, and glycolipids. Unlike duck IAVs, chicken H5N2 IAV was found preferentially to bind to fucosylated SAα2,3Gal receptors, which were detected on the epithelial cells of chicken trachea but absent from duck colon (67). This observation helps to explain why chickens were not experimentally infected with viruses isolated from ducks. The equine H3N8 IAVs were also reported to preferentially bind to Neu5Acα2-3Galβ1-4 (6-O-HSO_3) GlcNAcβ (Su-3′SLN) and Neu5Acα2-3Galβ1-4 (Fucα1-3)(6-O-HSO_3)GlcNAcβ (Su-SLe^x), suggesting equine H3N8 IAVs may originate from terrestrial poultry. Residues at positions 222 and 227 were identified to be associated with the differences of H5 avain IAVs binding to fucosylated α2, 3 sialic acid receptors (68). Although the HA of duck H5N2 IAV preferentially bind to non-fucosylated SAα2,3 Gal receptors, the

virions were able to bind to both non-fucosylated and fucosylated receptors through the NA (68). Taken together, the glycan receptors for IAVs are far more complex than the α, 2-3 and α, 2-6 linkages and the role of substructures of sialic acid receptors for IAVs host and tissue tropism is underrated.

Structure basis for binding specificity

The switch of their binding specificity from avian-like (SAα2,3 Gal) to human-like (SAα2,6Gal) receptors for avian IAVs was considered to be associated with the adaption of IAVs from avian to humans. The SAα2,3 Gal linkage has limited conformational flexibility and forms a cone-like topology (58). Conversely, the presence of C6-C5 bond provides additional conformational flexibility to SAα2,6 Gal linkage and enables it to adopt an umbrella-like topology (58).

The RBS of IAV consists of four structural features, 130 loop, 150 loop, 190 helix, 220 loop, and the conserved amino acids at positions 98, 153, 183, and 195 form a hydrogen-bonded network that serves as the basis of the binding pocket (69, 70).

The residues at positions 135, 136, 137, and 145 form major interactions with the sias moiety of the Neu5Acα2-3Gal linkage receptors. In contrast, the residue at positions 137, 145, 190, 226, and 228 forms the predominant interactions with the sias moiety of Neu5Acα2-6 Gal linkage receptors (58, 71).

Subtle changes of residues in the 220 loop affect the virus interactions to IAV receptors with different glycosidic linkages. For example, the avian IAVs possess Q226 preferentially binding to SAα2,3 Gal whereas human IAVs possess L226 with referential binding to SAα2,6 Gal receptors. The Q226 would not interact with the SAα2,6 Gal, whereas L226 allows to wide the site and forms a hydrogen bond with the α2-6 linked

sias (72-74). Similarly, a G225D mutation can completely switch binding specificity of an avian H6N1 virus from SAα2,3 Gal to SAα2,6Gal (75). Residues at position 186, 187, 190, 193, and 226 were shown to interact with the Gal linked to sialic acid via a α2-3 linkage, whereas residues at position 190, 222, 225, and 226 were shown to interact with the Gal linked to sias via a α2-6 linkage (58).

The residues at the 190 helix of the RBS can also affect the glycan binding specificity of IAVs. For example, the pandemic strain A/New York/1/1918 differs from the consensus avian H1N1 virus sequence by a E190D, which enables the virus to bind to both SAα2,3 Gal and SAα2,6Gal glycans (76, 77). Recently, de Vries *et al* (78) reported that three mutations (V186K/G, K193T, and G228S) switch binding preference of H7N9 IAV from avian like SAα2,3 Gal to human like SAα2,6Gal. A number of other amino acid residues such as 189, 194, 198, 211, 216, and 222 may also affect the architecture of the receptor-binding site and change the virus binding specificity and intensity to sialylated glycans (79-81).

NA and viral glycan binding

The NA polypeptide chain comprises 470 amino acid residues and consists of cytoplasmic, transmembrane, and stem domains. The NA of IAV belongs to the exo-sialidase family and cleaves the α-ketosidic linkage between synthetic substrates and the terminal sialic acid residues in oligosaccharides, glycoproteins, glycolipids, or colominic acids (82). The NA of IAV resembles a mushroom shape on the virion surface, with a molecular mass of 60 KDa for the monomer and 240 KDa for the tetramer (82). NA is indicated to be relevant in several stages during the IAV life cycle. First, NA helps the attachment of the virus to airway epithelium by cleaving sias from respiratory tract

mucins (83); second, NA plays an important role in the fusion of viral and cell membranes as the HA and NA activity needs to be balanced to allow effective infection (84, 85); third, NA promotes the budding of newly formed virions by cleaving sias from cell surfaces and thus preventing their aggregation (84).

Although the functions of two surface glycoproteins HA and NA are to some extent contradictory, sialic acid binding activity was also reported for the NA protein. Laver *et al* (86) first reported that the NA of subtype N9 subtype of NA has hemagglutination activity and antibody to either HA or NA alone was incapable of inhibiting hemagglutination by the virus. Later, Webster *et al* (87) suggested that the hemagglutinin activity of N9 was associated with a second Neu5Ac binding site away from the catalytic site. Exchange of amino acids in the loops at positions 368-370 and positions 399-403 of N9 into those corresponding positions at the N2 by site-directed mutagenesis made N2 acquire the hemagglutination activity as shown by N9 (88). Furthermore, it was suggested that the haemadsorbing of N9 is not sensitive to the neuraminidase activity of N9 because the erythrocytes agglutinated by N9 cannot be released by N9 neuraminidase activity (89). Similarly, N1 of A/FPV/Rostock/34 virus (H7N1) expressed in SF9 cells showed hemadsorption activity cannot be blocked by 2-deoxy-2,3-dehydro-N-acetylneuraminic acid, suggested that the N1 has a receptor binding site distinct from the catalytic site (90). The haemagglutinating activity of N1 was abolished when the amino acids on the two loops at putative binding sites were exchanged (90). However, the attempts to resolve the sialic acid binding site by X-ray diffraction were not successful until 1997, when Varghese *et al* (91) resolved the X-ray structure of a complex of N9 from A/tern/Australia/G70C/75 with Neu5Ac and found the

Neu5Ac binds to N9 in the chair conformation and the residues that interact with the N9 second binding site were the most conserved among all avian strains, but not in humans and swine strains.

Mutations near the active site can change the binding property of NA to sialic acid receptors. Lin *et al* (92) reported that human H3N2 isolates circulating between 2005 to 2009 has a D151G substitution, which enables the viruses to agglutinate erythrocytes in an oseltamivir-sensitive manner. The D151G mutation of human H3N2 viruses was later found to enabled NA binding to both 3′ sialyllactosamine (3'SLN) and 6′ sialyllactosamine (6'SLN), with even higher affinities than the corresponding affinities of HA to those sialosides (93). The introduction of the D151G mutation into both N1 and N2 reduced the neuraminidase activity and enabled preferential binding to α, 2-3 sialic acid receptors (93). Gulati *et al* (94) reported that the N2 of human H3N2 viruses isolated in 2006, 2010 and 2012 could bind to SAα2,3Gal receptors. Moreover, the entry of human H3N2 viruses isolated in 2010 and 2012 into Madin-Darby canine kidney cells was inhibited by oseltamivir (94). Interestingly, a G147R mutation in N1 enables a HA binding-deficient mutant to use N1 as the receptor binding protein (95). The G147R mutation of N1 occurs at low frequencies in human pandemic and seasonal H1N1 strains and this mutation did not impair the replicative capacity of virus in cell culture or mice (96). Recently, Donald *et al* (97) reported that the N9 of human H7N9 virus has an active haemadsorption site and can bind to sialic acid receptors, particularly human-like α SAα2,6Gal receptors, which enhanced the overall binding affinities of human H7N9 viruses to human-like receptors. Those studies suggested that NA could play important

roles in the virus attachment, entry, and evolution. However, the receptor binding function of NA in IAV biology is underrated and needs further research.

Knowledge Gaps

NA as a receptor binding protein

It is widely accepted that the functions of two surface glycoproteins HA and NA are to some extent contradictory: HA interacts with sialic acid receptors on host cells and mediates the virus attachment and entry whereas NA cleaves the α-ketosidic linkage between the terminal sialic acid residues and the remaining substructures. However, as aforementioned, increasing evidence suggest that NAs, especially N1, N2 and N9, have a receptor binding site distinct from or close to the catalytic site. Furthermore, the NA may even have higher affinity to sias than HA, and thus could broaden the spectrum of receptors that virus can bind to. Outstanding questions remain: 1) What is the role of NA in the attachment and entry of IAVs? 2) What are the physical interactions between HA and NA on the virion surface? 3) How do IAVs preserve a functional balance of HA and NA activities and maintain a good fitness during virus evolution?

Quantification of the binding affinities between virus and glycans

Currently, the conventionally used methods to characterize the receptor binding properties of HA/NA proteins of IAV include: agglutination of erythrocytes modified with desialylation and resialylation (98), solid-phase enzyme-linked assays (99), glycan microarray (100, 101), and bio-layer interferometry (102, 103). The erythrocytes are modified by desialylation and resialylation to express a specific glycosidic linkage and then tested with their agglutinations with viruses. For the solid-phase enzyme-linked

assays, biotinylated glycans are adsorbed on solid-phase ELISA plates coated with streptavidin and then viruses incubated on plates followed by detection with specific antibodies. The bio-layer interferometry uses streptavidin sensors to capture the biotinylated glycans followed by the binding to virus. Then the interference patterns of the sensor surfaces are recorded in real time. The glycan microarrays that have large number of synthetic sialosides printed onto NHS-coated glass slides were demonstrated to be useful in characterizing binding specificity and intensity of either proteins or whole virus particles to a large set of glycans.

However, the methodologies described above have obvious drawbacks (Table 1), which limit their applications in global influenza surveillance. For example, the erythrocytes agglutination assay cannot detect the binding specificity beyond the glycosidic linkages and the substantial variability is observed among batches of RBS. The solid-phase enzyme-linked assays and bio-layer interferometry cannot detect the binding properties of viruses in a high-throughput manner. Moreover, the glycan microarray and bio-layer interferometry require the purification of viruses, which is considerably time-consuming, and therefore not suitable for use in large-scale screening. The propagation of viruses in eggs or on cells during virus purification may introduce mutations on the receptor binding site, leading to different results compared to the clinical isolates. Lastly, the glycan analogues used in those arrays are not similar to those presented on the epithelium of the respiratory tract. A high-throughput method for large-scale analysis of the binding properties of IAV, which would be cost effective, simple, highly sensitive, label free and without the need for virus purification, is needed for the global influenza surveillance.

Predicting the binding properties of IAV

Computational models for quantifying antigenic distance of IAV and thus for improving influenza vaccine design have been reported by several groups (104-108). A computational model for predicting the binding specificity and intensity of IAV will assist in vaccine strain selection and global influenza surveillance. Studies have reported the calculation of HA-receptor binding by integrating molecular modeling, free energy simulation, *ab initio* fragment molecular orbital, and molecular dynamics (109-115). However, those studies lack a systematic evaluation and also neglected the role of NA in the binding properties of IAV. Moreover, it has been difficult to assess the effect of hallmark residues on the binding properties of circulating IAVs (116). In summary, there is an urgent need to develop an effective nucleic acid sequence-based and robust computational prediction models for quantifying IAV binding specificity and intensity to sialic acid receptors.

Substructures of sialic acid receptors affect tissue and host tropism

As mentioned above, the substructures of sialic acid receptors such as the linkages between Neu5Acα2-3Gal-disaccharide and GlcNAc, and the fucosylation and sulfation of GlcNAc affect virus binding properties as well as tissue and host tropism. Modification of GlcNAc is always observed in *O*-glycans. The receptors for human IAV in ferrets were reported to be *O*-linked sialylated glycans, which are predominantly distributed in the submucosal glands. The infection of such cells facilitates the efficient airborne transmission of virus by easily making the virus encapsulated into respiratory droplets (117). Entry of the SAα2,3Gal binding preference IAV was reported to require host fibronectin, suggesting that host proteins may also be involved in the receptor

binding process of IAV, which remains largely unknown. Outstanding questions are: 1) What is the distribution and intensity of modified sialic acid receptors among tissues and among hosts? 2) What is the role of those substructures in the adaption and host switch of IAVs? 3) What is the role of *O*-glycans in the IAV attachment and entry? 4) Besides *N*-glycan and *O*-glycan, what other kinds of host proteins are involved in the binding of IAV to host receptors and how are they regulated?

Objectives of this Dissertation

The hemagglutinin of IAVs binding to host carbohydrate receptors is one of the key determinants for viral host tropism. The residues forming the receptor-binding pockets of the hemagglutinin are known to affect the receptor binding affinity and subsequently the replication ability and transmissibility of IAV. However, most reports have been focusing on individual receptor binding sites (i.e. residues 226 and 228) and their impacts on two types of sialic acid receptors (i.e. avian-like receptor α 2,3-linked sialic acid (SAα2,3Gal) and human-like receptor α 2,6-linked sialic acid (SAα2,6Gal). There is lack of a systematic characterization of residues, especially, the synergetic effects of these residues on receptor binding and specificity. This project proposes to identify the mutations across the receptor binding pockets affecting receptor binding properties by developing and applying a novel and high-throughput error-prone PCR (epPCR)-based reverse genetic system, and to quantify and predict the receptor binding avidity and replication ability in target cells using machine learning. The construction of the computational model would link the virus genotype with the phenotypes such as receptor binding and replication properties. The computational model would help to

speed up the vaccine seed strain selection, field strain surveillance and to expand our knowledge on natural history of IAV across hosts.

We hypothesize that 1) *In vitro* breeding of a diverse influenza virus variant library by epPCR would enhance the process for generating high-yield strains; 2) Mutations in the receptor binding sites of IAV would change the binding and replication properties of the virus. With the use of epPCR-based reverse genetic system that we developed, we can generate a large scale of receptor binding site mutants, from which the high yield mutants would be selected; 3) The mutations in the HA proteins of equine H3N8 IAV affect viral binding properties and are important in adapting virus from horse to dogs, and dogs would have a unique set of glycan receptors determining canine influenza virus' host tropism; 4) Mutations in the receptor binding sites of IAV hemagglutinin would affect the influenza vaccine-associated characteristics such as receptor binding properties, replication, and antigenicity of the virus. There would be specific sub-regions in the receptor binding pockets of IAV hemagglutinin determining 1) virus binding to specific glycan receptors; 2) virus replication ability in hosts; 3) virus antigenicity, and these regions may not be same.

The objectives of this dissertation are to: 1) Develop and validate a novel and high throughput error-prone PCR (epPCR)-based reverse genetics system for rapid mutant generation and phenotypic characterization; 2) Characterize the molecular mechanisms determining high-yields and thermostability of influenza vaccines; 3) Characterize the molecular mechanisms affecting H3N8 equine IAVs adaption to dogs; and 4) Develop and validate a computational model for IAV glycan receptor binding specificity that can help predict the receptor binding specificities and antigenic properties, as well as help

quantify the vaccine yields by using genomic sequences. This model will be useful in optimizing selection of effective (i.e. high yield and antigenically-matched) influenza vaccine candidates directly from clinical samples. The genetic signatures to be identified will help optimize vaccine seeds for vaccine production.

Table 1. Comparison of methods to characterize the binding properties of HA/NA proteins of IAV

	Technique			
	Modified erythrocytes	Solid-phase ELISA	Glycan microarray	Bio-layer interferometry
Label free	Yes	Yes	No	Yes
Sensitivity	Low	High	High	High
Time	Hours	Days	Days	Hours
Quantitative	Poor	Good	Moderate	Good
High throughput	No	No	Yes	No
Cost/Sample	Moderate	Moderate	High	High
Label free	Yes	Yes	No	Yes
Purify of virus	No	Depends[a]	Yes	Yes
Real-time results	No	No	No	Yes
Reproducibility	Poor	Good	Moderate	Good
Native glycans	No	No	No	No

a. It depends on the grow efficiently of virus in chicken embryonic eggs or mammalian cell lines.

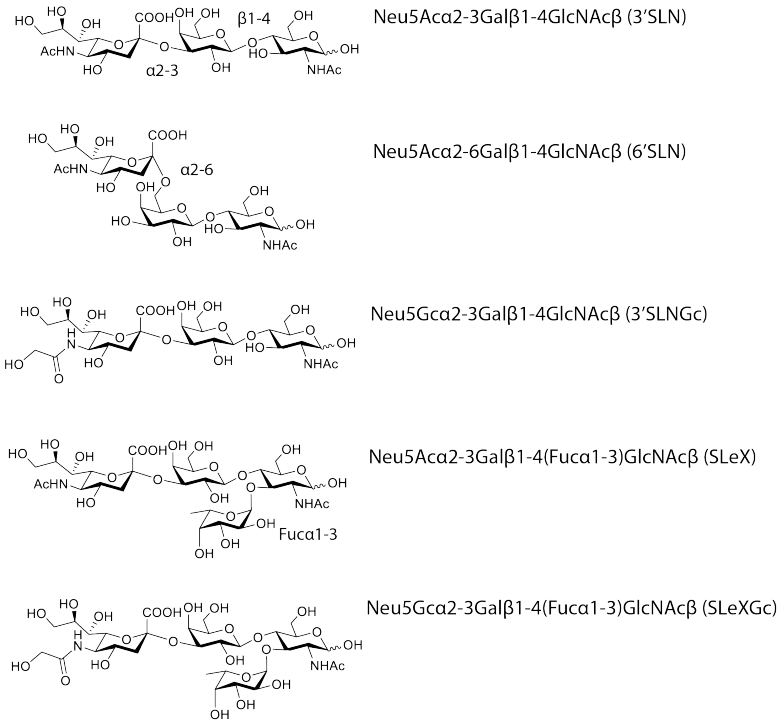


Figure 1. Representative glycans structures for influenza receptors.

CHAPTER II

ERROR-PRONE PCR-BASED MUTAGENESIS STRATEGY FOR RAPIDLY GENERATING HIGH-YIELD INFLUENZA VACCINE CANDIDATES

Vaccination is the primary strategy for the prevention and control of influenza outbreaks. However, the manufacture of influenza vaccine requires a high-yield seed strain, and the conventional methods for generating such strains are time consuming. In this study, we developed a novel method to rapidly generate high-yield candidate vaccine strains by integrating error-prone PCR, site-directed mutagenesis strategies, and reverse genetics. We used this method to generate seed strains for the influenza A(H1N1)pdm09 virus and produced six candidate strains. The yield of the candidate strains in eggs was up to 64.6-fold higher than that for the wild-type strain. We used a mouse model to assess the efficacy of two of the six candidate strains as a vaccine seed virus: both strains provided complete protection in mice against lethal challenge, thus validating our method. Results confirmed that the efficacy of these candidate vaccine seed strains was not affected by the yield-optimization procedure.

Introduction

Influenza viruses cause occasional pandemics and seasonal epidemics, thus presenting continuous challenges to public health. Vaccination is the primary strategy used to slow and stop transmission of the virus and to reduce the effect of the disease. A high-yield vaccine seed strain is required for vaccine manufacturing and, thus, is key to a successful influenza vaccination program. However, it is not uncommon that strains recommended as influenza vaccine strains by the World Health Organization (WHO) have low yields in chicken embryonic eggs (hereafter referred to as eggs) (118). Low-yield strains must be engineered into high-yield strains, a process that can be time consuming and, thus, significantly delay vaccine production. Almost three months were required for the WHO collaborative laboratories and vaccine companies to generate suitable vaccine strains to produce vaccine against the 2009 pandemic influenza virus, A(H1N1)pdm09 (119, 120). Without this delay, the vaccine could have lessened the effect of the pandemic's second wave of infection and reduced the number of associated deaths. Thus, the ability to rapidly generate high-yield vaccine strains is a critical factor in preventing and controlling influenza outbreaks.

Two common strategies have been used to generate influenza vaccine seed viruses: 1) co-infection of a cell or an egg with two viruses, one with high-yield features and another with antigenically matched hemagglutinin and neuraminidase (121, 122); and 2) reverse genetics to generate a virus with antigenically matched hemagglutinin and neuraminidase, and internal genes from another virus as backbone, which help improve yields (123-126). During the past decades, the co-infection method has been used widely to generate vaccine seed viruses. The disadvantage of co-infection is that the

reassortment events cannot be well controlled and may lead to viruses with undesired traits (127, 128). Thus, instead of co-infection, egg adaption is often used to improve seed virus growth in eggs (129-132). In contrast to the co-infection method, the reverse genetics method can define the genomic constellation (123-126). However, it is difficult to generate a high-yield seed strain if the hemagglutinin and neuraminidase are not compatible with the other six gene segments from the backbone virus, and this method does not readily allow for selecting different combinations of gene segments that may result in higher yield viruses. In recent years, a recombinant technique has been used to increase the yields of influenza vaccine candidates (133-138). These reassortants, which have chimeric hemagglutinin or neuraminidase, had considerably more seed virus growth in eggs or Madin-Darby canine kidney (MDCK) cells. However, these methods usually require weeks to generate a desired strain with high yields.

In this study, we developed a novel error-prone PCR (epPCR)–based mutagenesis approach to rapidly generate high-yield influenza vaccine strains.

Materials and Methods

Cells and viruses

MDCK cells and human embryonic kidney (293T) cells (both from American Type Culture Collection, Manassas, VA) were used for propagation and culture of influenza virus. The cells were maintained at 37°C with 5% CO_2 in Dulbecco's Modified Eagle Medium (GIBCO/BRL, Grand Island, NY) supplemented with 5% fetal bovine serum (FBS) (Atlanta Biologicals, Lawrenceville, GA), penicillin–streptomycin, and amphotericin B (GIBCO/BRL, Grand Island, NY). The wild-type strain of

A/California/04/09 (H1N1) virus (CA/04) was used in the mutant library construction (Figure 2); mouse-adapted CA/04 (139) was used for challenge experiments.

All viruses generated by reverse genetics were propagated in MDCK cells and cultured at 37°C with 5% CO_2 in Opti-MEM medium (GIBCO/BRL, Grand Island, NY) supplemented with 1 μg/ml of TPCK (*N*-tosyl-L-phenylalanine chloromethyl ketone)–Trypsin (Sigma-Aldrich, St. Louis, MO), penicillin–streptomycin, and amphotericin B (GIBCO/BRL, Grand Island, NY). Virus titers were determined by 50% tissue culture infectious dose ($TCID_{50}$) in MDCK cells.

Primers

Four primers were used in this study: 1) 130loop_F, 5'-TCA TGG CCC AAT CAT GAC TCG AAC-3'; 2) 190helix_F, 5'-TGG GGC ATT CAC CAT CCA TCT ACT-3'; 3) 190helix_R, 5'-AAC ATA TGT ATC TGC ATT CTG ATA-3'; and 4) 220loop_R, 5'-TAG TGT CCA GTA ATA GTT CAT TCT-3'. Primers 130loop_F and 190helix_R were used to amplify the sequence covering the 130-loop and 190-helix of the receptor-binding site in the hemagglutinin gene of influenza virus. Primers 190helix_F and 220loop_R were used to amplify the sequence covering the 190-helix and 220-loop of the receptor-binding site in the hemagglutinin gene of influenza virus.

epPCR-based mutagenesis method to generate hemagglutinin mutants of influenza virus

We used the GeneMorph II Random Mutagenesis Kit (Agilent Technologies, Santa Clara, CA) according to the manufacturer's instructions to perform epPCR. The resulting products were used as primers to perform site directed mutagenesis. The epPCR amplification mixture contained 17.75 μl of water, 5 μl of 10× buffer, 1 μl of 2.5 mM

deoxyribonucleotide triphosphates, 1 μl of 10 μmol primer (each), 2 μl of hemagglutinin plasmid of CA/04 (10 pg/μl), and 1 μl of Mutazyme II DNA polymerase (Agilent Technologies, Santa Clara, CA). The parameters of the epPCR were as follows: one cycle at 95°C for 5 min, followed by 30 cycles at 94°C for 1 min, 50°C for 1 min, and 72°C for 2 min, and then one cycle at 72°C for 10 min. Site-directed mutagenesis was performed by using the QuikChange II Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA) according to the manufacturer's instructions. In brief, the PCR mixture included 36 μl of water, 5 μl of 10× buffer, 1 μl of 2.5 mM deoxyribonucleotide triphosphates, 2 μl of epPCR product (25 ng/μl), 2 μl of hemagglutinin plasmid (10 ng/μl), and 1 μl of PfuUltra High-Fidelity DNA Polymerase (Agilent Technologies, Santa Clara, CA) (2.5 U/μl). After digestion with *Dpn*I (a restriction enzyme that only cleaves methylated DNA) at 37°C for 1 h, the PCR product (2 μl) was transfected into XL1-Blue Supercompetent Cells (Agilent Technologies, Santa Clara, CA). The transformed cells were directly inoculated onto LB (Luria Bertani) agar plates or into LB medium.

Generation of reassortant viruses by reverse genetics

The reassortant viruses were rescued by transfection in co-cultured 293T cells and MDCK cells, as described elsewhere (140), using mutated hemagglutinin gene, unmodified neuraminidase gene of CA/04, and six internal genes from influenza A/PR/8/1934(H1N1) virus. At day 3 after transfection, the media and cells were inoculated into 10-day-old eggs (0.2 ml/egg). Ninety-six hours after inoculation, we collected the allantoic fluids from the eggs for viral titration.

Genomic sequencing

Virion RNA and cDNA were prepared as previously described (140). Sequencing was performed at the Life Sciences Core Laboratories Center at Cornell University (Ithaca, NY) by using the automated 3730 DNA Analyzer (Applied Biosystems, Foster City, CA), which utilizes Big Dye Terminator chemistry and AmpliTaq-FS DNA Polymerase. Deep sequencing was performed by the Beijing Genomics Institute (Shenzhen, China), using a MiSeq sequencer (Illumina, Shanghai, China).

Growth kinetics in MDCK cells

To determine the growth kinetics of viruses, we inoculated the MDCK cells with influenza virus at a multiplicity of infection of 0.001 $TCID_{50}$ and incubated the cells in 5% CO_2 at 37°C for 1 h. The inocula were removed and washed two times with phosphate-buffered saline (PBS). The cells were then incubated (37°C in 5% CO_2) in Opti-MEM I (GIBCO, Grand Island, NY) containing TPCK–trypsin (1 μg/ml). At specified time points after inoculation, 200 μl of supernatants were collected, aliquoted, and stored at -70°C until use. Virus titers in supernatants collected at the different time points were determined by $TCID_{50}$ in MDCK cells.

Western blot assay

The Western blot assay was used to compare the yields of the vaccine candidates with those of the wild-type strain. The cell supernatants were lysed and then analyzed by Western blot assay. Western blots were developed by using horseradish peroxidase–conjugated goat anti-mouse IgG (Santa Cruz Biotechnology, Santa Cruz, CA) and an enhanced chemiluminescence detection system (ECL GE Healthcare, Pittsburgh, PA) and

then exposed to X-ray film. The primary antibodies used in the Western blot assay were influenza nucleoprotein monoclonal antibodies from the Biodefense and Emerging Infections Research Resources Repository (Manassas, VA).

Quantitative reverse transcription PCR (qRT-PCR)

qRT-PCR was used to determine the level of viral RNA in the allantoic fluid of eggs inoculated with an influenza virus. Viral RNA (50 µl total) was isolated from 200 µl of sample by using a Gene JET Viral RNA Purification Kit (Thermo Fisher Scientific, Pittsburgh, PA). Influenza virus–specific Primer Uni12 (5′-AGCAAAAGCAGG-3′) and 10 µl of template RNA were used in the cDNA synthesis (total volume of 25 µl); the cDNA synthesis was carried out using SuperScript III Reverse Transcriptase (Invitrogen, Grand Island, NY) according to the manufacturer's instructions. Primers 130loop_F and 190helix_R were used in the qPCR, which consisted of 5 µl of Master Mix (Applied Biosystems, Foster City, CA), 500 nM each primer, 2 µl of target cDNA, and sufficient diethylpyrocarbonate (DEPC)-treated water to make a volume of 10 µl. The cDNA was amplified by 40 two-step PCR cycles (3 s at 95°C for denaturation of the DNA, 20 s at 60°C for primer annealing and extension). qPCR amplifications were measured by using the Stratagene Mx3005P qPCR System (Agilent Technologies, Santa Clara, CA). As a sample standard, we used the pHW 2000 plasmid vector with the hemagglutinin gene of CA/04, which was serially diluted 10-fold to generate the standard curve. The amplification results were shown as Log_{10} copies/µl.

Total protein quantification

Viruses were purified from the allantoic fluids of 11-day-old embryonated chicken eggs by low-speed clarification (4000 rpm, 20 min, 4 °C) followed by ultracentrifugation through a cushion of 30%-60% sucrose in a 70Ti rotor (Beckman Coulter, Fullerton, California) (37,500 rpm, 3 h, 4 °C). The virus band was collected and purified through a cushion of 30% sucrose in a 70Ti rotor (37,500 rpm, 3 h, 4 °C). The virus pellet was resuspended in 200ul PBS and the total amount of purified virion proteins was determined by Quant-iT^{TM} protein assay kit (Invitrogen, USA).

Mouse vaccination and challenge

To assess the antigenicity and protective efficacy of high-yield mutant viruses, we vaccinated a group of 6-week-old female BALB/c mice (Harlan Laboratories, Indianapolis, IN) by intramuscular administration of 128 hemagglutination (HA) units (in 50-µl volumes) of a formaldehyde-inactivated vaccine candidate, mutant number 81 (n = 10 mice) or mutant number 88 (n = 10 mice). Two weeks later, we administered a booster vaccine. Mock-vaccinated mice (n = 10) received a volume of PBS equal to the amount of vaccine administered to the vaccinated mice. Five mice that were not vaccinated and not challenged served as environmental controls. Two weeks after the booster vaccination, mice were anesthetized and challenged by intranasal inoculation with mouse-adapted CA/04 at 10× the 50% lethal dose (LD_{50}). Serum samples were collected from mice before challenge and tested by using an HA inhibition (HI) assay. To determine lung virus titers, we euthanized three mice at day 4 after challenge. Lungs were homogenized and resuspended in 1 ml of sterile PBS, and virus titers were determined in MDCK cells. The samples were also stained with hematoxylin and eosin stain for

pathologic examination. The survival rate, clinical signs, and body weight of the remaining mice were monitored for 14 days after the challenge. All animal studies were approved by the Office of Regulatory Compliance Institutional Biosafety Committee and the Institutional Animal Care and Use Committee of Mississippi State University.

HA and HI assays

HA and HI assays were performed by using a 0.5% suspension of turkey red blood cells as described in the WHO Manual on Animal Influenza Diagnosis and Surveillance (http://www.who.int/csr/resources/publications/influenza/whocdscsrncs20025rev.pdf).

Structural modeling

Crystal structures of the hemagglutinin protein (Protein Databank [PDB] accession nos. 3LZG, 3UBN, and 3UBQ of the A(H1N1)pdm09 virus and the binding sites of carbohydrates 6'-Sialyl-N-acetyllactosamine (6SLN) and 3'-Sialyl-N-acetyllactosamine (3SLN) to this protein were obtained from PDB (141). Structural simulation of amino acid mutations was performed on the hemagglutinin by using the FoldX platform with its empirical force field (142) with crystal waters under the following conditions: temperature of 298K, pH 7, 0.05 ion strength. Chimera (143) was used to visualize the binding structures and measure contact distances. PoseScore, which was designed for ranking near-native ligand–protein interacting structures (144), was used to estimate the likeness of the protein–glycan binding avidities of the wild-type and mutant to that of the native virus. The computational analysis of mutation effects on HA–glycan bindings was focused on the mutants with 133 (H3 numbering) location only

because the other mutations are not located in the receptor-binding sites of hemagglutinin and did not modify the PoseScores of the hemagglutinin of the wild-type virus.

Accession numbers

The Crystal structures of the hemagglutinin protein were downloaded from Protein Data bank with accession numbers: 3LZG, 3UBN, and 3UBQ.

Results

epPCR-based mutagenesis approach

In this study, we developed an epPCR-based mutagenesis strategy for generating high-yield influenza vaccine candidates (Figure 2). In this method, the epPCR products are used directly as primers for site-directed mutagenesis, and individual cloned plasmids are then extracted. Viruses are rescued by using the individual selected plasmids, and the rescued viruses are screened for a phenotype of high yield in eggs.

To validate our method, we used the hemagglutinin of CA/04 as the template to generate high-yield vaccine candidates. We first applied epPCR to breed the regions of receptor-binding sites: one region covered the 130 loop and 190 helix of the hemagglutinin gene (referred to as 130loop+190helix), and the other region covered the 190 helix and 220 loop of the hemagglutinin gene (referred to as 190helix+220loop). The epPCR products (234 bp for 130loop+190helix; 171 bp for 190helix+220loop) were subsequently used as primers in site-directed mutagenesis.

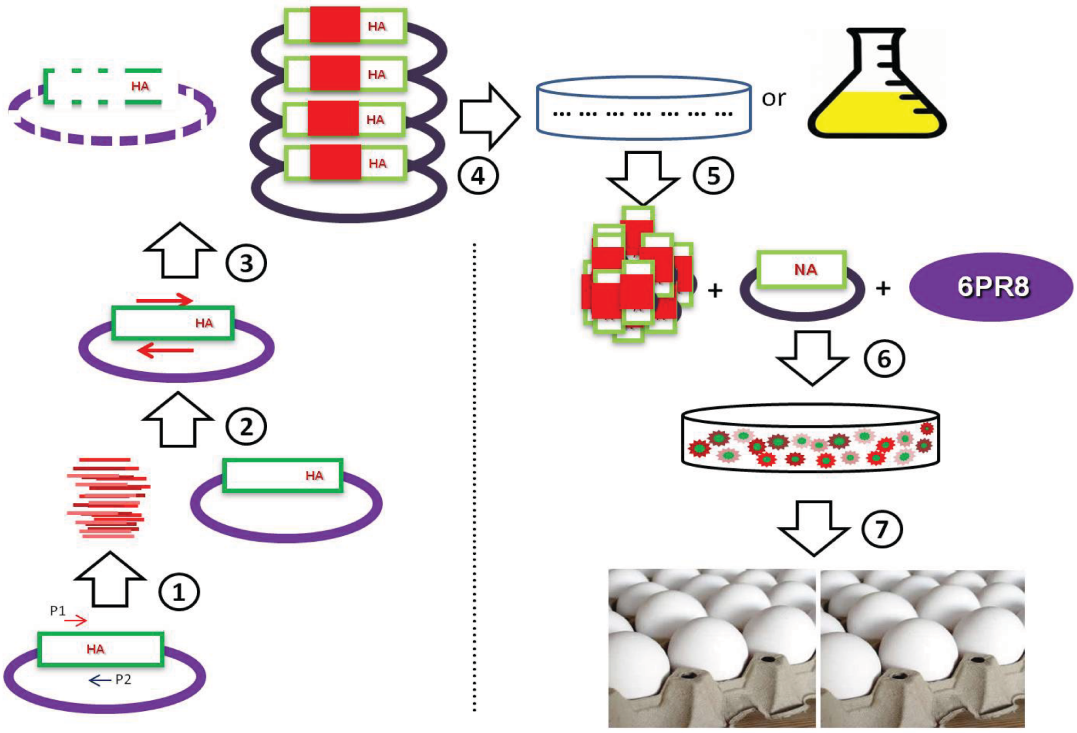


Figure 2. Overall strategy for generating a hemagglutinin mutant library. Based on this method, chicken embryonic egg–propagated seed viruses with high yield can be selected for influenza vaccine candidate strains. The 130loop + 190helix or 190helix + 220loop in hemagglutinin gene of influenza A/California/04/09 (H1N1) virus are amplified by error-prone PCR (epPCR) by using a GeneMorph II Random Mutagenesis Kit (Agilent Technologies, Santa Clara, CA) (step 1). epPCR products are then used as primers in the site-directed mutagenesis (step 2). After DpnI digestion at 37°C for 1 h (step 3), the PCR product is transformed into XL1-Blue Supercompetent Cells (Agilent Technologies, Santa Clara, CA), and the cells are inoculated onto LB (Luria Bertani) plates or into LB medium (step 4). The individual plasmid or plasmid library is then extracted (step 5) and used with 6 internal genes from A/Puerto Rico/8/1934(H1N1) and the neuraminidase gene from CA/04 to rescue viruses (step 6). Rescued viruses are inoculated into 10-day-old chicken embryonic eggs to generate high-yield vaccine seed. The steps are indicated by circled numbers in the figure.

Broad diversity of the hemagglutinin mutants generated by the epPCR-based mutagenesis strategy

To evaluate the profile of the hemagglutinin mutant library, we sequenced a total of 292 clones (142 for 130loop+190helix; 150 for 190helix+220loop). Of the 292 clones, 111 carried at least one amino acid mutation while maintaining the correct open-reading

frame. Of these 111 hemagglutinin mutants, 66 were derived from the 130loop+190helix epPCR products, and 45 were derived from the 190helix+220loop epPCR products. Ninety-eight of these 111 hemagglutinin mutants had a single amino acid mutation, eight had double mutations, and five had triple mutations. Except for two clones with the same mutations, all hemagglutinin mutants carried different mutations. The mutation profile analysis demonstrated that the mutated sites broadly spanned the entire target region from the 130loop to the 220loop. Of note, some of these mutations (V135I，A137T, P140L，H141Q，S146G， V155I， K156E， N159K， G173R， V176I， A189T， T200S，K212T, I219V) were observed in the hemagglutinin gene of the A(H1N1)pdm09 virus field strains.

To further demonstrate the diversity of the mutant library, we used next-generation sequencing to sequence the plasmid library generated by 130loop+190helix epPCR products. The readings showed that 63.8% of the sequences had nucleotide changes, and 83.9% of those led to effective amino acid mutations with correct open-reading frames. The mutations were distributed among the 130loop and the 190helix at both the nucleotide and amino acid levels (Figure 3).

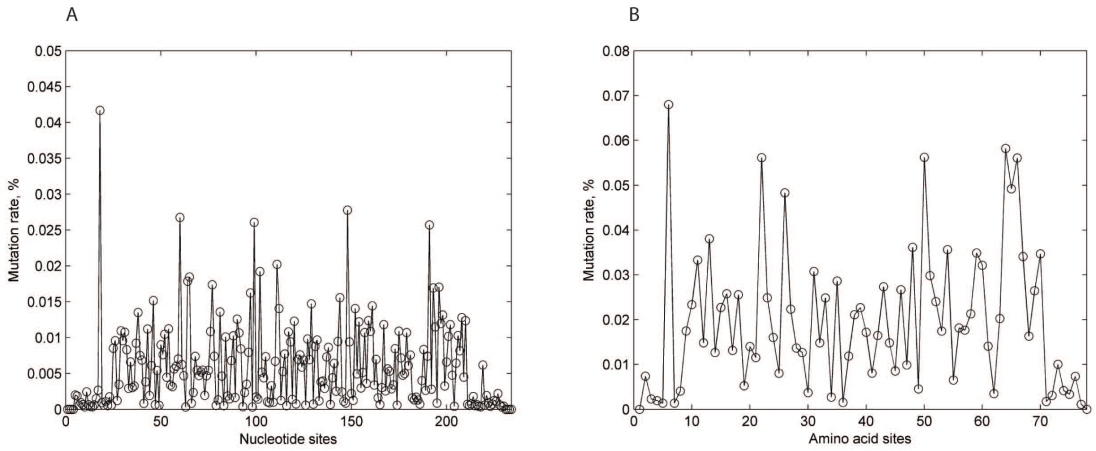


Figure 3. Profile of 130loop+190helix hemagglutinin plasmid library sequenced by using the Miseq sequencer (Illumina, Shanghai, China).

There were 238,961 readings in the library. Of those readings, 63.8% had mutations at the nucleotide level (A), and 83.9% of those led to effective amino acid mutations with correct open-reading frames (B). The mutation rate on each site demonstrates that mutations were distributed almost across the entire 130-loop and 190-helix, demonstrating that the plasmid library generated by this error-prone PCR strategy has a broad diversity.

Rapid generation of higher yield strains by using the epPCR-based mutagenesis strategy

To determine if seed strains with high yield could be directly rescued from the 111 mutants, we recovered the highest-yield viruses by transfection, inoculated them into 10-day-old eggs, and subsequently determined HA titers for comparison with the HA titers for the wild-type virus. The wild-type virus had an HA titer of 32, and six mutants (nos. 22, 58, 79, 81, 88, and 114) had an HA titer of at least 128 in passage one (Table 2). Of note, the HA titers of the wild-type virus remained stable at 32 in passages two and three, but the titers of mutant viruses number 81 and 88 increased to 512 after passage two and three, respectively (Table 2).

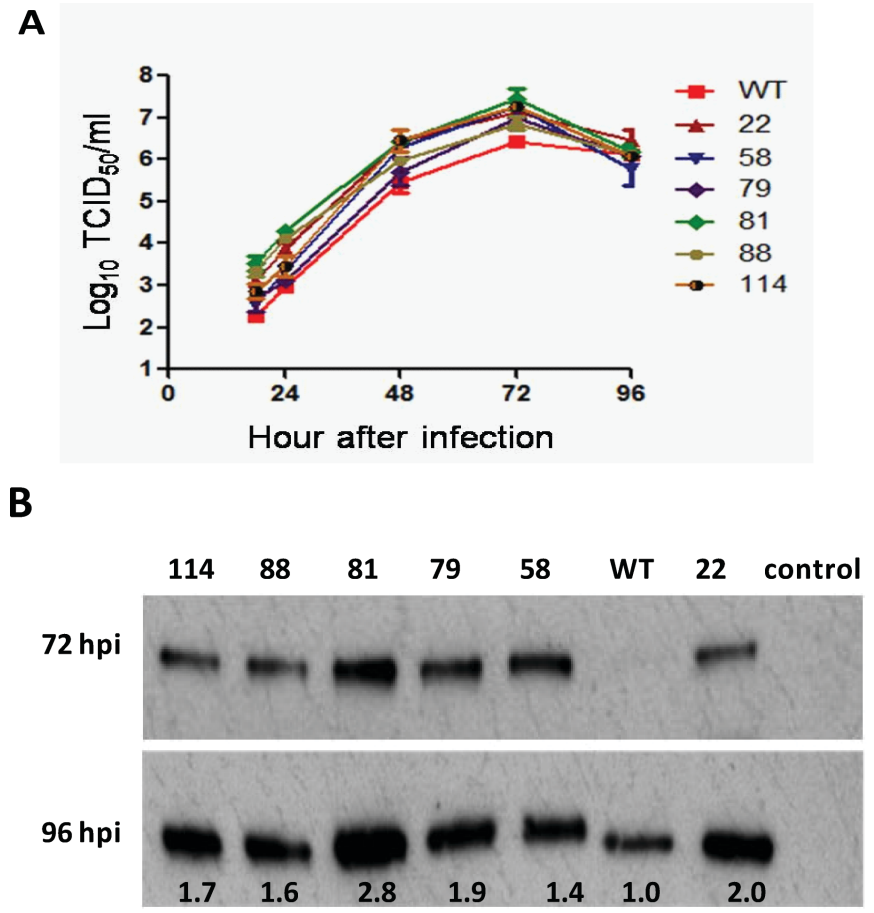


Figure 4. Characterization of the high-yield influenza A(H1N1)pdm09 vaccine candidates generated by using the error-prone PCR–based mutagenesis strategy. Viruses were inoculated into Madin-Darby canine kidney cells at a multiplicity of infection of 0.001. (A) Growth curves for wild-type (WT) and mutant viruses (nos. 22, 58, 79, 81, 88, and 144) measured by 50% tissue culture infectious dose ($TCID_{50}$; values shown below columns) at various hours after infection (hpi). (B) Western blot showing the nucleoprotein expression level for the WT and mutant viruses (viruses shown above columns).

Viral growth kinetics demonstrated that these six hemagglutinin mutants replicated more efficiently than the wild-type virus in MDCK cells (Figure 4A). The peak HA titers of these viruses reached 128–512 in MDCK cells, but the peak titer was only 32

for the wild-type virus. Among these six mutants, numbers 22 and 81 yielded the highest HA titer (512) in MDCK cells. The peak $TCID_{50}$ of these mutants was 0.4–1.0 log_{10} $TCID_{50}$ higher than that for wild-type virus (Figure 4A).

The Western blot assay results for cell lysate tested 72 h and 96 h after inoculation confirmed these results. Nucleoprotein expression for the mutant viruses was stronger than that for the wild-type strain. The nucleoprotein expression level for mutants number 22, 58, 79, 81, 88, and 114 were 2.0-, 1.4-, 1.9-, 2.8-, 1.6-, and 1.7-fold higher, respectively, than that for the wild-type virus (Figure 4B).

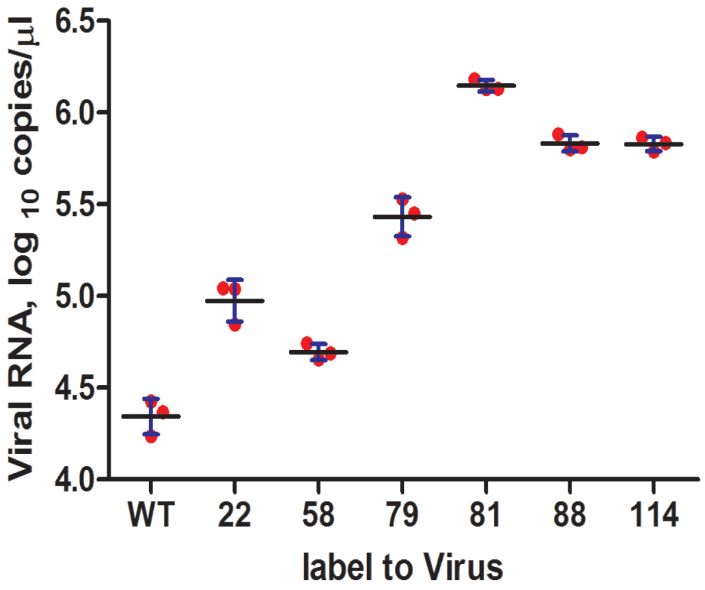


Figure 5. Quantification of wild-type (WT) and mutant viruses propagated in eggs as determined by using a hemagglutinin gene–specific quantitative reverse transcription PCR method.
Results are expressed as the median (horizontal bars) RNA copy number (1 µl cDNA) ± SD (vertical bars).

Figure 6. Total protein quantification of wild-type (WT) and mutant viruses that purified from the allantoic fluids of 11-day-old embryonated chicken eggs. Results are expressed as the median (horizontal bars) protein concentration ±SD (vertical bars).

Consistent with our HA assay results, qPCR results for the six selected mutants showed higher yields than those for the wild-type virus. On average, viral RNA levels in allantoic fluid of eggs inoculated with mutants number 22, 58, 79, 81, 88 and 114 were 4.3-, 2.3-, 12.3-, 64.6-, 30.9-, and 30.9-fold (copies/μl) higher, respectively, than levels in allantoic fluid of eggs inoculated with the wild-type strain (Figure 5). Meanwhile, total protein concentrations of mutants number 22, 58, 79, 81, 88 and 114 were 1.90-, 2.33-, 1.99-, 2.50-, 1.30- and 1.97-fold (μg/ml) higher, respectively, than that of the wild-type strain (Figure 6). These results further confirmed that, compared with the wild-type virus, the six selected mutants generated by the epPCR-based mutagenesis strategy grew more efficiently in eggs.

One-way HI assays using post-infection ferret antiserum raised against CA/04 showed that the homologous HI titers of 3,200 for CA/04 (Table 2). The heterologous HI

titers of the six mutant viruses were 3,200 or 1,600, demonstrating that these candidates had few antigenic differences from the wild-type strain.

Table 2. Characterization of influenza A(H1N1)pdm09 vaccine candidates generated by using an error-prone PCR–based mutagenesis strategy

Virus	HA titer[a], passage no.			HI titer[b], passage no.			Mutation (residue no.)[c]
	1	2	3	1	2	3	
Wild type	32	32	32	3,200	3,200	3,200	–
Mutant,[d] no.							
22	128	256	256	3,200	3,200	3,200	K133N (130), V178M (175)
58	128	256	128	3,200	3,200	3,200	K149E (146)
79	128	128	256	3,200	3,200	1,600	K157I (154)
81	128	128	512	1,600	1,600	1,600	K133I (130)
88	256	512	256	1,600	1,600	3,200	K157I (153), I169T (166)
114	256	128	256	3,200	3,200	3,200	K212T (209)

[a] HA, hemagglutination.
[b] HI, hemagglutination inhibition. Titers were determined by using ferret serum (anti-CA/04).
[c] Mutations in the hemagglutinin protein sequence of the mutant, H3 numbering; the residue numbers are residues in the hemagglutinin of influenza A(H1N1)pdm09 virus.
[d] Reverse genetics–derived viruses were generated by using hemagglutinin of wild-type influenza A/California/04/09 (H1N1) virus (CA/04) virus or hemagglutinin mutants derived from error-prone PCR, neuraminidase of CA/04, and six other gene segments (polymerases PA [polymerase acidic subunit], PB1 [polymerase basic subunit 1], and PB2 and nucleoprotein, matrix protein, and nonstructural protein) from influenza A/PR/8/34(H1N1) virus.

High-yield vaccine candidates providing full protection against lethal challenge in mice

To confirm that our mutagenesis approach had not affected subsequent vaccine efficacy, we selected the two mutants with the highest HA titers (mutants no. 81 and 88) as vaccine candidates and evaluated their efficacy in a mouse model. Serum samples collected from vaccinated mice 2 weeks after the booster vaccine was administered exhibited HI titers substantially higher than those for the mock-vaccinated group (Table 2). After viral challenge, virus replication (reaching levels up to $10^{5.4}$ $TCID_{50}$/ml) was detected in the lungs of all mock-vaccinated mice; no virus was detected in any of the vaccinated mice (Table 3). Mice in the mutant number 81– and 88–vaccinated groups did not lose weight, but the mock-vaccinated mice were lethargic and rapidly lost weight (Figure 7A). All vaccinated mice survived the lethal challenge; all mock-vaccinated mice died 7 days after challenge (Figure 7B).

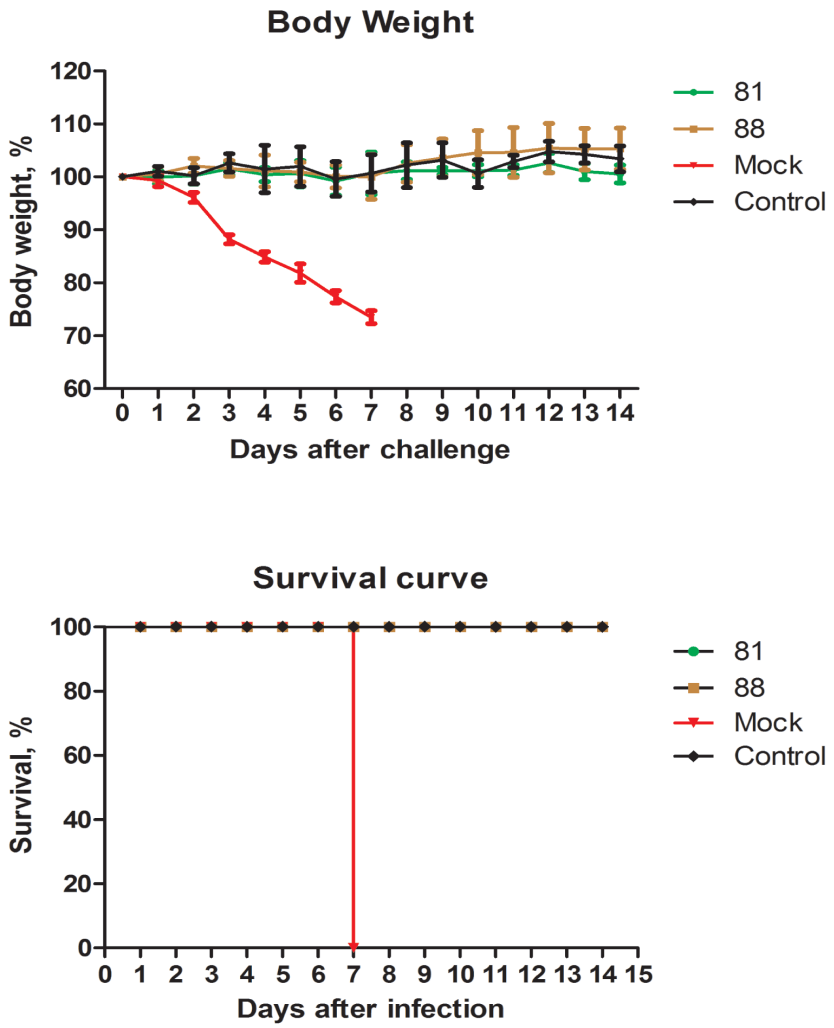


Figure 7. Protective effect of high-yield vaccine candidates in mice challenged with mouse-adapted influenza A/California/04/09 (H1N1) virus.

Groups of control; mock-vaccinated; and mutant number 81–, 88–vaccinated mice were intranasally inoculated with 10 × the 50% lethal dose of virus, after which their body weights (A) and survival times (B) were monitored for 14 days. Results are shown as the mean ± SD in each group (A).

Table 3. Pathogenesis and immunologic responses in vaccinated mice challenged with mouse-adapted influenza A/California/04/09 (H1N1) virus

Vaccine group	$Log_{10}TCID_{50}$/ml, mean ± SD[a]	log_2HI titer, mean ± SD[b]
Mutant no. (mutation)		
81 (K133I)	Below detection limit	8.32 ± 0.70
88 (K157I, I169T)	Below detection limit	8.12 ± 0.45
Mock	5.4 ± 0.1	Below detection limit

[a] Groups of BALB/c mice were inoculated intranasally under light anesthesia with 10 × the 50% lethal dose of mouse-adapted CA/04 virus. Three mice from each group were euthanized on day 4 after virus challenge, and virus titers in lungs were determined by $TCID_{50}$ (50% tissue culture infectious dose) in MDCK cells.
[b] Serum samples were collected 2 weeks after the booster vaccine was administered, and antibody response levels were measured by using the HI (hemagglutination inhibition) assay.

Photomicrographs of hematoxylin and eosin–stained lung sections are shown in Figure 8. Mock-vaccinated mice displayed acute, diffuse, necrotizing bronchitis and bronchiolitis 4 days after challenge (Figure 8C). No histopathologic changes were observed in the mice vaccinated with mutants number 81 or 88 (Figure 8A,B), demonstrating effective protection of both vaccine candidates from the lethal mouse-adapted CA/04 challenge. Thus, our results indicate that the high-yield candidate vaccines generated by our epPCR-based mutagenesis strategy did not have altered antigenicity and could serve as potential seed strains for the manufacture of influenza vaccines.

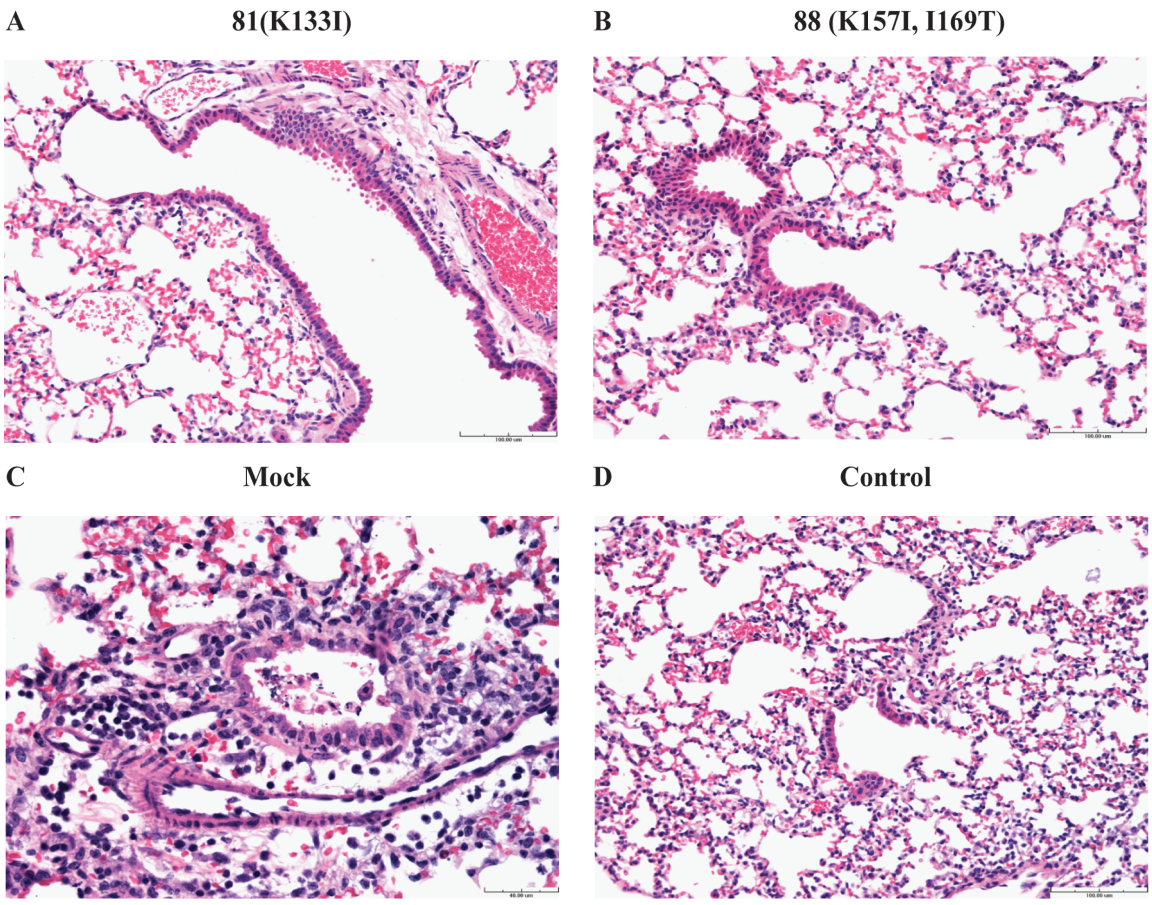


Figure 8. Histopathologic findings in hematoxylin and eosin–stained lung samples from vaccinated and mock-vaccinated mice challenged with mouse-adapted influenza A(H1N1)pdm09 virus.

Four days after challenge, groups of mutant number 81–vaccinated (A), mutant number 88–vaccinated (B), and mock-vaccinated (C) mice were euthanized, and lungs were collected for histopathologic examination; non-vaccinated, non-challenged mice served as environmental controls (D). A, B, and D) Scale bar =100 μm; C) scale bar = 40 μm.

Receptor-binding models of improved-yield mutants

Among the six mutation sites, only mutants number 22 and 81 had residue 133 located in the receptor-binding sites of hemagglutinin (Figure 9). These crystal structures of the bindings between the hemagglutinin of influenza A(H1N1)pdm09 and glycan analogs 6SLN or 3SLN were used in computational mutation simulations to study the effects of amino acid modifications on hemagglutinin-receptor bindings.

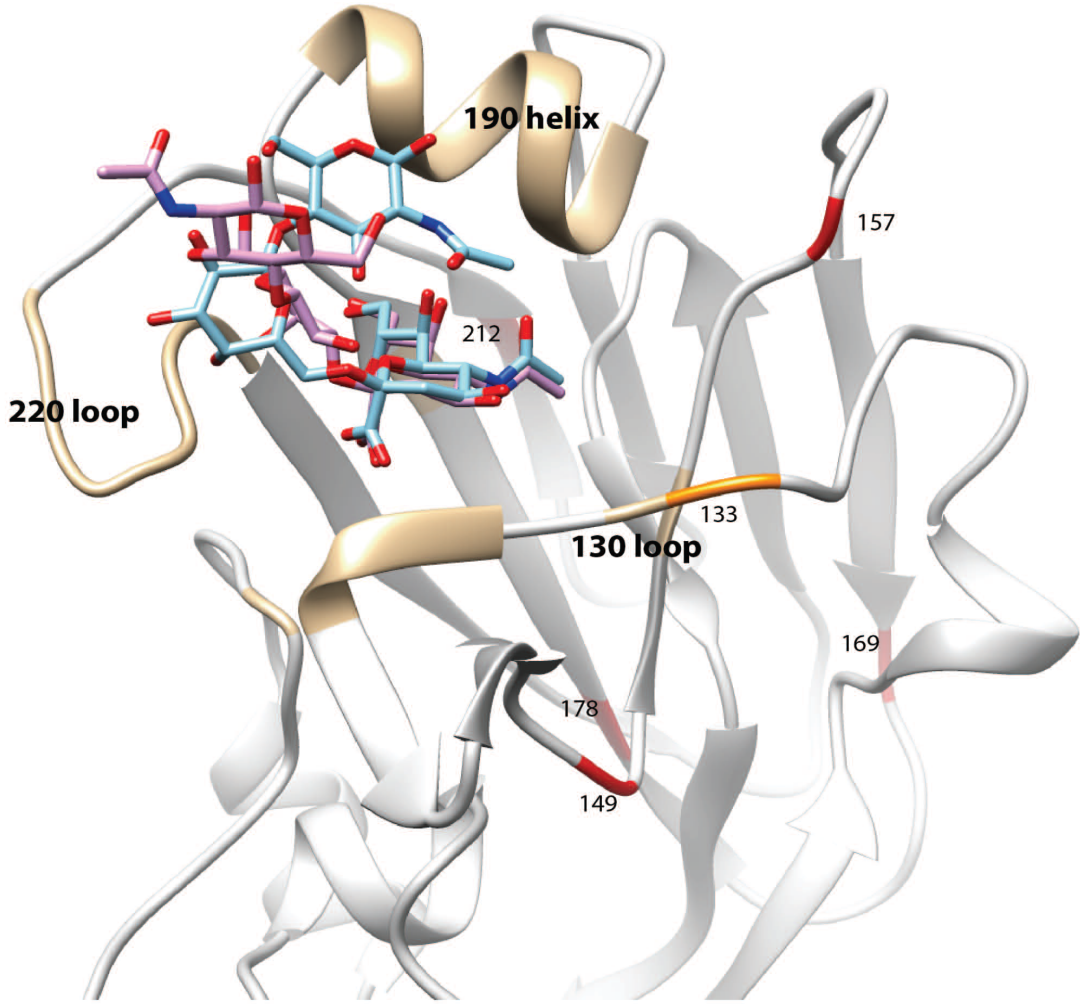


Figure 9. The locations of six mutations in the crystal structure of the hemagglutinin (HA) of influenza A(H1N1)pdm09 virus.

The HA was adapted from the structural template of the HA of A(H1N1)pdm09 virus (Protein Databank [PDB] 3LZG) interacting with glycans 6SLN (from PDB 3UBN) and 3SLN (from PDB 3UBQ). The HA protein structure is shown in light grey; receptor-binding residues on the 130-loop, 190-helix, and 220-loop are shown in gold. The mutation location 133 is in orange to distinguish it from other non-binding site mutation locations indicated in red (149, 157, 169, 178, and 212). Cyan indicates 6SLN and magenta indicates 3SLN; both are superposed as their native binding poses against the native HA structures.

As measurements of binding modifications caused by mutations, we calculated the distances between the oxygen atom of a water molecule (which formed a hydrogen

bond between amino acid residue 133 and glycan receptors) and the nearest atoms of sialic acid and the distances between the oxygen atom of a water molecule and residue 133. Moreover, as a quantification of hemagglutinin–glycan binding avidities, we obtained the PoseScore (144) for structural interactions between glycans and all wild-type and computationally mutated hemagglutinin structures (Figure 10).

When the hemagglutinin bound to 3SLN, the wild-type structure with a polar and positively charged lysine at residue 133 had a distance of 4.57 Å to the linker water molecule, and the whole interaction had a PoseScore of -6.32 (Figure 10B). The K133N mutant had a polar, but neutral, asparagine at residue 133, which was only 2.23 Å from the linker water molecule, and its PoseScore wass -7.04 (Figure 10D). The K133I mutant had a nonpolar and neutral isoleucine at residue 133, which was 2.95 Å to the linker water molecule, and its PoseScore was -6.88 (Figure 10F). These results showed that both the contact distances and pose binding scores were decreased by substituting a charged lysine with neutral asparagine or isoleucine. Thus, given that smaller contact distances and lower PoseScores suggest stronger interactions between glycans and proteins, mutations K133N and K133I could lead to increased 3SLN bindings compared with what the wild-type strain does. These stronger hemagglutinin–3SLN bindings for mutants number 22 and 81, compared with that for the wild type strain, may contribute to the higher replication efficiencies of these two mutants in eggs and MDCK cells.

Of interest, the structural simulation showed that both the contact distances and pose binding scores between hemagglutinin and 6SLN were also decreased by substituting a charged lysine with neutral asparagine or isoleucine. When the hemagglutinin bound to 6SLN, the wild-type structure with a polar and positively

charged lysine at residue 133 had a distance of 4.41 Å to the linker water molecule and the whole interaction had a PoseScore of -11.83 (Figure 10A). The K133N mutant had a polar but neutral asparagine, which wass only 2.04 Å from the linker water molecule, and its PoseScore wass -12.18 (Figure 10C). The K133I mutant had a nonpolar and neutral isoleucine at residue 133, which was 2.82 Å from the linker water molecule, and its PoseScore was -14.42 (Figure 10E). These results suggested that mutations K133N and K133I could lead to stronger 6SLN bindings in mutants number 88 and 81 than in the wild-type virus.

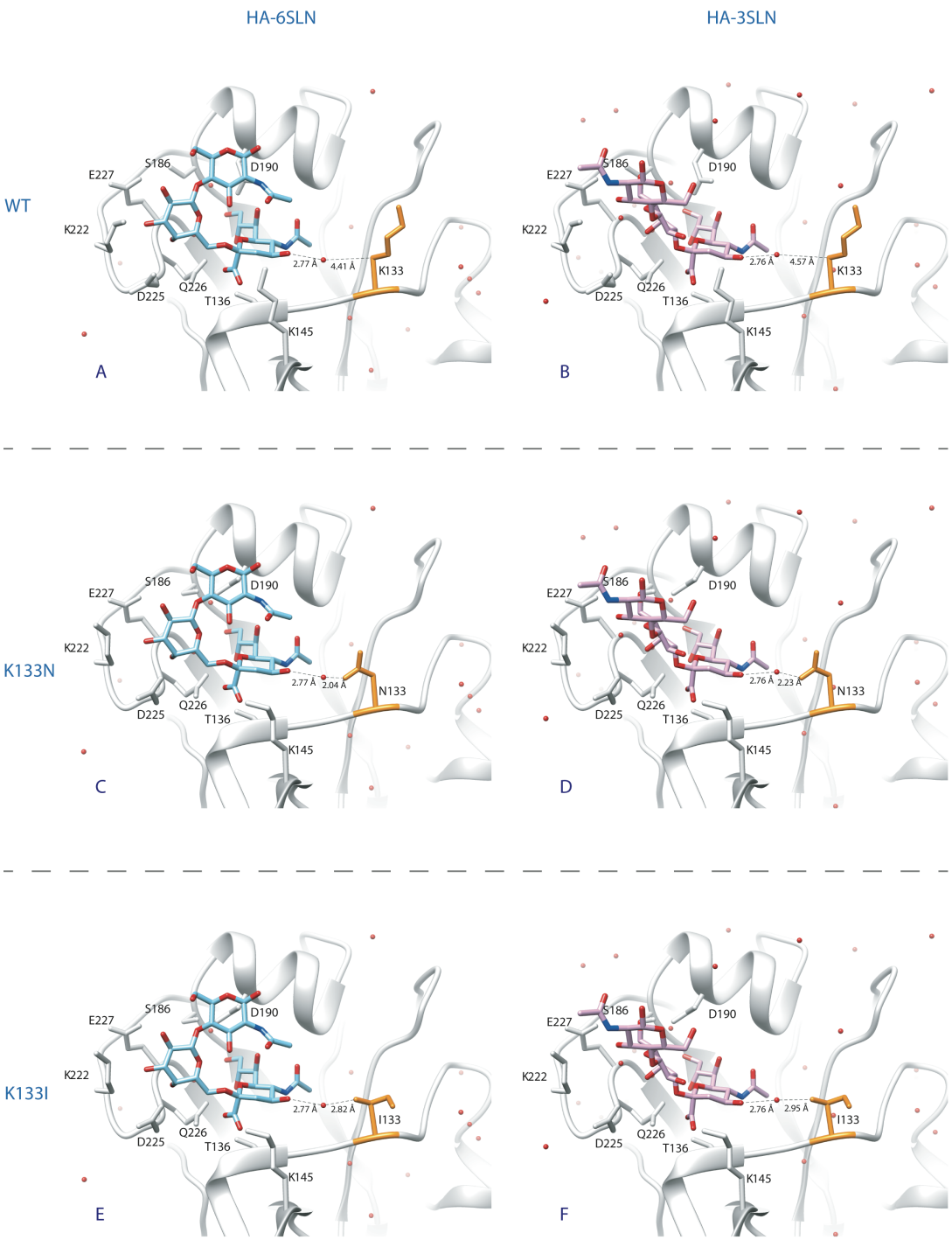


Figure 10. The three-dimensional structures of the hemagglutinin of the wild–type (WT) influenza A/California/04/09 (H1N1) virus and mutant viruses (nos. 22 and 81) in contact with human-like receptor analog 6SLN (panels A, C, and E) and avian-like receptor analog 3SLN (panels B, D, and F).

The protein structures of hemagglutinin are shown in light grey; residue 133 is in orange. The side chains of interacting residues on the receptor-binding sites are labeled with

residue names and locations. The single letter amino acid annotations were used together with H3 numbering for all binding residues. 6SLN is shown in cyan, and 3SLN is shown in magenta. Red dots indicate the oxygen atoms from all water molecules that are in contact with hemagglutinin side chains. The distances (in angstroms) between the water molecule and the nearest atom on both the protein and ligand sides are indicated by dashed lines. **(**A) WT hemagglutinin with lysine at position 133 in contact with 6SLN. (B) WT hemagglutinin with lysine at position 133 in contact with 3SLN. (C) K133N mutant hemagglutinin with asparagine at position 133 in contact with 6SLN. (D) K133N mutant hemagglutinin receptor-binding sites with asparagine at position 133 in contact with 3SLN. (E) Mution K133I at the hemagglutinin receptor-binding site with isoleucine at position 133 in contact with 6SLN. (F) Mutant K133I with isoleucine at position 133 in contact with 3SLN. Simulations on the hemagglutinin were performed by using the FoldX empirical force field (142), and the structure was visualized by using Chimera (143); PoseScore (144) was used to estimate the likeness of the WT and mutant protein–glycan binding avidities to that of the native virus.

Discussion

Conventional site-directed mutagenesis strategies depend on virus variants that are manually selected and usually have a single or a few mutations; egg-adaptation strategy depends on the virus variants generated by error in natural viral RNA replication (viral quasispecies) resulting from the lack of proofreading capacity in viral RNA polymerase. The small pool of virus variants produced by these approaches is a bottleneck in the process of selecting high-yield seed strains for influenza vaccine production. Thus, we hypothesized that the *in vitro* breeding of a diverse influenza virus variant library by epPCR would enhance the process for generating high-yield strains. The epPCR method has been widely used to generate libraries of mutant proteins for phenotype selection (145-149). By integrating epPCR, site-directed mutagenesis, and reverse genetics, we developed a novel strategy for rapidly generating high-yield vaccine candidates. Our data show that this strategy allows for the rapid generation of high-yield vaccine candidates for the A(H1N1)pdm09 virus. Whole-virus inactivated vaccines

derived from two high-yield candidates that we generated provided full protection against lethal A(H1N1)pdm09 virus challenge in mice.

Recently, Imai and Wu also used epPCR strategy to generate influenza mutant library for phenotype selection (150, 151). In their studies, the epPCR was used to amplify the partial HA or whole NA genes, and the amplified fragment or gene need to be digested with restriction enzyme and then subcloned to influenza reverse genetics vector for virus rescuing. Unlike their approaches for generating influenza mutant library, our strategy combined the epPCR and site mutagenesis approach efficiently, and we used epPCR product as primers for site mutagenesis. Therefore, through two single PCRs, one epPCR and another site mutagenesis PCR, we can target any region of influenza gene to generate plasmid mutant library without designing any restriction enzyme site and omitting the subcloning step. Thus, our approach developed here can be used for efficiently generating diverse mutant libraries for phenotype selection for influenza virus or other pathogen.

The mechanism(s) by which the mutations in the hemagglutinin gene improve virus replication are poorly understood. Previous studies suggested that the functional balance between receptor binding of the hemagglutinin and the receptor-releasing property of the neuraminidase is critical for efficient replication of influenza viruses (85, 152). Mutations near the receptor-binding sites of the hemagglutinin are known to affect the replication abilities of viruses (119, 120, 153, 154). In this study, we constructed a hemagglutinin plasmid library targeted to the receptor-binding sites of influenza A(H1N1)pdm09 virus. By using the cloned mutant hemagglutinin genes, we successfully

rescued 43 viruses without subsequent egg passages. Of these 43 viruses, six mutants had improved yields in eggs.

Among the mutations in the hemagglutinin of the six selected vaccine candidates in our study (Table 2), mutations K157I and K212T were previously reported in the high-yield A(H1N1)pdm09 vaccine strains generated in eggs (119). Identification of these two previously reported mutations, which improved the yields of A(H1N1)pdm09 virus, validates the effectiveness of our strategy. Our strategy is further validated by the fact that we also found some novel sites in the hemagglutinin that improved yields without changing virus antigenicity. For example, K133, which was found in two high-yield mutants, is absent in the hemagglutinin of the contemporary human seasonal influenza A(H1N1) virus and is highly conserved in both the A(H1N1)pdm09 virus and its precursor, the hemagglutinin gene of the 1918 pandemic H1N1 influenza virus. The K133N HA mutation was predicted to broadly change the electrostatic potentials surrounding the receptor-binding domain (155). As shown in receptor-binding models (Figure 10), the substitution of a positively charged lysine in position 133 by a polar uncharged asparagine or hydrophobic isoleucine could contribute to the increase in binding avidity and be responsible for the improved yield of these mutants in eggs and MDCK cells.

A previous study predicted the important role of K133 and K149 in receptor-binding affinity and in the virulence of A(H1N1)pdm09 virus (156). Substitution of K133R or K149R in “swine-like” A(H1N1)pdm09 virus reduced its pathogenicity (156). The mutants with K133N or K149E mutations that were identified in this study have higher yield properties and may have lower pathogenicity compared with that of the

wide-type strain; which would ensure a safer vaccine production. In addition, two of six selected mutants with improved yields harbor K157I mutations. It has been reported that mutants with K157E mutations are preferred for replication of A/California/7/2009(H1N1) virus in MDCK cells and eggs (120). Our results also indicate that the presence of a isoleucine residue with hydrophobic site chain in position 157 of hemagglutinin improves the growth ability of the A(H1N1)pdm09 virus.

Concerns have been raised that even single amino acid substitution near receptor-binding sites may change the antigenicity of influenza viruses (157). However, mutations near the receptor-binding sites of the hemagglutinin, such as L119I, A186D, N125D, and D127E, have been used as seed viruses for the A(H1N1)pdm09 vaccine (120). The HI assay results in our study indicate that the six high-yield candidate vaccine strains had few antigenic changes from the CA/04 wild-type strain (Table 2). For example, although K133 was recently documented to be involved in B cell epitopes recognized by monoclonal antibodies (158-160), in our study, mutation K133N did not alter the antibody-binding avidities of K133N mutant to anti-A(H1N1)pdm09 virus ferret serum; this finding has also been reported elsewhere (155). Moreover, inactivated vaccines based on two testing candidates that yielded the highest titers in eggs provided full protection against lethal challenge in experiments with mice. These results provide further evidence that the epPCR-based mutagenesis strategy combined with serologic assays, such as the HI assay, could be used for rapid generation of influenza seed viruses.

In summary, we report the use of a novel epPCR-based site-directed mutagenesis strategy to rapidly generate a diverse influenza hemagglutinin plasmid library and the application of this method to generate high-yield candidate strains for production of

A(H1N1)pdm09 virus vaccine candidates. The mutations associated with high yield in our study could facilitate a quick response to future outbreaks caused by A(H1N1)pdm09 virus and its variants. Moreover, this novel strategy could also be applied to the breeding of other functional regions of a gene to rapidly generate diverse mutant libraries for gain-of-function studies for influenza and potentially other pathogens.

CHAPTER III

A Y161F HEMAGGLUTININ SUBSTITUTION INCREASES THERMOSTABILITY AND IMPROVES YIELDS OF 2009 H1N1 INFLUENZA A VIRUS IN CELLS

Vaccination is the primary strategy for influenza prevention and control. Yet egg-based vaccines, the predominant production platform, have several disadvantages including the emergence of viral antigenic variants that can be induced during egg passage. These limitations have prompted development of cell-based vaccines which themselves are not without issue. Most importantly, vaccine seed viruses often do not grow efficiently in mammalian cell lines. Here we aimed to identify novel high-yield signatures for influenza viruses in continuous Madin-Darby canine kidney (MDCK) and Vero cells. Using influenza A(H1N1)pdm09 virus as the testing platform and an integrating error-prone PCR strategy, we identified a Y161F mutation in the hemagglutinin (HA) that not only enhanced the infectivity of the resultant virus by more than 300-fold, but also increased its thermostability without changing its original antigenic properties. Vaccine produced from the Y161F mutant fully protected mice against wild-type A(H1N1)pdm09 lethal challenge. Compared with A(H1N1)pdm09, the Y161F mutant had significantly higher avidity for avian-like and human-like receptor analogs. Of note, introduction of the Y161F mutation into the HA of seasonal H3N2 influenza A virus (IAV) and canine H3N8 IAV also increased yields and thermostability in MDCK cells and chicken embryotic eggs. Thus, residue F161 plays an important role in determining viral growth and thermostability, which could be harnessed to optimize IAV vaccine seed viruses.

Introduction

Influenza A viruses (IAVs) cause seasonal outbreaks and occasional pandemic outbreaks among humans and pose challenges to public health. The viruses responsible for four pandemics have been characterized: one each in 1918 and 2009 caused by H1N1 IAVs, one in 1957 caused by an H2N2 IAV, and one in 1968 caused by an H3N2 IAV (4). These pandemic outbreaks varied in impact, but each resulted in substantial mortality in a short time. Compared with pandemic outbreaks, seasonal influenza outbreaks are typically milder but still cause approximately 200,000 hospitalizations and 36,000 deaths each year in the United States alone (161).

Vaccination has been the most efficient and economic strategy for preventing influenza virus infection and controlling the spread of disease (162). Three types of virus-based influenza vaccines, inactivated vaccines, live-attenuated vaccines, and recombinant hemagglutinin (HA), are licensed in the United States, with egg-produced vaccine being the dominant source (127, 163). The egg-based platform for vaccine production has been used since the 1950s (164), but it has several disadvantages: first, passage of seed viruses in eggs can result in undesired egg-adapting mutations in the HA that can lead to changes in viral antigenicity (165-168); second, due to reactogenicity concerns, egg grown vaccine is contraindicated for those with egg allergies (169); and third, rapidly scaling up egg production is not easily achievable . Cell-based vaccine production platforms do not have the same limitations (170). All continuous cell lines, including Madin-Darby canine kidney (MDCK) cells and African green monkey kidney–derived Vero cells, must be

certified before being approved by regulatory authorities for use in the production of influenza vaccines (171-173).

A high-yield vaccine seed strain is required for timely vaccine manufacture and is thus a critical component of a successful influenza vaccination program. Unfortunately, it is not uncommon that the vaccine strains recommended by the World Health Organization (WHO) have less than desirable yields in eggs, cells, or both (174, 175). For example, the 2009 H1N1 pandemic seed strain was a low-yield strain, and it required almost 3 months for the WHO collaborative laboratories and vaccine companies to engineer the selected strain to meet the criteria required for vaccine production. Because of this delay, vaccine-derived immunity among the population arose after the peak of the second wave of the 2009 H1N1 pandemic (176). Therefore, quickly generating a high-yield vaccine seed virus is critical for rapid vaccine production and, thus, for effective influenza prevention and control. Adaptation of viruses to cells by multiple passaging or development of high-yield reassortant seeds using reverse genetics has been shown to be an effective way to increase yield of vaccine seed viruses (177-180). In addition, other studies have been performed to improve virus yields in cells by modifying the virus or the cell line. For example, Suphaphiphat et al. (181) showed that mutations S186P and L194I in the receptor binding site (RBS) of the A/California/04/09 (CA/04) H1N1 HA increased growth of the virus by more than 10-fold in MDCK cells, and Hamamoto et al. (182) reported that MDCK cells engineered with stable knockdown of interferon regulatory factor 7 increased IAV yields.

Thermostability is also important for vaccine quality, especially in low-income countries that lack the infrastructure to maintain a low and stable temperature during

vaccine transportation (183). The reduced thermostability of live attenuated vaccine for the 2009 H1N1 pandemic virus, may have been responsible for its restricted replication in vaccinated persons (184). Mutations in HA protein can improve thermostability; for example, mutations S133N, T189A, N198D, and L226Q in the RBS of influenza HA were reported to be associated with a significant increase in thermostability of an H9 IAV (185).

The objective of this study was to randomly introduce mutations into HA and screen for those that improved IAV thermostability and yields in MDCK and Vero cells. Such studies are important to optimize the preparation of cell-based influenza vaccine.

Materials and Methods

Cells and viruses

Human embryonic kidney (293T) cells, MDCK cells, and Vero cells were purchased from American Type Culture Collection (Manassas, VA). Cells were maintained at 37°C with 5% CO_2 in Dulbecco's Modified Eagle Medium (GIBCO/BRL, Grand Island, NY) supplemented with 5% fetal bovine serum (Atlanta Biologicals, Lawrenceville, GA) and penicillin–streptomycin (Invitrogen, Carlsbad, CA). The HA gene of CA/04 was cloned into the vector pHW2000 and used as template for construction of the mutant library; mouse-adapted CA/04 (139) was used for challenge experiments in mice. Influenza A/Texas/50/2012(H3N2) (TX/50) and A/canine/Iowa/13628/2005 (H3N8) virus (canine-H3N8) were used to validate the identified molecular markers.

The viruses generated by reverse genetics were propagated in MDCK cells and cultured at 37°C with 5% CO_2 in Opti-MEM medium (GIBCO/BRL, Grand Island, NY)

supplemented with 1 μg/ml of TPCK (*N*-tosyl-L-phenylalanine chloromethyl ketone)-Trypsin (Sigma-Aldrich, St. Louis, MO) and penicillin–streptomycin (Invitrogen, Carlsbad, CA). Virus titers were determined by $TCID_{50}$ in MDCK cells.

The viral total protein yield in eggs was tested as described by Adamo et al (133). Briefly, 10-day-old chicken embryotic eggs were infected with an influenza virus and then incubated at 37°C for 72 h. The allantoic fluid of the infected eggs was collected for virus purification and quantification of protein concentrations as described below.

Extractions of RNA and plasmids

RNA was extracted by using an RNeasy Mini Kit (QIAGEN, Valencia, CA); the plasmids used for transfection were prepared by using the GeneJET Plasmid Miniprep kit (Thermo Scientific, Waltham, MA).

Mutant generation using epPCR-based reverse genetics strategy

The mutant library with random mutations in HA RBSs was generated by epPCR as previously described (186). In brief, the randomly mutated short sequences (about 200 nucleotides) were used as primers in the site-directed mutagenesis with HA-pHW2000, leading to plasmids with random mutations in the HA RBSs. This strategy can avoid the need for a labor-intensive gene cloning process, and the HA-pHW2000 with mutations can be used directly in generating vaccine candidates (186). One day before transfection, 293T cells and MDCK cells were co-cultured in 24-well plates, using a 20:1 ratio of 293T cells to MDCK cells. The cell cultures were transfected with 125 ng of each of eight plasmids: a mutant HA gene, NA gene of CA/04, and 6 internal genes (PB2, PB1, PA, NP, M, and NS) of influenza A/PR/8/1934(H1N1) virus. Transfection was done

using TransIT-LT (Mirus, Madision, WI) according to the manufacturer's instructions. In brief, TransIT-LT transfection reagent was mixed with DNA at 2.5 μl /1 μg, incubated at room temperature for 20 min, and then added to the cells. After 24 h, Opti-MEM medium (GIBCO/BRL) supplemented with 1 μg/ml of TPCK-trypsin (Sigma-Aldrich, St. Louis, MO) was added to the cells. After 72 h of incubation, supernatants were collected and titrated in MDCK cells.

For phenotype comparison, we generated the wild-type reassortant virus (rg-wt) containing the wild-type HA and NA genes from CA/04 and six internal genes from A/Puerto Rico/8/34 (H1N1). The reassortant rg-TX/50 with HA and NA from TX/50 and six internal genes from A/Puerto Rico/8/34 (H1N1) virus and the reassortant rg-H3N8 with HA and NA from canine-H3N8 and six internal genes from A/Puerto Rico/8/34 (H1N1) virus were also rescued by reverse genetics.

Site-directed mutagenesis

The QuikChange II Site-Directed Mutagenesis Kit (Agilent Technologies, Santa Clara, CA) was used to create specific mutations in the HA gene. We used forward primer 5′- CCACTTAAACTTCAAATTCCCAGCATTGAACGTG-3′ and reverse primer 5′- CACGTTCAATGCTGGGAATTTGAAGTTTAAGTGG-3′ to generate mutation Y161F in HA of TX/50, and we used forward primer 5′-CAAAATCTGGAAGCTCTTTCCCCACATTGAATGTGAC-3′ and reverse primer 5′-GTCACATTCAATGTGGGGAAAGAGCTTCCAGATTTTG-3′ to generate mutation Y161F in HA of canine-H3N8. To ensure the absence of unwanted mutations, Eurofins (Louisville, KY) used Sanger sequencing to completely sequence all constructs.

Growth kinetics

To determine the growth kinetics of viruses, we inoculated the MDCK cells with a testing virus at an MOI of 0.001 and then incubated the cells in 5% CO_2 at 37°C for 1 h. The inocula were then removed and cells washed twice with phosphate-buffered saline (PBS). Opti-MEM I (GIBCO, Grand Island, NY) containing TPCK-trypsin (1 μg/ml) was added to the cells, which were then incubated in 5% CO_2 at 37°C. At specified time points after inoculation, 200 μl of supernatant was collected from the incubated cells, aliquoted, and stored at −70°C until use. Virus titers in supernatant collected at the different time points were determined by $TCID_{50}$ in MDCK cells.

Viral protein purification and protein concentration quantification

Viruses were purified from the cell supernatant or allantoic fluid by low-speed clarification (2,482 × *g*, 20 min, 4°C) to remove debris and then ultracentrifuged through a gradient of 30%–60% sucrose in a 70Ti Rotor (Beckman Coulter, Fullerton, CA) (100,000 × *g*, 3 h, 4°C). The virus band was collected and purified through a cushion of 30% sucrose in a 70Ti Rotor (100,000 × *g*, 3 h, 4°C). The virus pellet was resuspended in 200 μl of PBS, and the total amount of purified virion protein was determined by using the Pierce BCA Protein Assay Kit (Thermo Scientific, Rockford, IL).

HA and HI assays

HA and HI assays were performed by using 0.5% turkey erythrocytes as described by the WHO Global Influenza Surveillance Network Manual for the Laboratory Diagnosis and Virological Surveillance of Influenza (187). Guinea pig, chicken, horse, turkey, and dog (beagle) erythrocytes were obtained from Lampire

Biological Products (Everett, PA). The erythrocytes were washed three times with 1 × PBS (pH 7.2) and diluted to 0.5% for chicken, beagle, and turkey erythrocytes, 0.75% for guinea pig erythrocytes, and 1% for horse erythrocytes.

Western blot analysis

The Western blot analysis was used to compare the influenza virus specific protein yields of a wild type virus and those of a testing mutant. The monoclonal antibody specific for influenza A virus nucleoprotein (NR-43899) and H1 (NR-42019) were obtained through BEI Resources Repository (https://www.beiresources.org). The Western blots were developed by using HRP conjugated anti-mouse IgG antibody (Sigma-Aldrich, St. Louis, MO) and DAB substrate kit (Thermo Fisher Scientific, Waltham, MA) as manufacturer described. Bands were visualized by chemiluminescence using ChemiDoc imaging system (Bio-Rad, Hercules, CA) and analyzed by ImageJ (National Institutes of Health, Bethesda, MD).

Glycan microarray and data analyses

The 66 *N*-glycans (188) were printed on *N*-hydroxysuccinimide–derivatized slides as previously described (189). All glycans were printed in replicates of 6 in a subarray, and 8 subarrays were printed on each slide. All glycans were prepared at a concentration of 100 μM in phosphate buffer (100 mM sodium phosphate buffer, pH 8.5). The slides were fitted with an 8-chamber adapter (Grace Bio-Labs, Bend, OR) to separate the subarray into individual wells for assay. Before assay, slides were rehydrated for 5 min in TSMW buffer (20 mM Tris-HCl, 150 mM NaCl, 0.2 mM $CaCl_2$, 0.2 mM $MgCl_2$, and 0.05% Tween-20) and blocked for 30 min in TSMWB buffer (TSMW buffer with 1%

BSA). Viruses were purified by sucrose density gradient ultracentrifugation and titrated to about 1.0 $\times 10^5$ HAU/ml. To 150 μl of virus, we added 15 μl of 1.0 M sodium bicarbonate (pH 9.0) and then incubated the mixture with 25 μg of Molecular Probes Alexa 488 Succinimidyl Esters (NHS esters) (Thermo Fisher Scientific, Inc., Waltham, MA) for 1 h at 25°C. After overnight dialysis against a 7 KDa Slide-A-Lyzer MINI Dialysis Devices (Thermo Fisher Scientific, Inc., Waltham, MA) to remove excess Alexa 488 dye, viruses were checked by HA assay and then bound to glycan array. Labeled viruses were incubated on glycan microarray at 4°C for 1 h, washed, and centrifuged briefly before being scanned with an InnoScan 1100 AL Mircroarray Scanner (Inopsys, Toulouse, France).

Viruses-glycan receptor binding assay

Two biotinlyated glycan analogs, carbohydrates 3'-sialyl-*N*-acetyllactosamine (3′SLN) representing SA2,3GA and 6'-sialyl-*N*-acetyllactosamine (6′SLN) representing SA2,6GA, were purchased from GlycoTech (Gaithersburg, MD). The glycan stocks were reconstituted at 1 mg/ml in 50% glycerol - PBS (vol/vol) solution according to the manufacturer's instructions and were stored at 4°C until use. The viral particles in wild-type reassortant virus bearing HA_{161Y} (rg-wt) and a mutant virus bearing HA_{161F} (rg-Y161F) were determined using a Virocyt 2100 virus counter (ViroCyt, Boulder, CO). The kinetics buffer (PBS pH 7.4 with 0.01% bovine serum albumin and 0.002% Tween-20) containing neuraminidase inhibitors (10 μM zanamivir hydrate and 20 μM oseltamivir phosphate) was used to titrate the biotinylated glycan analogs and viruses during the binding assay (190). Binding of viruses (at 1 pM/virus) to the biotinylated glycan analogs was performed in an Octet RED96 biolayer interferometer equipped with

streptavidin biosensor tips (PALL FortéBIO, Menlo Park, CA) according to the manufacturer's assay protocol: 1) Biosensor coating with biotinylated glycan analogs X 300 sec; 2) Virus association X 1,200 sec ; and 3) Dissociation in the kinetics buffer with neuraminidase inhibitors X1,000 sec. The entire measurement cycle was maintained at 30°C with orbital shaking at 1,000 rpm.

Analyses of virus thermostability

Purified viruses were diluted in PBS to 128 HAU, and dispensed by 120 µl into 0.2-ml, thin-walled PCR tubes (USA Scientific, Ocala, FL). Tubes were placed in a Gradient Veriti 96-Well Thermal Cycler #9902 (Life Technologies, Camarillo, CA). The temperature range was set at 51.5°–63.0°C. Tubes were heated for 40 min, then transferred to ice. Control samples containing 120 µL of virus were incubated for 40 min at 0°C. Virus content in each sample was determined by HA assay using a 0.5% suspension of turkey erythrocytes. Each virus sample was analyzed three times for thermostability.

Animal experiments

We intramuscularly inoculated two groups of 6-week-old female BALB/c mice (Harlan Laboratories, Indianapolis, IN) with 15 µg (in a 50-µl volume) of a formaldehyde-inactivated vaccine candidate or wild-type virus (n = 10 mice/group). Specifically, the viruses used as vaccine candidate were inactivated in 0.025% formalin at 4 °C for three days and then confirmed by three blind passages on MDCK cells. Two weeks later, we administered a booster vaccine with the same amount of immunogen. A group of mock-vaccinated mice (n = 10) received an equal volume of PBS. A group of

mice serving as environmental controls (n = 5), were not vaccinated or challenged. Two weeks after the booster vaccination, mice were anesthetized and challenged by intranasal inoculation with mouse-adapted CA/04 (139) at 10 times the 50% lethal dose. Serum samples were collected from mice before challenge and tested by HI assays. To determine lung virus titers, we euthanized five mice at day 4 after challenge. Lungs (n = 3) were homogenized and resuspended in 1 ml of sterile PBS, and virus titers were determined in MDCK cells. The lung samples were also fixed in formalin and stained with hematoxylin and eosin stain for pathologic examination. We monitored the clinical signs, survival rate, and body weight of the remaining mice (n = 5) for 14 days after challenge. The mice were euthanized on 14 days after challenge.

Biosafety and animal handling

All laboratory and animal experiments were conducted under BSL-2 conditions, with investigators wearing appropriate protective equipment, and in compliance with protocols approved by the Institutional Animal Care and Use Committee of Mississippi State University.

Structural modeling

Crystal structures of the HA protein of the A(H1N1)pdm09 virus and the binding sites of 6′SLN and 3′SLN to this protein were obtained from PDB (Protein Databank: accession numbers 3LZG, 3UBN, and 3UBQ, respectively). Structural simulation of amino acid mutations was performed on the HA by using the computer algorithm FoldX (http://foldxsuite.crg.eu) with its empirical force field with crystal waters under the following conditions: temperature of 298K, pH 7, 0.05 ion strength. The binding

structures and measure of contact distances were visualized by using Chimera (http://www.cgl.ucsf.edu/chimera/). PoseScore, which was designed for ranking near-native ligand–protein interacting structures, was used to estimate the likeness of the protein–glycan binding avidities of the wild-type and mutant viruses to that of the native virus. PoseScore scores typically range from −100 to 100; the lower the score, the lower the binding affinity. The computational analysis of the effect of mutation on HA–glycan bindings was focused on the mutants with 161 (H3 numbering) location.

Genomic sequences, molecular characterization, and statistical analyses

HA sequences of IAVs were downloaded from the database of the Influenza Virus Resource on January 30, 2017. Multiple sequence alignments were conducted by using MUSCLE software (191). The mutations were identified by using Bioedit software (192). Survival curves were calculated by the Kaplan-Meier method, and significance was analyzed with the log-rank test using Graphpad Prism 5 software (http://www.graphpad.com/scientific-software/prism/).

Results

Generation of RBS variants of CA/04 and assessment of their growth characteristics in cells

To identify mutations associated with high-yields of CA/04/PR8 reassortant viruses, a cDNA library carrying random mutations in the CA/04 HA RBS was generated by epPCR and screened by Sanger sequencing. Hereafter, we will use the term rg to refer to mutants from epPCR. For example, rg-Y161F denotes a reassortant virus from epPCR with a Y161F mutation. Each mutated plasmid, together with the neuraminidase plasmid of CA/04 and internal genes (PB2, PB1, PA, NP, M, NS) plasmids of PR8, was used to

rescue virus by a reverse genetics approach. As a result, we obtained a total of eight mutants: rg-D130E, rg-P140T, rg-L154F-K156Q, rg-S160T, rg-Y161F, rg-K174E, rg-S188I, and rg-Y201H (Table 4).

Table 4. Characterization of MDCK cell grown receptor binding site mutants generated by error-prone PCR-based mutagenesis strategy

Mutation[a]	HI[b]	$TCID_{50}$[c]	HA titers[d]				
			Guinea pig	Chicken	Horse	Turkey	Dog[e]
rg-wt	640	5.749	4	8	<2	16	8
P140T	640	5.91 ± 0.29	4	4	<2	32	4
S188I	640	5.91 ± 0.29	4	16	<2	16	4
S160T	320	6.08 ± 0.29	4	2	<2	16	8
Y201H	320	6.08 ± 0.29	16	8	<2	64	16
D130E	320	6.249	16	16	<2	128	32
K174E	320	6.249	16	16	<2	64	32
L154F-K156Q	640	6.249	16	4	<2	128	64
Y161F	640	8.249	256	1,024	<2	512	128

[a]Viruses that carry mutations at the receptor binding site of wild-type influenza A/California/04/09 (H1N1) virus (CA/04) were generated by using an error-prone–based reverse genetic system. rg-wt, CA/04 mutant.
[b]HI, hemagglutination inhibition. Titers were determined by using ferret serum (anti-CA/04).
[c]$TCID_{50}$, 50% tissue culture infectious dose. The virus titers were determined by $TCID_{50}$ assay in MDCK cells.
[d]HA,hemagglutination assays against types of red blood cells were performed using standard procedures.
[e]Beagle.

Analysis of viral growth kinetics showed that replication efficiencies of these 8 mutants varied greatly in MDCK cells. Among the mutants, rg-D130E, rg-K174E, rg-L154F-K156Q, and rg-Y161F increased virus titers compared with the wild-type reassortant virus rg-wt. Mutants rg-S160T, rg-P140T, rg-S188I, and rgY201H grew to similar titers as did rg-wt (Table 4). Among all mutants, rg-Y161F had the highest virus

titer at 8.249 log (50% tissue culture infectious dose [$TCID_{50}$]), which was >300-fold higher than that of the wild-type virus in MDCK cells. To assess the stability for the mutation, the rg-Y161F mutant was passaged on MDCK cells for three times. The rg-Y161F mutant at the third passage was sequenced to confirm no additional mutations across HA, NA and six internal gene segments.

To determine if the mutations at the RBS altered HA antigenicity, we subjected the panel of eight RBS CA/04 mutants to a hemagglutination inhibition (HI) assay, using ferret antisera against CA/04. Both turkey and guinea pig erythrocytes were used in the HI assays, and the results were identical. In summary, the HI titers of four reverse genetic variants, rg-Y161F, rg-L154F-K156Q, and rg-P140T, were equivalent to that of the rg-wt virus. The remaining four mutants, rg-D130E, rg-K174E, rg-S160T, and rg-Y201H, had 2-fold lower HI titers compared with that of the rg-wt. Thus, all of the RBS mutants were antigenically similar to the parental CA/04 virus despite the presence of an altered virus growth property. Due to its preferred growth and unaltered antigenic characteristics, rg-Y161F was selected as a candidate vaccine virus for further studies.

In addition, we compared the HI titers for the vaccine candidate rg-Y161F and the wild type of CA/04 against a panel of 18 human sera. Results showed that there were no more than 2-fold change between the HI titers of rg-Y161F and those of CA/04 wild type virus (Table 5), confirming that mutation Y161F did not alter the antigenicity of CA/04 virus.

Table 5. Serological responses of wild type and 161F mutant against a panel of human sera using HI assays.

Sera[a]	rg-wt	rg-Y161F
	HI titer[b]	
AJP125	80	40
AJP126	80	80
AJP128	40	40
AJP225	160	160
AJP226	160	160
AJP228	20	20
CAJP251	160	160
CAJP252	160	160
CAJP261	80	80
CAJP262	80	80
CAJP271	80	80
CAJP272	80	40
CAJP281	80	80
CAJP282	160	80
CAJP291	160	160
CAJP292	160	160
CAJP301	80	80
CAJP302	160	80

[a] panel of 18 human sera randomly collected from a vaccine efficacy study.
[b] HI, hemagglutination inhibition. Titers were determined by using turkey red blood cells under standard procedures.

Growth properties of rg-Y161F

To evaluate the replication efficiencies of rg-Y161F, we characterized its growth kinetics alongside rg-wt in MDCK and Vero cells. We infected cells with viruses at a multiplicity of infection (MOI) of 0.001 (MDCK cell infection) or 0.01 (Vero cell infection) and determined the growth kinetics of the viruses for up to 96 h in MDCK cells and 120 h in Vero cells. In MDCK cells, the virus titers of rg-Y161F reached $10^{8.66}$ $TCID_{50}$/ml at 72 h after infections, >300-fold higher than the highest virus titer from rg-wt (Figure 11A). In Vero cells, the titers of rg-Y161F reached $10^{8.25}$ $TCID_{50}$/ml at 96 h after infections, a titer >100–fold higher than the highest virus titer from rg-wt (Figure

11B). The total viral protein of rg-Y161F in cells reached a mean titer of 1236.2 µg/ml, 2.07-fold higher than that of wild-type CA/04 ($p = 0.008$) (Figure 11C). These results suggest that mutation Y161F facilitates the replication efficiency of CA/04 in MDCK and Vero cells.

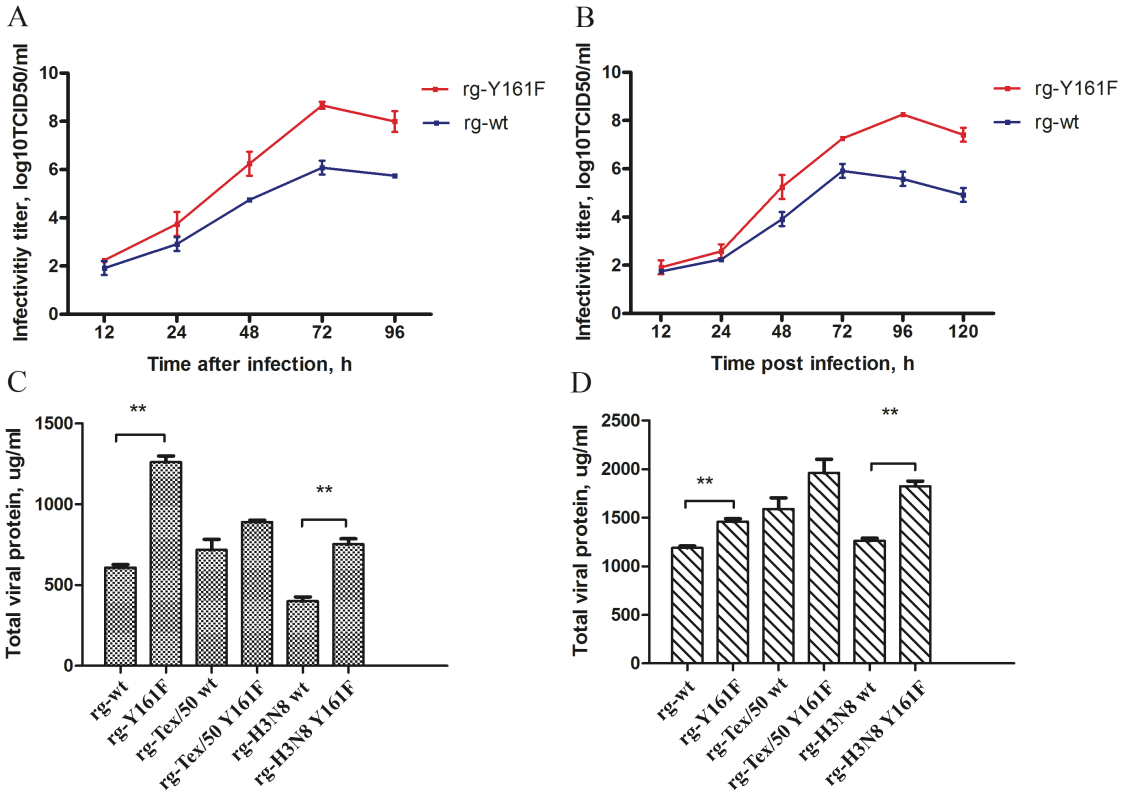


Figure 11. Growth properties of wild-type (WT) and Y161F mutant viruses. In Madin-Darby canine kidney cells (A) and Vero cells (B). Each data point represents the mean virus yield (log_{10} $TCID_{50}$/ml) from three individually infected wells ± the standard deviation. $TCID_{50}$, 50% tissue culture infectious dose. Total protein of viruses propagated in Madin-Darby canine kidney cells (C) and 10-day-old embryonated chicken eggs (D). Viruses were purified from cell supernatant or allantoic fluid by low-speed clarification and then subjected to sucrose density-gradient centrifugation. The virus band was collected and purified through a cushion of 30% sucrose. The virus pellet was resuspended in 200 µl of phosphate-buffered saline, and the total amount of purified virion proteins was determined by using a Pierce BCA Protein Assay Kit (Thermo Scientific, Rockford, lL).

Impact of HA RBS mutations on virus binding to erythrocytes

We next wanted to explore possible mechanisms for the increased yields of rg-Y161F by examining its interaction with host receptors. Due to their unique glycan receptor profiles (i.e., types and distributions of alpha-2,3-linked sialic acid on galactose [SA2,3GA] and alpha-2,3-linked sialic acid on galactose [SA2,6GA]), erythrocytes from various hosts have often been used to characterize receptor binding properties for influenza viruses through hemagglutination assays (HA assays). We used erythrocytes from guinea pig, chicken, horse, turkey, and dog (beagle) to compare the glycan profiles of the full panel of eight mutants. As shown in Table 4, all mutants and the wild-type (rg-wt) virus agglutinated erythrocytes from guinea pig, chicken, turkey, and beagle to different extents, but they did not agglutinate those from horse. The eight mutants could be separated into three groups: 1) those with increased HA titer against guinea pig, chicken, turkey, and beagle erythrocytes (rg-D130E, rg-K174E, and rg-Y161F mutants); 2) those with an hemagglutination pattern similar to that for wild-type virus (rg-S160T, rg-P140T, and rgS188I mutants); and 3) those that had increased HA titer against guinea pig, turkey, and beagle erythrocytes but no change in HA titers against chicken erythrocytes (rg-L154F-K156Q and rg-Y201H mutants). Among the eight mutants, it was striking that rg-Y161F had the highest HA titers (range, 128–1,024 HA units [HAU]) to erythrocytes from guinea pig, chicken, turkey, and beagle.

Effect of Y161F mutation on the receptor binding

To further explore the molecular mechanisms of the rg-Y161F high-growth yield phenotype, we characterized its receptor binding profile by using an *N*-linked glycan isoform microarray. The microarray consisted of 66 chemoenzymatically synthesized and

purified *N*-glycans (188) (Figure 12). As shown in Figure 13, both rg-wt and rg-Y161F bound predominantly to *N*-glycans terminating with Neu5Ac. The rg-wt virus showed binding to α2,6-Neu5Ac–linked glycans (N0x3, N113, NN144, N213, N223, and N244), whereas the Y161F mutant showed a preference for binding to α2,3-Neu5Ac–linked glycans (N0x2, N112, N122, N134, N212, N222, N234, N0x5, N115, N125, N155, N215, N225, and N255) and relatively weaker binding to α2,6-Neu5Ac–linked glycans. Neither rg-wt or rg-Y161F were observed with noticeable binding toward α2,6-Neu5Gc–terminated glycans (N013G and N023G) or α2,3-Neu5Gc–terminated glycans (N012G, N022G, N015G, and N025G) (Table 6).

Table 6. Glycan-binding affinity of rg-wt and rg-Y161F

Glycan	rg-Y161F		rg-WT	
	Mean RFU	ST	Mean RFU	ST
N000	188.492	550.5156	285	701.4905
N001	707.6438	327.5512	718	647.6747
N002	359.4128	756.848	9784	6650.103
N003	36156.92	1352.438	4101	1750.521
N004	4082.929	709.2189	2341	333.9741
N005	10568.33	3183.352	37495	3439.082
N010	2003.127	1170.928	279	368.8189
N011	-6.38956	331.721	40	193.4308
N012	87.85645	256.7783	1704	663.6608
N012G	1375.353	211.4849	3054	966.775
N013	7033.308	1224.56	1098	201.3879
N013G	107.0251	390.9631	1463	389.2667
N014	1060.667	220.0813	331	310.5335
N015	3314.584	813.7164	14582	2740.829
N015G	5498.216	1443.446	6669	6423.093
N020	1182.069	188.3146	236	727.8254
N021	1453.625	383.826	82	149.2909
N022	5320.906	755.8239	2670	685.0261
N022G	456.8535	325.5246	3025	711.258
N023	21654.22	1986.137	1394	579.4759
N023G	777.929	284.4424	1682	3332.606
N024	3212.351	829.6893	2533	999.7127
N025	6335.249	937.4352	15833	2407.865
N025G	239.6085	152.6565	2612	530.023
N026	249.1928	95.88622	730	267.9652
N030	178.9077	112.3649	94	140.3139
N031	3563.777	1083.908	1032	533.767
N032	3998.267	339.778	3654	2067.069
N033	13274.31	1664.171	3113	721.1124
N034	1472.794	365.1906	2500	917.1672
N035	13330.22	1665.253	18559	2067.877
N040	1218.809	624.9239	698	412.6959
N041	1258.743	180.6654	801	394.284
N042	830.6428	292.9735	2392	871.2684
N043	7586.005	1134.389	4347	3866.418
N044	-142.1677	239.8647	980	615.5946
N045	3739.49	395.2908	22341	2375.834

Table 6 (continued)

Glycan	rg-Y161F		rg-WT	
N049	2231.554	687.8912	2637	1119.127
N050	75.07733	225.1201	396	349.6206
N051	619.7873	326.0887	519	288.3692
N052	2450.396	237.1405	5243	2329.654
N053	6280.938	757.6721	6048	464.0661
N054	1621.351	439.9279	2405	718.5867
N055	6367.197	492.0465	21395	5372.608
N110	4210.72	1093.158	1929	583.4338
N111	1177.276	309.5498	1606	391.5552
N112	1388.132	213.1503	5276	1010.739
N113	4657.989	805.3834	1835	300.0338
N114	691.6699	180.6393	1747	680.6205
N115	2158.074	594.1063	15231	1578.711
N122	1047.888	740.8103	2926	1313.506
N123	1057.472	141.4559	836	395.9283
N124	1431.261	359.1119	1960	306.9908
N125	2273.086	228.571	9348	1251.171
N133	11023.59	486.8432	17598	3034.698
N134	5560.515	573.4557	15579	2390.632
N135	5995.005	1074.86	12655	1902.752
N144	4792.17	971.6005	2658	615.2907
N145	7766.51	2068.815	14714	1666.691
N155	6670.701	615.475	14761	3575.685
N210	6141.964	1427.767	9351	2169.871
N211	4458.315	465.7474	7704	1117.586
N212	3717.126	690.5637	3216	531.5538
N213	5090.882	717.456	2059	535.9692
N214	1314.652	147.5842	2649	833.5652
N215	4864.053	702.9951	15558	3083.07
N222	4202.733	760.8467	12091	2769.343
N223	5563.709	1585.746	6272	417.2571
N224	929.681	425.931	1595	737.1837
N225	2763.485	272.7199	6804	844.0257
N233	5712.267	564.436	9637	1302.378
N234	5375.217	327.8162	11365	2309.938
N235	4597.289	1014.55	25374	12768.65
N244	9734.495	2026.691	10064	1478.048

Table 6 (continued)

Glycan	rg-Y161F		rg-WT	
N245	7924.652	1231.199	12093	1442.352
N255	5079.7	759.0126	10519	4117.714
N6000	122.999	253.1477	987	644.012
N6030	1057.472	327.948	2049	706.8542
N6211	2848.146	475.4761	3637	1002.466
N6212	9988.479	1237.444	15156	3172.203
NC1	1191.653	418.669	1632	423.1982
NC2	142.1677	150.9403	720	287.198
PC	2309.826	262.7612	1980	500.1264
SG3LNnT	5250.621	1317.346	6814	1574.936
LNnT	1544.676	394.8941	2180	907.6503

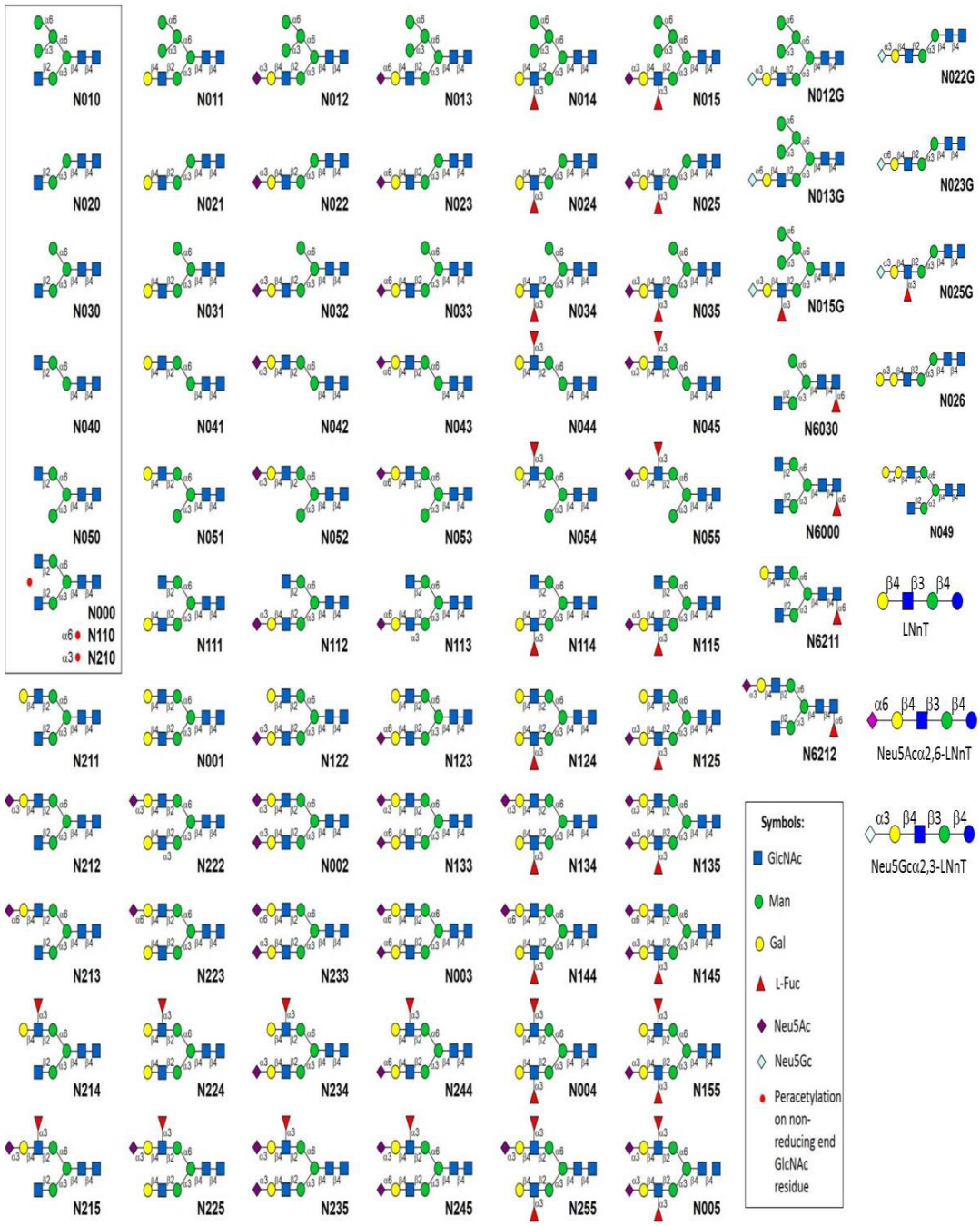


Figure 12. Structures of chemoenzymatically synthesized *N*-linked glycans on the isoform microarray.

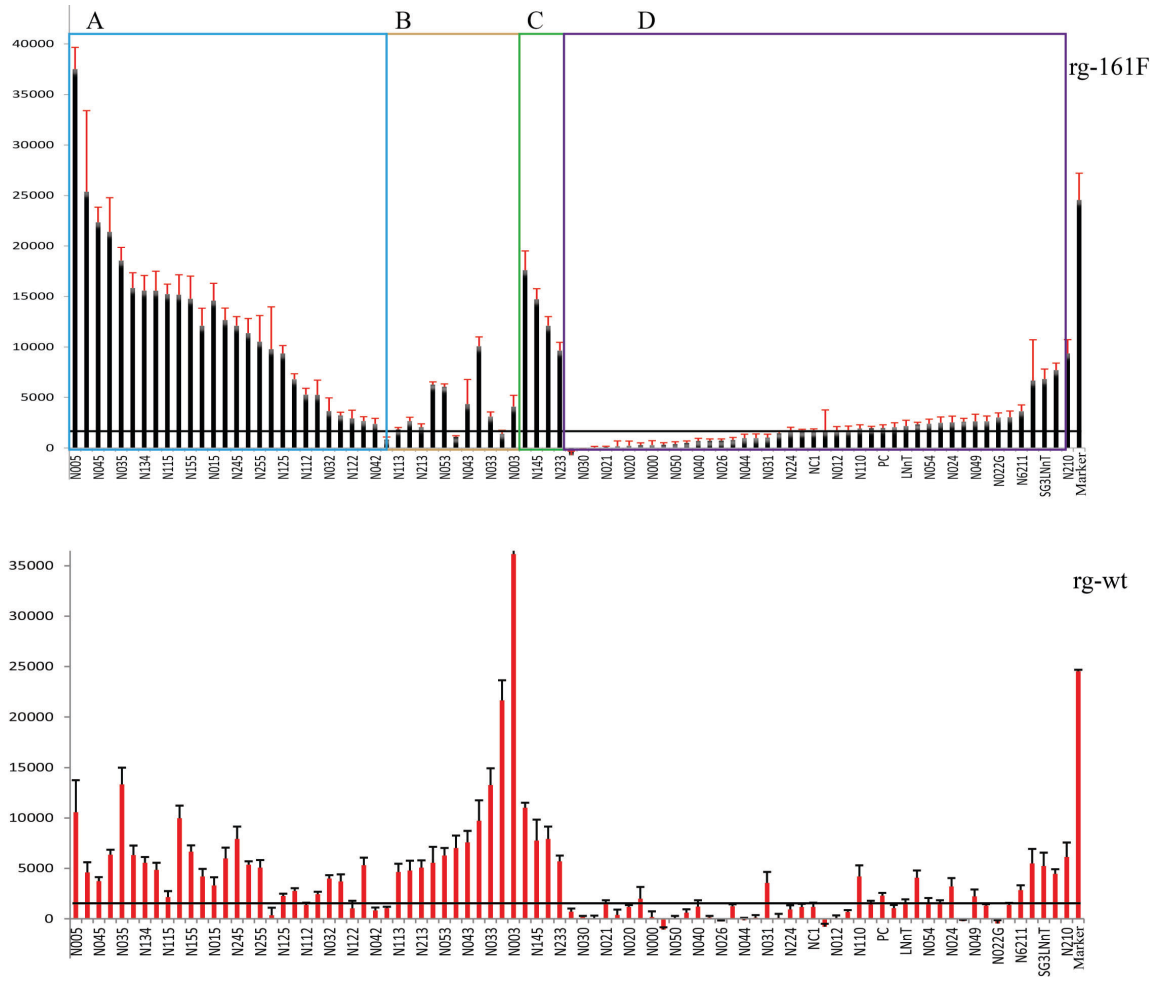
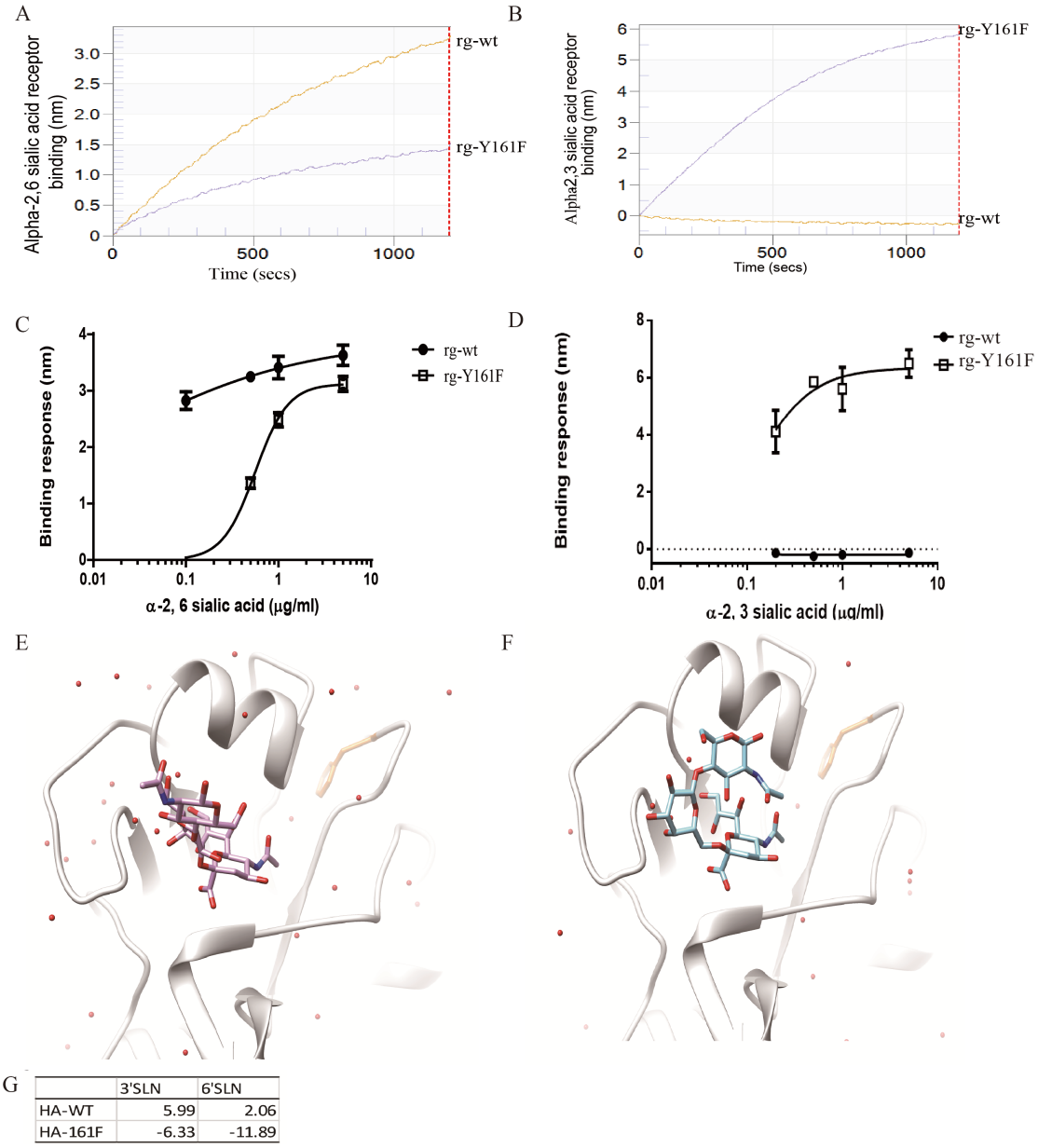


Figure 13. Receptor binding specificity of wild-type and Y161F hemagglutinin 1 mutant viruses analyzed by glycan microarray analysis.

A–D represent different categories of glycans: A, alpha-2,3 sialic acid glycans; B, alpha-2,6 sialic acid glycans; C, alpha-2,3 and alpha-2,6 acid glycans; D, non–sialic acid glycans. Vertical bars denote the fluorescent binding signal intensity. The signal for the marker on the array was derived from the reaction between the anti-Human IgG, Cy3 (0.01 mg/ml) and anti-Human IgG, Alexa647 (0.01 mg/ml), which serve as a positive control in glycan microarray.

To confirm the binding profiles revealed by glycan array, we performed glycan binding assays to characterize the dynamics and avidity of virus binding to two glycan analogs, 3′SLN and 6′SLN. The representative binding plots for each glycan analog are shown in Figure 14A and Figure 14B. Both rg-wt and rg-Y161F showed strong binding to 6’SLN, although rg-Y161F exhibited relatively weaker binding affinity than rg-wt

(<1.2 fold, Figure 14B). Conversely, rg-Y161F showed much stronger binding to 3′SLN; whereas rg-wt had no detectable binding at the same 3′SLN concentration range tested (Figure 14D). These results confirmed the glycan array profiles that the Y161F mutation in HA dramatically increased binding affinity to 3'SLN while retaining strong binding affinity to 6'SLN.



G

	3'SLN	6'SLN
HA-WT	5.99	2.06
HA-161F	-6.33	-11.89

Figure 14. Glycan binding specificity of virus by Bio-Layer Interferometry (fortéBIO, Menlo Park, CA).

A, B) Representative binding curve of H1N1 CA WT and Y161F toward alpha 2,6–linked sialic acid (SA2,6GA) and alpha 2,3–linked sialic acid (SA2,3GA) receptor, respectively. The streptavidin-coated biosensors were first preloaded with biotin-labeled sialic acid receptors, followed by the 1 pM each virus binding for 1,200 seconds in a standard kinetic buffer with neuraminidase inhibitors (zanamivir hydrate and oseltamivir phosphate). C, D) Sialic acid receptor concentrations were titrated from 0.1 to 5 μg/ml (SA2,6GA) or 0.2 to 5 μg/ml (SA2,3GA) when loading with the biotin-labeled receptors. Each individual concentration was performed three times independently, and the binding response unit (nm) was recorded at the 1,196 second time point (4 seconds before the

start of dissociation).Three-dimensional structures of the hemagglutinin of wild-type (HA-WT) influenza A/California/04/09 (H1N1) virus and mutant viruse (HA-161F) in contact with avian-like receptor analog carbohydrate 3'-sialyl-*N*-acetyllactosamine (3′SLN) (E) and human-like receptor avian 6'-sialyl-*N*-acetyllactosamine (6′SLN) (F). G) The calculated PoseScore for viruses.

Structural mechanism of increased rg-Y161F binding to 3'SLN and 6'SLN

Crystal structure modeling was performed to characterize the effect of the Y161F mutation on binding affinity between HA and two testing glycan analogs, 3'SLN (Figure 14E) and 6'SLN (Figure 14F). For 3'SLN, the PoseScore of the wild-type CA/04 HA was 5.99, whereas that for the Y161F HA was −6.63. For 6'SLN, the PoseScores for wild-type CA/04 HA and the Y161F mutant were 2.06 and −11.89, respectively (Figure 14G). These results suggest that the Y161F mutation leads to increased binding of CA/04 to 3'SLN and 6'SLN.

Effect of the Y161F mutation on replication efficiency of other IAVs

To understand the naturally occurring molecular polymorphisms at residue 161 of HA, we compared a total of 59,016 HA sequences covering all 18 documented HA subtypes (H1–H18) of IAVs. The results showed that Y161 is conserved in H1–H5, H8, H9, H11, H13, H14, and H16 IAV subtypes, whereas F161 is conserved in H7, H10, H12, and H15 IAV subtypes (Figure 15A). Phylogenetically, H7, H10, and H15 are group 2 HAs, and H12 is a group 1 HA (Figure 15B).

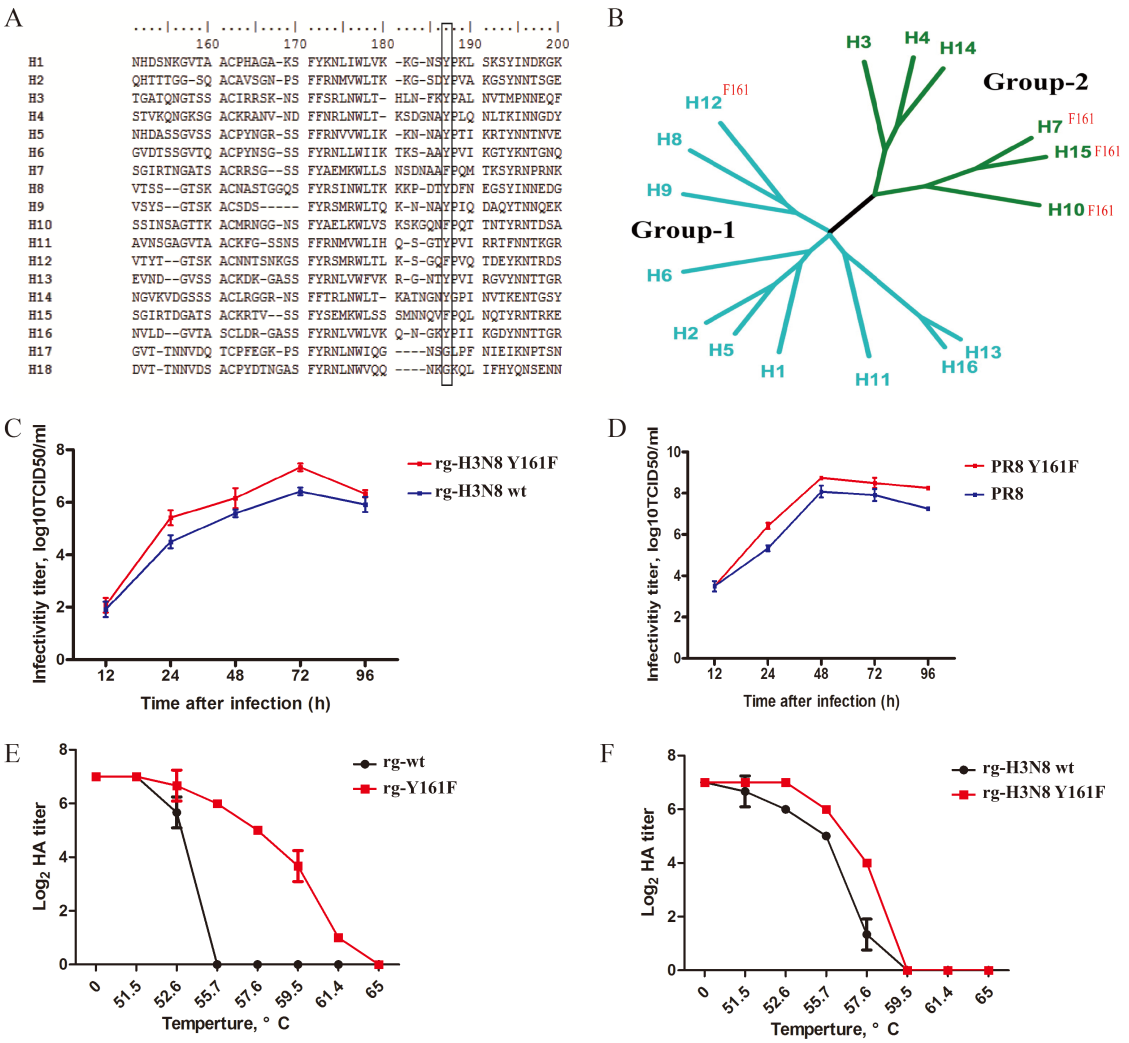


Figure 15. Effect of 161F mutation on growth and thermostability of influenza A virus

(A) Sequence alignment of hemagglutinin 1 (HA1) from 18 different influenza A virus HA subtypes (H1–H18). Residues 161 (H3 numbering) are indicated by vertical rectangle. (B) Y161 is conserved in subtypes H1–H5, H8, H9, H11, H13, H14, and H16. F161 is conserved in subtypes H7, H10, H12, and H15. Each residue is numbered according to the H3 HA numbering. C) Growth properties of canine subtype H3N8 influenza virus (cH3N8) and its hemagglutinin 1 (HA1) F161 mutant virus (cH3N8 Y161F) and of (D) influenza A/Puerto Rico/8/34 (H1N1) virus (PR8) and its HA1 F161 mutant virus (PR8 Y161F) in Madin-Darby canine kidney cells. Each data point represents the mean virus yield (log10 TCID50/ml) from three individually infected wells ± the standard deviation. E) Effect of 161F mutation (H1N1 161) on the thermostability of influenza A/California/04/09 (H1N1) virus (H1N1 WT). F) Effect of 161F mutation (H3N8 161F) on the thermostability of influenza A/canine/Iowa/13628/2005 (H3N8) virus (H3N8). The viruses with equal HA titers were incubated at the indicated temperatures for 40 minutes and then titers were determined. TCID50, 50% tissue culture infectious dose.

To test whether theY161F mutation would increase growth yields in IAVs other than CA/04, we generated 161F mutants for two additional strains: TX/50 and canine-H3N8. Analysis of growth kinetics in MDCK cells at an MOI of 0.001 showed that the TX/50 161F mutant generated viral titers of $10^{3.17}$, $10^{5.92}$, $10^{7.20}$, and $10^{7.00}$ TCID50/ml at 12, 24, 48, and 72 h, respectively. This finding compared with $10^{2.33}$, $10^{5.25}$, $10^{7.08}$, and $10^{7.00}$ TCID50/ml at 12, 24, 48, and 72 h, respectively, for the wild type TX/50 virus. The TX/50 161F mutant in MDCK cells generated a mean total viral protein titer of 882.2 mg/ml, 1.15-fold higher than that for the TX/50 wild-type virus (Figure 11C). When subjected to ferret antisera in an HI assay, the TX/50 wild-type virus and 161F mutant both had mean HI titers of 1:1280.

Analysis of growth kinetics in MDCK cells at an MOI of 0.001 showed that the 161F mutant of canine-H3N8 had the highest titer ($10^{7.249}$TCID50/ml) at 72 h after infection; this titer was about 10-fold higher than that generated by the canine-H3N8 wild-type virus (Figure 15C). The mean HI titers of wt and rg canine-H3N8 to ferret antisera was 640 and 533, respectively. The total viral protein of the canine-H3N8 161F mutant reached a mean titer of 777.2 μg/ml, which was 1.85-fold higher than that of canine-H3N8 wild-type virus ($p = 0.02$) (Figure 11C). We further compared influenza virus specific protein yields by Western blot analysis using NP specific monoclonal antibody, and results showed that the Y161F mutation increased the influenza specific NP protein yields for rg-CA/04, rg-Tex/50, and rg-H3N8 by 42%, 18%, and 20%, respectively. By using H1 specific monoclonal antibody, results showed that the Y161F mutation increased the HA protein yields for rg-CA/04 by 39% (Figure 16).

The results suggest that the Y161F mutant can enhance replication efficiency but does not change the antigenicity of TX/12 and canine-H3N8 viruses.

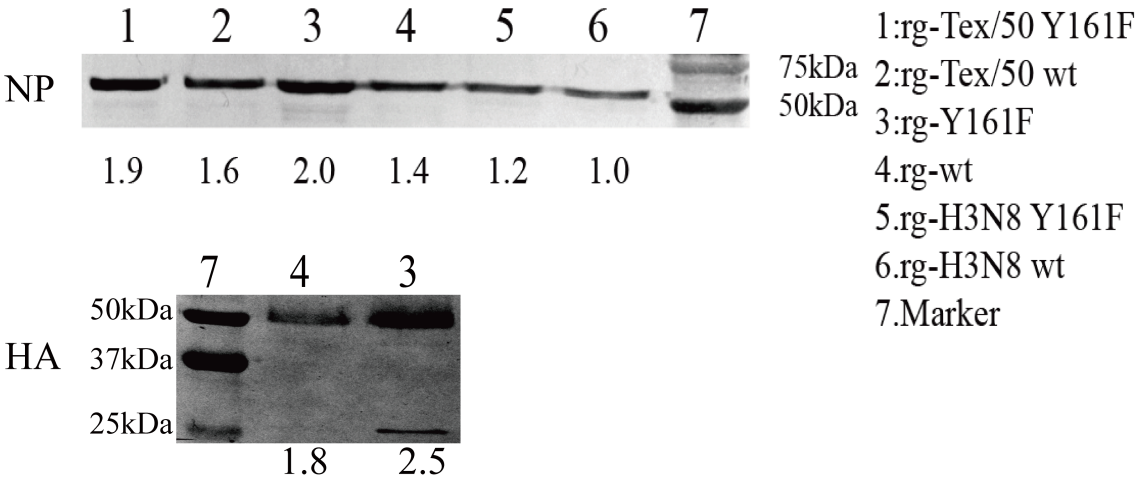


Figure 16. Western blot showing the NP and HA protein expression level for the wild type and mutant viruses. The bands were analyzed by ImageJ software.

Effect of the Y161F mutation on replication efficiency of IAVs in eggs

While we were able to show an elevated growth phenotype of HAY161F containing viruses in cells, we next sought to determine its impact on egg growth. To do this we quantified the total protein yields of wt and Y161F mutant H1N1, H3N2, and H3N8 IAVs in eggs. The total viral protein of the rg-Y161F mutant reached a mean titer of 1,460.2 μg/ml, 1.22-fold higher than that of wild-type CA/04. Similarly, for TX/50, the Y161F mutation conferred a 1.23-fold increase in total viral protein (1,589.2 μg/ml). The total viral protein of the canine-H3N8 Y161F mutant reached a mean titer of 1,824.2 μg/ml, 1.45 fold higher than that of wild-type canine-H3N8 (Figure 11D). Thus, the Y161F mutation also increased viral replication efficiency of three H1N1, H3N2, and H3N8 IAVs tested in eggs.

Impact of the Y161F mutation on viral thermostability

Another desirable property of an influenza vaccine virus is an increased stability. To determine whether the Y161F mutation correlated with changes in viral thermostability, purified viruses were diluted to 128 HAU/50 μl and incubated at a series of high temperatures (51.5°–65°C) for 40 min, and the integrity of the HA protein was then detected by an HA assay using 0.5% turkey erythrocytes. The CA/04 rg-wt virus showed a precipitous drop in HA titer (from 128 to 2 HAUs) after 40 min of incubation at 55.7°C. In contrast, rg-Y161F maintained an HA titer of 64 HAU at 55.7°C which did not drop until 59.5°C. The mutant virus maintained an HA titer of 2 HAU even at 61.4°C. The rg-wt virus completely lost its hemagglutination ability at 55.7°C (Figure 15E). Similar phenotypes were observed in the 161F mutants for canine-H3N8. The canine-H3N8 Y161F mutant had an HA titer of 16 HAU when incubated at 57.6°C for 40 min; this titer was 8-fold higher than that of wild-type H3N8 virus (Figure 15F). For TX/50, the wild-type and Y161F viruses maintained titers of 8 HAU and 16 HAU, respectively, at 61.4°C (data not shown). Taken together, the results show that the Y161F mutation conferred higher viral temperature stability on viruses.

High-yield vaccine candidate protected mice against lethal challenge

Although we had shown that the Y161F mutation had no impact on HAI titers, we wanted to directly confirm that there was not an associated loss in vaccine efficacy. To do this we prepared inactivated whole virus vaccines from rg-wt and rg-Y161F CA/04 viruses and evaluated their efficacy in a mouse model. Mice were administered vaccine or PBS (as a mock vaccine, and 2 weeks later we collected blood samples for testing. All vaccinated mice had seroconverted and their HI titers were substantially higher than those

of the mock-vaccinated mice (Table 7); mice vaccinated with rg-wt vaccine and rg-Y161F vaccine had log2(HI) titers of 7.65 ± 0.57 and 7.32 ± 0, respectively. The heterologous HI titers were indistinguishable (p = 0.3739) from homologous titers, again demonstrating the antigenic similarly of wt and mutant viruses.

Table 7. Immunologic and pathogenic responses in mice challenged with mouse-adapted influenza A/California/04/09 (H1N1)

Vaccine group	$Log_{10}TCID_{50}$/ml, mean ± SD[a]	Log_2HI titer, mean ± SD[b]	log_2HI titer, mean ± SD[c]
WT	Below detection limit	7.65 ± 0.57	7.32±0
Mutant	Below detection limit	7.32±0	7.32±0
PBS[d]	5.28 ± 0.14	Below detection limit	Below detection limit

[a]Groups of BALB/c mice were inoculated intranasally with 10× the 50% lethal dose of mouse-adapted CA/04 virus under light anesthesia. Three mice from each group were euthanized on day 4 after virus challenge, and virus titers in lungs were determined by $TCID_{50}$ (50% tissue culture infectious dose) in MDCK cells.
[b]Serum samples were collected before challenge, and antibody response levels against the wild-type (WT) virus were measured by using the hemagglutination inhibition (HI) assay.
[c]Serum samples were collected before challenge, and antibody response levels against the immunogen mutant were measured by using the HI assay.
[d]PBS, phosphate-buffered saline mick-infection.

Following challenge, a high level of virus replication (up to 105.45 TCID50) was observed in mock-vaccinated mice, but no virus was detected in mice vaccinated with rg-wt or rg-Y161F derived vaccine. Mock-vaccinated mice exhibited signs of inactivity and lethargy, had ruffled hair, and rapidly lost weight following challenge (Figure 17A). In

contrast, mice vaccinated with rg-wt or rg-Y161F derived vaccines did not exhibit any detectable clinical signs. All vaccinated mice survived, but by post-challenge day 6, all mock-vaccinated mice (n = 5) had lost 25% of their pre-experiment body weight and were euthanized (Figure 17B).

Results from histopathologic analyses showed that mice immunized with rg-wt or rg-Y161F derived vaccine had no apparent pathologic changes (Figure 17C and Figure 17D). However, the mock-vaccinated mice exhibited severe bronchiolitis, and their bronchioles showed necrosis and some attenuated regenerative epithelial cells along the basement membrane (Figure 17E).

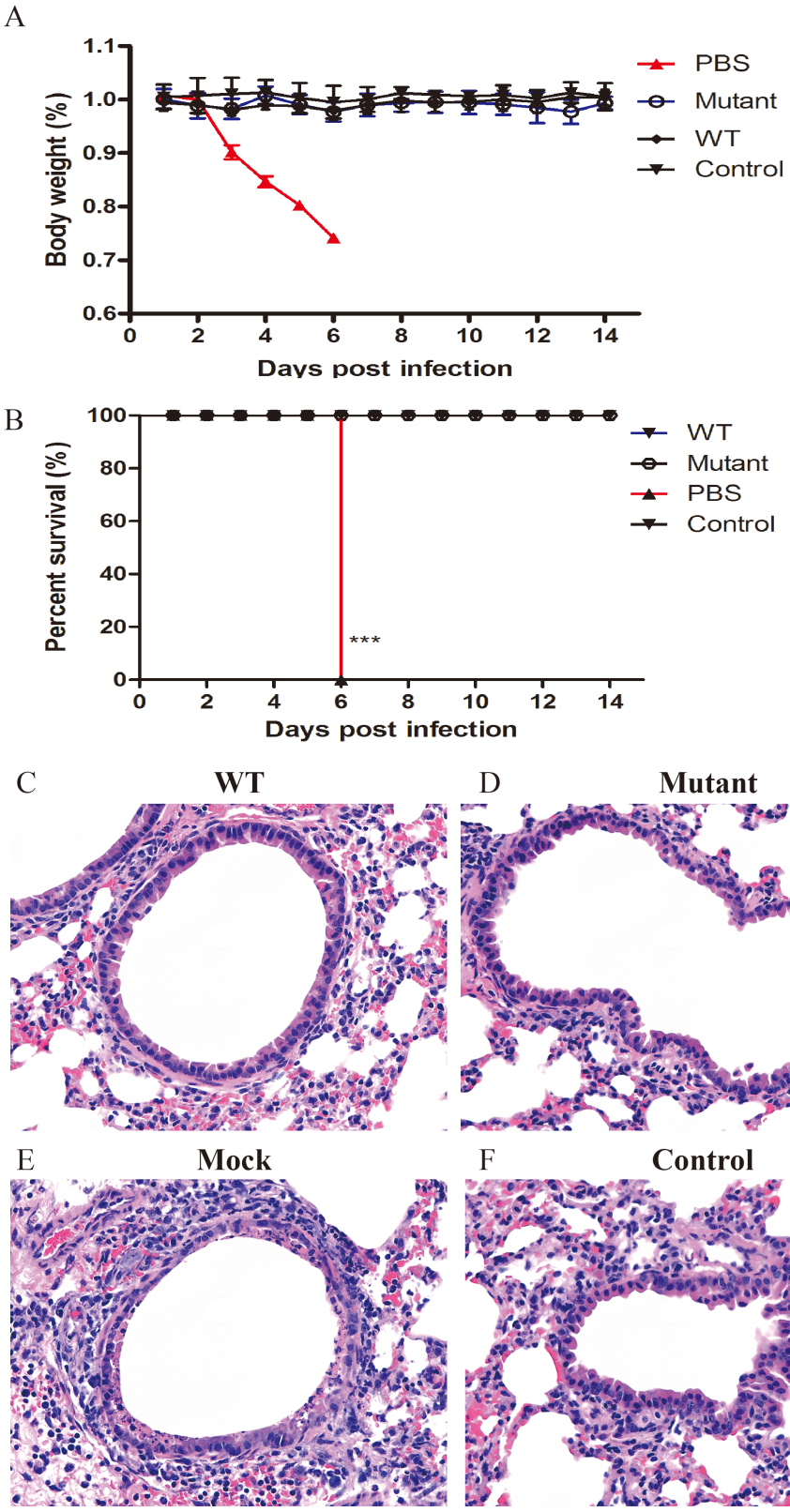


Figure 17. Weight loss and survival among vaccinated mice challenged with a lethal dose (LD) of influenza A/California/04/09 (H1N1) virus (CA/04).

Mice were vaccinated with wild-type (WT) or Y161F (Mutant) virus vaccine or mock-vaccinated with phosphate-buffered saline (PBS) and then challenged with 10 LD50 of mouse-adapted CA/04 by intranasal inoculation and monitored daily for 2 weeks. A) Percent change in body weight; each point represents the mean body weight of 5 mice per group. B) Percent surviving mice after challenge. ***, $p<0.0001$ for PBS group versus WT group, as calculated by GraphPad Prism 5 software (http://www.graphpad.com/scientific-software/prism/). Histopathologic changes in hematoxylin and eosin–stained lung samples from groups of mice vaccinated with rg-CA/04 wild-type (WT) (C) or rg-CA/04 161F (Mutant) (D) virus vaccine; mock-vaccinated with phosphate buffered saline (E); or serving as controls (F). Samples were collected 4 days after vaccinated and mock-vaccinated mice were challenged with 10 LD50 of mouse-adapted influenza A/California/04/09 (H1N1) virus. Magnification ×200.

In summary, our results from experiments in mice suggest that the Y161F mutation in HA did not alter the viral antigenicity of CA/04 or efficiency of the vaccine in mice.

Discussion

An effective influenza vaccination is dependent on a number of factors, not the least of which are high yielding and stable vaccine viruses. A high yield of virus is critical to vaccine manufacturing, and thermostability of viral antigens is critical for vaccine shelf life which can be compromised during transportation and storage. In this study we screened a group of randomly generated CA/04 mutants carrying substitutions at the HA RBS. While a number of these mutants displayed enhanced features, one, containing an Y161F change, had increased thermostability and the highest viral yields. This mutation was also able to impart these properties on a seasonal H3N2 and a canine H3N8 virus showing that the effect was not subtype-specific. Compared to those on CA/04, the effects of Y161F on the seasonal H3N2 and canine H3N8 viruses we tested were relatively less effective. One possible explanation is that, unlike the CA04 wild type

virus which replicates poorly in MDCK cells ($10^{5.75}$ TCID50/ml at 72 h), TX/50 and canine-H3N8 wild type viruses have a high replication efficiency with $10^{7.20}$ TCID50/ml and $10^{6.50}$ TCID50/mL at 72 h, respectively. Furthermore, the Y161F mutation did not change the antigenicity of the H1N1, H3N2, and H3N8 viruses tested. Animal experiments further showed that the Y161F change in CA/04 did not have measurable impact on the efficacy of inactivated whole-virus vaccines containing it. These results highlight the application potential of the HA 161F signature in influenza vaccine manufacture.

The initial step in viral infection is the binding of HA to the sialic acid receptors on the epithelial cell surface (193). This interaction is mediated by the RBS which is located at the globular head of the HA and consists of the 130 and 220 loops and the 190 helix. Mutations at the RBS have long been known to affect the yield of a vaccine strain which was our rationale for targeting it. For example, mutation L194P increased the yield of an A/England/611/07 (H3N2) 6 + 2 reassortant virus (153), and single or double mutations at 191 (194 in H3), 197 (200), 222 (225), and 223 (226) increased replication of A/California/7/09 (H1N1) in eggs (120). Mutations at residue 186 and 194 in the HA of an A(H1N1)pdm09 virus have also been shown to improve viral titers in MDCK cells and eggs (181). Avian virus yields can also be improved by targeting the RBS and the double mutation of N133D/G198E in the HA have been reported to increase H7N9 viral yields (194). The challenge of targeting the RBS for improving virus yields is that some mutations which lead to improved growth also alter antigenicity. For example, mutation G144E was shown to increase the yield of B/Victoria/504/2000, but the antigenic properties of the virus were also changed (154). Single amino acid changes at positions

119 (122 in H3), 153 (156), 154 (157), and 186 (189) could increase the yield of A/California/7/09 (H1N1) in eggs, but mutations at residues 153 (156 in H3), 154 (157), and 155 (158) drastically altered viral antigenicity (119).

The natural plasticity of the RBS for accepting substitutions was highlighted by Yasugi et al. (195). These authors used Roche 454 sequencing to directly sequence nasal specimens from three patients infected with A(H1N1)pdm09 virus. They found the virus' HAs showed high levels of amino acid diversity, with polymorphisms ranging from 3.45–8.59% for a K119N substitution, 1.01–4.99% for a N125D substitution, 0.74–21.49% for a D222G substitution, and 2.39-4.64% for a Q223R substitution (195). The percentages of the K119N, N125D, D222G, and Q223R mutations reached up to 60.4%, 96.7%, 85%, and 95.8%, respectively, after egg adaptation of the primary specimens (195). Similarly, mutations K119N and D222G were also found in the high-yield, egg-adapted A(H1N1)pdm09 virus vaccine strain NIBRG-121xp, and a Q223R mutation was found in another high-yield, egg-adapted A(H1N1)pdm09 strain NYMC-181A (119). These findings demonstrate that selection or generation of an HA variant, especially in RBS of HA, can in some cases be rapidly achieved. However, the genetic features for high-yield property are still not fully understood and generation of high-yielding viruses using classical virologic techniques is sometimes more challenging. In this study we therefore opted for a random mutagenesis approach which generated eight mutants, of which four (D130E, K174E, L154F-K156Q, Y161F) increased viral yields in cells without changing antigenic properties (Table 4). Among these mutants, Y161F had the largest increase in viral replication efficiencies in both MDCK and Vero cells (Figure 11).

Binding to the host cell is the first step of influenza virus infection. Thus, the presence of favorable receptors on a specific cell is one of the key factors determining host and tissue tropisms of IAV. Most studies related to influenza receptors have classified the sialic acid receptors into two groups on the basis of positions of the sialic acid–galactose linkage: SA2,3GA or SA2,6GA. Both SA2,3GA and SA2,6GA are present in any single cell type, but their distributions varies based on the types of cells. For example, SA2,3GA and SA2,6GA are present in both MDCK and Vero cells, but SA2,3GA is considerably more abundant than SA2,6GA in both cell lines (196-198). However, in chicken erythrocytes, SA2,6GA is more abundant than SA2,3GA (55). Thus, an ideal high-yield vaccine candidate would have high binding affinities to both SA2,3GA and SA2,6GA (199). This double binding was the phenotype the Y161F mutant in multiple assays, providing a plausible explanation for the increased yields in MDCK and Vero cells.

The HA RBS is a member of the lectin superfamily, and the specificity of RBS contributes to the host range of IAVs. Results of our structure modeling showed that residue 161 locates at the top of the RBS (Figure 14). Others have previously reported that an Y161A substitution in the H5N1 HA changed the receptor binding preference from Neu5Ac to N-glycolylneuraminic acid (Neu5Gc), however, and the mutation Y161A did not increase virus replication efficiency on MDCK cells (200). In addition, the results of that study showed that the Y161A mutants were best in viral replication and plaque forming ability. Our findings suggest that the binding preference of virus is changed by introducing the F substitution at residue 161. This substitution in CA/04 accommodates virus binding to both SA2,3GA and SA2,6GA and is responsible for the

acquisition of specificity to SA2,3GA receptors. The Phe side chain lacks the O4 hydroxyl group present on the Tyr shortening the distance between the oxygen atom of a water molecule and residue 161; and this potentially facilitates acquisition of the viral specificity to SA2,3GA.

Environmental factors such as temperature and pH have been reported to affect airborne transmission of IAV, possibly by affecting viral thermostability (201). Thus, viral thermostability could affect the quality of and transmissibility of influenza viruses. A T318I substitution in HA increased its stability and affected the binding property of an reassortant H5 HA/H1N1 influenza virus (150). Thus, identification of amino acid substitutions that increase viral thermostability would also be important risk assessment factors for emerging IAVs, such as those of subtypes H5 and H7 (202). Moreover, in agreement with our observations, Watanabe et al. (203) reported that HA thermostability was correlated with viral replication and glycan receptor binding of H5N1 viruses. However, further studies are needed to interpret the molecular mechanism of increased HA thermostability by mutations.

In conclusion, our study shows that the mutation Y161F in the RBS of the A(H1N1)pdm09 HA significantly increased viral yield in MDCK and Vero cells by promoting virus binding to both SA2,3GA and SA2,6GA without altering the original antigenicity. The results suggest that Y161F mutation in HA might have potential in generating seed virus of high yield for cell-based influenza vaccine development and production.

CHAPTER IV

MUTATION W222L AT THE RECEPTOR BINDING SITE OF HEMAGGLUTININ COULD FACILITATE VIRAL ADAPTION FROM EQUINE INFLUENZA A(H3N8) VIRUS TO DOGS

An outbreak of respiratory disease caused by the equine-origin influenza A(H3N8) virus was first detected in dogs in 2004 and since then, has been enzootic among dogs. Currently, the molecular mechanisms underlying host adaption of this virus from horses to dogs are unknown. Here, we have applied quantitative binding, growth kinetics, and immunofluorescence analyses to elucidate these mechanisms. Our findings suggest that a substitution of W222L in the hemagglutinin of the equine-origin A(H3N8) virus facilitated its host adaption to dogs. This mutation increased binding avidity of the virus specifically to receptor glycans with *N*-glycolylneuraminic acid (Neu5Gc) and sialyl Lewis X (SLe^X) motifs. We've demonstrated these motifs are abundantly located in the submucosal glands of dog trachea. Our findings also suggest that in addition to the type of glycosidic linkage (e.g., α 2,3-linkage or α2,6-linkage), the type of sialic acid (Neu5Gc or 5-N-acetyl neuraminic acid) and the glycan substructure (e.g., SLe^X) also play an important role in host tropism of influenza A viruses.

Introduction

Influenza A viruses (IAVs) can infect a wide range of species, including humans, poultry, wild birds, pigs, horses, mink, marine mammals, and dogs (204-206). IAV was first isolated from dogs in the United States in January 2004 during an outbreak of respiratory disease among dogs in Florida. Epidemiologic evidence along with antigenic characterization and genetic analyses of the isolated subtype A(H3N8) IAV suggested that the virus had "jumped" from horses to dogs (206). The virus gained the ability to transmit between dogs and became enzootic among dog populations in the United States.

IAV transmission from birds to humans, or even between different mammalian hosts, is not uncommon, but such transmission is usually transient and causes only sporadic cases or small outbreaks with limited spread. It has been postulated that the adaption of avian IAVs to humans requires an intermediate mammalian host, such as the pig, that harbors both avian-like receptor saccharides containing α2,3-linked sialic acid-galactose (SA2,3Gal) and human-like receptor saccharides containing α2,6-linked sialic acid-galactose (SA2,6Gal) in the respiratory tract (207, 208). However, the glycosidic linkages alone cannot explain interspecies transmission of IAV between mammalian hosts or between mammalian and avian hosts. For example, receptor saccharides containing SA2,3Gal (26, 209, 210) are predominant in the respiratory tract of dogs and horses, but equine IAV (EIV) cannot be transmitted freely between dogs and horses, as evidenced by the fact that the North American-lineage A(H3N8) EIV was only recently introduced into dogs and subsequently became enzootic among them (211).

The adaption of an influenza virus to a new host usually requires mutation(s) at one or multiple virus gene segments (212, 213). It has been suggested that the A(H3N8)

EIV and A(H3N8) canine IAV (CIV) have minimal biological differences, and their cross-species transmission and adaption may be mediated by subtle changes in virus biology (214). Sequence analyses suggested that A(H3N8) CIV, compared with its precursor North American lineage A(H3N8) EIV, had mutations N54K, N83S, W222L, I328T, and N483T in the hemagglutinin (HA) (Table 8) and a few additional mutations in neuraminidase and other internal gene segments (215, 216). A difference in receptor binding between the wild-type A(H3N8) EIV (wt-eH3N8) and the wild-type A(H3N8) CIV (wt-cH3N8) has also been suggested (217).

The molecular mechanism for how A(H3N8) IAV adapts from equine to canine remains unclear, largely because the types of sialosides and their relative distributions and abundance in the respiratory tract of dogs and horses have not been fully elucidated. We aim to gain an understanding of the molecular mechanisms that enabled A(H3N8) IAV to switch from horses to dogs and to provide insight into how this emerging IAV adapted to a new host.

Table 8. Predominant residues at the receptor binding sites of H3N2 (canine, avian), and H3N8 (canine, equine) influenza viruses.

Canine(H3N2)	Avian (H3N2)	Canine (H3N8)	Equine (H3N8)		
Y	Y	Y	Y	98	
G	G	G	G	134	130 Loop
G	G	R	R/G	135	130 Loop
S	S	S	S	136	130 Loop
G	G	G	G	137	130 Loop
A	A	A	A	138	130 Loop
W	W	W	W	153	
H	H	H	H	183	
N	N	N	N/T	188	190 Loop
Q	Q	Q/N	Q/K	189	190 Loop
E	E	E	E	190	190 Loop
Q	Q	Q	Q	191	190 Loop
T	T	T	T	192	190 Loop
S	S	K	K	193	190 Loop
L	L	L	L	194	190 Loop
T	T	T	T	195	190 Loop
P	P	P	P	221	190 Loop
L	W	L	W	222	190 Loop
V	V	V/I	V	223	220 Loop
R	R	R	R	224	220 Loop
G	G	G	G	225	220 Loop
Q	Q	Q	Q	226	220 Loop
S	S	S	S	227	220 Loop
G	G	G	G	228	220 Loop

Materials and Methods

Cells and viruses

MDCK, A549, and DF-1 cells were purchased from American Type Culture Collection (Manassas, VA). The cells were maintained at 37°C with 5% CO_2 in Dulbecco's Modified Eagle Medium (GIBCO/BRL, Grand Island, NY) supplemented with 10% fetal bovine serum (FBS; Atlanta Biologicals, Lawrenceville, GA) and penicillin (100 units/ml)–streptomycin (100 μg/ml). Influenza viruses, A/canine/Iowa/13628/2005(H3N8), Miami-eH3N8, and Pennsylvania-eH3N8 maintained in our laboratory were used in the study; the viruses were propagated and titrated in the MDCK cells to determine their $TCID_{50}$.

Gene cloning, site-directed mutagenesis, and virus rescue

All gene segments of A/canine/Iowa/13628/2005(H3N8) virus were cloned into the phw2000 vector, as previously described (218). The QuikChange II Site-Directed Mutagenesis Kit (Agilent Technologies, Santa Clara, CA) was used to create specific mutations in the HA gene by using the following primers: forward 5′-TGGGGAAAATATGCAACAACTCATATAGAATTCTAGATGG-3′ and reverse 5′-CCATCTAGAATTCTATATGAGTTGTTGCATATTTTCCCCA-3′ were used to generate mutation K54N in HA of A/canine/Iowa/13628/2005 (H3N8) virus; forward 5′-GCCCTTCAGTATGAGAATTGGGACCTCTTTATAG-3′ and reverse 5′-CTATAAAGAGGTCCCAATTCTCATACTGAAGGGC-3′ were used to generate mutation S83N in HA of A/canine/Iowa/13628/2005 (H3N8) virus; and forward 5′-CGAATCTAGACCGTGGGTCAGAGGTCAATC-3′ and reverse 5′-GATTGACCTCTGACCCACGGTCTAGATTCG-3′ were used to generate mutation

L222W in HA of A/canine/Iowa/13628/2005 (H3N8) virus. In brief, the site-directed mutagenesis PCR amplification mixture contained 38.5 µl of water, 5 µl of 10× reaction buffer, 1 µl of 2.5 mM dNTP mix, 1.25 µl (100 ng/µl) of each primer, 2 µl (5 ng/µl) of wild-type (wt) HA segments of A/canine/Iowa/13628/2005 (H3N8) plasmid, and 1 µl of *PfuUltra* HF DNA polymerase (2.5 U/µl). The parameters of the site-directed mutagenesis PCR were as follows: one cycle at 95°C for 30 s, followed by 16 cycles at 95°C for 30 s, 55°C for 1 min, and 68°C for 5 min. The PCR products were then digested with *Dpn*I at 37°C for 1 h, the PCR product (2 µl) was transfected into XL1-Blue Supercompetent Cells (Agilent Technologies, Santa Clara, CA). The transformed cells were directly inoculated onto Luria broth agar plates. Then the plasmids were extracted, and the mutant viruses were rescued by plasmid-based reverse genetics methods, as described by Hoffmann et al. (219). All the viruses were sequenced to confirm no unwanted mutations.

Replication kinetics

To determine multistep growth curves, we infected MDCK, A549, and DF-1 cells with viruses at a multiplicity of infection of 0.01. After 1 h of incubation at 37°C, the cells were washed twice with warm PBS and overlaid with infection medium (Opti-MEM; GIBCO/BRL, Grand Island, NY) supplied with 1 µg/ml of TPCK-treated trypsin and penicillin (100 units/ml)–streptomycin (100 µg/ml)). Supernatants were collected at 12, 24, 48, and 72 h after infection and stored at −80°C for titration by $TCID_{50}$ assay in MDCK cells.

Trachea collection and preparation

Trachea tissues were obtained within 2 h from adult dogs following euthanasia for non-respiratory related illness. Whole tracheas were transported in pre-warmed wash medium consisting of a 1:1 mixture of DMEM and Bronchial Epithelial Cell Growth Medium (BEGM; Lonza, Walkersville, MD) supplemented with penicillin (1,000 U/ml; Sigma, St. Louis, MO), streptomycin (500 µg/ml; Sigma, St. Louis, MO). Once in the laboratory, tracheas were kept at 37°C in a 5% CO_2 humidified incubator. Over 3 h, four to six washes were performed by immersing the tissues in fresh, warm wash medium.

Preparation of the tracheal explants

After the washing steps, we removed the surrounding connective tissue exterior to the tracheal cartilage, opened the tracheas, and cut them lengthwise into four strips. Each segment was then cut into approximately 5 mm explants by Biopsy punch (IntegraMiltex, York, PA). The explant slices were placed on a 24-well transwell plate (Corning, Corning, NY) supplied with 200 µl of explant growth medium consisting of a 1:1 mixture of DMEM and BEGM supplemented with penicillin (200 U/ml), streptomycin (200 µg/ml), amphotericin B (5 µg/ml; Lonza, Walkersville, MD), and a BEGM SingleQuots Kit (Lonza, Walkersville, MD) (1 kit in 500 ml of 1:1 mixture of DMEM and BEGM). Explants were maintained at 37°C in a 5% CO_2 humidified incubator; the apical and basal growth media were changed every four h to remove any secreted cytokines and inflammatory mediators. Before being infected, the trachea explants were washed two times with 2× sterile PBS. Then 200 µl of infection medium containing BEBM with 0.5% bovine serum albumin (BSA), penicillin (200 U/ml), streptomycin (200 µg/ml), and amphotericin B (5 µg/ml) was added into the apical chamber, and the tissue was allowed

to equilibrate for 30 min in the incubator. After removal of the infection medium, 400 $TCID_{50}$ each of wt-cH3N8, Pennsylvania-eH3N8, Miami-eH3N8, rg54N-cH3N8, rg83N-cH3N8, and rg222W-cH3N8 mutants was added to the apical chamber of the well and incubated for 1 h at 37°C. After three washes with 1× sterile PBS, 500 µl of fresh growth medium was added to the basal chamber, and 200 µl of infection medium was added to the apical membrane. Infection medium was removed from the apical chamber at different timepoints, and the viral titers were determined by $TCID_{50}$ assay in MDCK cells.

Consortium for Functional Glycomics (CFG) glycan array

Viruses binding to a wide range of glycan analogs were evaluated by CFG glycan array. Viruses were first purified through a sucrose cushion at 100,000 × *g* for 3 h. The purified viruses were labeled with desiccated Alexa Fluor 488 NHS Ester (Succinimidyl Ester; Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. After dialysis with the Slide-A-Lyzer MINI Dialysis devices (Thermo Scientific, Rockford, IL), the Alexa Fluor 488–labeled viruses were transferred to clean tubes and stored at 4°C until used in glycan microarray hybridizations. In glycan hybridization, the version 5.0 glycan slides (CFG) were used as described (220). The binding image was read in a Perkin-Elmer ProScanArray scanner and analyzed using ImaGene 6.0 image analysis software (BioDiscovery Inc., El Segundo, CA). Relative fluorescence unit (RFU) data were normalized by adjusting the total RFU to the same level across all experiments. A threshold of 2,000 RFU was used to floor the samples; only glycans with at least 2,000 RFU were analyzed statistically. The Wilcoxon signed rank-sum test was used to

compare the glycan-binding patterns among rg-wt–cH3N8, rg-54N–cH3N8, rg-83N–cH3N8, and rg-222W–cH3N8 mutants.

N-Glycan isoform microarray

The 83 N-glycans (188) were printed on *N*-hydroxysuccinimide (NHS)–derivatized slides as described previously(221). All glycans were printed in replicates of six in a subarray, and eight subarrays were printed on each slide. All glycans were prepared at a concentration of 100 pM in phosphate buffer (100 mM sodium phosphate buffer, pH 8.5). The slides were fitted with an eight-chamber adapter to separate the subarray into individual wells for assay. Before the assay, slides were rehydrated for 5 min in TSMW buffer (20 mM Tris-HCl, 150 mM NaCl, 0.2 mM CaCl2, and 0.2 mM MgCl2, 0.05% Tween). Viruses are purified by sucrose density gradient ultracentrifugation and titrated to about 10^5 hemagglutination units/ml. Then 15 μl of 1.0 M sodium bicarbonate (pH 9.0) was added to 150 μl of virus, and the virus was incubated with 25 μg of Alexa Fluor 488 NHS Ester (Succinimidyl Ester; Invitrogen, Carlsbad, CA) for 1 h at 25°C. After overnight dialysis to remove excess Alexa 488, viruses HA titer were checked and then bound to glycan array. Labeled viruses were incubated on the slide at 4°C for 1 h, washed, and centrifuged briefly before being scanned with an InnoScan 1100 AL fluorescence imager (Innopsys, Carbonne, France).

Biolayer interferometry

Two biotinylated glycan analogs (3′SLN: Neu5Acα2-3Galβ1-4GlcNAcβ and 6′SLN: Neu5Acα2-6Galβ1-4GlcNAcβ) linked to 30-KDa polymers containing 20% mol sugar and 5% mol biotin (GlycoTech, Gaithersburg, MD) were used to represent α2,3-

linked sialic acid (3′SLN) and α2,6-linked sialic acid (6′SLN). The stock solution (1 mg/ml) was prepared in 1× PBS (v/v) according to manufacturer's instructions. Biotinylated glycans Neu5Acα2-3Galβ1-4(Fucα1-3)GlcNAcβ (SLe^X), Neu5Gcα2-3Galβ1-4GlcNAcβ 3′SLN (Gc), and Neu5Gcα2-3Galβ1-4(Fucα1-3)GlcNAcβ ($SLe^{X\,(Gc)}$) were synthesized. All glycan analogs and viruses were further diluted into the kinetics buffer (pH 7.4; PBS solution containing 0.01% BSA and 0.002% Tween-20) with neuraminidase inhibitors (10 μM zanamivir hydrate and 10 μM oseltamivir phosphate) for binding analysis. The binding affinities of viruses (100 pM) to glycan analogs (ranging from 0 to 0.5 μg/ml for 3′SLN and 6′SLN; 0 to 0.5μM for SLe^X, Gc, and $SLe^{X\,(Gc)}$) were determined using an Octet RED96 interferometer equipped with streptavidin biosensors (Pall ForteBio LLC, Fremont, CA). In summary, RSL of the biosensor was calculated at the end of a 5- to 10-min loading, and the binding signal was measured at 25°C in a 20-min association step and 20-min dissociation step, with orbital shaking at 1,000 × rpm. The response of virus binding to a certain glycan loading was recorded at the end of the association step. The normalized response was calculated by dividing the maximum response of each glycan. The fractional saturation of the biosensor surface was calculated by using the Hill equation as reported by Xiong et al. (102). $RSL_{0.5}$ is the relative sugar loading on the streptavidin biosensor when the fractional saturation of the biosensor surface equals to 0.5. Given two testing viruses and a specific glycan, a higher binding response and a smaller value of $RSL_{0.5}$ a stronger binding avidity a virus will be; given two testing glycans and a specific virus, a higher binding response and a smaller value of $RSL_{0.5}$, a stronger binding avidity a virus will be.

Detection of SLeX and Neu5Gc glycans in horse and dog trachea

Normal horse and dog tracheal tissues were fixed in 10% neutral buffered formalin for 30 h at room temperature. The tissues were paraffin-embedded and sectioned at 5μm and then mounted on 3-aminopropyltrethoxy-silane–coated slides (Sigma, St. Louis, MO). The slides were then deparaffinized in xylene and rehydrated through graded alcohols. Antigen was then retrieved using Target Retrieval Solution (Dako, Carpinteria, CA), after which the slides were blocked by blocking reagent in 5% BSA for 1h at room temperature and then washed three times with PBST (PBS with 0.05% Tween 20). For detection of sialyloligosaccharides reactive with SLe^X antibody, the sections were incubated overnight at 4°C with 250 μl of Alexa Fluor 488 anti-human Sialyl Lewis X (dimeric) Antibody (1:100, BioLegend, San Diego, CA). Sections were washed three times each with PBST and PBS, then slides were counterstained with 4′,6-diamino-2-phenylindole, dihydrochloride (DAPI; Thermo Scientific, Rockford, IL). For detection of sialyloligosaccharides reactive with Neu5Gc, slides were incubated overnight at 4°C with Anti-Neu5Gc antibody (1:400, BioLegend, San Diego, CA). The slides were washed three times each in PBST and PBS and then incubated at room temperature for 1h with Goat anti-Chicken IgY (H+L) Secondary Antibody, Alexa Fluor 594 (Thermo Scientific, Rockford, IL) diluted 1:100 in PBS with 1% BSA. The sections were counterstained with DAPI for 3 min at room temperature. Then, after three washes with PBS, the sections were mounted on coverslips with ProLong Gold Antifade reagent (Thermo Scientific, Rockford, IL) and evaluated with a fluorescence microscope (Nikon, Tokyo, Japan). Photos were taken with a digital microscope camera (Olympus, Tokyo, Japan).

Sequences alignment and statistical analyses

HA sequences of H3N8 (equine and canine) and H3N2 (canine and avian) influenza virus were extracted from the influenza research database (http://www.fludb.org). The H3N2 CIVs included 120 isolates recovered from 2006 to 2018 and the H3N8 CIVs contained 347 isolates from 2003 to 2018. The multiple sequence alignments and the identification of the mutations were identified by using Bioedit software (192). Analysis of variance or unpaired *t* test was used to compare viral titers and binding to IAV glycan receptors.

Results

Growth Kinetics

To evaluate the effect of mutations K54N, S83N, and L222W in the HA on the growth properties of A(H3N8) IAV, we compared the growth kinetics in canine trachea explants for the two EIV strains (i.e., one prototype EIV, A/equine/Miami/63(H3N8) [abbreviated as Miami-eH3N8]), and a wild type field strain of EIV, A/equine/Pennsylvania/1/2007 (H3N8) [abbreviated as Pennsylvania-eH3N8]), wt-cH3N8, and three wt-cH3N8–derived HA mutants (54N, 83N, and 222W). No significant difference ($p > 0.05$) was found between infectivity titers for 54N-cH3N8 and wt-cH3N8 in canine trachea explants (Figure 18A). In contrast, the growth ability of cH3N8 in the explant tissues was significantly hampered by mutations S83N and L222W. wt-cH3N8 reached an average titer of $10^{5.249}$ 50% tissue culture infective dose ($TCID_{50}$)/ml at 72 h after infection; this titer was 26.1-fold higher than that for 83N-cH3N8 ($p < 0.05$) and 21.5-fold higher than that for 222W-cH3N8 ($p < 0.05$). Of interest, titers for both Miami-eH3N8 and Pennsylvania-eH3N8 in canine trachea explants were similar to those for

222W-cH3N8 ($p > 0.05$). No virus infectivity was detected in the mock-infected canine trachea explants.

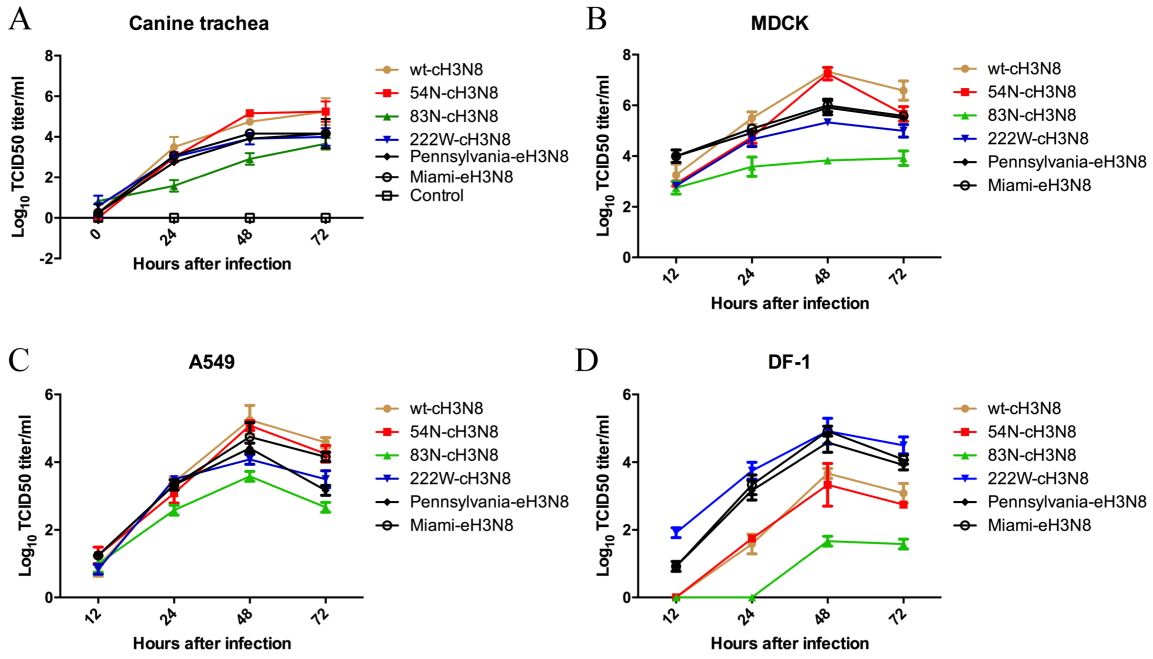


Figure 18. Growth properties of equine influenza A virus, canine influenza A virus, and mutant viruses derived from canine influenza A viruses in canine trachea explants (A), MDCK cells (B), A549 cells (C), and DF-1 cells (D).

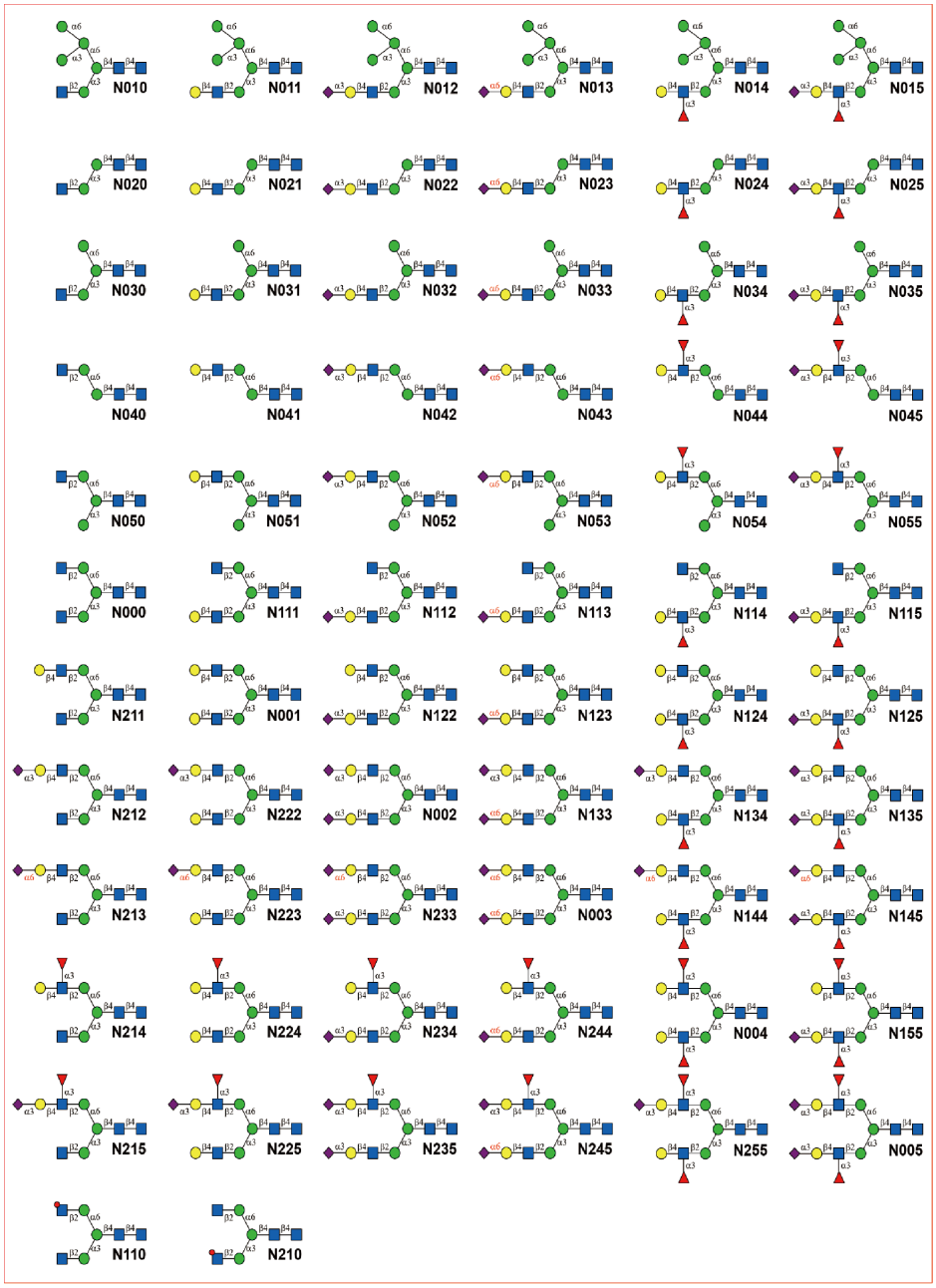
Trachea explants from dogs were infected with 400 $TCID_{50}$ of each virus. Cells were infected in triplicate with each virus at a multiplicity of infection of 0.01 and incubated at 37°C. Supernatant fluids were collected at indicated time points, and the viral titers were determined by $TCID_{50}$ assay in MDCK cells. Each data point represents the mean virus yield (log_{10} $TCID_{50}$/ml) from three individually infected wells ± standard deviation (vertical bars).

We further evaluated the growth properties of wt-cH3N8, three mutants, and wt-eH3N8 on Madin-Darby canine kidney (MDCK) cells, human lung carcinoma (A549) cells, and chicken fibroblast (DF-1) cells. Compared with wt-cH3N8, mutants 83N-cH3N8 and 222W-cH3N8 exhibited poorer growth in MDCK (Figure. 18B) and A549

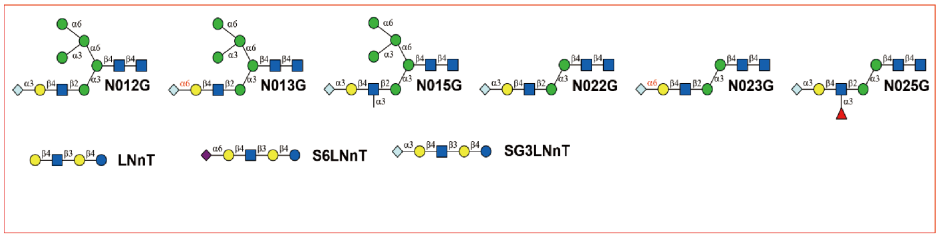
cells (Figure 18C), and 83N-cH3N8 had the poorest growth among the three mutants tested. Of interest, similar to the results in canine trachea explants, the replication kinetics for Miami-eH3N8 and Pennsylvania-eH3N8 in MDCK and A549 cells were similar to those for 222W-cH3N8 ($p > 0.05$) (Figure 18B and 18C). In DF-1 cells, wt-cH3N8, 54N-cH3N8, and 83N-cH3N8 showed limited replication ability compared with the other three testing viruses (i.e., 222W-cH3N8, Miami-eH3N8, and Pennsylvania-eH3N8), which reached mean titers of $10^{4.83}$, $10^{4.58}$, and $10^{4.91}$ $TCID_{50}$/ml, respectively (Figure 18D).

Receptor Binding Avidity and Specificity

To determine the glycan structures specific for CIV, we generated four reassortant viruses using reverse genetics (rg) (rg-wt–cH3N8, rg-54N, rg-83N, and rg-222W), which have HA from cH3N8 (wild type HA, HA with mutation K54N, HA with mutation S83N, and HA with mutation L222W, respectively), neuraminidase from wt-cH3N8, and six internal genes from A/PR/8/1934(H1N1), and then compared the glycan binding profiles of these four reassortant viruses using an array of 611 glycans from the Consortium of Functional Glycomics (CFG; http://www.functionalglycomics.org), and an array of 83 *N*-linked glycan isoforms (hereafter referred to as isoform microarray). To compare the glycan binding profiles of the CIVs with those of their precursor EIVs, we also determined that the glycan binding profiles of the wild type EIV field strain Pennsylvania-eH3N8 using the isoform microarray. Glycans on the CFG array have diverse terminal structures and a variety of spacer arms; glycans on the isoform microarray have the same base structures and spacer arms but different terminal structures (Figure 19).


N010
N011
N012
N013
N014
N015
N020
N021
N022
N023
N024
N025
N030
N031
N032
N033
N034
N035
N040
N041
N042
N043
N044
N045
N050
N051
N052
N053
N054
N055
N000
N111
N112
N113
N114
N115
N211
N001
N122
N123
N124
N125
N212
N222
N002
N133
N134
N135
N213
N223
N233
N003
N144
N145
N214
N224
N234
N244
N004
N155
N215
N225
N235
N245
N255
N005
N110
N210


N6030
N6000
N6211
N6212
N026
N049


N012G
N013G
N015G
N022G
N023G
N025G
LNnT
S6LNnT
SG3LNnT


Gal
Glc
GlcNAc
Man
L-Fuc
Neu5Ac
Neu5Gc
KDN

Figure 19. Glycan structures on the N-glycan isoform microarray. GlcNAc, *N*-acetylglucosamine; Man, mannose; Gal, galactose; L-Fuc, L-fucose; Neu5Ac, *N*-acetyl neuraminic acid; Neu5Gc, *N*-glycolylneuraminic acid.

Results from the CFG array showed that rg-wt–cH3N8 virus and all three mutants preferred binding to glycans containing α2,3-linkage sialic acids over those containing α2,6-linked sialic acids (Figure 20A). Binding specificity of the rg-54N–cH3N8 mutant to some SA2,3GA-like glycans was slightly increased over that of the rg-wt–cH3N8 virus.

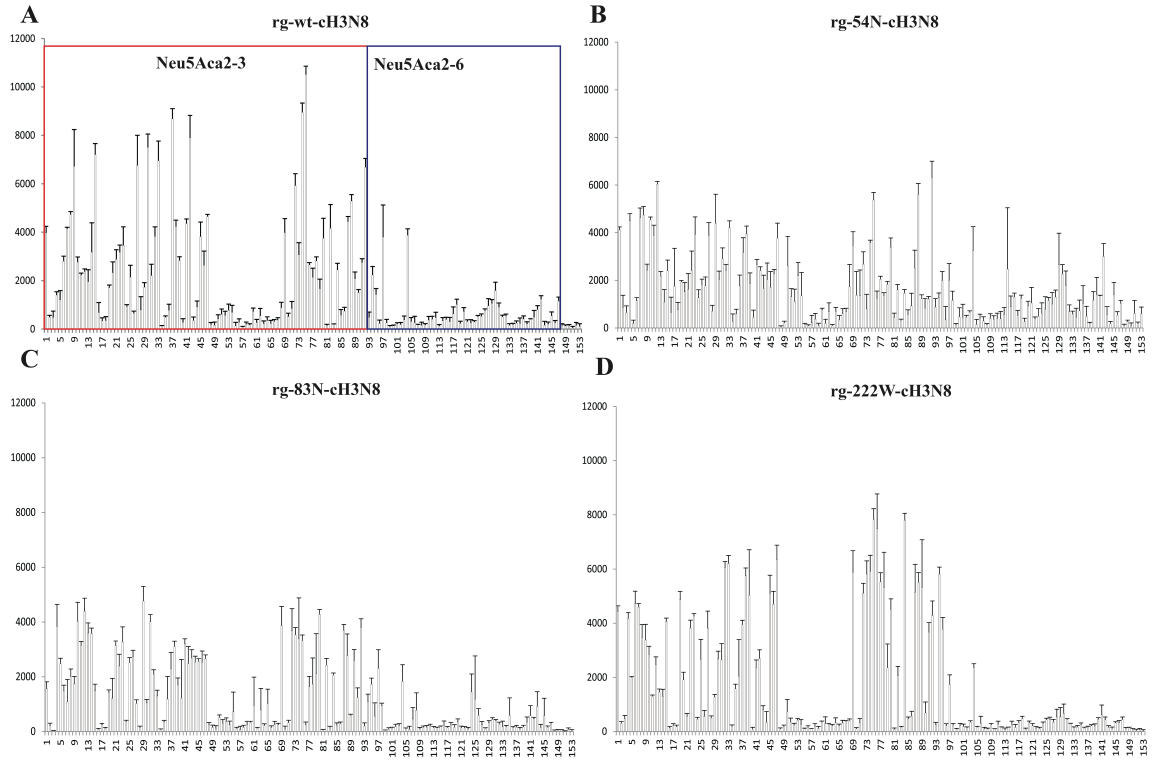


Figure 20. Binding profile of four influenza virus mutants to sialic acid glycans on the Consortium of Functional Glycomics (http://www.functionalglycomics.org) glycan array.

Different categories of glycans on the array are highlighted as follows: red (No. 1-92), Neu5Acα2,3 glycans; Blue (No. 93-147), Neu5Acα2,6 glycans No. 148-153 glycans are non-sialic acid glycans; Dashed line indicates minimum relative fluorescence units of 2,000. Vertical bars indicate standard deviation.

Consistent with results from the CFG array, results from the isoform microarray showed that all the testing viruses bound to glycans containing branched α2,3-linked sialic acids but not α2,6-linked sialic acids (Figure 21A-E). Of interest, the data from this isoform microarray demonstrated that rg-wt–cH3N8 virus had stronger binding affinity to the α2,3-linked sialic acids with core fucose than those without. For example, the binding affinity of rg-wt–cH3N8 virus to N012 was about 7-fold less than that to N015, which is the fucosylated form of N012; rg-wt–cH3N8 virus showed strong binding signals to both linear (N025G) and branched (N015G) α2,3-linked, fucosylated *N*-glycolylneuraminic acid (Neu5Gc) but not to linear (N022G) or branched (N012G) α2,3-linked, non-fucosylated Neu5Gc (Figure 21A). Compared with the binding signals for rg-wt–cH3N8, the binding signals for all three mutants to the fucosylated α2,3-linked *N*-acetyl neuraminic acid (Neu5Ac) or Neu5Gc were reduced to different degrees (Figure 21B-D). In contrast, Pennsylvania-eH3N8 showed moderate binding responses to α2,3-linked sialic acids, and the responses of the Pennsylvania-eH3N8 to α2,3-linked sialic acids with core fucose were similar to those without (Figure 21E). Of note, the binding responses of Pennsylvania-eH3N8 to α2,3-linked sialic acids with core fucose are much weaker than those of rg-wt–cH3N8. The Pennsylvania-eH3N8 showed moderate or low responses to fucosylated Neu5Gc (N015G and N025G), linear non-fucosylated Neu5Gc (N022G), and

branched non-fucosylated Neu5Gc (N012G), which are similar to cH3N8-222W but much weaker than those by rg-wt–cH3N8 (Figure 21E).

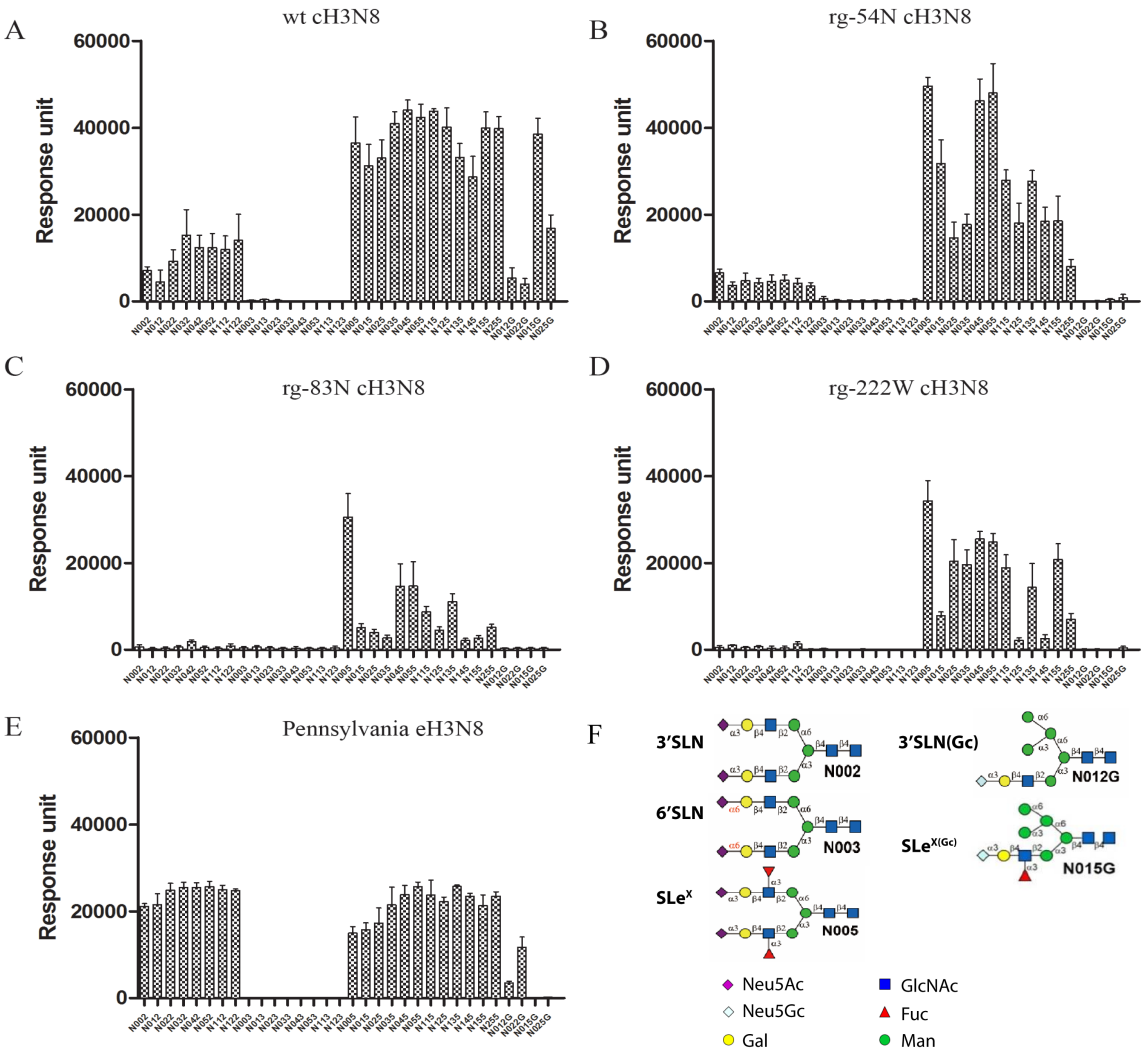


Figure 21. *N*-glycan microarray binding profiles of canine influenza A virus (CIV) and CIV-derived mutant viruses to representative linear and branched glycans. A, wt–cH3N8. B, rg-54N–cH3N8. C, rg-83N–cH3N8. D, rg-222W–cH3N8. E, Pennsylvania-eH3N8. Viruses are indicated at the top of figure panels. Data are the mean fluorescent binding signal intensity ± standard deviation (vertical bars). F, representative glycan structures on the isoform N-glycan microarray. N002-N122, Neu5Acα2,3 glycans; N003-N123, Neu5Acα2,6 glycans; N005-N255, Fucosylated Neu5Acα2,3 glycans; N12G-N22G, Neu5Gcα2,3 glycans; N015G-N025G, Fucosylated Neu5Gcα2,3 glycans. Neu5Ac, *N*-acetyl neuraminic acid; Neu5Gc, *N*-glycolylneuraminic acid. The structure details of all those glycans were listed in Figure 19.

To confirm these results from glycan microarray analyses and further determine the distinct patterns derived from three mutations in CIV, we used biolayer interferometry analyses to quantify the binding avidity and specificity of wt-cH3N8 and the three mutants to five glycan analogs [Neu5Acα2-3Galβ1-4GlcNAcβ (3'SLN), Neu5Acα2-6Galβ1-4GlcNAcβ (6'SLN), Neu5Acα2-3Galβ1-4[Fucα1-3]GlcNAcβ (SLe^X), Neu5Gcα2-3Galβ1-4GlcNAcβ (3'SLN(Gc)), and Neu5Gcα2-3Galβ1-4[Fucα1-3]GlcNAcβ ($SLe^{X(Gc)}$)] (Figure 22A). For data interpretation, wt-eH3N8 was also included as a comparison. In agreement with glycan array analyses, these analyses showed that all tested viruses bound to 3′SLN (Figure 22C) and SLe^X glycan analogs (Figure 22D), both of which contain α2,3-linked Neu5Ac, but not to 6′SLN (Figure 22B), which contains α2,6-linked Neu5Ac. The rg-wt–cH3N8 showed strong binding affinity to $SLe^{X(Gc)}$ and 3′SLN(Gc), two glycan analogs with α2,3-linked Neu5Gc (Figure 22E); the Pennsylvania-eH3N8 showed decent binding responses to 3′SLN(Gc) but limited binding responses to $SLe^{X(Gc)}$ (Figure 22E). Compared to Pennsylvania-eH3N8, the laboratory adapted prototype CIV Miami-eH3N8 had much weaker binding responses to α2,3-linked Neu5Gc (Figure 22E).

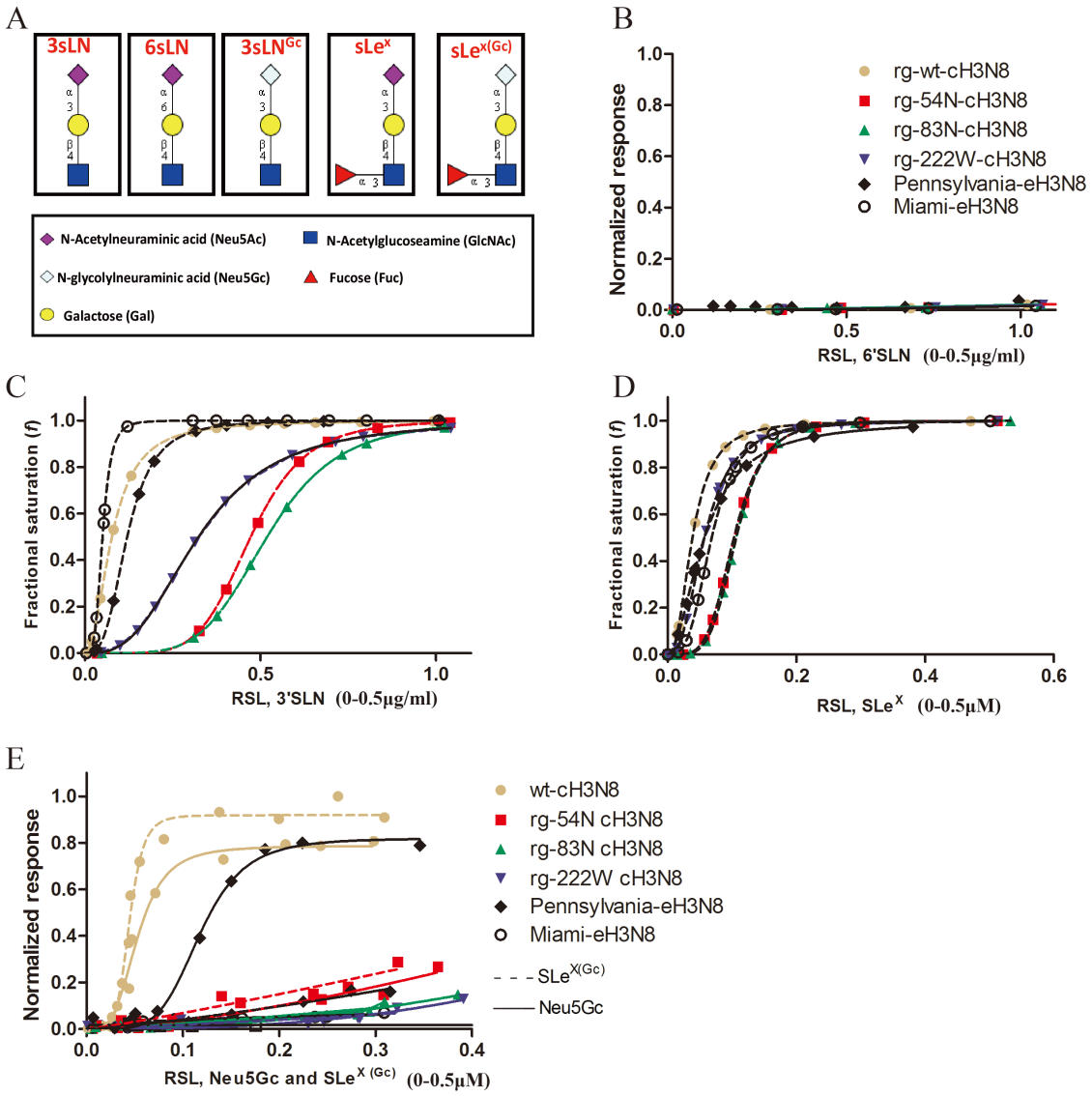


Figure 22. Glycan binding specificity of canine influenza A virus (CIV) and CIV-derived mutant viruses

Binding of viruses to biotinylated α2,6–linked sialic acid (6′SLN), α2,3–linked sialic acid (3′SLN), Neu5Acα2-3Galβ1-4(Fucα1-3)GlcNAcβ (SLe^X), Neu5Gcα2-3Galβ1-4GlcNAcβ 3′SLN(3′SLN Gc), and Neu5Gcα2-3Galβ1-4(Fucα1-3)GlcNAcβ ($SLe^{X\,(Gc)}$) glycan analogs as determined by Bio-Layer Interferometry (Pall ForteBio LLC, Fremont, CA). A, Structures of the glycan analogs. B, Normalized response of virus to 6′SLN by dividing the maximum of wild-type (wt) virus response to 3′SLN. C, Fractional saturation of viruses to 3′SLN. D, Fractional saturation of viruses to SLe^X. E, Normalized response of viruses to 3′SLN Gc and $SLe^{X\,(Gc)}$. Values on the *y* axes represent the response of viruses to glycans after the association step. Normalized response data were determined by dividing the maximum response to each glycan or fractional saturation of the sensor at each relative sugar loading (RSL) at a fixed virus concentration of 100 pM. The $RSL_{0.5}$ is the relative sugar loading on the streptavidin biosensor when the fractional saturation of the biosensor surface equals to 0.5. Given two testing viruses and a specific glycan, the

higher binding response and the smaller value of $RSL_{0.5,}$ the stronger binding avidity a virus will have; given two testing glycans and a specific virus, the higher binding response and the smaller value of $RSL_{0.5}$, the stronger binding avidity a virus will have.

To determine distinct patterns of binding among the three mutants (i.e., rg-54N-cH3N8, rg-83N-cH3N8, and rg-222W-cH3N8), rg-wt–cH3N8, and Pennsylvania-eH3N8 against these testing glycan analogs, we quantified 50% relative sugar loading concentration ($RSL_{0.5}$), which is the relative sugar loading on the streptavidin biosensor when the fractional saturation of the biosensor surface equals to 0.5. Given two testing viruses and a specific glycan, the smaller value of RSL0.5, the stronger binding avidity a virus will have. Results suggested that among three mutants, RSL0.5 of Pennsylvania-eH3N8 to these testing glycan analogs reassemble to those of rg-222W-cH3N8 but not to those of the other two mutants. Specifically, for 3′SLN, rg-wt-cH3N8, rg-54N–cH3N8, and rg-83N–cH3N8 had $RSL_{0.5}$ of 0.0771, 0.4726, and 0.5194, respectively, whereas Pennsylvania-eH3N8 and rg-222W–cH3N8 had 0.1207 and 0.3239, respectively; for SLe^X, wt-cH3N8, rg-54N–cH3N8, and rg-83N–cH3N8 had $RSL_{0.5}$ of 0.0385, 0.1030, and 0.1066, respectively, whereas Pennsylvania-eH3N8, rg-222W–cH3N8 had 0.0572 and 0.0573, respectively; for $SLe^{X(Gc)}$, wt-cH3N8, rg-54N–cH3N8, and rg-83N–cH3N8 had $RSL_{0.5}$ of 0.1446, 0.3097, and 0.3851, respectively, whereas Pennsylvania-eH3N8 and rg-222W–cH3N8 had 1.332 and 1.074, respectively. In summary, rg-222W–cH3N8 and Pennsylvania-eH3N8 had a similar glycan binding pattern, and mutation L222W of HA decreased binding affinities of A(H3N8) CIV to glycans with α2,3-linked, fucosylated Neu5Gc.

SLeX and Neu5Gc Glycan Distribution

Based on the receptor binding results, we hypothesized that α2,3-linked fucosylated Neu5Gc is the receptor that determines the host tropism of A(H3N8) CIV, and that α2,3-linked fucosylated Neu5Gc expression in the dog trachea would differ from expression in horse trachea. To test this hypothesis, we used immunofluorescence to detect SLe^X and Neu5Gc in trachea from dog and horse, to determine glycan receptor distribution. The canine submucosal glands were extensively and strongly immunopositive, indicating high SLe^X glycan expression (Figure 23A, E). Although the highest level of SLe^X glycans was mainly found in the dog submucosal glands, staining was also present along the surface of the respiratory epithelium (Figure 23F). In contrast, in the horse trachea, no SLe^X was detected in the ciliated cells, non-ciliated cells, or submucosal glands (Figure 23B, G). On the other hand, Neu5Gc glycans were widely distributed in both dog and horse trachea, including the ciliated and non-ciliated epithelial cells and submucosal glands (Figure 23A, B). These results demonstrate high expression of of Neu5Gc in horse trachea but not SLe^X, whereas dog trachea extensively expresses both SLe^X and Neu5Gc.

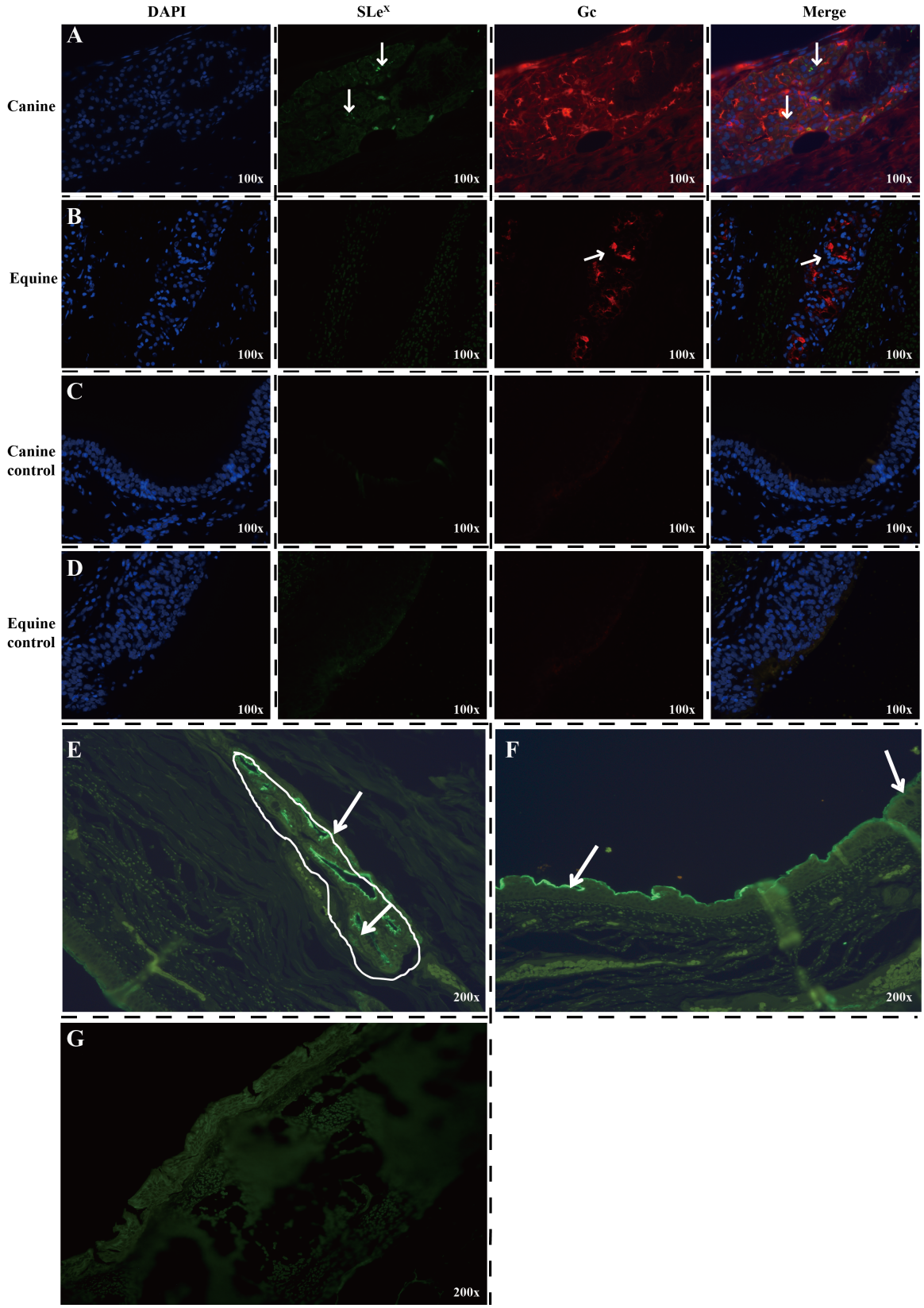

DAPI
SLe^X
Gc
Merge
A
Canine
B
Equine
C
Canine control
D
Equine control
E
F
G
100x
200x

Figure 23. Immunofluorescence detection of SLe^X (green) and Neu5Gc glycans (red) in dog (panels A, C, E, F) and horse trachea (panels B, D, G).
A and E. SLe^X glycans were mainly found in the epithelial cells of submucosal glands in dogs. B. No SLe^X glycans were detected in the ciliated cells, non-ciliated cells, or submucosal glands of the horse trachea. Neu5Gc glycans were widely distributed in both dog and horse trachea, including the ciliated and non-ciliated epithelial cells and submucosal glands. F. Expression of SLe^X on the canine surface tracheal respiratory epithelium. G. Detection of SLe^X on the horse tracheal respiratory epithelium and submucosal glands. Original magnification ×100 (panels A, B, C, D) or ×200 (panels E, F, G). The SLe^X glycan was detected by using Alexa Fluor 488 anti-human Sialyl Lewis X (dimeric) antibody (green), and the Neu5Gc glycans by using Alexa Fluor 594 Anti-Neu5Gc antibody (red). The arrows indicate immunopositive staining.

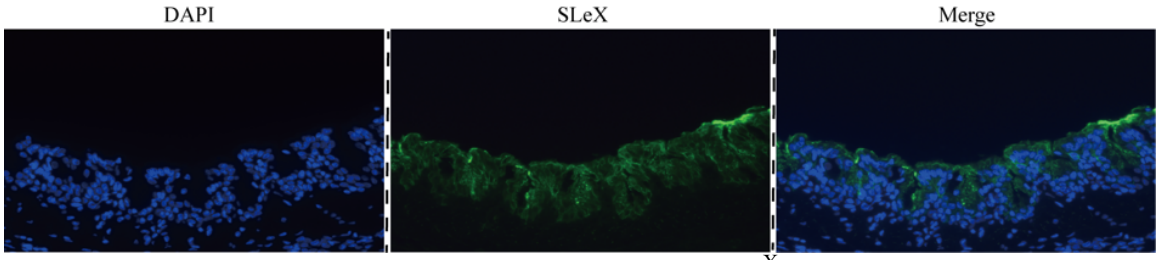


Figure 24. Immunofluorescence assay detection of SLe^X glycans in chicken trachea. The nuclei were stained with DAPI (blue). The SLe^X glycan was detected by using Alexa Fluor 488 anti-human Sialyl Lewis X (dimeric) antibody (green). The original magnification was ×200.

Discussion

To understand the molecular mechanisms of the host adaption of an emerging influenza virus, it is necessary to characterize the virus receptor binding properties and the types and distributions of host glycan receptors, two key factors determining successful introduction and adaption of an IAV to a new host (222). In this study, we provide evidence supporting that A(H3N8) CIV prefers glycans with a SLe^X epitope and that SLe^X was differentially distributed in the trachea of the dog *versus* the horse. Our comparative analyses, combined with our previous observations (33), suggests that

mutation W222L of HA increases the binding ability of A(H3N8) CIV to SLe^X and Gc epitopes, and that gaining the $SLe^{X(Gc)}$ binding ability could have facilitated the adaption of the equine-origin A(H3N8) virus to dogs.

Since the first outbreak of A(H3N8) IAV among dogs in 2004, this virus has become enzootic among the US dog population (216). Prior epidemiologic and genetic studies demonstrated that this emerging A(H3N8) CIV originated in horses, and was most closely related to the Florida sub-lineage strains (206). Interspecies transmission of A(H3N8) EIV from horses to dogs was observed when dogs were put in close contact with experimentally infected horses (223); however, A(H3N8) CIV–infected dogs cannot transmit CIV to horses (224), and A(H3N8) CIV replicates inefficiently in experimentally infected equine respiratory epithelial cells (225).This suggests that CIV and EIV are not exchangeable between equine and canine species and A(H3N8) CIV has been well adapted in dogs. Nevertheless, A(H3N8) EIV and CIV have similar replication abilities in canine trachea primary cell lines (214) and similar growth patterns in MDCK cells, canine A72 cells, Norden Laboratory feline kidney cells, ferret Mpf cells, human A549 cells, and equine EQKD cells (214). However, differences in glycan binding preferences were observed between A(H3N8) CIV, Eurasia-lineage A(H3N8) EIV, and North America-lineage A(H3N8) EIV. Like Collins et al. (217), we observed that, compared with the Florida sub-lineage A(H3N8) EIV, A(H3N8) CIV had higher avidity to SLe^X, supporting that, CIV and EIV are not exchangeable between equine and canine species.

Only three mutations (K54N, S83N, and L222W) were observed between the head of HA protein of A(H3N8) CIV and that of A(H3N8) EIV. It is plausible that one of the mutations could facilitate virus adaption from horse to dogs. von Grotthuss and

Rychlewski, who used computational simulation in a prior study (226), suggested that N54K could affect the glycosylation of HA and play a role in host adaption of A(H3N8) CIV. Our data demonstrated that mutations S83N and L222W, but not K54N, have a significant effect on the growth properties and binding properties of A(H3N8) CIV, a finding that does not support the hypothesis of von Grotthuss and Rychlewski (226) (Figs. 1 and 3). It is striking that mutant cH3N8-222W has a similar growth pattern as that of wt-eH3N8. Compared to cH3N8-222W, the mutant cH3N8-83N reduced significantly replication abilities of A(H3N8) CIV not only in canine trachea tissues but also in in three cell lines, including MDCK, A549, and DF-1 cells. A prior study reported that the infection of enzootic H3N8 CIV could lead to the quasispecies in the position 83 of CIV HA, suggesting this position was likely to be associated with a viral fitness cost (227). Of interests, further analyses of HA sequences of H3N8 CIVs showed that there are polymorphisms at position 83 of CIV HA (1.4% N; 0.3% T; 0.3% R) (data not shown) and that, in contrast, all H3N8 CIVs have L at position 222 compared to the conserved W at this position for H3N8 EIVs (Table 8). This finding suggests that mutation W222L, but not N83S, could facilitate host adaption of A(H3N8) IAV from horses to dogs. Mutation W222L is located in the 220-loop of the HA receptor binding site, and the importance of residues in the 220-loop in determining the receptor specificity has been reported in other IAV subtypes (131, 228).

Previous studies demonstrated that the substructure of the carbohydrate receptors determine influenza host range and tissue tropisms. The abundance of Neu5Gcα2,3Gal moiety (a type of SAα2,3GA) is important for EIV to replicate in equine respiratory epithelium cells (26) as well as for CIV to replicate in canine respiratory epithelium cells

(209, 210). Such results were confirmed by viral binding data from the CFG array, which showed that wt-cH3N8 and wt-eH3N8 prefer SAα2,3GA, including Neu5Gcα2,3Gal. However, such data still cannot explain the potential host factor(s) leading to host adaption of A(H3N8) from horses to dogs. Internal complexity of glycan structures below the sialic acid can influence glycan interaction with HA (229, 230); thus, host adaption cannot be understood without first understanding the glycan substructures of NeuGcα2,3Gal. The data from the isoform microarray and biolayer interferometry analyses demonstrated clear patterns of A(H3N8) CIV binding on a set of glycans with fucosylated α2,3-linked glycans (Figs. 2 and 3), and, of interest, a single L222W mutation dismissed that preference so that cH3N8-222W decreased binding to Neu5Gcα2,3Gal with the fucosylated motif (Fig. 3). Thus, mutation W222L seems to have increased the A(H3N8) binding preference from Neu5Gcα2,3Gal to $SLe^{X(Gc)}$, a Neu5Gcα2,3Gal with a fucosylated motif. Of note, a prior study suggested that amino acid residues at positions 222 and 227 of HA could affect IAV binding to glycans with the SLe^{X} motif (68).

Immunofluorescent staining further demonstrated the distribution of SLe^{X} in the trachea (i.e., virus entry sites) of dogs and the lack of SLe^{X} motifs in the tracheal tissue of horses. These results supported our hypothesis that SLe^{X} motifs could be one of the key factors driving host adaption of A(H3N8) EIV in dogs; specifically, host adaption to dogs could have been driven by the presence of mutation W222L in the virus HA protein. In addition, the distribution of SLe^{X} motifs is not only host-specific but also tissue-specific. SLe^{X} glycans were primarily expressed in canine tracheal submucosal glands with some expression noted in ciliated epithelial cells (Fig. 4). As a comparison, we probed SLe^{X} in

chicken trachea; results showed that, unlike in dog trachea, SLe^X glycans in chicken trachea were extensively expressed in both the ciliated epithelial cells and the submucosal glands (Figure 24). The tracheobronchial submucosal glands play multiple roles during infection by producing mucus to inactivate the virus, but may also be infected and thus enhance the infection (231). The receptors for human IAV in ferrets was reported to be *O*-linked sialylated glycans, which are predominantly distributed in the submucosal glands, and the infection of such cells facilitates the efficient airborne transmission of virus by easily making the virus encapsulated into respiratory droplets (117). The L222W mutation was shown to increase binding of H3N8 CIV to SLe^X glycans, which are present in abundancein the submucosal glands, thereby potentially facilitating the airborne transmission of the H3N8 CIV.

Neu5Ac and Neu5Gc are two major sialic acids present in mammalian cells; they play important roles in the recognition of influenza virus during the initial step of viral infection. Neu5Gc has been reported to be present in horses (26), dogs (33), pigs (34), and mice (35), but not in humans. Of interest, although it has been reported that horse trachea expresses 90% of the α2,3-linked Gc receptors present in horses (26), our results demonstrate that Pennsylvania-eH3N8 binds moderately to 3′SLN(Gc), a finding that is consistent with that in a previous study (44). However, compared with those of Pennsylvania-eH3N8, the laboratory adapted EIV prototype strain Miami-eH3N8 showed much weaker binding responses to 3′SLN(Gc); such low responses of Miami-eH3N8 to 3′SLN(Gc) might have been caused by the extensive passage of this virus in embryonated chicken eggs, which lacks Neu5Gc. Suzuki et al. (26) reported the presence of Gc glycans in equine trachea but not chicken trachea, a finding that was also observed in our

study. Our immunofluorescence staining also showed strong signal for Neu5Gc glycans in dog trachea (Fig. 4). The interaction of Neu5Ac or Neu5Gc glycans to the receptor binding site of viruses could affect the host tropism of those viruses. It has been reported that the large hydrophobic side chain of F75 of two human polyomaviruses (BK polyomavirus and JC polyomavirus) would clash with the glycolyl hydroxyl group and prevent binding of those two viruses to receptors terminating in Neu5Gc (232). Similarly, the L222W mutation introduced a large hydrophobic side chain and possibly caused the loss of binding to Neu5Gc glycans.

In this study, we only characterized the distribution and abundance of two individual glycan motifs, SLe^X and Gc, in dog and horse trachea by using immunofluorescenct staining. One potential limitation of this study was that we were unable to test the $SLe^{X(Gc)}$-like glycans due to the unavailablity of $SLe^{X(Gc)}$-specific antibody; thus, the abundance of this type of glycan substructure in the respiratory tracts of dogs and horses is unknown.

Like H3N8 CIV, avian H3N2 virus has been identified and is enzootic in canine populations in Southeast Asia (233, 234); furthermore, it has been shown that avian H3N2 virus was transmitted to the US dog population (235). A previous study suggested that the W222L mutation of HA facilitated the viral binding affinity of H3N2 CIV to Neu5Gc and improved its replication ability in canine trachea primary cells (33). Interestingly, similar to W222L of HA in H3N8 CIV (Fig. 2), the W222L mutation of HA increased binding affinities of H3N2 CIV to $SLex^{(Gc)}$ (33, 236).

In summary, these findings suggest that mutation W222L facilitates the host adaption of avian-origin H3N2 and equine-origin A(H3N8) IAVs in dogs by increasing

the virus's ability to bind to receptors with Neu5Gc and/or SLeX, which are widely distributed in canine trachea tissues.

CHAPTER V

SEQUENCE BASED INFLUENZA VACCINE STRAIN SELECTION USING SYSTEMS BIOLOGY

Introduction

Although the mechanisms remain largely unknown, the receptor binding property of influenza A virus (IAV) is clearly one of the key factors that affect the virus host and cell tropisms. The initial step of the IAV infection is the binding of the HA protein to the sialylated glycans on the host cells. The HA protein forms trimers, each of the monomers has a relatively conserved receptor binding site (RBS) that functions to engage the virus with the sialic acid (SA) glycan receptors on the host cells. The IAV was reported to have the ability to agglutinate erythrocytes (24), and the agglutination ability was found to be dependent on the types of SA on host cells (25). Typically, avian IAV prefers SAs that are linked to galactose in α2,3-linkage (SA2,3Gal) whereas human IAV binds to galactose with a α2,6-linkage (SA2,6Gal). The binding specificity and the binding affinity of an IAV to the glycan receptors, which can be expressed differentially on different cells and different hosts, are two of the key factors determining the virus cell and host tropisms.

The RBS is located at the globular head of the IAV HA protein and consists of 130-loop, 150 loop, 190 helix, and 220-loop (193). Mutations in the HA RBS are well documented to affect IAV binding affinity and specificity to SA2,3Gal and/or SA2,6 Gal. For example, Q226L and G228S, two mutations at the 220-loop of the HA RBS, switched the binding specificity of both subtype H2 and subtype H3 avian IAVs from SA2,3Gal to SA2,6Gal (237); a single mutation G225D at the 220-loop of the HA RBS

can switch virus binding specificity of subtype H6N1 avian IAV from SA2,3 Gal to SA2,6Gal; mutations V186G/K-K193T-G228S or V186N-N224K-G228S of the HA RBS can switch the receptor specificity of subtype H7N9 IAV from SA2,3 Gal to SA2,6Gal (78). On the other hand, the mutation D222G in the HA RBS of 2009 H1N1 IAV decreased virus binding affinity to SA2,6 Gal but increased that to SA2,3 Gal (238).

In addition to virus binding affinity and virus binding specificity, the mutation on the HA RBS can change antigenic properties and growth properties of IAV. During viral evolution, IAV mutates its HA RBS in order to adjust virus receptor binding avidity, and these mutations at the RBS can occasionally lead to antigenic changes of the virus (239). For example, a single mutation, N145K, at the HA RBS of subtype H3N2 human IAV altered both virus binding affinity and virus antigenicity (240). Mutations N145K and Y159N in the HA RBS of subtype H3N2 swine IAV were reported to drive the antigenic evolution in US swine (241). Mutations at positions 119, 153, 154, and 186 in the HA RBS of the 2009 H1N1 IAV improved virus replication ability in both cells and embryonated chicken eggs (120).

Vaccination is the primary option to counteract and reduce the impacts of influenza outbreaks (242). A high yield vaccine seed strain is required for timely vaccine manufacture and is thus a critical component of a successful influenza vaccine campaign. As described above, the sequences and structures of the HA RBS of IAV affect virus receptor binding property, virus replication efficiency, and virus antigenic properties. However, given the sequence of HA RBS, the virus receptor binding property, virus replication efficiency, or antigenicity these features are still difficult to be predicted as a priori. In this study, we will integrate both bench and computational approaches to

identify the key features determining these three key biological properties for an influenza vaccine strain. We aim to generate a large scale of mutants using epPCR (Chapter II), to characterize three influenza vaccine specific phenotypes (i.e., receptor binding specificity, replication efficacy in both MDCK cells and embryonated chicken eggs, and antigenic properties), and then to identify and validate the key genetic signatures, including genetic features in both viruses (i.e. position and residue compositions at RBS) and hosts (glycan substructures) by applying a machine-learning model in these phenotypic datasets. At last, a computational model is proposed to score the vaccine candidacy based on a query sequence of the HA RBS by integrating their weights on those phenotypes (i.e. receptor binding specificity, replication efficacy in both MDCK cells and embryonated chicken eggs, and antigenicity). This model can be used to screen influenza viruses in influenza surveillance to select influenza vaccine candidates, which shall meet the following criteria that (1) their antigenic properties are similar to the prototype vaccine antigen; (2) their yield capability in commonly used vaccine production platforms will meet the criteria required for vaccine manufacturing; (3) the phenotypes are stable upon at least three passages, and; (4) they remain low pathogenic yet efficacious (as a vaccine antigen) in humans.

Materials and Methods

Cells and viruses

Human embryonic kidney (293T) cells, Madin-Darby canine kidney (MDCK) cells were purchased from American Type Culture Collection (Manassas, VA). Cells were maintained at 37°C with 5% CO_2 in Dulbecco's Modified Eagle Medium (GIBCO/BRL, Grand Island, NY) supplemented with 5% fetal bovine serum (Atlanta

Biologicals, Lawrenceville, GA) and penicillin–streptomycin (Invitrogen, Carlsbad, CA). The HA gene of A/California/04/2009 (H1N1) (abbreviated as CA/04) was cloned into the vector pHW2000 and used as template for construction of the mutant library.

The viruses generated by reverse genetics were propagated in MDCK cells and cultured at 37°C with 5% CO_2 in Opti-MEM medium (GIBCO/BRL, Grand Island, NY) supplemented with 1 μg/ml of TPCK (*N*-tosyl-L-phenylalanine chloromethyl ketone)-Trypsin (Sigma-Aldrich, St. Louis, MO) and penicillin–streptomycin (Invitrogen, Carlsbad, CA). Virus titers were determined by $TCID_{50}$ in MDCK cells.

Construction of plasmid library and rescue of mutants

The mutant plasmid library with random mutations in HA RBSs was generated by epPCR strategy as described previously (186). Briefly, four primers were used in the generation of HA-pHW2000 RBSs mutant library: 1) 130loop_F, 5'-TCA TGG CCC AAT CAT GAC TCG AAC-3'; 2) 190helix_F, 5'-TGG GGC ATT CAC CAT CCA TCT ACT-3'; 3) 190helix_R, 5'-AAC ATA TGT ATC TGC ATT CTG ATA-3'; and 4) 220loop_R, 5'-TAG TGT CCA GTA ATA GTT CAT TCT-3'. The epPCR product (2 μl) was transfected into XL1-Blue Supercompetent Cells (Agilent Technologies, Santa Clara, CA). The transformed cells were directly inoculated onto LB (Luria Bertani) agar plates and the clones were propagated in 5ml LB media. The clones generated from the RBSs mutant library was confirmed by Sanger sequencing with the sequencing primer 5'-GAA CGT GTT ACC CAG GAG ATT-3'. Mutant viruses were rescued by plasmid based reverse genetics with the NA genes from CA/04 and six internal genes from influenza A/Puerto Rico/8/34 (H1N1) as described previously in Chapter I. For phenotype comparison, we generated the wild-type reassortant virus, rg-wt, with wild-type HA and

NA genes from CA/04 and six internal genes from influenza A/Puerto Rico/8/34 (H1N1) virus by reverse genetics approaches.

Virus sequencing

Viral RNA (80 µl total) was isolated from 200 µl of sample by using a 5X MagMAX™ Pathogen RNA/DNA kit (Thermo Fisher Scientific, Pittsburgh, PA) according to the manufacture's instruction. Influenza virus – specific primer Uni12 (5'-AGCAAAAGCAGG-3') (218) and 10 µl of each isolated RNA were used in the cDNA synthesis (total volume of 25 µl); the cDNA synthesis was carried out using SuperScript III Reverse Transcriptase (Invitrogen, Grand Island, NY) according to the manufacturer's instructions. The HA segment of mutants were amplified with the primers CA/04_HA_F, 5'-ATGAAGGCAATACTAGTAGTTCTGC-3' and CA/04_HA_R, 5'-TTAAATACATATTCTACACTGTAGAGACC-3' by using Phusion High-Fidelity PCR Kit (Thermo Fisher Scientific, Pittsburgh, PA) according to the manufacturer's instructions. The PCR products (50 µl) were purified by GeneJET PCR Purification kit (Thermo Fisher Scientific, Pittsburgh, PA) according to the manufacturer's instructions. The HA sequences of mutants were confirmed by Sanger sequencing.

Evaluation of viral growth

To evaluate the effect of mutations on the growth properties of viruses, we inoculated the MDCK cells with each testing influenza virus at a multiplicity of infection of 0.001 $TCID_{50}$ and incubated the cells in 5% CO_2 at 37°C for 1 h. The inocula were removed and cells washed two times with phosphate-buffered saline (PBS). The cells were then incubated (37°C in 5% CO_2) in Opti-MEM I (GIBCO, Grand Island, NY)

containing TPCK–trypsin (1 μg/ml). At 48h after inoculation, 200 μl of supernatants were collected, aliquoted, and stored at -70°C until use. To evaluate the growth properties of mutants on eggs, 9-day old SPF chicken eggs were inoculated 200 $TCID_{50}$ of each virus and incubated under 34 °C for 48 h. The titer of harvested viruses in both MDCK cells and SPF chicken eggs were determined by $TCID_{50}$ in MDCK cells.

Virus concentration and purification

Viruses for the glycan array analysis were purified as previously described(221). Briefly, viruses were purified from the cell supernatant or allantoic fluid by low-speed clarification (2,482 × *g*, 30 min, 4°C) to remove debris and then followed by ultracentrifugation through a cushion of 30%–60% sucrose in a 70Ti Rotor (Beckman Coulter, Fullerton, CA) (100,000 × *g*, 3 h, 4°C). The virus pellet was re-suspended in 100 μl of PBS and stored at −80°C until use.

Glycan microarray

To identify unique substructures bound specific sets of mutants, an isoform glycan array with 75 isoform glycans were printed on *N*-hydroxysuccinimide (NHS)–derivatized slides as described previously (221). The 75 glycans were selected to represent four different unique substructures, including *N*-glycans, Asn-linked *N*-glycans, Gangliosides, Thr-linked O-mannosyl glycans (Figure 25). The structures of these 75 glycans are shown in Figure 25. These glycans on the isoform microarray have the same base structures and spacer arms but different terminal structures (Figure 25), whereas glycans on the CFG (Consortium of Functional Glycomics) array have diverse terminal structures and a variety of spacer arms.

The isoform glycans were printed in replicates of four in a subarray, and sixteen subarrays were printed on each slide. All glycans were prepared at a concentration of 100 pM in phosphate buffer (100 mM sodium phosphate buffer, pH 8.5). The slides were fitted with a 16-chamber adapter to separate the subarray into individual wells for assay. The unreacted NHS groups on the slides are blocked with 50 mM ethanolamine in 50 mM sodium borate buffer (pH 9.2) at 4 °C for 1 hour and then the slides were rinsed with water. Before the assay, slides were rehydrated for 5 min in TSMW buffer (20 mM Tris-HCl, 150 mM NaCl, 0.2 mM $CaCl_2$, and 0.2 mM $MgCl_2$, 0.05% Tween). Viruses are purified by sucrose density gradient ultracentrifugation and titrated to about 32, 000 hemagglutination units/ml. Then 10 µl of 1.0 M sodium bicarbonate (pH 9.0) was added to 80 µl of virus, and the virus was incubated with 10 µg of Alexa Fluor 488 NHS Ester (Succinimidyl Ester; Invitrogen, Carlsbad, CA) for 1 h at 25°C. After overnight dialysis to remove excess Alexa 488, viruses HA titer were checked and then bound to glycan array. Labeled viruses were incubated on the slide at 4°C for 2 h, washed, and centrifuged briefly before being scanned with an InnoScan 1100 AL fluorescence imager (Innopsys, Carbonne, France).

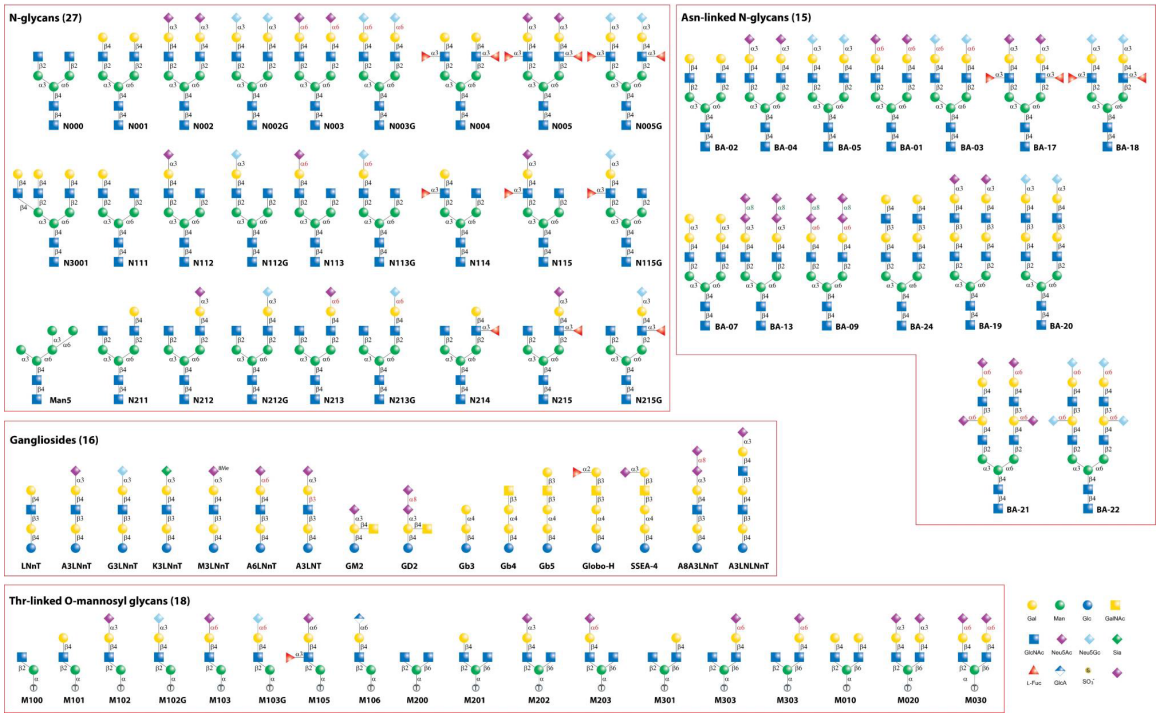


Figure 25. The sequences of synthetic glycans on the isoform glycan microarray.

Generation of ferret antisera

Ferrets antisera were generated in 6 to 8-week-old male or female ferrets, which were tested seronegative against CA/04 H1N1, A/Switzerland/9715293/2013 (H3N2), and A/Hong Kong/4801/2014 (H3N2). Each ferret was inoculated intranasally with 10^6 $TCID_{50}$ of the wild type strain or a testing mutant virus. Ferret sera were collected at 21 days post-inoculation and subjected for antigenic phenotyping using serological assays.

Antigenic phenotype determined by haemagglutination inhibition (HI) assays

Haemagglutination (HA) and HI assays were performed by using 0.5% turkey erythrocytes as described by the WHO Global Influenza Surveillance Network Manual for the Laboratory Diagnosis and Virological Surveillance of Influenza (187). Turkey erythrocytes were obtained from Lampire Biological Products (Everett, PA). The turkey

erythrocytes were washed three times with 1 × PBS (pH 7.2) before use and then diluted to 0.5% in 1 × PBS (pH 7.2).

Biolayer interferometry

Two biotinylated glycan analogs (3′SLN: Neu5Acα2-3Galβ1-4GlcNAcβ and 6′SLN: Neu5Acα2-6Galβ1-4GlcNAcβ) linked to 30-KDa polymers containing 20% mol sugar and 5% mol biotin (GlycoTech, Gaithersburg, MD) were used to represent α2,3-linked sialic acid (3′SLN) and α2,6-linked sialic acid (6′SLN). Biotinylated glycans Neu5Acα2-3Galβ1-4(Fucα1-3)GlcNAcβ (SLe^X) were synthesized. The stock solution (1 mg/ml) was prepared in 1× PBS (v/v) according to manufacturer's instructions. All glycan analogs and viruses were further diluted into the kinetics buffer (pH 7.4; PBS solution containing 0.01% BSA and 0.002% Tween-20) with neuraminidase inhibitors (10 μM zanamivir hydrate and 10 μM oseltamivir phosphate) for binding analysis. The binding affinities of viruses (100 pM) to glycan analogs were determined using an Octet RED96 interferometer equipped with streptavidin biosensors (Pall ForteBio LLC, Fremont, CA). In summary, RSL of the biosensor was calculated at the end of a 5- to 10-min loading, and the binding signal was measured at 25°C in a 20-min association step and 20-min dissociation step, with orbital shaking at 1,000 × rpm. The response of virus binding to a certain glycan loading was recorded at the end of the association step. The normalized response was calculated by dividing the maximum response of each glycan.

Machine learning model

The in-house machine learning method (105, 243-246) was used to identify key sequence features determining the phenotypes described above, including receptor

binding, antigenicity, replication efficiencies. Specifically, a multi-task sparse learning was used to identify key features associated with these phenotypes and a quantitative function was developed to quantify distances for each phenotype on the basis of their HA protein sequences. Each dataset was assigned with a single task in sparse learning, and the features were derived to determine each of these phenotypes. In addition, the features determining receptor binding, antigenicity, virus replication efficiencies were compared to determine 1) mutations at the HA RBS affecting virus antigenic properties; 2) the mutations at the HA RBS increasing and decreasing the yields in cells and/or eggs; 3) the glycan substructures (i.e. glycan motif) in the arrays associated with the yields in cells and/or eggs; 4) the glycan substructures (i.e. glycan motif) in the arrays associated with the mutations in the HA RBS.

Protein motifs determining virus yields and antigenicity using machine learning

In the machine-learning method, we investigated impacts of individual residues as well as the substructure motif on the HA protein structure on the phenotypes. To evaluate the impacts of the substructure motif on the HA protein structure, a sliding-window of residue was used as a single feature to integrate the synergistic effects of multiple residues in a substructure motif with neighboring residues. The sliding windows could be either overlapped or non-overlapped, and the size of windows was optimized from a range of 6 to 14 by maximizing the predicting performance.

Specifically, we formulated the phenotype-genotype problem as a sparse learning model. Mathematically, the sparse learning model is a linear regression loss function that subject to regularization. The LASSO model was adopted with a least square loss and ℓ_1 norm. We considered genetic difference matrix among HA protein sequences as X,

phenotype difference as Y, sample number as N. Then the objective of our model is to solve:

$$min\frac{1}{N}\|Y - X\theta\|_2^2, \text{ Subject to } \|\theta\|_1 \leq t$$

Where t is a predefined regularization parameter, θ denotes the numerical weights of individual features (either a single residue or a group of neighboring residues). Absolute value of weight indicates the impact of mutation(s) on a specific feature to phenotypes (e.g. receptor binding, antigenicity, and replication efficiencies). The larger the weight is, the greater the impact would be. Positive weights indicate a potential positive impact on the phenotype; otherwise, negative.

The prediction performance of the model was evaluated by root mean square error (RMSE):

$$\text{RMSE} = \sqrt{\frac{1}{\text{n}}\sum_{\text{i=1}}^{\text{n}}(\hat{\text{y}}_\text{i} - \text{y}_\text{i})^2}$$

A lower RMSE indicated a better prediction performance.

Glycan microarray data pre-processing and analyses.

All the glycan microarray data were normalized to [0,1] by columns (Viruses) before any statistical analysis or machine learning analysis was performed. Specifically, each value of binding affinity was normalized by the follow equations:

$$Normalized\ N_{i,j} = \frac{log_{10}\left(N_{i,j}\right) - min(log_{10}\ (N_{.j}))}{max\left(log_{10}\left(N_{.j}\right)\right) - \ min(log_{10}\left(N_{.j}\right))}$$

Values lower than 100 (including negative values) were set to 100 before data normalization. A value of 2,000 was used as the minimum threshold to determine whether a virus has bounding affinity to a glycan.

Definition of glycan substructures as glycan motifs

Based on the glycan's terminal structures, a total of 30 glycan substructure features were extracted as glycan motifs from glycan sequences of 75 synthetic glycans on the isoform glycan microarray. Each glycan motif was considered as an individual feature, and each glycan is transformed as an array (dimension n = 30) of one or multiple glycan motifs. In the glycan motif array, each feature was coded as 1 if a specific feature is present, otherwise, as 0.

Glycan motif identification using machine learning

Because it would be difficult to integrate the data from individual experiments, we introduced a multi-task learning framework to overcome this challenge. Multi-task Lasso model is a conventional multi-task sparse learning model and is frequently used in feature selection problems. Multi-task Lasso model trained least square loss with L1/L2 mixed-norm as regularizer. It solves the following objective:

$$min\frac{1}{N}\|Y - X\theta\|_2^2 + \alpha\|\theta\|_{2,1}$$

$$\|\theta\|_{2,1} = \sum_{i=1}^{n}\|\theta_i\|_2 = \sum_{i=1}^{n}\sqrt{\sum_{j=1}^{m}\theta_{i,j}^2}$$

Where $\theta \subset \mathbb{R}^{n\times m}$, θ_i is the i-th row while θ_j is the j-th column.

Multi-task Lasso model considers each column (virus) as an individual task, and is able to run all tasks at the same time. Each feature (glycan motif) will be assigned a numerical weight, which indicate the impact of the corresponding features on the glycan binding. The larger the weight is, the more important the glycan motif will be.

To identify unique glycan motifs associated with the yielding phenotypes, we divide 200 viruses into six different datasets: 1) array data with high yield egg mutants, which has a titer >3.16 $x10^6$ TCID50/ml; 2) array data with medium yield egg mutants, which has a titer between 3.16 $x10^4$ TCID50/ml and 3.16 $x10^6$ TCID50/ml ; 3) array data with low yield egg mutants, which has a titer <3.16 $x10^4$ TCID50/ml ; 4) array data with high yield cell mutants, which has a titer >1.52 $x10^6$ TCID50/ml ; 5) array data with medium yield cell mutants, which has a titer between 1.52 $x10^4$ TCID50/ml and 1.52 $x10^6$ TCID50/ml ; 6) array data with low yield cell mutants, which has a titer <1.52$x10^4$ TCID50/ml . Then six individual machine learning processes were performed. The glycan motifs from these learning were compared to address the following questions: 1) what are the specific glycan motifs responding to high yields in eggs? 2) What are the specific glycan motifs responding to low yields in eggs? 3) What are the specific glycan motifs responding to high yields in cells? 4) What are the specific glycan motifs responding to low yields in cells? 5) What are the specific glycan motifs responding to high yields in both eggs and cells? 6) What are the specific glycan motifs responding to low yields in both eggs and cells? 7) What are the unique glycan motifs responding to the yields in eggs from those from cells?

Prediction of vaccine candidacy

The vaccine candidacy could be predicted using features and their weights to be determined by aforementioned machine learning. Given the sequences of a testing virus, a linear scoring function is proposed to predict the difference of growth property between the testing virus and a virus has experimental data (e.g. WT). Then the value of vaccine candidacy ($\hat{y}$) could be calculated by the summary of the aforementioned difference ($x\theta$) and value of the known virus (y_0).

$$\hat{y} = x\theta + y_0$$

To avoid potential bias, average value from multiple predictions will be more precise.

$$\hat{y}_{ave} = \frac{1}{N}\sum_{i=1}^{N}(x_i\theta + y_i)$$

Public sequence data, sequence alignment, and molecular characterization

HA sequences of subtype H1 IAVs from humans (including both the seasonal H1N1 IAVs from 1977-2009 and the 2009 H1N1 IAVs from 2009 to 2018), swine, and avian were extracted from the Influenza Research Database (http://www.fludb.org). The multiple sequence alignments and the identification of the mutations were identified by using Bioedit software (192). The positions were matched those corresponding mutations at the HA protein of subtype H3N2 IAVs.

Visualization of proteins structure

The three dimensional structure of HA protein was modeled using the HA structure template of CA/04 (PDB: 3LZG). The mutations were visualized using Chimera software.

Results

Features of RBS mutant library

A total of 200 mutants, each of which carries one to three random mutations in the HA RBS of CA/04, were rescued successfully from 826 plasmids with confirmed mutations, and the mutations were then confirmed by Sanger sequencing. The mutations cover the region from position 122 to 244 (H3 numbering) of HA1 of CA/04 (Figure 26A). Among those mutants, 123 mutants had one mutation, 61 mutants two mutations, and 16 mutants three mutations (Figure 26B). There were 119 mutants carries mutation(s) at 130-190 loop of the RBS of CA/04; 51 mutants carries mutation(s) at the 190-220 of the RBS of CA/04; 30 mutants carries mutations covering both 130-190 and 190-220 of the RBS of CA/04. The virus rescue did not succeed for the other 626 mutant plasmids after three trials, and mutations at these failing plasmids include positions 134, 139, 147, 155, 168, 170, 177, 179, 180, 181, 183, 185, 191, 209 and 215, which were found highly conserved in the receptor binding pocket (Figure 26C, D).

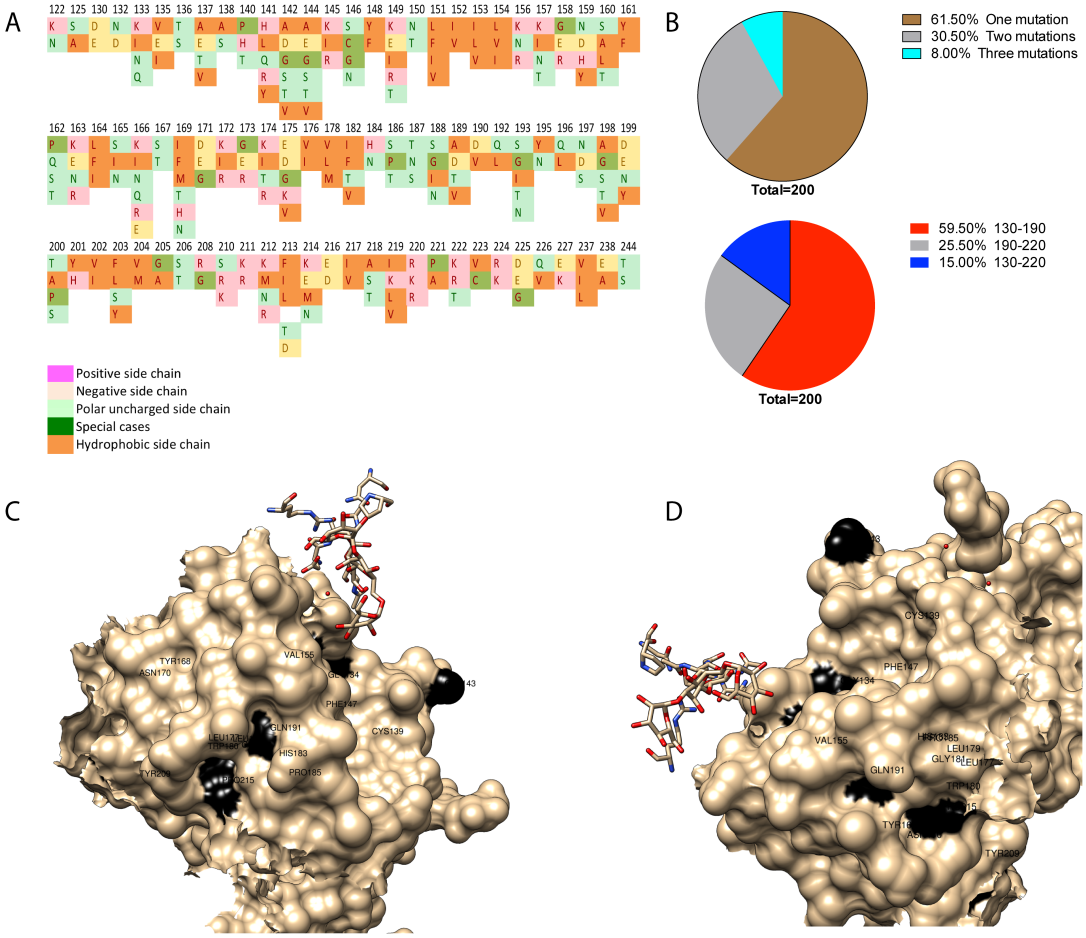


Figure 26. Features of the receptor binding site mutants of influenza A(H1N1)pdm09 virus.

A, the list of mutations on the HA receptor binding site of influenza A(H1N1)pdm09 virus. The amino acids are grouped with colors by: amino acids with positive charged side chain (R, H, K), amino acids with negative charged side chain (D, E), amino acids with polar uncharged side chain (S, T, N, Q), Special cases (C, U, G, P); amino acids with hydrophobic side chain (A, I, L, M, F, W, Y, V); B, the summary of numbers of receptor binding site mutants of CA/04. C, The locations of conserved sites in top view (C) and side view (D) of the crystal structure of the hemagglutinin (HA) of influenza A(H1N1)pdm09 virus. The conserved residues (dark black) on the receptor-binding sites are labeled with residue names and locations. The single letter amino acid annotations were used together with H3 numbering for all binding residues. The HA was adapted from the structural template of the HA of A(H1N1)pdm09 virus (Protein Databank [PDB] 3LZG).

Antigenicity of RBS mutants

The HI titer of each RBS mutant against the CA/04 ferret antisera was recorded (Table 9). Four mutants (S160A S188N R220K, K172R, T200S, S206T) had a HI titer of 1:160 against the CA/04 WT ferret antisera, which was at least four-fold less comparing to the homologous titer of CA/04 WT. Of note, the D130E, S193T, A198S mutants had a HI titer of 1:10 against the CA/04 WT ferret antisera, suggesting significant antigenic changes due to these mutations at the RBS. Among those mutations, residues 160, 188, 193, and 200 locate at positions overlapping RBS and antigenic site B; residue 172 locates at position overlapping RBS and antigenic site D; and residue 206 locates at position close to antigenic site D. The mutations that could alter the antigenicity of CA/04 H1H1 were majorly located in the antigenic site B, suggesting the immunodominance of antigenic site B over site A of HA of CA/04 H1N1.

On the other hand, the highest HI titer for the RBS mutants was 1:1280. None of the mutants had > 2-fold increase over the homologous HI titer.

Table 9. Serological responses of wild type and receptor binding site mutants of CA/04 against ferret anti-CA/04 WT sera using HI assays

HI[a]	M1[b]	M2	M3	HI	M1	M2	M3
10	D130E	S193T	A198S	640	D171E		
160	S160A	S188N	R220K	640	K149E		
160	D130E	K172R		640	A142T		
160	T200S			640	L151V		
160	S206T			640	L154V		
320	A138S			640	H141L	I152V	
320	P162Q			640	K133Q		
320	S125A	D130E		640	S146C		
320	D130E	A142V		640	D130E	K149I	
320	K174I			640	I169F		
320	S167T	A189T		640	V178L		
320	H141R			640	D130E	G173R	
320	K166I			640	S186T	G158R	
320	E175K			640	S167T		
320	H141Y	N159D		640	N150T		
320	N159D	K166I		640	A142S		
320	T187N			640	S186T		
320	K156N			640	S167T	I169H	S193N
320	S193G			640	D130E	S146G	S165I
320	E216D			640	E238A		
320	K212R	V237I	E227K	640	A189V		
320	V202I			640	K166R		
320	D130E			640	A142D	L154I	
320	F203S			640	S188N		
320	K211R			640	D190V		
320	T200P			640	K157N	I169F	
320	D130E	K212M		640	D130E	T187S	
320	D225E			640	H141Y		
320	A198S			640	K145R		
320	S188N	I217V		640	A189T		
320	V237I	E227K		640	I182F		
320	A198T			640	I152L		
320	K157T			640	K156R		
320	D130E	S188N		640	S146G		
320	K122N	S160T		640	K163R		
320	N132D	G173R		640	K166N		
320	K157I	D225G		640	S165I	E175V	

Table 9 (continued)

HI[a]	M1[b]	M2	M3	HI	M1	M2	M3
320	K157I	I169T		640	K157I		
320	D130E	G158E	D225G	640	A142D		
320	D130E	G158E	A198V	640	K172R	K174R	
320	D130E	F203Y		640	I169T		
320	D130E	A189V		640	G173E		
640	A142V			640	A144S		
640	Q192L			640	I152V	I169N	
640	D130E			640	K149I		
640	E175D			640	T136S		
640	S131L			640	P140T		
640	K172R			640	K166I	D171G	E175D
640	S160T			640	S193I		
640	L151I	L164I		640	D130E	R224K	
640	A137V			640	D130E	I219L	
640	A137V	K166E		640	D130E	Y201H	
640	V237I			640	D130E	S165N	
640	F203L			640	D130E	S210R	
640	D130E	I219V		1280	K166Q		
640	F203Y			1280	K172I		
640	D130E	D199E		1280	P221A		
640	A198V	V237L	T244S	1280	K133N		
640	D199E	K214E		1280	A144G		
640	N197D	D225G		1280	I182V		
640	A198G			1280	K157N		
640	K212N			1280	I169M	D130E	D199Y
640	F213L			1280	K149T		
640	F213I			1280	S146N		
640	I219L	V237I		1280	K174T		
640	V237L			1280	S146G	K174R	
640	S188I			1280	P162S		
640	R208G			1280	P140H		
640	D199N	F203Y	V204M	1280	V178M		
640	V237I	S165N		1280	A189D	I219R	
640	K222R			1280	H141L		
640	D130E	I219L		1280	K149R		
640	K222T	D225G		1280	L151F	Q226V	
640	S186P	K222T		1280	A144T	T187S	

Table 9 (continued)

HI[a]	M1[b]	M2	M3	HI	M1	M2	M3
640	K214E			1280	G173R		
640	N197S			1280	I182T		
640	A218S			1280	K145I		
640	A189V	E216D		1280	S188G		
640	D199E			1280	A142G		
640	G205A			1280	K163E		
640	A198V			1280	K163N		
640	K214M			1280	E175G		
640	I219L			1280	V135E	D199E	
640	I219V			1280	A137E		
640	K214N			1280	Q196L		
640	T200A			1280	A144T		
640	S160L	K212D		1280	I169M	G173E	
640	A144V			1280	K133Q	P140H	
640	H141R	K149E		1280	V135I		
640	K133N	V178M	D225G	1280	K212M		
640	A144E	K163N	I219K	1280	K212N	A218T	
640	V176I			1280	I217V		
640	N132D	K163N	H184N	1280	A189D	Q192L	
640	K133I			1280	A189D	A218S	
640	A137T	H141Q		1280	D130E	A189T	
640	D130E	K212T		1280	D130E	V223C	D225E
				640	WT		

a. HI, hemagglutination inhibition;
b. Mutations (H3 numbering) on the receptor-binding site of CA/04.

Replication efficiency of RBS mutants

We evaluated the effect of RBS mutations on the replication efficiency of 2009 pandemic H1N1 virus in both MDCK cells and SPF chicken eggs. The results suggested 14 mutants had more than 10 fold increased $TCID_{50}$ titer compared to the WT in MDCK cells. The N159D, K166I mutant reached highest $TCID_{50}$ titer in MDCK cells, which is about 100 fold higher compared to the WT. A total of 29 mutants had more than 10 fold decrease in $TCID_{50}$ titer compared to that of the WT in cells. In contrast, 33 mutants had

more than a 10 fold increase in $TCID_{50}$ titer compared to the WT in SPF chicken eggs. The D130E, S193T, A198S mutants reached the highest $TCID_{50}$ titer in SPF chicken eggs, which is about 850 fold higher compared to the WT. A total of 19 mutants had > 10 fold decreased $TCID_{50}$ titer compared to that of the WT in eggs.

Key signatures determining yields from machine learning

We applied machine learning models to identify the key signatures determining the virus yield. The overall growth data of the RBS mutants shown in Figure 27 suggested high degree of correlation between the three samples of each mutant that grew in both cells and eggs. Thus, we used the geometrical average value of the $TCID_{50}$ titer for each mutant in the following machine learning analysis. The results identified key signatures at positions 122, 125, 161, 133, 166, 172, 182, 198, 214, and 218 are responsible for increasing virus yield in chicken eggs; positions 137, 190, 206 are responsible for decreasing virus yield in chicken eggs; residues 122, 161 were responsible for increasing virus yield in MDCK cells; residues 137, 146, 178, 199, 200, and 206 were responsible for decreasing virus yield in MDCK cells (Figure 28). The structure modeling results suggested that the residues increasing the virus yield were mainly located in the surface area of the HA trimer, while the residues responsible for decreased virus yield were mainly located in the interior of HA trimer (Figure 29B).

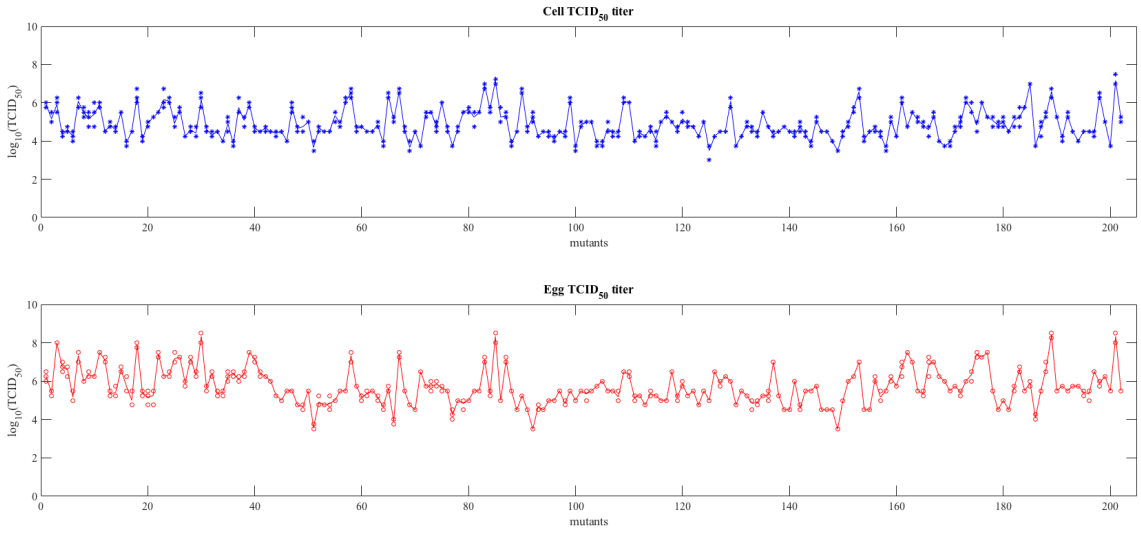


Figure 27. Overall growth efficiency of RBS mutants in MDCK cells and SPF chicken eggs.

The X axis represents the mutant number and the Y axis represents the $\log_{10}$ $TCID_{50}$ titer/ml. The blue line links the geometrical average of the 3 MDCK cell grown samples for each RBS mutant. The red line links the geometrical average of the 3 SPF chicken egg grown samples for each RBS mutant.

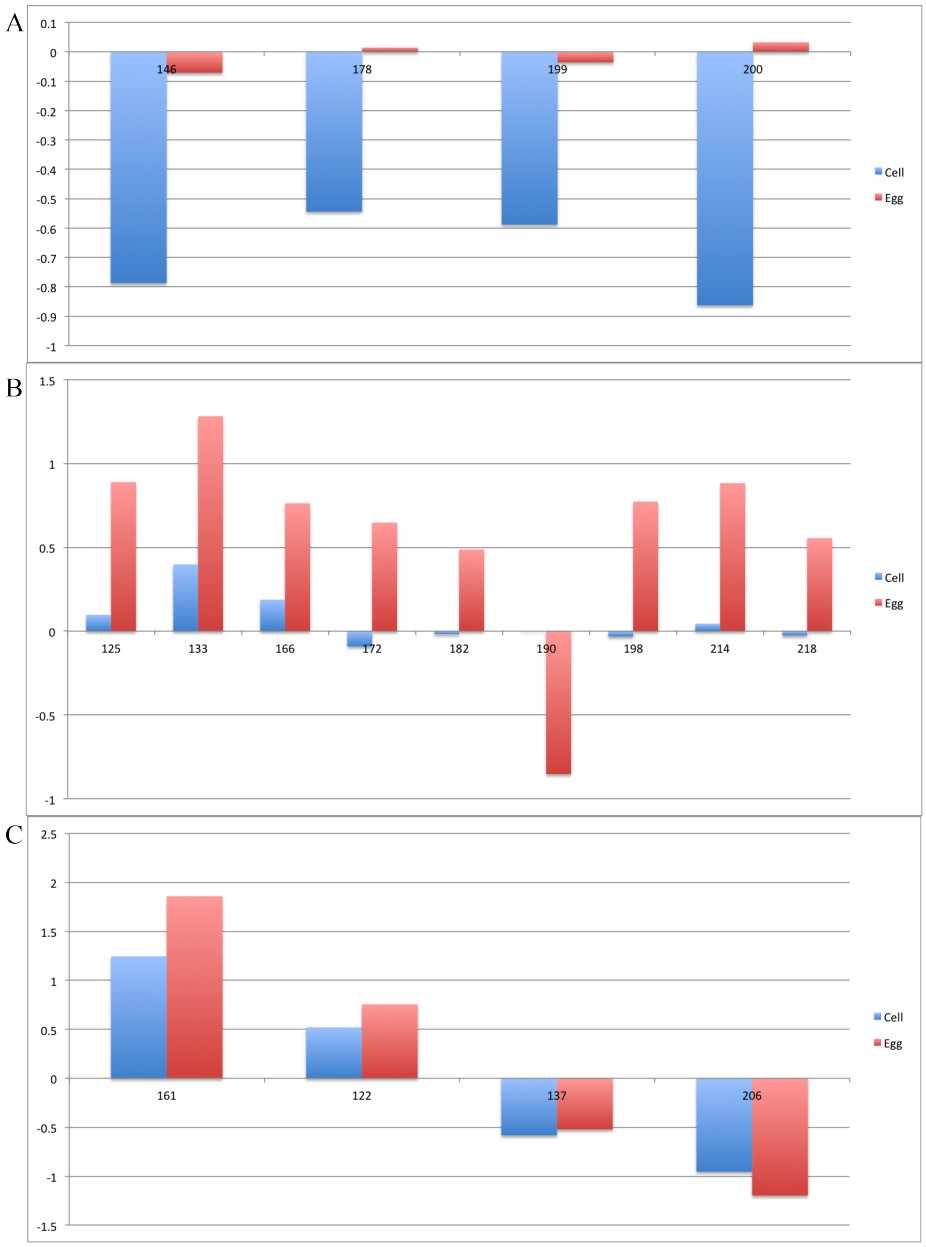


Figure 28. Key residues for the growth of CA/04 H1N1 in cells and eggs
A, important residues for the growth of CA/04 H1N1 in cells only; B, important residues for the growth of CA/04 H1N1 in eggs only; C, important residues for the growth of CA/04 H1N1 in both cells and eggs.

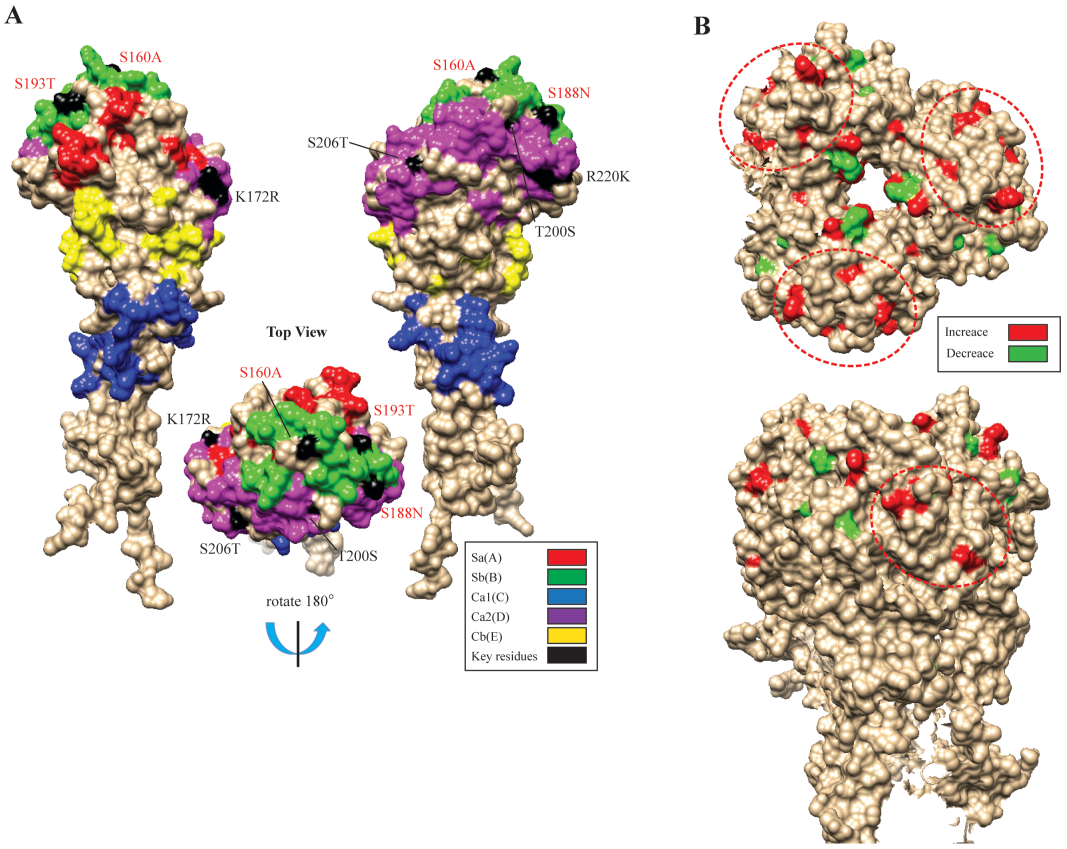


Figure 29. Key residues on the HA receptor binding site that affect the antigenicity and yield of 2009 pandemic H1N1.

A, antigenic structure of key residues would affect the antigenicity of virus. The antigenic sites Sa (A), Sb (B), Ca1(C), Ca2 (D), and Cb (E) are marked with red, green, blue, purple, and yellow, respectively. The key residues that affect the antigenicity of 2009 pandemic H1N1 are marked with dark black. B, schematic surface diagram of residues affecting the virus yield in cells and eggs. The residues that increase or decrease the virus yield are marked with red and green, respectively. All the sequences mentioned in the text are based on H3 numbering. The HA was adapted from the structural template of the HA of A(H1N1)pdm09 virus (Protein Databank [PDB] 3LZG).

Glycan binding affinities of RBS mutants

The binding specificity and intensity of the RBS mutants was determined by glycan microarray, which had 75 representative glycans and were classified into 30 features (Table 10) based on the glycan structures. The overall binding intensities of

mutants plotted with the glycans are shown as a heat map in Figure 30. As expected, the tested mutants showed overall strong bindings to Neu5Acα2-6Gal terminated glycans such as: N003, N113, N213, BA-01, BA-19, BA-21, A6LNnT, M103, M303, and M030. In addition, the tested mutants had an overall stronger preference for binding to Neu5Acα2-6Galβ terminated Gangliosides (A6LNnT) than Neu5Acα2-6Galβ terminated Thr-linked O-mannosyl glycans (M103). The viruses binding to N-glycans and Asn-linked N-glycans with the same terminal structures (eg. N003 and BA-01) did not differ significantly. A higher binding response of viruses to glycans with multiple Neu5Acα2-6Gal terminal structures than glycans with single Neu5Acα2-6Gal terminal structure was observed.

We examined the receptor binding profiles of the high yield mutants in both MDCK cells and eggs to further investigate the potential role of glycan binding in the mechanism of variations in virus yield. For example, the top two mutants (D103E, S193T, A198S; I169F) with the highest $TCID_{50}$ titer in chicken eggs were $10^{2.93}$ and $10^{2.88}$ fold higher than the wild type. The glycan array results suggest that both mutants (D130E, S193T, A198S and I169F) gained the ability to bind 3'SLN and SLe^X terminated glycans. For example, the ratio of WT to 6'SLN (N003): 3'SLN (N002): SLe^X (N005) was 1:0.027:0.034. In contrast, the ratio of for the D130E, S193T, A198S and I169F mutants were 1: 0.64:1.02 and 1: 0.33: 0.36. The results suggested a more balanced binding ratio of α2-3 sialic acid glycans to α2-6 sialic acid glycans would significantly increase H1N1 virus yield in eggs.

Further analysis of virus binding to non-sialic acid glycans such as N001 suggested the potential role of non-sialic acid glycans during virus infections. For

example, the top 3 mutants (K133N; H141Y, N159D; and I182V) with highest binding signal to N001 all had an increased $TCID_{50}$ titer in MDCK cells and/or chicken eggs. The K133N and I182V mutation increased the virus $TCID_{50}$ titer in chicken eggs by $10^{1.89}$ and $10^{2.43}$ fold, respectively. The H141Y, N159D and I182 mutations increased virus $TCID_{50}$ titer in MDCK cells by $10^{1.75}$ and $10^{1.26}$ fold, respectively.

We noticed that the increase of virus binding to Neu5Acα2-3Gal terminated gangliosides may increase the virus yield. The $TCID_{50}$ titers of K214E mutant were$10^{1.46}$ and $10^{1.5}$ fold higher than the WT in MDCK cells and chicken eggs, respectively. The glycan array binding results suggested that the K214E mutant had moderate binding signal to A3LNnT but not to Neu5Acα2-3Gal terminated N-glycans (N002) or SLe^X glycans (N005). Additionally, the K214E mutant had increased binding to Neu5Gcα2-6Gal terminated glycan (N213G), which may responsible for the increase of virus yield in MDCK cells.

Table 10. Features of the glycans on the glycan microarray

No#	Glycan features
1	AEAB labed core pentasaccharides
2	Asn linked core pentasacchrides
3	Thr linked core mannose
4	Core lactose
5	High-Man
6	GlcNAc(b1-2)
7	Gal(b1-4)GlcNAc
8	Gal(a1-3)Gal(b1-4)GlcNAc
9	Neu5Ac(a2-3)Gal(b1-4)GlcNAc
10	Neu5Ac(a2-6)Gal(b1-4)GlcNAc
11	Neu5Gc(a2-3)Gal(b1-4)GlcNAc
12	Neu5Gc(a2-6)Gal(b1-4)GlcNAc
13	Neu8Me(a2-3)Gal(b1-4)GlcNAc
14	Kdn(a2-3)Gal(b1-4)GlcNAc
15	Neu5Ac(a2-8)Neu5Ac(a2-3)Gal(b1-4)GlcNAc
16	Neu5Ac(a2-8)Neu5Ac(a2-6)Gal(b1-4)GlcNAc
17	GlcA(a1-6)Gal(b1-4)GlcNAc
18	Gal(b1-4)[Fuc(a1-3)]GlcNAc
19	Neu5Ac(a2-3)Gal(b1-4)[Fuc(a1-3)]GlcNAc
20	Neu5Ac(a2-6)Gal(b1-4)[Fuc(a1-3)]GlcNAc
21	Neu5Gc(a2-3)Gal(b1-4)[Fuc(a1-3)]GlcNAc
22	Gal(a1-4)
23	GalNAc(b1-3)Gal(a1-4)
24	Gal(b1-3)GalNAc(b1-3)Gal(a1-4)
25	Fuc(a1-2)Gal(b1-3)GalNAc(b1-3)Gal(a1-4)
26	Neu5Ac(a2-3)Gal(b1-3)GalNAc(b1-3)Gal(a1-4)
27	Neu5Ac(a2-3)[GalNAc(b1-4)]Gal(b1-4)Glc-AEAB
28	Neu5Ac(a2-8)Neu5Ac(a2-3)[GalNAc(b1-4)]Gal(b1-4)Glc-AEAB
29	Length
30	Size

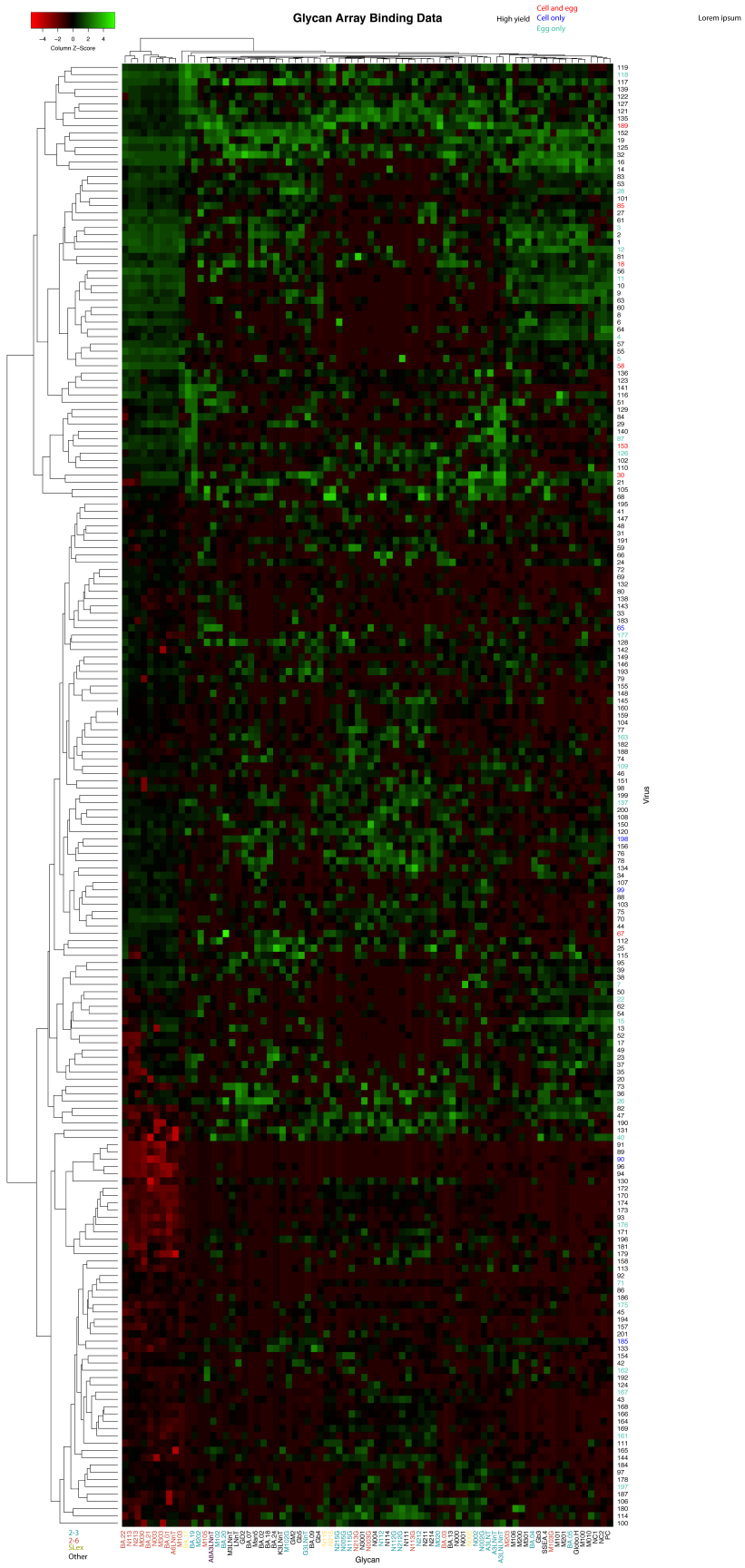

Glycan Array Binding Data
High yield
Cell and egg
Cell only
Egg only
Lorem ipsum
Column Z-Score
−4
−2
0
2
4
Virus
Glycan
2-3
2-6
SLex
Other

Figure 30. Heat map illustration of binding intensity of viruses bound to glycans on the glycan array.
The normalized binding signals were log2-transformed. Each row represents a receptor binding site mutant of CA/04 and each column represents an individual glycans. The linkage hierarchical clustering was performed by Hierachical Clustering Explorer 3.0 to generate the heat map and tree structure.

Glycan substructures determining yield signatures in eggs and/or cells

We used machine learning models to investigate the weight of glycan terminal structures on virus receptor binding properties. The results suggested that Neu5Acα2-3Gal, Asn linked core pentasacchrides, Core lactose, AEAB labed core pentasaccharides, and Thr linked core mannose are the most important structure motifs determining the overall virus binding of 2009 pandemic H1N1 virus (Figure 31).

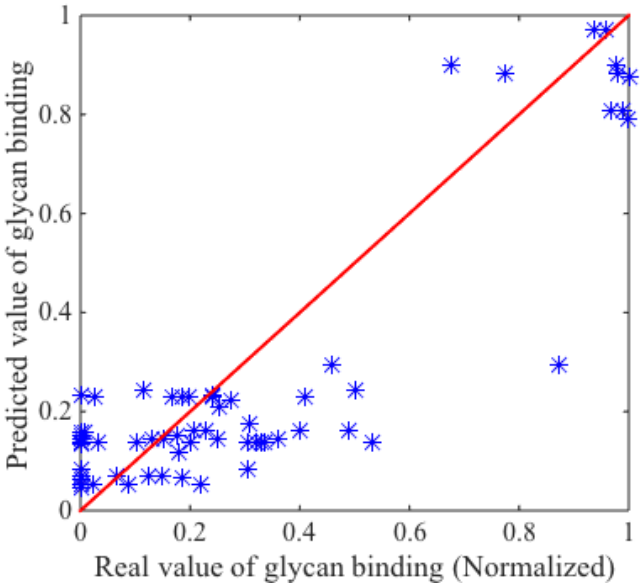


Figure 31. Prediction of virus glycan binding by machine learning
The X axis represents the real value of normalized glycan binding signal; The Y axis represents the predicted value of glycan binding signal. The red line represents the reference axis for the prediction. The blue dots represent the binding signals from the experiments.

Validation of the key residues in the HA RBS affecting virus antigenic properties

To further determine the effect of the D130E, S193T, and 198S mutations on the antigenicity of CA/04 H1N1, we generated ferret antisera against this mutant and then performed a two-way HI assay. The homologous HI titers of each D130E, S193T, and A198S mutants had a HI titer of 1:1280, which was 8 fold higher than that of the WT. The two-way HI assay further confirmed the significant antigenic change of the virus by mutations D130E, S193T, and A198S.

Among 199 mutant viruses, ten mutants have at least one mutation among the positions 130, 193, and 198 of the HA protein. To further determine the types of mutation(s) responsible for the antigenic change, all ten mutants with mutations in 130, 193, and 198 were included in the two-way HI assay (Table 11). The results of two-way

HI assays showed that the D130E, A198S, A198G, A198V, and A198T mutants had a HI titer of < 2-fold change compared to that of the WT virus, suggesting a limited role for positions 130 and 198 in virus antigenic properties. However, the S193N and S193I had the same HI titer as the homologous titer of the WT virus. Those results suggest an important role of S193T in affecting virus antigenicity of CA/04 H1N1.

Table 11. Serological responses of the 130, 193, and 198 mutants against ferret anti-CA/04 WT and CA/04 D130E, S193T, and A198S mutant sera using HI assays

Viruses[a]	(CA/04 WT)[b]	HI (D130E, S193T, A198S)[c]
WT	640	160
D130E, S193T, A198S	10	1280
D130E G158E A198V	320	320
S167T I169H S193N	640	320
S193I	640	320
S193G	320	160
A198V V237L T244S	640	160
A198G	640	160
A198S	320	160
A198T	320	320
A198V	640	320

[a]Mutants generated by ep-PCR based reverse genetic approaches that contains mutations at residue 130, 193, or 198.
b. hemagglutination inhibition titers of mutants determined by using ferret serum (anti-CA/04).
c. hemagglutination inhibition titers of mutants determined by using ferret serum (anti-CA/04 D130E, S193T, A198S.

Validation of the key residues affecting receptor binding properties

To test the binding profiles of the D130E, S193T, A198S mutants, we performed biolayer interferometry analyses to characterize the dynamics and avidity of virus binding to five representative glycan analogs, [Neu5Acα2-3Galβ1-4GlcNAcβ (3'SLN), Neu5Acα2-6Galβ1-4GlcNAcβ (6'SLN), Neu5Acα2-3Galβ1-4[Fucα1-3]GlcNAcβ (SLe^X), Neu5Gcα2-3Galβ1-4GlcNAcβ (3'SLN(Gc)), and Neu5Gcα2-3Galβ1-4[Fucα1-3]GlcNAcβ ($SLe^{X(Gc)}$)]. Compared to the CA/04 WT that had no detectable binding to 3′SLN (see chapter III), the D130E, S193T, A198S mutant gained binding avidity to 3'SLN (Response=2.919) and SLe^X (Response=2.449). The mutant had a 1.61 fold less binding avidity to 6'SLN (Response=1.818) compare to that of 3'SLN, suggesting that the D130E, S193T, A198S mutations changed binding preference of CA/04 H1N1 to 3'SLN over 6'SLN (Figure 32A). For the α2-3 linked sialic acids, the mutants had a 1.19-fold higher binding affinity to the 3'SLN than the SLe^X. Similar to the WT, the mutant had no detectable binding to 3'SLN(Gc) and $SLe^{X(Gc)}$). These results indicated that the D130E, S193T, A198S mutation in HA increased the binding affinity to 3'SLN and switched the binding preference of virus to 3'SLN to 6'SLN. The structure modeling analysis showed that positions 130, 193, and 198 locates at the globular head of RBS. The D130E and A198S mutations might function together to change the binding avidity and specificity of the CA/04 H1N1 (Figure 32B).

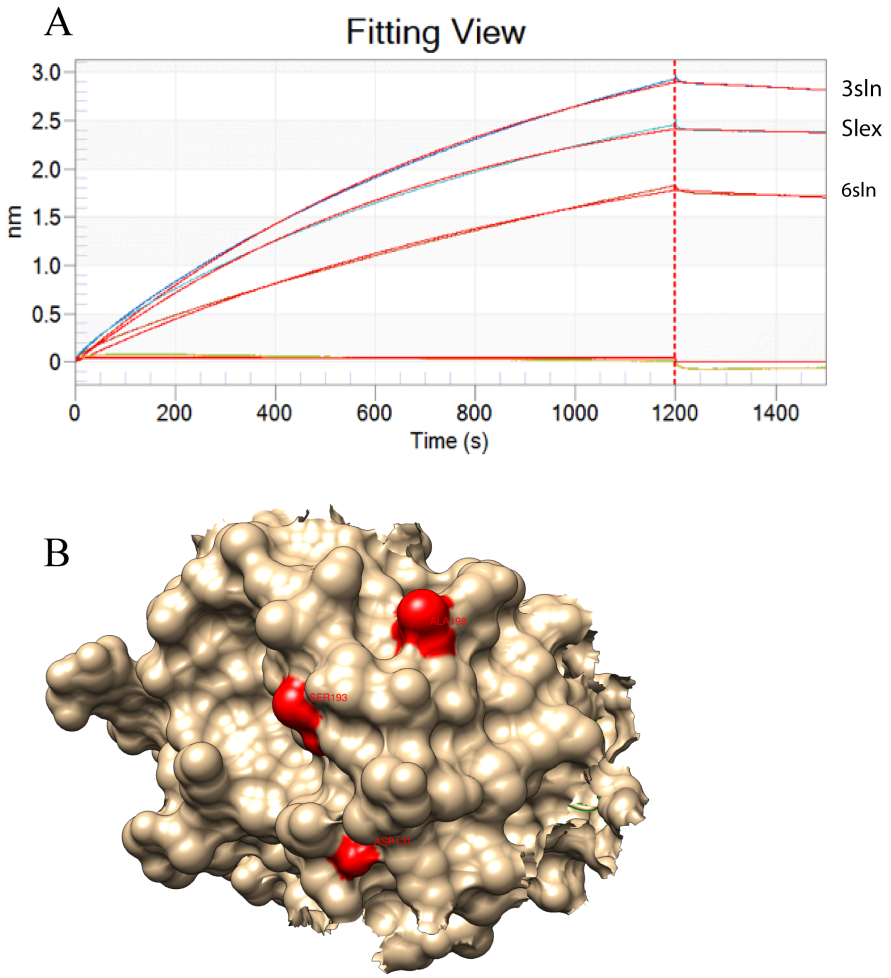


Figure 32. Glycan binding specificity of the D130E, S193T, A198S mutant of CA/04 by Bio-Layer Interferometry (fortéBIO, Menlo Park, CA).

A. Binding of viruses to biotinylated α2,6–linked sialic acid (6′SLN), α2,3–linked sialic acid (3′SLN), Neu5Acα2-3Galβ1-4(Fucα1-3)GlcNAcβ (SLe^X) , Neu5Gcα2-3Galβ1-4GlcNAcβ (3'SLN(Gc)), and Neu5Gcα2-3Galβ1-4[Fucα1-3]GlcNAcβ ($SLe^{X(Gc)}$)] glycan analogs as determined by Bio-Layer Interferometry (Pall ForteBio LLC, Fremont, CA). The streptavidin-coated biosensors were first preloaded with biotin-labeled sialic acid receptors, followed by the 1 pM each virus binding for 1,200 seconds in a standard kinetic buffer with neuraminidase inhibitors (zanamivir hydrate and oseltamivir phosphate). C, D) Sialic acid receptor concentrations were titrated to 1 µg/ml (3'SLN and 6'SLN) or 1uM (SLe^X) when loading with the biotin-labeled receptors. The binding response unit (nm) was recorded at the 1,196 second time point (4 seconds before the start of dissociation). B. Three-dimensional view of the D130E, S193T, A198S mutant of influenza A/California/04/09 (H1N1) virus.

Evolutionary analysis of residues 130, 193 and 198 in H1 IAVs

Sequence alignment of position 130 and 198 from sequences of avian H1N1, 1918 pandemic H1N1, and 2009 pandemic H1N1 suggested that 130E, 198T are predominant in waterfowl H1N1 whereas the 1918 pandemic H1N1 had 130E and 198A, and the 2009 pandemic H1N1 had 130D and 198A (Table 12). Previous studies suggested that G225D would switch binding preference of the H1N1 and H6N1 viruses from Neu5Acα2-3Gal to Neu5Acα2-6Gal (75, 247). In this study, we identified residues 198 and 130, which may play critical role in the host adaption of avian H1N1 virus to humans. The 1918 pandemic H1N1 was hypothesized to originate from avian H1N1 without adaptions in mammalian hosts such as pigs. The T198A mutation may change the binding specificity of the avian H1N1 virus from α2-3 to α2-6. Additionally, the HA of the 2009 pandemic H1N1 which originated from the 1918 pandemic H1N1 gained a mutation at E130D which caused a loss of binding to Neu5Acα2-3Gal. Our results suggest a novel evolutionary pathway for the adaption of avian H1N1 in obtaining binding specificity for humans and therefore increasing the probability of causing pandemic outbreaks in humans.

The HA sequences of North American swine H1N1 viruses from 1977 to 2018 were analyzed since the HA of the 2009 pandemic H1N1 virus originated from the H1N1 swine North America lineage (Figure 33). The results suggest that the North American swine H1N1 only had 130E between the years of 1977 to 2000. The 130D was observed in 2000, and then 130D became increasingly prevalent and was predominant in the North America swine population. Similar patterns were found in residue 193 of the North American swine H1N1 viruses. The North American swine H1N1 only had 193T

between the years of 1977 to 1999. The 193S was first observed in the year of 2000 and then co-circulated with the 193T. As we observed that S193T could change the antigenicity of the H1N1 viruses, the North American swine H1N1 may have experienced a drastic change in antigenicity change around the years of 2001-2008. In contrast, the sequence analysis suggested that 198A was predominant in the North American swine H1N1 during the years of 1977-2018, suggesting that 198A was introduced into the swine population in earlier times. Most likely, the T198A mutation occurred before 1918 and then facilitated the outbreak of the 1918 pandemic H1N1 in humans.

Table 12. Sequence alignment of residues 130 and 198 from avian H1N1, 1918 pandemic H1N1, and 2009 pandemic H1N1 virus.

Viruses	130	191	192	193	194	195	196	197	198	225	Binding
Waterfowl H1N1	E	Q	Q	S	L	Y	Q	N	T	G	2,3>2,6
1918 pH1N1	E	Q	Q	S	L	Y	Q	N	A	D	2,6>2,3
2009 pH1N1	D	Q	Q	S	L	Y	Q	N	A	D	2,6
Turkey H1N1	D	Q	Q	S	L	Y	Q	N	A	D	?

The HA sequences were extracted from the Influenza Research Database (https://www.fludb.org). Residues are numbered through H3 numbering. Residues 130 and 198 are marked with red.

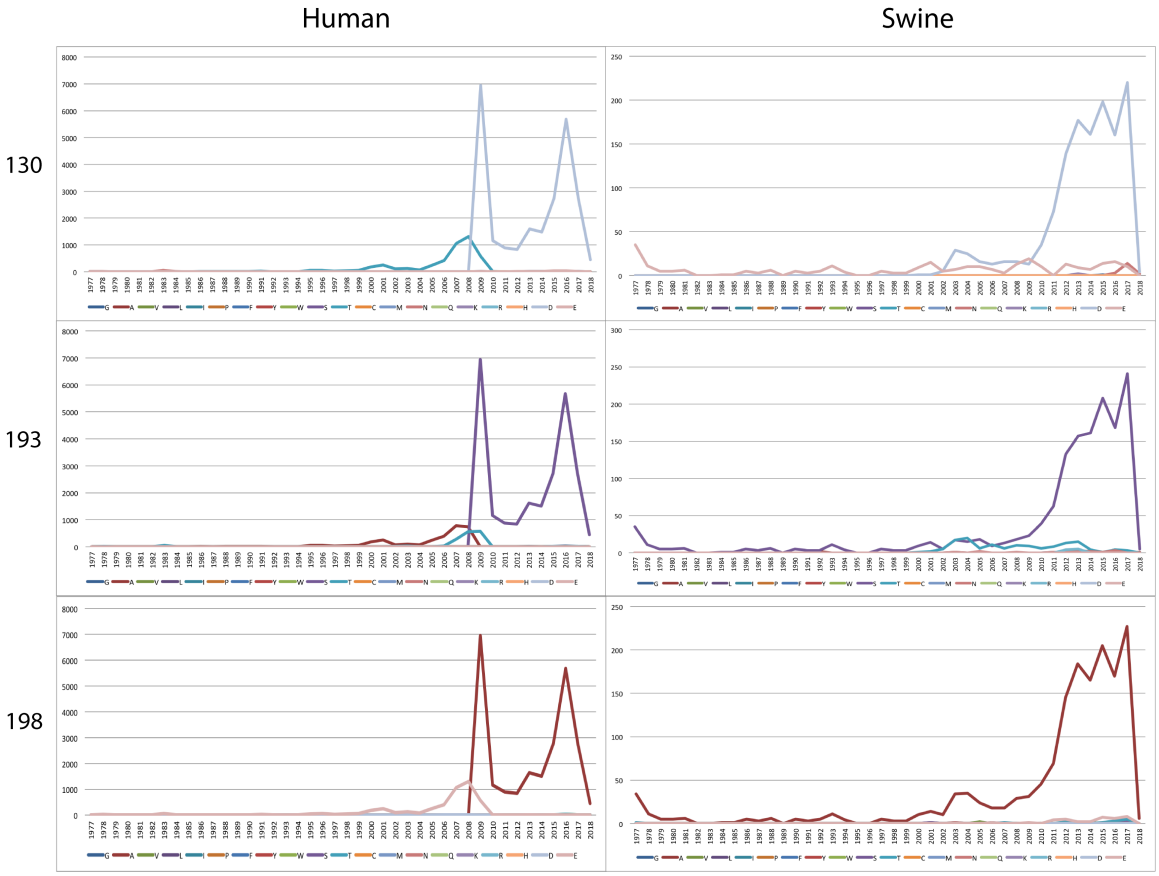


Figure 33. Evolution of amino acid positions 130, 193, and 198 in the receptor binding site of human seasonal/pandemic H1N1 and swine H1N1 influenza viruses. The HA sequences of human seasonal/pandemic H1N1 and swine H1N1 influenza viruses from 1977-2018 were extracted from the Influenza Research Database (https://www.fludb.org). The numbers of each amino acid (Y axis) were plotted as a function of calendar year (X axis) of isolation. Each amino acid is indicated by a different color.

Discussion

The HA plays important role in the biology of IAV, including mediating virus entry through the attachment to sialic acid receptors on the host cells and serving as the major antigen of IAV. Mutations in the RBS that locate at the globular head of the HA has been shown to affect virus binding specificity, antigenicity, and yield. The biological properties of an IAV such as antigenicity and growth properties are critical for the success of influenza vaccine seed strain. In this study, we performed large scale of mutagenesis in the RBS of the HA of 2009 pandemic H1N1 and identified critical residues on the RBS of pandemic H1N1, which would affect virus antigenicity, yield and binding specificity. The novel biomarkers identified in this study would help the influenza vaccine seed strain selection and expand the knowledge of H1N1 virus evolution and host adaption.

The RBS of HA are be highly conserved among different IAVs, especially the same HA subtypes of avian IAVs (248). In this study, we identified that mutations on residues 134, 139, 147, 155, 168, 170, 177, 179, 180, 181, 183, 185, 191, 209 and 215 were deleterious among those RBS mutants. For example, sequence analysis of HA from H1-H14 IAV showed that residues 134G, 139C, 147F, 170N, 180W, and 183H were 100% conserved from H1 to H14 of IAV. The residues 181G, 185P, and 215P were 92.8% conserved from H1 to H14 of IAV except residue 181A for H12, residue 185S for H7, and residue 215L for H13 (data not shown). Our results indicate that residue C was conserved among all position (7, 45, 58,70, 93, 139, 278, 282, 306, 467,474, 478, and 551) in the HA of IAV from H1 to H14. Residue C (7, 45, 58,70, 93, 139, 278, 282, 306, 467,474, 478, and 551) forms the disulfide linkages, which is critical in forming the structure of HA1. In addition, hydrophobic aromatic residues such as F and W were

found to form the base of the RBS, which was also reported by Xu et al (72). Our study supported that the highly conserved amino acids form the framework of the HA RBS and that mutations at those sites reduced virus's potency and are associated with virus fitness.

This study identified that the mutations S193T, S160A, S188N, R220K, K172R, T200S, S206T were associated with the antigenicity of H1N1 IAV. The results suggest the immunodominance of antigenic site B over antigenic site A for HA of CA/04 H1N1. In a prior study, Popova et al (249) reported the immunodominance of antigenic site B over site A for a recent seasonal H3N2 human IAV. More recently, Wang et al (250) reported that the amino acids in the antigenic site B determine antigenic difference between a human H3N2 variant and an ancestral seasonal H3N2 influenza virus. The heavy glycosylation of HA1 of contemporary H3N2 human IAV masked the virus and leading to the immunodominance of the RBS over other antigenic sites. The RBS has to maintain its structures in order to bind sialic acid receptors and thus is exposed to the neutralizing antibodies. However, the CA/04 HA only has one glycosylation site, suggesting that the immunodominance of antigenic site B may not be directly associated with the extent of glycosylation of HA.

Residue 193 is located at antigenic site B and has been reported to have a strong antigenic impact on human H3N2 IAV(157, 251). Of note, our results suggested that residue 193 had a strong antigenic impact on the 2009 pandemic H1N1 virus. Residue 193 showed a high level of diversity between subtypes of HA. Moreover, residue 193 was reported to be located at the epitope recognized by mouse monoclonal antibodies (251). Those findings suggest that residue 193 plays an important role in the antigenic changes of both human H1N1 and H3N2 IAV and thus drive viral evolution. Threonine

differs from serine by a substitution of hydrogen to a methyl group. The substitution of S to T at position 193 increased the bulkiness around that area, which may make it more difficult for residue 193 to adopt an alpha-helical conformation in the 190-helix of the RBS and therefore alter the recognition of targeting neutralizing antibodies.

This study identified 13 mutations that would improve virus replication efficiency in both eggs and cells without changing virus antigenicity. For example, mutation N159D improved virus yield in both MDCK cells and eggs. It has been reported that N159D functioned together with Q226R and switched the binding preference of the 2009 pandemic H1N1 virus from Neu5Acα2-6Gal to Neu5Acα2-3Gal(252). Residue 159 is located at the very top of the RBS. It has been reported that N159D may increase the local negative charge at the globular head, which would decrease the affinity of binding to the human receptor but increase affinity for the avian receptor. Therefore, residue 159 may be involved in host adaption of avian IAV to humans. Interestingly, we identified that the K166Q HA mutation significantly improved virus yield in chicken eggs but not in MDCK cells. Our sequence analysis suggested K166Q mutation occurred during the 2012-2013 seasons and then 166Q quickly become dominant in pandemic isolates (>99%) after the year 2013. It has been reported that the K166Q HA mutant of the 2009 pandemic H1N1virus is antigenically similar to the wild type strain, which was also observed in this study. Linderman et al reported that the K166 epitope is shielded by a glycosylation site located at residue 129 that was present in swine H1N1 IAV circulating after 1985(253). A considerable homology in the vicinity of K166 was observed between the 1918 pandemic H1N1 and the 2009 pandemic H1N1(254). Our study supports the observation that the selective pressure on HA was aimed to improve viral fitness and

replication to produce a large number of progeny, rather than only avoiding the host immune recognition.

We identified one mutant virus, which possessed the D130E, S193T, and A198S HA mutations, had switched the binding preference of 2009 pandemic H1N1 from the human receptor to the avian receptor. Our sequence analysis suggested that residue 130E was conserved in duck H1N1 isolates. The D130E mutation has been shown to increase virus replication efficiency in human lung epithelial cells and in mouse lungs (255). In addition, North American swine H1N1 was found to have 130E only between the years of 1977 to 2000. The 130D residue started to emerge around 2000, and then became increasingly prevalent and now dominant in the North American swine population. Our results support the hypothesis that the E130D mutation facilitated the emergence of the 2009 pandemic, through the mechanism of decreasing the virus binding affinity to avian receptors. It has been widely considered that a single mutation at D225G enables virus binding to both Neu5Acα2-6Gal to Neu5Acα2-3Gal(77, 256). Here, we report a novel biomarker, A198T, which enables the binding of the 2009 pandemic H1N1to both Neu5Acα2-6Gal to Neu5Acα2-3Gal. Our results suggest that the T198A mutation changed the binding preference of avian H1N1 from Neu5Acα2-3Gal to Neu5Acα2-6Gal. The T198A mutation may have occurred before 1918 and played important role in the outbreak of 1918 pandemic H1N1. Sequence analysis showed that the incidence of 225G is 13.6 fold higher than the 198T, suggesting that 198A is a better human adapted biomarker than 225D for H1N1 IAV.

A high yield and antigenic match vaccine seed virus is critical for the success of the influenza vaccine production. Optimization of influenza vaccine seed selections is

urgently need in the influenza vaccine industry. Currently, the mechanisms for determining the IAV yield remains largely unknown. Besides the virus RNP activity, the mutations on the receptor binding site of IAVs clearly could improve the virus growth in host cells, either by expanding the receptor targets for IAVs or by enhancing the balance of HA and NA activity. At least four key glycan substructure features associated with vaccine yields were elucidated from this study: 1). increasing the binding ratio of α2-3 sialic acid glycans to α2-6 sialic acid glycans to expand the binding receptor targets; 2) increasing of binding ratio of Sle^X glycans to α2-6 sialic acid glycans might increase the virus yield, especially in eggs, which may express high levels of SLe^X; 3) increasing virus binding to the non-sialic acid glycans might increase yield; 4) increasing virus binding to Neu5Gcα2-6Gal might increase the yield in MDCK cells, which might express α2-6 linked Gc glycans as a functional receptor for virus entry.

In summary, this study systematically evaluated the effect of mutations on the RBS of H1N1 IAV to modulate virus antigenicity, growth, and binding properties. The residues identified in this study, which had a great impact on virus antigenicity and growth efficiency, would help the influenza vaccine seed strain selection. The novel biomarkers that related to the receptor binding specificity of H1N1 IAV expanded the knowledge of H1N1 IAV evolution and host adaptions.

CHAPTER VI

CONCLUSIONS

The studies in this dissertation have generated following conclusions:

First, a novel method was developed and validated for rapidly producing influenza virus mutants by using an error-prone PCR–based mutagenesis strategy. This method targets a specific region of a gene for rapidly generating a mutant library; then the desired mutants are selected by phenotype screening of the library. Selecting seed virus for influenza vaccine production is time consuming, and candidate strains must be generated rapidly, produce at high yields, and have unaltered antigenicity. This novel method was used to develop six influenza vaccine candidates that had unaltered antigenicity and produced yields in chicken embryonic eggs that were higher than those produced by the wild-type strain. Animal studies using two of the candidate strains demonstrated that both strains provided effective protection and thus could be potential vaccine candidates. This error-prone PCR–based mutagenesis method has the potential to increase vaccine yields and could also be applied to gain-of-function studies for influenza virus and other pathogens.

Second, although a promising complement to current egg-based influenza vaccines, cell-based vaccines have one big challenge; high-yield vaccine seeds for production. In this study, we identified a molecular signature- the Y161F mutation in hemagglutinin (HA) that resulted in increased virus growth in Madin-Darby canine kidney cells and Vero cells, the two commonly used cell lines in influenza vaccine manufacturing. The Y161F mutation not only increased HA thermostability, but also enhanced its binding affinity to α2,6 and α2,3-linked Neu5Ac. These results suggest that a vaccine strain bearing the Y161F mutation in HA could potentially increase vaccine yields in mammalian cell culture systems.

Third, IAVs cause a significant burden on human and animal health, and the mechanisms for interspecies transmission of IAVs are far from being understood. The findings from this study suggest that an equine-origin A(H3N8) IAV with the mutation W222L at its hemagglutinin increased binding to canine-specific receptors with sialyl Lewis X and Neu5Gc motifs and, therefore may have facilitated viral adaptation from horses to dogs. These findings suggest that in addition to the glycosidic linkage (e.g., α2,3-linked and α2,6-linked), the substructure in the receptor saccharides (e.g., sialyl Lewis X and Neu5Gc) could present an interspecies transmission barrier for IAVs and drive viral mutations to overcome such barriers.

Fourth, we revealed the immunodominance of antigenic site B over site A for HA of CA/04 H1N1. The residue at position193 was found to have a great impact on the antigenicity of CA/04 H1N1. The high yield signatures identified in this dissertation would improve the vaccine seed strain selection. This dissertation identified novel biomarkers, residues at positions 130 and 198, which can affect the binding specificity of

H1N1 IAV, which expanded the knowledge of the 1918 pandemic H1N1 and the 2009 pandemic H1N1 emergence and evolution.

Overall, this dissertation has systematically evaluated the effects of the receptor binding sites of IAV on virus antigenicity, yields, and receptor binding properties and the key signatures were identified. This novel information will be useful in understanding influenza cell and host tropisms and in selecting effective vaccine strains in influenza surveillance.

REFERENCES

1. Hause BM, Collin EA, Liu R, Huang B, Sheng Z, Lu W, Wang D, Nelson EA, Li F. 2014. Characterization of a novel influenza virus in cattle and Swine: proposal for a new genus in the Orthomyxoviridae family. MBio 5:e00031-14.

2. Hause BM, Ducatez M, Collin EA, Ran Z, Liu R, Sheng Z, Armien A, Kaplan B, Chakravarty S, Hoppe AD, Webby RJ, Simonson RR, Li F. 2013. Isolation of a novel swine influenza virus from Oklahoma in 2011 which is distantly related to human influenza C viruses. PLoS Pathog 9:e1003176.

3. Tong S, Zhu X, Li Y, Shi M, Zhang J, Bourgeois M, Yang H, Chen X, Recuenco S, Gomez J, Chen LM, Johnson A, Tao Y, Dreyfus C, Yu W, McBride R, Carney PJ, Gilbert AT, Chang J, Guo Z, Davis CT, Paulson JC, Stevens J, Rupprecht CE, Holmes EC, Wilson IA, Donis RO. 2013. New world bats harbor diverse influenza A viruses. PLoS Pathog 9:e1003657.

4. Taubenberger JK, Morens DM. 2010. Influenza: the once and future pandemic. Public Health Rep 125 Suppl 3:16-26.

5. Chen W, Calvo PA, Malide D, Gibbs J, Schubert U, Bacik I, Basta S, O'Neill R, Schickli J, Palese P, Henklein P, Bennink JR, Yewdell JW. 2001. A novel influenza A virus mitochondrial protein that induces cell death. Nat Med 7:1306-12.

6. Wise HM, Foeglein A, Sun J, Dalton RM, Patel S, Howard W, Anderson EC, Barclay WS, Digard P. 2009. A complicated message: Identification of a novel PB1-related protein translated from influenza A virus segment 2 mRNA. J Virol 83:8021-31.

7. Priore SF, Kauffmann AD, Baman JR, Turner DH. 2015. The Influenza A PB1-F2 and N40 Start Codons Are Contained within an RNA Pseudoknot. Biochemistry 54:3413-5.

8. Jagger BW, Wise HM, Kash JC, Walters KA, Wills NM, Xiao YL, Dunfee RL, Schwartzman LM, Ozinsky A, Bell GL, Dalton RM, Lo A, Efstathiou S, Atkins JF, Firth AE, Taubenberger JK, Digard P. 2012. An overlapping protein-coding region in influenza A virus segment 3 modulates the host response. Science 337:199-204.

9. Muramoto Y, Noda T, Kawakami E, Akkina R, Kawaoka Y. 2013. Identification of novel influenza A virus proteins translated from PA mRNA. J Virol 87:2455-62.

10. Lamb RA, Lai CJ, Choppin PW. 1981. Sequences of mRNAs derived from genome RNA segment 7 of influenza virus: colinear and interrupted mRNAs code for overlapping proteins. Proc Natl Acad Sci U S A 78:4170-4.

11. Lamb RA, Choppin PW, Chanock RM, Lai CJ. 1980. Mapping of the two overlapping genes for polypeptides NS1 and NS2 on RNA segment 8 of influenza virus genome. Proc Natl Acad Sci U S A 77:1857-61.

12. Kochs G, Garcia-Sastre A, Martinez-Sobrido L. 2007. Multiple anti-interferon actions of the influenza A virus NS1 protein. J Virol 81:7011-21.

13. Calder LJ, Wasilewski S, Berriman JA, Rosenthal PB. 2010. Structural organization of a filamentous influenza A virus. Proc Natl Acad Sci U S A 107:10685-90.

14. Bouvier NM, Palese P. 2008. The biology of influenza viruses. Vaccine 26 Suppl 4:D49-53.

15. Smirnov Yu A, Kuznetsova MA, Kaverin NV. 1991. The genetic aspects of influenza virus filamentous particle formation. Arch Virol 118:279-84.

16. Hughey PG, Roberts PC, Holsinger LJ, Zebedee SL, Lamb RA, Compans RW. 1995. Effects of antibody to the influenza A virus M2 protein on M2 surface expression and virus assembly. Virology 212:411-21.

17. Bourmakina SV, Garcia-Sastre A. 2003. Reverse genetics studies on the filamentous morphology of influenza A virus. J Gen Virol 84:517-27.

18. Lakdawala SS, Lamirande EW, Suguitan AL, Jr., Wang W, Santos CP, Vogel L, Matsuoka Y, Lindsley WG, Jin H, Subbarao K. 2011. Eurasian-origin gene segments contribute to the transmissibility, aerosol release, and morphology of the 2009 pandemic H1N1 influenza virus. PLoS Pathog 7:e1002443.

19. Nemeroff ME, Barabino SM, Li Y, Keller W, Krug RM. 1998. Influenza virus NS1 protein interacts with the cellular 30 kDa subunit of CPSF and inhibits 3'end formation of cellular pre-mRNAs. Mol Cell 1:991-1000.

20. Steinhauer DA. 1999. Role of hemagglutinin cleavage for the pathogenicity of influenza virus. Virology 258:1-20.

21. Horimoto T, Nakayama K, Smeekens SP, Kawaoka Y. 1994. Proprotein-processing endoproteases PC6 and furin both activate hemagglutinin of virulent avian influenza viruses. J Virol 68:6074-8.

22. Varki A, Lowe JB. 2009. Biological Roles of Glycans. *In* nd, Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, Etzler ME (ed), Essentials of Glycobiology, Cold Spring Harbor (NY).

23. Lundblad A. 2015. Gunnar Blix and his discovery of sialic acids. Fascinating molecules in glycobiology. Ups J Med Sci 120:104-12.

24. Hirst GK. 1941. The Agglutination of Red Cells by Allantoic Fluid of Chick Embryos Infected with Influenza Virus. Science 94:22-3.

25. Gottschalk A. 1957. Neuraminidase: the specific enzyme of influenza virus and Vibrio cholerae. Biochim Biophys Acta 23:645-6.

26. Suzuki Y, Ito T, Suzuki T, Holland RE, Jr., Chambers TM, Kiso M, Ishida H, Kawaoka Y. 2000. Sialic acid species as a determinant of the host range of influenza A viruses. J Virol 74:11825-31.

27. Gambaryan AS, Tuzikov AB, Piskarev VE, Yamnikova SS, Lvov DK, Robertson JS, Bovin NV, Matrosovich MN. 1997. Specification of receptor-binding phenotypes of influenza virus isolates from different hosts using synthetic sialylglycopolymers: non-egg-adapted human H1 and H3 influenza A and influenza B viruses share a common high binding affinity for 6'-sialyl(N-acetyllactosamine). Virology 232:345-50.

28. Rogers GN, Paulson JC. 1983. Receptor determinants of human and animal influenza virus isolates: differences in receptor specificity of the H3 hemagglutinin based on species of origin. Virology 127:361-73.

29. Chu VC, Whittaker GR. 2004. Influenza virus entry and infection require host cell N-linked glycoprotein. Proc Natl Acad Sci U S A 101:18153-8.

30. de Vries E, de Vries RP, Wienholts MJ, Floris CE, Jacobs MS, van den Heuvel A, Rottier PJ, de Haan CA. 2012. Influenza A virus entry into cells lacking sialylated N-glycans. Proc Natl Acad Sci U S A 109:7457-62.

31. Ablan S, Rawat SS, Blumenthal R, Puri A. 2001. Entry of influenza virus into a glycosphingolipid-deficient mouse skin fibroblast cell line. Arch Virol 146:2227-38.

32. Higa HH, Rogers GN, Paulson JC. 1985. Influenza virus hemagglutinins differentiate between receptor determinants bearing N-acetyl-, N-glycollyl-, and N,O-diacetylneuraminic acids. Virology 144:279-82.

33. Yang G, Li S, Blackmon S, Ye J, Bradley KC, Cooley J, Smith D, Hanson L, Cardona C, Steinhauer DA, Webby R, Liao M, Wan XF. 2013. Mutation tryptophan to leucine at position 222 of haemagglutinin could facilitate H3N2 influenza A virus infection in dogs. J Gen Virol 94:2599-608.

34. Suzuki T, Horiike G, Yamazaki Y, Kawabe K, Masuda H, Miyamoto D, Matsuda M, Nishimura SI, Yamagata T, Ito T, Kida H, Kawaoka Y, Suzuki Y. 1997. Swine influenza virus strains recognize sialylsugar chains containing the molecular species of sialic acid predominantly present in the swine tracheal epithelium. FEBS Lett 404:192-6.

35. Hedlund M, Tangvoranuntakul P, Takematsu H, Long JM, Housley GD, Kozutsumi Y, Suzuki A, Wynshaw-Boris A, Ryan AF, Gallo RL, Varki N, Varki A. 2007. N-glycolylneuraminic acid deficiency in mice: implications for human biology and evolution. Mol Cell Biol 27:4340-6.

36. Varki A. 2001. Loss of N-glycolylneuraminic acid in humans: Mechanisms, consequences, and implications for hominid evolution. Am J Phys Anthropol Suppl 33:54-69.

37. Chou HH, Takematsu H, Diaz S, Iber J, Nickerson E, Wright KL, Muchmore EA, Nelson DL, Warren ST, Varki A. 1998. A mutation in human CMP-sialic acid hydroxylase occurred after the Homo-Pan divergence. Proc Natl Acad Sci U S A 95:11751-6.

38. Irie A, Koyama S, Kozutsumi Y, Kawasaki T, Suzuki A. 1998. The molecular basis for the absence of N-glycolylneuraminic acid in humans. J Biol Chem 273:15866-71.

39. Hayakawa T, Aki I, Varki A, Satta Y, Takahata N. 2006. Fixation of the human-specific CMP-N-acetylneuraminic acid hydroxylase pseudogene and implications of haplotype diversity for human evolution. Genetics 172:1139-46.

40. Springer SA, Diaz SL, Gagneux P. 2014. Parallel evolution of a self-signal: humans and new world monkeys independently lost the cell surface sugar Neu5Gc. Immunogenetics 66:671-4.

41. Ng PS, Bohm R, Hartley-Tassell LE, Steen JA, Wang H, Lukowski SW, Hawthorne PL, Trezise AE, Coloe PJ, Grimmond SM, Haselhorst T, von Itzstein M, Paton AW, Paton JC, Jennings MP. 2014. Ferrets exclusively synthesize Neu5Ac and express naturally humanized influenza A virus receptors. Nat Commun 5:5750.

42. 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 344:1494-500.

43. Liu Y, Han C, Wang X, Lin J, Ma M, Shu Y, Zhou J, Yang H, Liang Q, Guo C, Zhu J, Wei H, Zhao J, Ma Z, Pan J. 2009. Influenza A virus receptors in the respiratory and intestinal tracts of pigeons. Avian Pathol 38:263-6.

44. Gambaryan AS, Matrosovich TY, Philipp J, Munster VJ, Fouchier RA, Cattoli G, Capua I, Krauss SL, Webster RG, Banks J, Bovin NV, Klenk HD, Matrosovich MN. 2012. Receptor-binding profiles of H7 subtype influenza viruses in different host species. J Virol 86:4370-9.

45. Takahashi T, Takano M, Kurebayashi Y, Masuda M, Kawagishi S, Takaguchi M, Yamanaka T, Minami A, Otsubo T, Ikeda K, Suzuki T. 2014. N-glycolylneuraminic acid on human epithelial cells prevents entry of influenza A viruses that possess N-glycolylneuraminic acid binding ability. J Virol 88:8445-56.

46. Angata T, Varki A. 2002. Chemical diversity in the sialic acids and related alpha-keto acids: an evolutionary perspective. Chem Rev 102:439-69.

47. Wasik BR, Barnard KN, Parrish CR. 2016. Effects of Sialic Acid Modifications on Virus Binding and Infection. Trends Microbiol 24:991-1001.

48. Klein A, Krishna M, Varki NM, Varki A. 1994. 9-O-acetylated sialic acids have widespread but selective expression: analysis using a chimeric dual-function probe derived from influenza C hemagglutinin-esterase. Proc Natl Acad Sci U S A 91:7782-6.

49. Schauer R, Schmid H, Pommerencke J, Iwersen M, Kohla G. 2001. Metabolism and role of O-acetylated sialic acids. Adv Exp Med Biol 491:325-42.

50. Schauer R. 2004. Sialic acids: fascinating sugars in higher animals and man. Zoology (Jena) 107:49-64.

51. Rogers GN, Herrler G, Paulson JC, Klenk HD. 1986. Influenza C virus uses 9-O-acetyl-N-acetylneuraminic acid as a high affinity receptor determinant for attachment to cells. J Biol Chem 261:5947-51.

52. Song H, Qi J, Khedri Z, Diaz S, Yu H, Chen X, Varki A, Shi Y, Gao GF. 2016. An Open Receptor-Binding Cavity of Hemagglutinin-Esterase-Fusion Glycoprotein from Newly-Identified Influenza D Virus: Basis for Its Broad Cell Tropism. PLoS Pathog 12:e1005411.

53. Baum LG, Paulson JC. 1990. Sialyloligosaccharides of the respiratory epithelium in the selection of human influenza virus receptor specificity. Acta Histochem Suppl 40:35-8.

54. Couceiro JN, Paulson JC, Baum LG. 1993. Influenza virus strains selectively recognize sialyloligosaccharides on human respiratory epithelium; the role of the host cell in selection of hemagglutinin receptor specificity. Virus Res 29:155-65.

55. Aich U, Beckley N, Shriver Z, Raman R, Viswanathan K, Hobbie S, Sasisekharan R. 2011. Glycomics-based analysis of chicken red blood cells provides insight into the selectivity of the viral agglutination assay. FEBS J 278:1699-712.

56. Ito T, Suzuki Y, Mitnaul L, Vines A, Kida H, Kawaoka Y. 1997. Receptor specificity of influenza A viruses correlates with the agglutination of erythrocytes from different animal species. Virology 227:493-9.

57. Makkoch J, Prachayangprecha S, Payungporn S, Chieochansin T, Songserm T, Amonsin A, Poovorawan Y. 2012. Erythrocyte binding preference of human pandemic influenza virus a and its effect on antibody response detection. Ann Lab Med 32:276-82.

58. Chandrasekaran A, Srinivasan A, Raman R, Viswanathan K, Raguram S, Tumpey TM, Sasisekharan V, Sasisekharan R. 2008. Glycan topology determines human adaptation of avian H5N1 virus hemagglutinin. Nat Biotechnol 26:107-13.

59. Ito T, Couceiro JN, Kelm S, Baum LG, Krauss S, Castrucci MR, Donatelli I, Kida H, Paulson JC, Webster RG, Kawaoka Y. 1998. Molecular basis for the generation in pigs of influenza A viruses with pandemic potential. J Virol 72:7367-73.

60. Suzuki Y. 2005. Sialobiology of influenza: molecular mechanism of host range variation of influenza viruses. Biol Pharm Bull 28:399-408.

61. Guo CT, Takahashi N, Yagi H, Kato K, Takahashi T, Yi SQ, Chen Y, Ito T, Otsuki K, Kida H, Kawaoka Y, Hidari KI, Miyamoto D, Suzuki T, Suzuki Y. 2007. The quail and chicken intestine have sialyl-galactose sugar chains responsible for the binding of influenza A viruses to human type receptors. Glycobiology 17:713-24.

62. Walther T, Karamanska R, Chan RW, Chan MC, Jia N, Air G, Hopton C, Wong MP, Dell A, Malik Peiris JS, Haslam SM, Nicholls JM. 2013. Glycomic analysis of human respiratory tract tissues and correlation with influenza virus infection. PLoS Pathog 9:e1003223.

63. Jia N, Barclay WS, Roberts K, Yen HL, Chan RW, Lam AK, Air G, Peiris JS, Dell A, Nicholls JM, Haslam SM. 2014. Glycomic characterization of respiratory tract tissues of ferrets: implications for its use in influenza virus infection studies. J Biol Chem 289:28489-504.

64. Muthana SM, Campbell CT, Gildersleeve JC. 2012. Modifications of glycans: biological significance and therapeutic opportunities. ACS Chem Biol 7:31-43.

65. Gambaryan A, Yamnikova S, Lvov D, Tuzikov A, Chinarev A, Pazynina G, Webster R, Matrosovich M, Bovin N. 2005. Receptor specificity of influenza viruses from birds and mammals: new data on involvement of the inner fragments of the carbohydrate chain. Virology 334:276-83.

66. Gambaryan AS, Tuzikov AB, Pazynina GV, Desheva JA, Bovin NV, Matrosovich MN, Klimov AI. 2008. 6-sulfo sialyl Lewis X is the common receptor determinant recognized by H5, H6, H7 and H9 influenza viruses of terrestrial poultry. Virol J 5:85.

67. Hiono T, Okamatsu M, Nishihara S, Takase-Yoden S, Sakoda Y, Kida H. 2014. A chicken influenza virus recognizes fucosylated alpha2,3 sialoglycan receptors on the epithelial cells lining upper respiratory tracts of chickens. Virology 456-457:131-8.

68. Hiono T, Okamatsu M, Igarashi M, McBride R, de Vries RP, Peng W, Paulson JC, Sakoda Y, Kida H. 2016. Amino acid residues at positions 222 and 227 of the hemagglutinin together with the neuraminidase determine binding of H5 avian influenza viruses to sialyl Lewis X. Arch Virol 161:307-16.

69. Wilson IA, Skehel JJ, Wiley DC. 1981. Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 A resolution. Nature 289:366-73.

70. Weis W, Brown JH, Cusack S, Paulson JC, Skehel JJ, Wiley DC. 1988. Structure of the influenza virus haemagglutinin complexed with its receptor, sialic acid. Nature 333:426-31.

71. Martin J, Wharton SA, Lin YP, Takemoto DK, Skehel JJ, Wiley DC, Steinhauer DA. 1998. Studies of the binding properties of influenza hemagglutinin receptor-site mutants. Virology 241:101-11.

72. Xu R, McBride R, Paulson JC, Basler CF, Wilson IA. 2010. Structure, receptor binding, and antigenicity of influenza virus hemagglutinins from the 1957 H2N2 pandemic. J Virol 84:1715-21.

73. Liu J, Stevens DJ, Haire LF, Walker PA, Coombs PJ, Russell RJ, Gamblin SJ, Skehel JJ. 2009. Structures of receptor complexes formed by hemagglutinins from the Asian Influenza pandemic of 1957. Proc Natl Acad Sci U S A 106:17175-80.

74. Ha Y, Stevens DJ, Skehel JJ, Wiley DC. 2003. X-ray structure of the hemagglutinin of a potential H3 avian progenitor of the 1968 Hong Kong pandemic influenza virus. Virology 309:209-18.

75. de Vries RP, Tzarum N, Peng W, Thompson AJ, Ambepitiya Wickramasinghe IN, de la Pena ATT, van Breemen MJ, Bouwman KM, Zhu X, McBride R, Yu W, Sanders RW, Verheije MH, Wilson IA, Paulson JC. 2017. A single mutation in Taiwanese H6N1 influenza hemagglutinin switches binding to human-type receptors. EMBO Mol Med 9:1314-1325.

76. Stevens J, Blixt O, Glaser L, Taubenberger JK, Palese P, Paulson JC, Wilson IA. 2006. Glycan microarray analysis of the hemagglutinins from modern and pandemic influenza viruses reveals different receptor specificities. J Mol Biol 355:1143-55.

77. Glaser L, Stevens J, Zamarin D, Wilson IA, Garcia-Sastre A, Tumpey TM, Basler CF, Taubenberger JK, Palese P. 2005. A single amino acid substitution in 1918 influenza virus hemagglutinin changes receptor binding specificity. J Virol 79:11533-6.

78. de Vries RP, Peng W, Grant OC, Thompson AJ, Zhu X, Bouwman KM, de la Pena ATT, van Breemen MJ, Ambepitiya Wickramasinghe IN, de Haan CAM, Yu W, McBride R, Sanders RW, Woods RJ, Verheije MH, Wilson IA, Paulson JC. 2017. Three mutations switch H7N9 influenza to human-type receptor specificity. PLoS Pathog 13:e1006390.

79. Stevens J, Blixt O, Tumpey TM, Taubenberger JK, Paulson JC, Wilson IA. 2006. Structure and receptor specificity of the hemagglutinin from an H5N1 influenza virus. Science 312:404-10.

80. Stevens J, Chen LM, Carney PJ, Garten R, Foust A, Le J, Pokorny BA, Manojkumar R, Silverman J, Devis R, Rhea K, Xu X, Bucher DJ, Paulson JC, Cox NJ, Klimov A, Donis RO. 2010. Receptor specificity of influenza A H3N2 viruses isolated in mammalian cells and embryonated chicken eggs. J Virol 84:8287-99.

81. Bradley KC, Galloway SE, Lasanajak Y, Song X, Heimburg-Molinaro J, Yu H, Chen X, Talekar GR, Smith DF, Cummings RD, Steinhauer DA. 2011. Analysis of influenza virus hemagglutinin receptor binding mutants with limited receptor recognition properties and conditional replication characteristics. J Virol 85:12387-98.

82. Varghese JN, Colman PM. 1991. Three-dimensional structure of the neuraminidase of influenza virus A/Tokyo/3/67 at 2.2 A resolution. J Mol Biol 221:473-86.

83. Matrosovich MN, Matrosovich TY, Gray T, Roberts NA, Klenk HD. 2004. Neuraminidase is important for the initiation of influenza virus infection in human airway epithelium. J Virol 78:12665-7.

84. Wagner R, Wolff T, Herwig A, Pleschka S, Klenk HD. 2000. Interdependence of hemagglutinin glycosylation and neuraminidase as regulators of influenza virus growth: a study by reverse genetics. J Virol 74:6316-23.

85. Mitnaul LJ, Matrosovich MN, Castrucci MR, Tuzikov AB, Bovin NV, Kobasa D, Kawaoka Y. 2000. Balanced hemagglutinin and neuraminidase activities are critical for efficient replication of influenza A virus. J Virol 74:6015-20.

86. Laver WG, Colman PM, Webster RG, Hinshaw VS, Air GM. 1984. Influenza virus neuraminidase with hemagglutinin activity. Virology 137:314-23.

87. Webster RG, Air GM, Metzger DW, Colman PM, Varghese JN, Baker AT, Laver WG. 1987. Antigenic structure and variation in an influenza virus N9 neuraminidase. J Virol 61:2910-6.

88. Nuss JM, Air GM. 1991. Transfer of the hemagglutinin activity of influenza virus neuraminidase subtype N9 into an N2 neuraminidase background. Virology 183:496-504.

89. Air GM, Laver WG. 1995. Red cells bound to influenza virus N9 neuraminidase are not released by the N9 neuraminidase activity. Virology 211:278-84.

90. Hausmann J, Kretzschmar E, Garten W, Klenk HD. 1995. N1 neuraminidase of influenza virus A/FPV/Rostock/34 has haemadsorbing activity. J Gen Virol 76 (Pt 7):1719-28.

91. Varghese JN, Colman PM, van Donkelaar A, Blick TJ, Sahasrabudhe A, McKimm-Breschkin JL. 1997. Structural evidence for a second sialic acid binding site in avian influenza virus neuraminidases. Proc Natl Acad Sci U S A 94:11808-12.

92. Lin YP, Gregory V, Collins P, Kloess J, Wharton S, Cattle N, Lackenby A, Daniels R, Hay A. 2010. Neuraminidase receptor binding variants of human influenza A(H3N2) viruses resulting from substitution of aspartic acid 151 in the catalytic site: a role in virus attachment? J Virol 84:6769-81.

93. Zhu X, McBride R, Nycholat CM, Yu W, Paulson JC, Wilson IA. 2012. Influenza virus neuraminidases with reduced enzymatic activity that avidly bind sialic Acid receptors. J Virol 86:13371-83.

94. Gulati S, Smith DF, Cummings RD, Couch RB, Griesemer SB, St George K, Webster RG, Air GM. 2013. Human H3N2 Influenza Viruses Isolated from 1968 To 2012 Show Varying Preference for Receptor Substructures with No Apparent Consequences for Disease or Spread. PLoS One 8:e66325.

95. Hooper KA, Bloom JD. 2013. A mutant influenza virus that uses an N1 neuraminidase as the receptor-binding protein. J Virol 87:12531-40.

96. Hooper KA, Crowe JE, Jr., Bloom JD. 2015. Influenza viruses with receptor-binding N1 neuraminidases occur sporadically in several lineages and show no attenuation in cell culture or mice. J Virol 89:3737-45.

97. Benton DJ, Wharton SA, Martin SR, McCauley JW. 2017. Role of Neuraminidase in Influenza A(H7N9) Virus Receptor Binding. J Virol 91.

98. Carroll SM, Higa HH, Paulson JC. 1981. Different cell-surface receptor determinants of antigenically similar influenza virus hemagglutinins. J Biol Chem 256:8357-63.

99. Gambaryan AS, Matrosovich MN. 1992. A solid-phase enzyme-linked assay for influenza virus receptor-binding activity. J Virol Methods 39:111-23.

100. Fukui S, Feizi T, Galustian C, Lawson AM, Chai W. 2002. Oligosaccharide microarrays for high-throughput detection and specificity assignments of carbohydrate-protein interactions. Nat Biotechnol 20:1011-7.

101. Stevens J, Blixt O, Paulson JC, Wilson IA. 2006. Glycan microarray technologies: tools to survey host specificity of influenza viruses. Nat Rev Microbiol 4:857-64.

102. Xiong X, Coombs PJ, Martin SR, Liu J, Xiao H, McCauley JW, Locher K, Walker PA, Collins PJ, Kawaoka Y, Skehel JJ, Gamblin SJ. 2013. Receptor binding by a ferret-transmissible H5 avian influenza virus. Nature 497:392-6.

103. Petersen RL. 2017. Strategies Using Bio-Layer Interferometry Biosensor Technology for Vaccine Research and Development. Biosensors (Basel) 7.

104. Liao YC, Lee MS, Ko CY, Hsiung CA. 2008. Bioinformatics models for predicting antigenic variants of influenza A/H3N2 virus. Bioinformatics 24:505-12.

105. Sun H, Yang J, Zhang T, Long LP, Jia K, Yang G, Webby RJ, Wan XF. 2013. Using sequence data to infer the antigenicity of influenza virus. MBio 4.

106. Steinbruck L, Klingen TR, McHardy AC. 2014. Computational prediction of vaccine strains for human influenza A (H3N2) viruses. J Virol 88:12123-32.

107. Neher RA, Bedford T, Daniels RS, Russell CA, Shraiman BI. 2016. Prediction, dynamics, and visualization of antigenic phenotypes of seasonal influenza viruses. Proc Natl Acad Sci U S A 113:E1701-9.

108. Yao Y, Li X, Liao B, Huang L, He P, Wang F, Yang J, Sun H, Zhao Y, Yang J. 2017. Predicting influenza antigenicity from Hemagglutintin sequence data based on a joint random forest method. Sci Rep 7:1545.

109. Li M, Wang B. 2006. Computational studies of H5N1 hemagglutinin binding with SA-alpha-2, 3-Gal and SA-alpha-2, 6-Gal. Biochem Biophys Res Commun 347:662-8.

110. Iwata T, Fukuzawa K, Nakajima K, Aida-Hyugaji S, Mochizuki Y, Watanabe H, Tanaka S. 2008. Theoretical analysis of binding specificity of influenza viral hemagglutinin to avian and human receptors based on the fragment molecular orbital method. Comput Biol Chem 32:198-211.

111. Sawada T, Hashimoto T, Tokiwa H, Suzuki T, Nakano H, Ishida H, Kiso M, Suzuki Y. 2008. Ab initio fragment molecular orbital studies of influenza virus hemagglutinin-sialosaccharide complexes toward chemical clarification about the virus host range determination. Glycoconj J 25:805-15.

112. Das P, Li J, Royyuru AK, Zhou R. 2009. Free energy simulations reveal a double mutant avian H5N1 virus hemagglutinin with altered receptor binding specificity. J Comput Chem 30:1654-63.

113. Jongkon N, Mokmak W, Chuakheaw D, Shaw PJ, Tongsima S, Sangma C. 2009. Prediction of avian influenza A binding preference to human receptor using conformational analysis of receptor bound to hemagglutinin. BMC Genomics 10 Suppl 3:S24.

114. Newhouse EI, Xu D, Markwick PR, Amaro RE, Pao HC, Wu KJ, Alam M, McCammon JA, Li WW. 2009. Mechanism of glycan receptor recognition and specificity switch for avian, swine, and human adapted influenza virus hemagglutinins: a molecular dynamics perspective. J Am Chem Soc 131:17430-42.

115. Cao Y, Koh X, Dong L, Du X, Wu A, Ding X, Deng H, Shu Y, Chen J, Jiang T. 2011. Rapid estimation of binding activity of influenza virus hemagglutinin to human and avian receptors. PLoS One 6:e18664.

116. Tharakaraman K, Raman R, Viswanathan K, Stebbins NW, Jayaraman A, Krishnan A, Sasisekharan V, Sasisekharan R. 2013. Structural determinants for naturally evolving H5N1 hemagglutinin to switch its receptor specificity. Cell 153:1475-85.

117. Jayaraman A, Chandrasekaran A, Viswanathan K, Raman R, Fox JG, Sasisekharan R. 2012. Decoding the distribution of glycan receptors for human-adapted influenza A viruses in ferret respiratory tract. PLoS One 7:e27517.

118. Gerdil C. 2003. The annual production cycle for influenza vaccine. Vaccine 21:1776-9.

119. Robertson JS, Nicolson C, Harvey R, Johnson R, Major D, Guilfoyle K, Roseby S, Newman R, Collin R, Wallis C, Engelhardt OG, Wood JM, Le J, Manojkumar R, Pokorny BA, Silverman J, Devis R, Bucher D, Verity E, Agius C, Camuglia S, Ong C, Rockman S, Curtis A, Schoofs P, Zoueva O, Xie H, Li X, Lin Z, Ye Z, Chen LM, O'Neill E, Balish A, Lipatov AS, Guo Z, Isakova I, Davis CT, Rivailler P, Gustin KM, Belser JA, Maines TR, Tumpey TM, Xu X, Katz JM, Klimov A, Cox NJ, Donis RO. 2011. The development of vaccine viruses against pandemic A(H1N1) influenza. Vaccine 29:1836-43.

120. Chen Z, Wang W, Zhou H, Suguitan AL, Jr., Shambaugh C, Kim L, Zhao J, Kemble G, Jin H. 2010. Generation of live attenuated novel influenza virus A/California/7/09 (H1N1) vaccines with high yield in embryonated chicken eggs. J Virol 84:44-51.

121. Kilbourne ED. 1969. Future influenza vaccines and the use of genetic recombinants. Bull World Health Organ 41:643-645.

122. Kilbourne ED, Murphy JS. 1960. Genetic studies of influenza viruses. I. Viral morphology and growth capacity as exchangeable genetic traits. Rapid in ovo adaptation of early passage Asian strain isolates by combination with PR8. J Exp Med 111:387-406.

123. Fodor E, Devenish L, Engelhardt OG, Palese P, Brownlee GG, Garcia-Sastre A. 1999. Rescue of influenza A virus from recombinant DNA. J Virol 73:9679-82.

124. Neumann G, Watanabe T, Ito H, Watanabe S, Goto H, Gao P, Hughes M, Perez DR, Donis R, Hoffmann E, Hobom G, Kawaoka Y. 1999. Generation of influenza A viruses entirely from cloned cDNAs. Proc Natl Acad Sci U S A 96:9345-50.

125. Nicolson C, Major D, Wood JM, Robertson JS. 2005. Generation of influenza vaccine viruses on Vero cells by reverse genetics: an H5N1 candidate vaccine strain produced under a quality system. Vaccine 23:2943-52.

126. Horimoto T, Murakami S, Muramoto Y, Yamada S, Fujii K, Kiso M, Iwatsuki-Horimoto K, Kino Y, Kawaoka Y. 2007. Enhanced growth of seed viruses for H5N1 influenza vaccines. Virology 366:23-7.

127. Fulvini AA, Ramanunninair M, Le J, Pokorny BA, Arroyo JM, Silverman J, Devis R, Bucher D. 2011. Gene constellation of influenza A virus reassortants with high growth phenotype prepared as seed candidates for vaccine production. PLoS One 6:e20823.

128. Ottmann M, Duchamp MB, Casalegno JS, Frobert E, Moules V, Ferraris O, Valette M, Escuret V, Lina B. 2010. Novel influenza A(H1N1) 2009 in vitro reassortant viruses with oseltamivir resistance. Antivir Ther 15:721-6.

129. Gambaryan AS, Robertson JS, Matrosovich MN. 1999. Effects of egg-adaptation on the receptor-binding properties of human influenza A and B viruses. Virology 258:232-9.

130. Robertson JS, Bootman JS, Newman R, Oxford JS, Daniels RS, Webster RG, Schild GC. 1987. Structural changes in the haemagglutinin which accompany egg adaptation of an influenza A(H1N1) virus. Virology 160:31-7.

131. Rogers GN, Paulson JC, Daniels RS, Skehel JJ, Wilson IA, Wiley DC. 1983. Single amino acid substitutions in influenza haemagglutinin change receptor binding specificity. Nature 304:76-8.

132. Widjaja L, Ilyushina N, Webster RG, Webby RJ. 2006. Molecular changes associated with adaptation of human influenza A virus in embryonated chicken eggs. Virology 350:137-45.

133. Adamo JE, Liu T, Schmeisser F, Ye Z. 2009. Optimizing viral protein yield of influenza virus strain A/Vietnam/1203/2004 by modification of the neuraminidase gene. J Virol 83:4023-9.

134. Harvey R, Guilfoyle KA, Roseby S, Robertson JS, Engelhardt OG. 2011. Improved antigen yield in pandemic H1N1 (2009) candidate vaccine viruses with chimeric hemagglutinin molecules. J Virol 85:6086-90.

135. Harvey R, Nicolson C, Johnson RE, Guilfoyle KA, Major DL, Robertson JS, Engelhardt OG. 2010. Improved haemagglutinin antigen content in H5N1 candidate vaccine viruses with chimeric haemagglutinin molecules. Vaccine 28:8008-14.

136. Pan W, Dong Z, Meng W, Zhang W, Li T, Li C, Zhang B, Chen L. 2012. Improvement of influenza vaccine strain A/Vietnam/1194/2004 (H5N1) growth with the neuraminidase packaging sequence from A/Puerto Rico/8/34. Hum Vaccin Immunother 8:252-9.

137. Jing X, Phy K, Li X, Ye Z. 2012. Increased hemagglutinin content in a reassortant 2009 pandemic H1N1 influenza virus with chimeric neuraminidase containing donor A/Puerto Rico/8/34 virus transmembrane and stalk domains. Vaccine 30:4144-52.

138. Gomila RC, Suphaphiphat P, Judge C, Spencer T, Ferrari A, Wen Y, Palladino G, Dormitzer PR, Mason PW. 2013. Improving influenza virus backbones by including terminal regions of MDCK-adapted strains on hemagglutinin and neuraminidase gene segments. Vaccine 31:4736-43.

139. Ilyushina NA, Khalenkov AM, Seiler JP, Forrest HL, Bovin NV, Marjuki H, Barman S, Webster RG, Webby RJ. 2010. Adaptation of pandemic H1N1 influenza viruses in mice. J Virol 84:8607-16.

140. Ye J, Sorrell EM, Cai Y, Shao H, Xu K, Pena L, Hickman D, Song H, Angel M, Medina RA, Manicassamy B, Garcia-Sastre A, Perez DR. 2010. Variations in the hemagglutinin of the 2009 H1N1 pandemic virus: potential for strains with altered virulence phenotype? PLoS Pathog 6:e1001145.

141. Xu R, Krause JC, McBride R, Paulson JC, Crowe JE, Jr., Wilson IA. 2013. A recurring motif for antibody recognition of the receptor-binding site of influenza hemagglutinin. Nat Struct Mol Biol 20:363-70.

142. Schymkowitz JW, Rousseau F, Martins IC, Ferkinghoff-Borg J, Stricher F, Serrano L. 2005. Prediction of water and metal binding sites and their affinities by using the Fold-X force field. Proc Natl Acad Sci U S A 102:10147-52.

143. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem 25:1605-12.

144. Fan H, Schneidman-Duhovny D, Irwin JJ, Dong G, Shoichet BK, Sali A. 2011. Statistical potential for modeling and ranking of protein-ligand interactions. J Chem Inf Model 51:3078-92.

145. Cherry JR, Lamsa MH, Schneider P, Vind J, Svendsen A, Jones A, Pedersen AH. 1999. Directed evolution of a fungal peroxidase. Nat Biotechnol 17:379-84.

146. Cirino PC, K. M. Mayer, and D. Umeno. . 2003. Generating mutant libraries using error-prone PCR. Methods Mol Biol 231:3-9.

147. Daugherty PS, Chen G, Iverson BL, Georgiou G. 2000. Quantitative analysis of the effect of the mutation frequency on the affinity maturation of single chain Fv antibodies. Proc Natl Acad Sci U S A 97:2029-34.

148. Wan L, Twitchett MB, Eltis LD, Mauk AG, Smith M. 1998. In vitro evolution of horse heart myoglobin to increase peroxidase activity. Proc Natl Acad Sci U S A 95:12825-31.

149. Zang H, Irimia A, Choi JY, Angel KC, Loukachevitch LV, Egli M, Guengerich FP. 2006. Efficient and high fidelity incorporation of dCTP opposite 7,8-dihydro-8-oxodeoxyguanosine by Sulfolobus solfataricus DNA polymerase Dpo4. J Biol Chem 281:2358-72.

150. Imai M, Watanabe T, Hatta M, Das SC, Ozawa M, Shinya K, Zhong G, Hanson A, Katsura H, Watanabe S, Li C, Kawakami E, Yamada S, Kiso M, Suzuki Y, Maher EA, Neumann G, Kawaoka Y. 2012. Experimental adaptation of an influenza H5 HA confers respiratory droplet transmission to a reassortant H5 HA/H1N1 virus in ferrets. Nature 486:420-8.

151. Wu NC, Young AP, Dandekar S, Wijersuriya H, Al-Mawsawi LQ, Wu TT, Sun R. 2013. Systematic identification of H274Y compensatory mutations in influenza A virus neuraminidase by high-throughput screening. J Virol 87:1193-9.

152. Wagner R, Matrosovich M, Klenk HD. 2002. Functional balance between haemagglutinin and neuraminidase in influenza virus infections. Rev Med Virol 12:159-66.

153. Hartgroves LC, Koudstaal W, McLeod C, Moncorge O, Thompson CI, Ellis J, Bull C, Havenga MJ, Goudsmit J, Barclay WS. 2010. Rapid generation of a well-matched vaccine seed from a modern influenza A virus primary isolate without recourse to eggs. Vaccine 28:2973-9.

154. Lugovtsev VY, Vodeiko GM, Levandowski RA. 2005. Mutational pattern of influenza B viruses adapted to high growth replication in embryonated eggs. Virus Res 109:149-57.

155. Li Y, Myers JL, Bostick DL, Sullivan CB, Madara J, Linderman SL, Liu Q, Carter DM, Wrammert J, Esposito S, Principi N, Plotkin JB, Ross TM, Ahmed R, Wilson PC, Hensley SE. 2013. Immune history shapes specificity of pandemic H1N1 influenza antibody responses. J Exp Med 210:1493-500.

156. Meroz D, Yoon SW, Ducatez MF, Fabrizio TP, Webby RJ, Hertz T, Ben-Tal N. 2011. Putative amino acid determinants of the emergence of the 2009 influenza A (H1N1) virus in the human population. Proc Natl Acad Sci U S A 108:13522-7.

157. Koel BF, Burke DF, Bestebroer TM, van der Vliet S, Zondag GC, Vervaet G, Skepner E, Lewis NS, Spronken MI, Russell CA, Eropkin MY, Hurt AC, Barr IG, de Jong JC, Rimmelzwaan GF, Osterhaus AD, Fouchier RA, Smith DJ. 2013. Substitutions near the receptor binding site determine major antigenic change during influenza virus evolution. Science 342:976-9.

158. O'Donnell CD, Vogel L, Wright A, Das SR, Wrammert J, Li GM, McCausland M, Zheng NY, Yewdell JW, Ahmed R, Wilson PC, Subbarao K. 2012. Antibody pressure by a human monoclonal antibody targeting the 2009 pandemic H1N1 virus hemagglutinin drives the emergence of a virus with increased virulence in mice. MBio 3.

159. Krause JC, Tsibane T, Tumpey TM, Huffman CJ, Basler CF, Crowe JE, Jr. 2011. A broadly neutralizing human monoclonal antibody that recognizes a conserved, novel epitope on the globular head of the influenza H1N1 virus hemagglutinin. J Virol 85:10905-8.

160. Whittle JR, Zhang R, Khurana S, King LR, Manischewitz J, Golding H, Dormitzer PR, Haynes BF, Walter EB, Moody MA, Kepler TB, Liao HX, Harrison SC. 2011. Broadly neutralizing human antibody that recognizes the receptor-binding pocket of influenza virus hemagglutinin. Proc Natl Acad Sci U S A 108:14216-21.

161. Thompson WW, Shay DK, Weintraub E, Brammer L, Bridges CB, Cox NJ, Fukuda K. 2004. Influenza-associated hospitalizations in the United States. JAMA 292:1333-40.

162. Muennig PA, Khan K. 2001. Cost-effectiveness of vaccination versus treatment of influenza in healthy adolescents and adults. Clin Infect Dis 33:1879-85.

163. Lambert LC, Fauci AS. 2010. Influenza vaccines for the future. N Engl J Med 363:2036-44.

164. Milian E, Kamen AA. 2015. Current and emerging cell culture manufacturing technologies for influenza vaccines. Biomed Res Int 2015:504831.

165. Gambaryan AS, Marinina VP, Tuzikov AB, Bovin NV, Rudneva IA, Sinitsyn BV, Shilov AA, Matrosovich MN. 1998. Effects of host-dependent glycosylation of hemagglutinin on receptor-binding properties on H1N1 human influenza A virus grown in MDCK cells and in embryonated eggs. Virology 247:170-7.

166. Katz JM, Naeve CW, Webster RG. 1987. Host cell-mediated variation in H3N2 influenza viruses. Virology 156:386-95.

167. Meyer WJ, Wood JM, Major D, Robertson JS, Webster RG, Katz JM. 1993. Influence of host cell-mediated variation on the international surveillance of influenza A (H3N2) viruses. Virology 196:130-7.

168. Schild GC, Oxford JS, de Jong JC, Webster RG. 1983. Evidence for host-cell selection of influenza virus antigenic variants. Nature 303:706-9.

169. Davies R, Pepys J. 1976. Egg allergy, influenza vaccine, and immunoglobulin E antibody. J Allergy Clin Immunol 57:373-83.

170. Perdue ML, Arnold F, Li S, Donabedian A, Cioce V, Warf T, Huebner R. 2011. The future of cell culture-based influenza vaccine production. Expert Rev Vaccines 10:1183-94.

171. Chu C, Lugovtsev V, Golding H, Betenbaugh M, Shiloach J. 2009. Conversion of MDCK cell line to suspension culture by transfecting with human siat7e gene and its application for influenza virus production. Proc Natl Acad Sci U S A 106:14802-7.

172. Halperin SA, Smith B, Mabrouk T, Germain M, Trepanier P, Hassell T, Treanor J, Gauthier R, Mills EL. 2002. Safety and immunogenicity of a trivalent, inactivated, mammalian cell culture-derived influenza vaccine in healthy adults, seniors, and children. Vaccine 20:1240-7.

173. Kistner O, Barrett PN, Mundt W, Reiter M, Schober-Bendixen S, Dorner F. 1998. Development of a mammalian cell (Vero) derived candidate influenza virus vaccine. Vaccine 16:960-8.

174. Lau SC, Scholtissek C. 1995. Abortive infection of Vero cells by an influenza A virus (FPV). Virology 212:225-31.

175. Nakamura K, Homma M. 1981. Protein synthesis in Vero cells abortively infected with influenza B virus. J Gen Virol 56:199-202.

176. Rambhia KJ, Watson M, Sell TK, Waldhorn R, Toner E. 2010. Mass vaccination for the 2009 H1N1 pandemic: approaches, challenges, and recommendations. Biosecur Bioterror 8:321-30.

177. Cobbin JC, Verity EE, Gilbertson BP, Rockman SP, Brown LE. 2013. The source of the PB1 gene in influenza vaccine reassortants selectively alters the hemagglutinin content of the resulting seed virus. J Virol 87:5577-85.

178. Hoffmann E, Krauss S, Perez D, Webby R, Webster RG. 2002. Eight-plasmid system for rapid generation of influenza virus vaccines. Vaccine 20:3165-70.

179. Hu W, Zhang H, Han Q, Li L, Chen Y, Xia N, Chen Z, Shu Y, Xu K, Sun B. 2015. A Vero-cell-adapted vaccine donor strain of influenza A virus generated by serial passages. Vaccine 33:374-81.

180. Liu M, Liu CG, Zhang Y, Shi WL, Wang W, Liu YY. 2012. Efficacy of a high-yield attenuated vaccine strain wholly derived from avian influenza viruses by use of reverse genetics. Vet Microbiol 161:43-8.

181. Suphaphiphat P, Franti M, Hekele A, Lilja A, Spencer T, Settembre E, Palmer G, Crotta S, Tuccino AB, Keiner B, Trusheim H, Balabanis K, Sackal M, Rothfeder M, Mandl CW, Dormitzer PR, Mason PW. 2010. Mutations at positions 186 and 194 in the HA gene of the 2009 H1N1 pandemic influenza virus improve replication in cell culture and eggs. Virol J 7:157.

182. Hamamoto I, Takaku H, Tashiro M, Yamamoto N. 2013. High yield production of influenza virus in Madin Darby canine kidney (MDCK) cells with stable knockdown of IRF7. PLoS One 8:e59892.

183. Brandau DT, Jones LS, Wiethoff CM, Rexroad J, Middaugh CR. 2003. Thermal stability of vaccines. J Pharm Sci 92:218-31.

184. O'Donnell CD, Vogel L, Matsuoka Y, Jin H, Subbarao K. 2014. The matrix gene segment destabilizes the acid and thermal stability of the hemagglutinin of pandemic live attenuated influenza virus vaccines. J Virol 88:12374-84.

185. Rudneva IA, Timofeeva TA, Ignatieva AV, Shilov AA, Ilyushina NA. 2016. Effects of hemagglutinin amino acid substitutions in H9 influenza A virus escape mutants. Arch Virol doi:10.1007/s00705-016-3038-x.

186. Ye J, Wen F, Xu Y, Zhao N, Long L, Sun H, Yang J, Cooley J, Todd Pharr G, Webby R, Wan XF. 2015. Error-prone pcr-based mutagenesis strategy for rapidly generating high-yield influenza vaccine candidates. Virology 482:234-43.

187. Organization WH. 2011. Manual for the laboratory diagnosis and virological surveillance of influenza. Geneva : World Health Organization,

188. Li L, Liu Y, Ma C, Qu J, Calderon AD, Wu B, Wei N, Wang X, Guo Y, Xiao Z, Song J, Sugiarto G, Li Y, Yu H, Chen X, Wang PG. 2015. Efficient Chemoenzymatic Synthesis of an N-glycan Isomer Library. Chem Sci 6:5652-5661.

189. Wu Z, Liu Y, Ma C, Li L, Bai J, Byrd-Leotis L, Lasanajak Y, Guo Y, Wen L, Zhu H, Song J, Li Y, Steinhauer DA, Smith DF, Zhao B, Chen X, Guan W, Wang PG. 2016. Identification of the binding roles of terminal and internal glycan epitopes using enzymatically synthesized N-glycans containing tandem epitopes. Org Biomol Chem 14:11106-11116.

190. Xiong X, Martin SR, Haire LF, Wharton SA, Daniels RS, Bennett MS, McCauley JW, Collins PJ, Walker PA, Skehel JJ, Gamblin SJ. 2013. Receptor binding by an H7N9 influenza virus from humans. Nature 499:496-9.

191. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792-7.

192. Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series 41:95-98.

193. Skehel JJ, Wiley DC. 2000. Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. Annu Rev Biochem 69:531-69.

194. Chen Z, Baz M, Lu J, Paskel M, Santos C, Subbarao K, Jin H, Matsuoka Y. 2014. Development of a high-yield live attenuated H7N9 influenza virus vaccine that provides protection against homologous and heterologous H7 wild-type viruses in ferrets. J Virol 88:7016-23.

195. Yasugi M, Nakamura S, Daidoji T, Kawashita N, Ramadhany R, Yang CS, Yasunaga T, Iida T, Horii T, Ikuta K, Takahashi K, Nakaya T. 2012. Frequency of D222G and Q223R hemagglutinin mutants of pandemic (H1N1) 2009 influenza virus in Japan between 2009 and 2010. PLoS One 7:e30946.

196. Govorkova EA, Murti G, Meignier B, de Taisne C, Webster RG. 1996. African green monkey kidney (Vero) cells provide an alternative host cell system for influenza A and B viruses. J Virol 70:5519-24.

197. Ito T, Suzuki Y, Takada A, Kawamoto A, Otsuki K, Masuda H, Yamada M, Suzuki T, Kida H, Kawaoka Y. 1997. Differences in sialic acid-galactose linkages in the chicken egg amnion and allantois influence human influenza virus receptor specificity and variant selection. J Virol 71:3357-62.

198. Seo SH, Goloubeva O, Webby R, Webster RG. 2001. Characterization of a porcine lung epithelial cell line suitable for influenza virus studies. J Virol 75:9517-25.

199. Chen Z, Zhou H, Kim L, Jin H. 2012. The receptor binding specificity of the live attenuated influenza H2 and H6 vaccine viruses contributes to vaccine immunogenicity and protection in ferrets. J Virol 86:2780-6.

200. Wang M, Tscherne DM, McCullough C, Caffrey M, Garcia-Sastre A, Rong L. 2012. Residue Y161 of influenza virus hemagglutinin is involved in viral recognition of sialylated complexes from different hosts. J Virol 86:4455-62.

201. Pica N, Bouvier NM. 2012. Environmental factors affecting the transmission of respiratory viruses. Curr Opin Virol 2:90-5.

202. Schrauwen EJ, Fouchier RA. 2014. Host adaptation and transmission of influenza A viruses in mammals. Emerg Microbes Infect 3:e9.

203. Watanabe Y, Arai Y, Daidoji T, Kawashita N, Ibrahim MS, El-Gendy Eel D, Hiramatsu H, Kubota-Koketsu R, Takagi T, Murata T, Takahashi K, Okuno Y, Nakaya T, Suzuki Y, Ikuta K. 2015. Characterization of H5N1 influenza virus variants with hemagglutinin mutations isolated from patients. MBio 6.

204. Morens DM, Taubenberger JK. 2010. Historical thoughts on influenza viral ecosystems, or behold a pale horse, dead dogs, failing fowl, and sick swine. Influenza Other Respir Viruses 4:327-37.

205. Anthony SJ, St Leger JA, Pugliares K, Ip HS, Chan JM, Carpenter ZW, Navarrete-Macias I, Sanchez-Leon M, Saliki JT, Pedersen J, Karesh W, Daszak P, Rabadan R, Rowles T, Lipkin WI. 2012. Emergence of fatal avian influenza in New England harbor seals. MBio 3:e00166-12.

206. Crawford PC, Dubovi EJ, Castleman WL, Stephenson I, Gibbs EP, Chen L, Smith C, Hill RC, Ferro P, Pompey J, Bright RA, Medina MJ, Johnson CM, Olsen CW, Cox NJ, Klimov AI, Katz JM, Donis RO. 2005. Transmission of equine influenza virus to dogs. Science 310:482-5.

207. Webby RJ, Webster RG. 2001. Emergence of influenza A viruses. Philos Trans R Soc Lond B Biol Sci 356:1817-28.

208. Webster RG, Govorkova EA. 2014. Continuing challenges in influenza. Ann N Y Acad Sci 1323:115-39.

209. Ning ZY, Wu XT, Cheng YF, Qi WB, An YF, Wang H, Zhang GH, Li SJ. 2012. Tissue distribution of sialic acid-linked influenza virus receptors in beagle dogs. J Vet Sci 13:219-22.

210. Daly JM, Blunden AS, Macrae S, Miller J, Bowman SJ, Kolodziejek J, Nowotny N, Smith KC. 2008. Transmission of equine influenza virus to English foxhounds. Emerg Infect Dis 14:461-4.

211. Gibbs EP, Anderson TC. 2010. Equine and canine influenza: a review of current events. Anim Health Res Rev 11:43-51.

212. Ince WL, Gueye-Mbaye A, Bennink JR, Yewdell JW. 2013. Reassortment complements spontaneous mutation in influenza A virus NP and M1 genes to accelerate adaptation to a new host. J Virol 87:4330-8.

213. Manz B, Schwemmle M, Brunotte L. 2013. Adaptation of avian influenza A virus polymerase in mammals to overcome the host species barrier. J Virol 87:7200-9.

214. Feng KH, Gonzalez G, Deng L, Yu H, Tse VL, Huang L, Huang K, Wasik BR, Zhou B, Wentworth DE, Holmes EC, Chen X, Varki A, Murcia PR, Parrish CR. 2015. Equine and Canine Influenza H3N8 Viruses Show Minimal Biological Differences Despite Phylogenetic Divergence. J Virol 89:6860-73.

215. Rivailler P, Perry IA, Jang Y, Davis CT, Chen LM, Dubovi EJ, Donis RO. 2010. Evolution of canine and equine influenza (H3N8) viruses co-circulating between 2005 and 2008. Virology 408:71-9.

216. Payungporn S, Crawford PC, Kouo TS, Chen LM, Pompey J, Castleman WL, Dubovi EJ, Katz JM, Donis RO. 2008. Influenza A virus (H3N8) in dogs with respiratory disease, Florida. Emerg Infect Dis 14:902-8.

217. Collins PJ, Vachieri SG, Haire LF, Ogrodowicz RW, Martin SR, Walker PA, Xiong X, Gamblin SJ, Skehel JJ. 2014. Recent evolution of equine influenza and the origin of canine influenza. Proc Natl Acad Sci U S A 111:11175-80.

218. Hoffmann E, Stech J, Guan Y, Webster RG, Perez DR. 2001. Universal primer set for the full-length amplification of all influenza A viruses. Arch Virol 146:2275-89.

219. Hoffmann E, Neumann G, Kawaoka Y, Hobom G, Webster RG. 2000. A DNA transfection system for generation of influenza A virus from eight plasmids. Proc Natl Acad Sci U S A 97:6108-13.

220. Bradley KC, Jones CA, Tompkins SM, Tripp RA, Russell RJ, Gramer MR, Heimburg-Molinaro J, Smith DF, Cummings RD, Steinhauer DA. 2011. Comparison of the receptor binding properties of contemporary swine isolates and early human pandemic H1N1 isolates (Novel 2009 H1N1). Virology 413:169-82.

221. Wen F, Li L, Zhao N, Chiang MJ, Xie H, Cooley J, Webby R, Wang PG, Wan XF. 2017. A Y161F Hemagglutinin Substitution Increases Thermostability and Improves Yields of 2009 H1N1 influenza A Virus in Cells. J Virol doi:10.1128/JVI.01621-17.

222. de Graaf M, Fouchier RA. 2014. Role of receptor binding specificity in influenza A virus transmission and pathogenesis. EMBO J 33:823-41.

223. Yamanaka T, Nemoto M, Tsujimura K, Kondo T, Matsumura T. 2009. Interspecies transmission of equine influenza virus (H3N8) to dogs by close contact with experimentally infected horses. Vet Microbiol 139:351-5.

224. Yamanaka T, Nemoto M, Bannai H, Tsujimura K, Kondo T, Matsumura T, Muranaka M, Ueno T, Kinoshita Y, Niwa H, Hidari KI, Suzuki T. 2012. No evidence of horizontal infection in horses kept in close contact with dogs experimentally infected with canine influenza A virus (H3N8). Acta Vet Scand 54:25.

225. Quintana AM, Hussey SB, Burr EC, Pecoraro HL, Annis KM, Rao S, Landolt GA. 2011. Evaluation of infectivity of a canine lineage H3N8 influenza A virus in ponies and in primary equine respiratory epithelial cells. Am J Vet Res 72:1071-8.

226. von Grotthuss M, Rychlewski L. 2006. Influenza mutation from equine to canine. Science 311:1241-2; author reply 1241-2.

227. Hoelzer K, Murcia PR, Baillie GJ, Wood JL, Metzger SM, Osterrieder N, Dubovi EJ, Holmes EC, Parrish CR. 2010. Intrahost evolutionary dynamics of canine influenza virus in naive and partially immune dogs. J Virol 84:5329-35.

228. Gamblin SJ, Haire LF, Russell RJ, Stevens DJ, Xiao B, Ha Y, Vasisht N, Steinhauer DA, Daniels RS, Elliot A, Wiley DC, Skehel JJ. 2004. The structure and receptor binding properties of the 1918 influenza hemagglutinin. Science 303:1838-42.

229. Nycholat CM, McBride R, Ekiert DC, Xu R, Rangarajan J, Peng W, Razi N, Gilbert M, Wakarchuk W, Wilson IA, Paulson JC. 2012. Recognition of sialylated poly-N-acetyllactosamine chains on N- and O-linked glycans by human and avian influenza A virus hemagglutinins. Angew Chem Int Ed Engl 51:4860-3.

230. 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.

231. van Riel D, Munster VJ, de Wit E, Rimmelzwaan GF, Fouchier RA, Osterhaus AD, Kuiken T. 2007. Human and avian influenza viruses target different cells in the lower respiratory tract of humans and other mammals. Am J Pathol 171:1215-23.

232. Stehle T, Khan ZM. 2014. Rules and exceptions: sialic acid variants and their role in determining viral tropism. J Virol 88:7696-9.

233. Song D, Kang B, Lee C, Jung K, Ha G, Kang D, Park S, Park B, Oh J. 2008. Transmission of avian influenza virus (H3N2) to dogs. Emerg Infect Dis 14:741-6.

234. Wang H, Jia K, Qi W, Zhang M, Sun L, Liang H, Du G, Tan L, Shao Z, Ye J, Sun L, Cao Z, Chen Y, Zhou P, Su S, Li S. 2013. Genetic characterization of avian-origin H3N2 canine influenza viruses isolated from Guangdong during 2006-2012. Virus Genes 46:558-62.

235. Voorhees IEH, Glaser AL, Toohey-Kurth K, Newbury S, Dalziel BD, Dubovi EJ, Poulsen K, Leutenegger C, Willgert KJE, Brisbane-Cohen L, Richardson-Lopez J, Holmes EC, Parrish CR. 2017. Spread of Canine Influenza A(H3N2) Virus, United States. Emerg Infect Dis 23.

236. Pulit-Penaloza JA, Simpson N, Yang H, Creager HM, Jones J, Carney P, Belser JA, Yang G, Chang J, Zeng H, Thor S, Jang Y, Killian ML, Jenkins-Moore M, Janas-Martindale A, Dubovi E, Wentworth DE, Stevens J, Tumpey TM, Davis CT, Maines TR. 2017. Assessment of Molecular, Antigenic, and Pathological Features of Canine Influenza A(H3N2) Viruses That Emerged in the United States. J Infect Dis 216:S499-S507.

237. Matrosovich M, Tuzikov A, Bovin N, Gambaryan A, Klimov A, Castrucci MR, Donatelli I, Kawaoka Y. 2000. Early alterations of the receptor-binding properties of H1, H2, and H3 avian influenza virus hemagglutinins after their introduction into mammals. J Virol 74:8502-12.

238. Belser JA, Jayaraman A, Raman R, Pappas C, Zeng H, Cox NJ, Katz JM, Sasisekharan R, Tumpey TM. 2011. Effect of D222G mutation in the hemagglutinin protein on receptor binding, pathogenesis and transmissibility of the 2009 pandemic H1N1 influenza virus. PLoS One 6:e25091.

239. Hensley SE, Das SR, Bailey AL, Schmidt LM, Hickman HD, Jayaraman A, Viswanathan K, Raman R, Sasisekharan R, Bennink JR, Yewdell JW. 2009. Hemagglutinin receptor binding avidity drives influenza A virus antigenic drift. Science 326:734-6.

240. Li Y, Bostick DL, Sullivan CB, Myers JL, Griesemer SB, Stgeorge K, Plotkin JB, Hensley SE. 2013. Single hemagglutinin mutations that alter both antigenicity and receptor binding avidity influence influenza virus antigenic clustering. J Virol 87:9904-10.

241. Lewis NS, Anderson TK, Kitikoon P, Skepner E, Burke DF, Vincent AL. 2014. Substitutions near the hemagglutinin receptor-binding site determine the antigenic evolution of influenza A H3N2 viruses in U.S. swine. J Virol 88:4752-63.

242. Harper SA, Fukuda K, Uyeki TM, Cox NJ, Bridges CB. 2004. Prevention and control of influenza: recommendations of the Advisory Committee on Immunization Practices (ACIP). MMWR Recomm Rep 53:1-40.

243. Cai Z, Ducatez MF, Yang J, Zhang T, Long LP, Boon AC, Webby RJ, Wan XF. 2012. Identifying antigenicity-associated sites in highly pathogenic H5N1 influenza virus hemagglutinin by using sparse learning. J Mol Biol 422:145-55.

244. Yang J, Zhang T, Wan XF. 2014. Sequence-based antigenic change prediction by a sparse learning method incorporating co-evolutionary information. PLoS One 9:e106660.

245. Han L, Zhang Y, Wan XF, Zhang T. 2016. Generalized Hierarchical Sparse Model for Arbitrary-Order Interactive Antigenic Sites Identification in Flu Virus Data. KDD 2016:865-874.

246. Han L, Li L, Wen F, Zhong L, Zhang T, Wan XF. 2018. Graph-Guided Multi-Task Sparse Learning Model: a Method for Identifying Antigenic Variants of Influenza A(H3N2) Virus. Bioinformatics doi:10.1093/bioinformatics/bty457.

247. Xu R, McBride R, Nycholat CM, Paulson JC, Wilson IA. 2012. Structural characterization of the hemagglutinin receptor specificity from the 2009 H1N1 influenza pandemic. J Virol 86:982-90.

248. Matrosovich MN, Gambaryan AS, Teneberg S, Piskarev VE, Yamnikova SS, Lvov DK, Robertson JS, Karlsson KA. 1997. Avian influenza A viruses differ from human viruses by recognition of sialyloligosaccharides and gangliosides and by a higher conservation of the HA receptor-binding site. Virology 233:224-34.

249. Popova L, Smith K, West AH, Wilson PC, James JA, Thompson LF, Air GM. 2012. Immunodominance of antigenic site B over site A of hemagglutinin of recent H3N2 influenza viruses. PLoS One 7:e41895.

250. Wang X, Ilyushina NA, Lugovtsev VY, Bovin NV, Couzens LK, Gao J, Donnelly RP, Eichelberger MC, Wan H. 2017. Amino Acids in Hemagglutinin Antigenic Site B Determine Antigenic and Receptor Binding Differences between A(H3N2)v and Ancestral Seasonal H3N2 Influenza Viruses. J Virol 91.

251. Lee MS, Chen JS. 2004. Predicting antigenic variants of influenza A/H3N2 viruses. Emerg Infect Dis 10:1385-90.

252. He L, Jiang K, Wu Q, Duan Z, Xu H, Liu J, Cui Z, Gu M, Wang X, Liu X, Liu X. 2014. Two amino acid substitutions in the haemagglutinin of the 2009 pandemic H1N1 virus decrease direct-contact transmission in guinea pigs. J Gen Virol 95:2612-7.

253. Linderman SL, Chambers BS, Zost SJ, Parkhouse K, Li Y, Herrmann C, Ellebedy AH, Carter DM, Andrews SF, Zheng NY, Huang M, Huang Y, Strauss D, Shaz BH, Hodinka RL, Reyes-Teran G, Ross TM, Wilson PC, Ahmed R, Bloom JD, Hensley SE. 2014. Potential antigenic explanation for atypical H1N1 infections among middle-aged adults during the 2013-2014 influenza season. Proc Natl Acad Sci U S A 111:15798-803.

254. Xu R, Ekiert DC, Krause JC, Hai R, Crowe JE, Jr., Wilson IA. 2010. Structural basis of preexisting immunity to the 2009 H1N1 pandemic influenza virus. Science 328:357-60.

255. Wormann X, Lesch M, Welke RW, Okonechnikov K, Abdurishid M, Sieben C, Geissner A, Brinkmann V, Kastner M, Karner A, Zhu R, Hinterdorfer P, Anish C, Seeberger PH, Herrmann A, Meyer TF, Karlas A. 2016. Genetic characterization of an adapted pandemic 2009 H1N1 influenza virus that reveals improved replication rates in human lung epithelial cells. Virology 492:118-29.

256. Zhang W, Shi Y, Qi J, Gao F, Li Q, Fan Z, Yan J, Gao GF. 2013. Molecular basis of the receptor binding specificity switch of the hemagglutinins from both the 1918 and 2009 pandemic influenza A viruses by a D225G substitution. J Virol 87:5949-58.