J. Biomedical Science and Engineering, 2010, 3, 227-240 **JBiSE**
doi:10.4236/jbise.2010.33031 Published Online March 2010 (http://www.SciRP.org/journal/jbise/).

Quantifying the effects of mutations on receptor binding specificity of influenza viruses

Wei Hu

Department of Computer Science, Houghton College, Houghton, USA.
Email: wei.hu@houghton.edu

Received 7 January 2010; revised 11 January 2010; accepted 15 January 2010.

ABSTRACT

Hemagglutinin (HA) of influenza viruses is a cylindrically shaped homotrimer, where each monomer comprises two disulfide-linked subdomains HA1 and HA2. Influenza infection is initiated by binding of HA1 to its host cell receptors and followed by the fusion between viral and host endosomal membranes mediated by HA2. Human influenza viruses preferentially bind to sialic acid that is linked to galactose by an α2,6-linkage (α2,6), whereas avian and swine influenza viruses preferentially recognize α2,3 or α 2,3/α2,6. For animal influenza viruses to cross host species barriers, their HA proteins must acquire mutations to gain the capacity to allow human-to-human transmission. In this study, the informational spectrum method (ISM), a bioinformatics approach, was applied to identify mutations and to elucidate the contribution to the receptor binding specificity from each mutation in HA1 in various subtypes within or between hosts, including 2009 human H1N1, avian H5N1, human H5N1, avian H1N1, and swine H1N2. Among others, our quantitative analysis indicated that the mutations in HA1 of 2009 human H1N1 collectively tended to reduce the swine binding affinity in the seasonal H1N1 strains and to increase that in the pandemic H1N1 strains. At the same time, they increased the human binding affinity in the pandemic H1N1 strains and had little impact on that in the seasonal H1N1 strains. The mutations between the consensus HA1 sequences of human H5N1 and avian H5N1 increased the avian binding affinity and decreased the human binding affinity in avian H5N1 while produced the opposite effects on those in human H5N1. Finally, the ISM was employed to analyze and verify several mutations in HA1 well known for their critical roles in binding specificity switch, including E190D/G225D in H1N1 and Q192R/ S223L/ Q226L/ G228S in H5N1.

Keywords: Binding Specificity; Discrete Fourier Transform; Electron-Ion Interaction Potential; Entropy; Hemagglutinin; Influenza; Informational Spectrum Method; Mutation; Receptor

1. INTRODUCTION

Influenza A viruses are classified into different subtypes based on the viral surface proteins hemagglutinin (HA) and neuraminidase (NA). The initial step in the influenza infection is the binding of HA to sialylated glycan receptors on the host cells. HA is also the primary target for the immune response in the infected host. Human and swine influenza viruses are derived from avian viruses, facilitated by regular close contact among humans, birds, and pigs [1]. The past three influenza pandemics, the Spanish flu (H1N1) in 1918, the Asian flu (H2N2) in 1957, and the Hong Kong flu (H3N2) in 1968, all had arisen from a reassortment from avian, swine, and human viruses. The current 2009 influenza pandemic was caused by a swine-origin H1N1 virus. The adaptation of the virus to a new host entails the compatibility between host and virus genetic requirements to allow efficient replication and sustained transmission. The host barrier for influenza viruses to transmit in humans is multigenic, however, the receptor specificity of HA proteins is a key determinant.

The binding preference of influenza virus HAs affects the host specificity for infection. In general, human influenza viruses bind preferentially to α2,6 receptors, avian influenza viruses tend to bind to α2,3 receptors, and swine influenza viruses can bind to either α2,6 receptors or both α2,6 and α2,3 receptors, primarily based on differences in the amino acids in the HA receptor binding domain (RBD). The RBD of HA in various influenza subtypes has three structural elements in common, one α-helix (190-helix) and two loops (130-loop and 220-loop). Different hosts express diverse SA isomers, *i.e.*, α2,3 linkages in the gut of waterfowl, α2,3 and α2,6 linkages in the lung and intestinal epithelium of chickens, and α2,6 linkages in the upper respiratory epithelium and α2,3 and α2,6 linkages in the lower respiratory epithelium of humans [2].

A change of binding preference is essential for cross species transfer, which involves mutations in HA to alter its glycan receptor preference [3]. It is hypothesized that

to facilitate efficient human transmission, mutations in HA are required to increase α2,5 binding and at the same time to decrease α2,3 binding [4]. A variety of mutations that can shift receptor preference of HA proteins have been identified.

In a study of receptor specificity of influenza A/H5 viruses [5], all but two isolates exhibited high affinity to α2,3 receptors. The two isolates with a unique S223N change in HA demonstrated decreased affinity to α2,3 and moderate affinity to α2,6 receptors. Another study showed that introduction of the mutation Q192R enhanced the binding of HA in H5N1 to α2,6 receptors, and introduction of both mutations Q192R and S223N increased the binding preference significantly. Residue 192 is close to the 190 helix, and residue 223 is part of the 220 loop, where it is feasible for them to influence the binding affinity [6]. As for the H2, H3, H4, and H9 HAs, two substitutions Q226L and G228S are mainly responsible for the switch from avian to human binding [3,7,8,9].

Residues 138, 186, 190, 194, 225, 226, and 228 can modulate the binding affinity of H1N1 HA proteins, and two residues 190 and 225 play key dominant roles in binding affinity [10,11]. The sequences of the RBD of avian H1 viruses maintain a Glu at position 190 and a Gly at 225 (H3 numbering), while the human H1 viruses generally have an Asp at both positions 190 and 225. It is known that E190D and G225D mutations in H1 viruses can shift binding patterns from avian to human type. In the five 1918 H1N1 HA sequences, three have a D190 and a D225 with α2,6 affinity, and two have a D190 and a G225 with mixed α2,6/α2,3 specificities [12]. In general, mutations D190/D225 favor α2,6 receptors in humans, D190/G225 like α2,6 and α2,3 receptors in swine, and E190/G225 prefer α2,3 receptors in avian [13]. The biochemical analysis in [14] quantified the multivalent HA-glycan interactions, and showed the effects of these mutations on glycan binding amplified by multivalency.

To date, the symptoms of 2009 H1N1 are mild. The fear is that the virus may continue to mutate to bring about another more lethal outbreak in the subsequent months as the 1918 Spanish flu. In [15] two representative 2009 H1N1 HA sequences, A/California/4/2009 and A/Hamburg/5/2009, were shown to bind to both α2,6 and α2,3 receptors with some minor differences in a carbohydrate microarray analysis, as predicted in [16]. There were three amino acid mutations between these HA sequences: S83P, A197T, and V321I, which might account for these differences in binding. These findings suggested that no major change in binding affinity is necessary for pandemic virus to acquire human binding patterns, and the dual binding to α2,6/α2,3 receptors is one contributor to the greater virulence of the pandemic virus than seasonal flu virus.

There were two other recent reports on the mutations of the 2009 H1N1 virus. The first report [17] located the potential mutations and strongly co-mutated positions in NA. The second report [18] focused on HA and the interaction between HA and NA. The mutations of HA in 2009 H1N1 were found and mapped to the 3D homology model of H1, and the mutations on the five epitope regions on H1 were identified. With help from the results of the first study, two co-mutation networks were uncovered, one in HA and one in NA, where each mutation in one network co-mutates with the mutations in the other network across the two proteins HA and NA. These two networks residing in HA and NA separately may provide a functional linkage between the mutations that can change the drug binding sites in NA and those that can affect the host immune response or vaccine efficacy in HA.

In references [19,20] the informational spectrum method (ISM) [21] was applied to investigate the interaction between HA and its receptors, which showed that HA1 of different flu subtypes encodes one highly conserved domain that might be determinants of HA binding affinity. The study in [22] extended the results in references [19,20] by identifying multiple domains in HA1 associated with each receptor interaction pattern. These conserved domains in HA1 might be used to identify new therapeutic targets for drug development.

In references [19,20] it was found that the consensus informational spectrum (CIS) of HA1 of influenza strains have the following characteristic dominant peaks at different IS frequencies as presented in Table 1. In this study, F(0.295) will be referred to as pandemic human H1N1 receptor interaction frequency, F(0.055) as swine receptor interaction frequency, F(0.076) as avian receptor interaction frequency, and F(0.236) as seasonal human H1N1 receptor interaction frequency. In addition to the dominant peak at IS frequencies in each subtype, there are secondary peaks at various IS frequencies [19,20,22].

Viral evolution can help influenza viruses surmount species barriers. Once adapted in a new host, they still need to continue their evolution to fit better in the new environment. In this study, we sought to investigate the effects of mutations in HA1, either within or between hosts, on binding preference shift through a quantitative analysis, the ISM. The analysis performed in this study was based on the observation that several influenza viruses display dual specific recognition of receptors with α2,6 or α2,3 linkages. Our goal was to utilize the ISM to uncover the amino acid polymorphisms in HA1 within or between hosts and to measure their contribution to the binding specificity switch quantitatively.

Table 1. Characteristic IS frequencies of HA proteins in 2009 H1N1, swine H1N1/H1N2, avian H5N1, and seasonal human H1N1.

Subtype	2009 H1N1	Swine H1N2/H1N1	Avian H5N1	Seasonal human H1N1
Frequency	F(0.295)	F(0.055)	F(0.076)	F(0.236)

2. MATERIALS AND METHODS

2.1. Sequence Data

All HA sequences were retrieved from the Influenza Virus Resource (http://www.ncbi/nlm.nih.giv/genomes/FLU/FLU.html) of the National Center for Biotechnology Information (NCBI) on November 20, 2009. Only the full length and unique sequences were selected. There were 450 HA sequences of human 2009 H1N1, 201 HA sequences of human H5N1 from 1979 to 2009, 1228 HA sequences of avian H5N1 from 1959 to 2009, 78 HA sequences of avian H1N1 from 1976 to 2008, and 83 HA sequences of swine H1N2 from 1980 to 2009. All the sequences used in the study were aligned with MAFFT [23].

2.2. Entropy

In information theory [24], entropy is a measure of disorder or randomness associated with a random variable. Let x be a discrete random variable that has a set of possible values $\{a_1, a_2, a_2, ...a_n\}$ with probabilities $\{p_1, p_2, p_3, ...p_n\}$ where $P(x = a_i) = p_i$. The entropy H of x is

$$H(x) = -\sum_i p_i \log p_i$$

In the current study, each of the n columns in a multiple sequence alignment of a set of HA sequences of N residues is considered as a discrete random variable x_i ($1 \le i \le N$) that takes on one of the 20 (n=20) amino acid types with some probability. $H(x_i)$ has its minimum value 0 if all the residues at position *i* are the same, and achieves its maximum if all the 20 amino acid types appear with equal probability at position *i*, which can be verified by the Lagrange multiplier technique. A position of high entropy means that the amino acids are often varied at this position. $H(x_i)$ measures the genetic diversity at position *i* in our current study. A brief overview of the extensive applications of entropy in sequence analysis, in particular the flu virus sequences, can be found in [17].

2.3. Important Sites in HA

Although there is a great variation due to high selection pressure in the HA1 sequences of various flu subtypes, the active site of HA1 is well conserved, which is located in a cleft composed of the residues 91, 150, 152, 180, 187, 191, and 192. The three amino acids at positions 187, 191 and 192 are a part of the 190 helix. The active site cleft of HA is formed by its right edge (131_GVTAA) and left edge (221_RGQAGR) (H1 numbering), which are also commonly referred to as the 130 loop and 220 loop, respectively [25,26].

3. RESULTS

3.1. Mutations within Hosts

3.1.1 2009 Human H1N1

After visual inspection of the alignment of 2009 human H1N1 HA1 sequences, there was either an Asp (single letter code D) at position 127 in the pandemic strains or a deletion at position 127 in the seasonal strains. Since Asp had the highest EIIP value of 0.128 [22], this deletion at position 127 might influence the DFT spectral distribution. Of the 450 HA1 sequences of 2009 human H1N1 collected, there were 345 pandemic H1N1 sequences with an Asp at position 127 and 105 seasonal H1N1 sequences with a deletion at position 127.

The CIS of the pandemic H1N1 HA1 sequences were plotted in **Figure 1(a)**, which have a dominant peak at frequency F(0.295) (pandemic human H1N1 binding), and the CIS of the seasonal H1N1 HA1 sequences were plotted in **Figure 1(b)**, which have a dominant peak at frequency F(0.055) (swine binding). According to the ISM concepts, this demonstrated the different receptor binding patterns of the 2009 pandemic and seasonal H1N1 strains. **Figure 2** illustrated the consensus sequences of the pandemic H1N1 strains (human binding preference) and seasonal H1N1 strains (swine binding preference), respectively.

There were 90 amino acid substitutions between the two consensus HA1 sequences of human and swine binding characteristics (**Figure 2** and **Table 2**). Based on

Table 2. Mutations between swine and human bindings in HA1 sequences of 2009 H1N1.

I3L	N35D	S36K	L43K	K45R	I47V	Q51H	N54K	S56N	V57I	L69S	I71S	S72T	K73A	E74S	K82T	P83S	N84S
P85S	E86D	H94D	A96I	E120T	-127D	S128T	V129N	T130K	S133T	S135A	S137P	N139A	E141A	S142K	R146K	L149I	T152V
G153K	N155G	G156N	L157S	N160K	A166I	N167D	E170G	V179I	P183S	N184T	I185S	V186A	K189Q	T190S	H193Q	T194N	E195A
N196D	S200F	V202G	H205R	R208K	T211K	K216I	I227M	L234V	T239K	I241T	N245T	I249V	A250V	L257M	S258E	G260N	F261A
N267I	N269D	A270T	M272V	D273H	K274D	D276N	A277T	K278T	Q283K	V295I	V298I	E302K	R308K	A310T	M314L	V315A	I321V

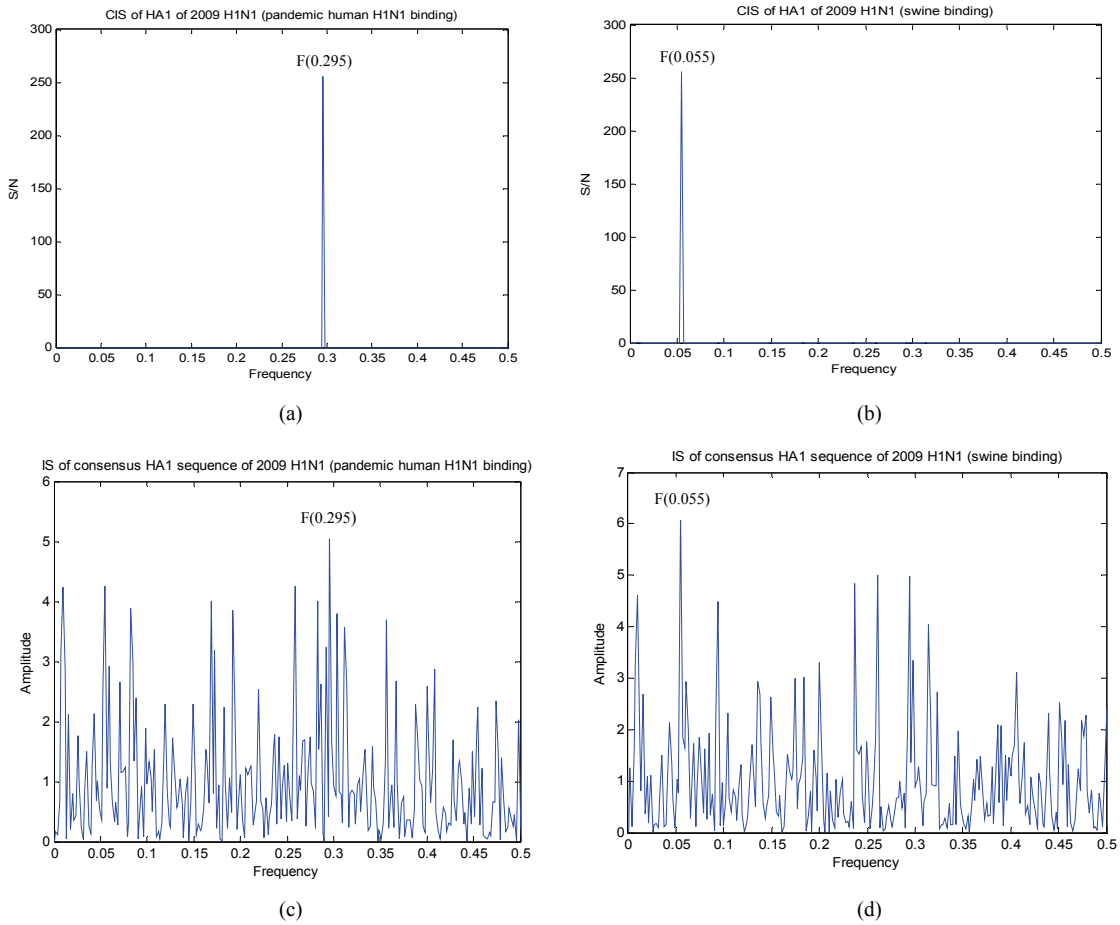


Figure 1. (a) CIS of HA1 of 2009 pandemic H1N1 (pandemic human H1N1 binding); (b) CIS of HA1 of 2009 seasonal H1N1 (swine H1N1 binding); (c) IS of HA1 of 2009 pandemic H1N1 (pandemic human H1N1 binding); (d) IS of HA1 of 2009 seasonal H1N1 (swine H1N1 binding).

the ISM theory, the mutations in HA1 that increased the amplitude of F(0.295) and decreased that of F(0.055) would contribute the switch of receptor binding affinity from swine to human type. The variation amount of the amplitudes of F(0.295) and F(0.055) was calculated for each of the 90 mutations applied to each consensus HA1 sequence, swine binding or human binding. The top 32 mutations that resulted in the amplitude change at frequency F(0.295) or F(0.055) (ΔA) more than 6% were listed in **Table 3**, suggesting that these mutations might be critical for modulating the binding preferences between swine and humans. In general, increasing the amplitude at one frequency F(0.295) or F(0.055) will decrease that at another frequency, but there were several exceptions. Three "hot spots", D94, D196, and D274, found in [19] contributed to the amplitudes at frequencies F(0.295) and F(0.055) with different amounts (**Table 3**). There were a mutation T152V at the binding site, and two mutations S133T and S135A at the right edge of the binding pocket, which were not listed in **Table 3** because their ΔA value was relatively small. **Table 3** also contained several mutations of interest, which were T130K and S137P near the right edge of the binding pocket, and P183S, N184T, I185S, H193Q, T194N, and N196D near the active site.

In [18], three networks of co-mutations in HA of 2009 H1N1 were uncovered. The first one had residues 269, 276, and 309, the second one had residues 34, 167, 195, and 268, and the third one had residues 129, 210, and 238, where each residue co-mutated with others in the same network. Two pairs of mutations N167D/E195A and N269D/D276N in **Table 2** were part of the aforementioned co-mutation networks discovered in [18]. Their individual and combined effects on ΔA were listed in **Table 4**, with the second pair having a much larger impact on the binding preference than the first. There were two clusters of mutations in **Table 2**, where the first was located at positions from 152 to 170, and the

second at positions from 257 to 278. The first cluster of mutations was contained in a pandemic human H1N1 receptor recognition domain (150:174) with the characteristic IS frequency at F(0.295) found in [22]. Prompted by this finding, we searched for a similar domain near the second cluster, and found a new domain (246:286) of swine binding characteristic with the IS dominant peak at frequency F(0.055) (**Figure 3**).

```
Swine Binding   DTICIGYHANNSTDTVDTVLEKNVTVTHSVNLLENSHNGKLCLLKGIAPLQLGNCSVAGW
Human Binding   DTLCIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDKHNGKLCKLRGVAPLHLGKCNIAGW
                **:*********************************:.****** *:*:***:**:*.:***

Swine Binding   ILGNPECELLISKESWSYIVEKPNPENGTCYPGHFADYEELREQLSSVSSFERFEIFPKE
Human Binding   ILGNPECESLSTASSWSYIVETSSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKT
                ******** * : .*******.***:*******.* **********************

                           Right edge
Swine Binding   SSWPNH-TVTGVSASCSHNGESSFYRNLLWLTGKNGLYPNLSKSYANNKEKEVLVLWGVH
Human Binding   SSWPNHDSNKGVTAACPHAGAKSFYKNLIWLVKKGNSYPKLSKSYINDKGKEVLVLWGIH
                ****** : .**:*:*.* * .***:**:**. *.. **:***** *:* ********:*

                                                          Left edge
Swine Binding   HPPNIVDQKTLYHTENAYVSVVSSHYSRKFTPEIAKRPKVRDQEGRINYYWTLLEPGDTI
Human Binding   HPSTSADQQSLYQNADAYVFVGSSRYSKKFKPEIAIRPKVRDQEGRMNYYWTLVEPGDKI
                **.. .**::**:. :*** * **:**:**.**** **********:******:****.*

Swine Binding   IFEANGNLIAPRYAFALSRGFGSGIINSNAPMDKCDAKCQTPQGAINSSLPFQNVHPVTI
Human Binding   TFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCNTTCQTPKGAINTSLPFQNIHPITI
                ***.***:.******:.*. ***** *::*:..*::.****:****:******:**:**

Swine Binding   GECPKYVRSAKLRMVTGLRNIPSIQSR
Human Binding   GKCPKYVKSTKLRLATGLRNVPSIQSR
                *:*****:*:***:.*****:******
```

Figure 2. Alignment of two consensus HA1 sequences of 2009 pandemic H1N1 (pandemic human H1N1 binding) and 2009 seasonal H1N1 (swine binding). The binding sites in HA are colored in red, the left and right edges of the binding cleft in blue.

Table 3. Changes of amplitudes of IS frequencies by top 32 mutations with large ΔA value in HA1 of 2009 H1N1.

	Mutating Consensus HA1 Sequence of Swine Binding Patterns		Mutating Consensus HA1 Sequence of Human Binding Patterns	
Mutations	ΔA[F(0.055)]%	ΔA[F(0.295)]%	ΔA[F(0.055)]%	ΔA[F(0.295)]%
N35D	11.025	10.814	-12.123	-12.608
K45R	-6.935	2.2999	8.1835	-6.5807
S56N	-8.347	-9.5275	10.448	0.093211
L69S	1.1681	6.0652	-0.70212	-8.5586
I71S	7.0299	-9.42	-7.5854	4.8054
E74S	8.719	-3.4947	-9.8488	9.1621
K82T	-6.9799	-4.8104	8.282	-5.4234
P83S	-7.2536	7.4623	8.6929	-1.423
N84S	-7.7848	0.88042	9.617	10.047
H94D	6.739	5.3039	-7.8366	14.058
E120T	-9.7084	2.2202	11.954	-6.9883
T130K	-5.6997	7.272	5.909	5.7416
S137P	7.0562	8.6072	-8.3647	4.9132
T152V	0.49754	-6.8741	-3.6229	-9.9879
L157S	-9.3159	-5.0672	10.114	-9.6759
P183S	6.6852	7.4566	-8.2446	2.4515
N184T	10.495	0.64815	-11.043	-10.482
I185S	9.2299	-10.327	-8.6728	3.045
H193Q	-6.1179	4.6131	6.9435	3.7993
T194N	10.605	-4.89	-9.7841	10.506
N196D	-8.4752	-0.04744	6.2191	16.521
N245T	-5.3848	8.7494	9.7128	-10.596
L257M	9.7014	-7.2577	-10.573	10.05
S258E	-8.2999	-8.3169	8.7374	4.3745
N269D	-9.1545	8.9086	6.7895	-12.176
M272V	-3.1046	-9.3909	7.0087	9.0809
D273H	-6.1335	7.448	11.253	-0.73203
K274D	7.7964	-10.052	-10.303	12.443
D276N	-12.437	-3.5437	13.913	9.3104
A277T	5.1287	-7.8075	-4.4654	6.8946
R308K	-0.59247	7.761	3.5286	-6.9255
M314L	-7.2075	5.2935	6.3119	-5.2079
Total	-27.0553	0.9764	40.4485	29.9315

Table 4. Changes of amplitudes of IS frequencies by mutations contained in the co-mutation networks in 2009 H1N1 discovered in [18].

	Mutating Consensus HA1 Sequence of Swine Binding Patterns		Mutating Consensus HA1 Sequence of Human Binding Patterns	
Mutations	ΔA[F(0.055)]%	ΔA[F(0.295)]%	ΔA[F(0.055)]%	ΔA[F(0.295)]%
N167D	0	0	0	0
E195A	-2.674	-3.7307	2.1705	-0.36167
N167D, E195A	-2.6740	-3.7307	2.1705	-0.3617
N269D	-9.1545	8.9086	6.7895	-12.176
D276N	-12.437	-3.5437	13.913	9.3104
N269D, D276N	-20.8003	4.3377	21.7617	-3.6715

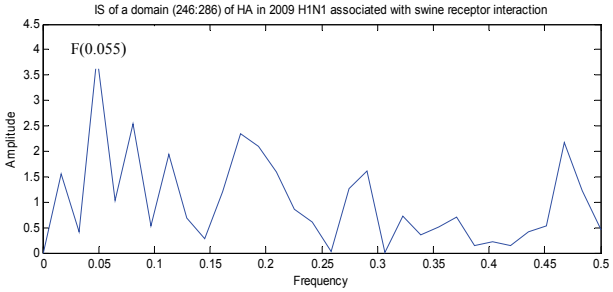
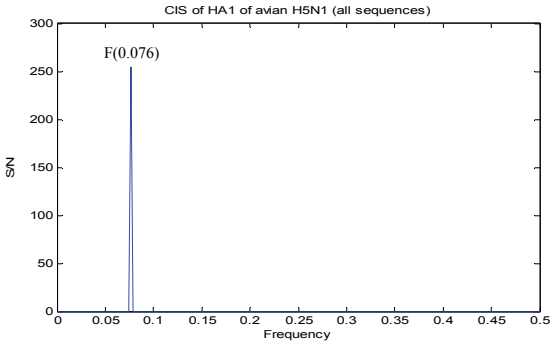


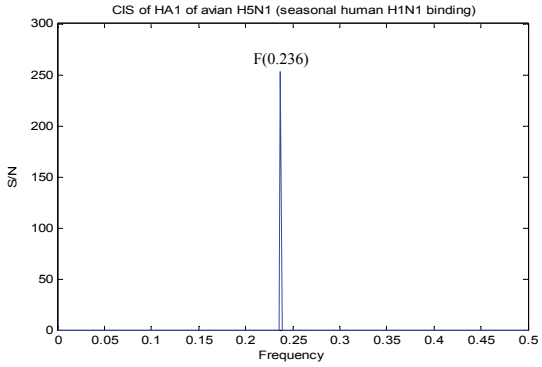
Figure 3. IS of one domain (246:286) of swine binding characteristic in HA1 of 2009 H1N1.

3.1.2. Avian H5N1

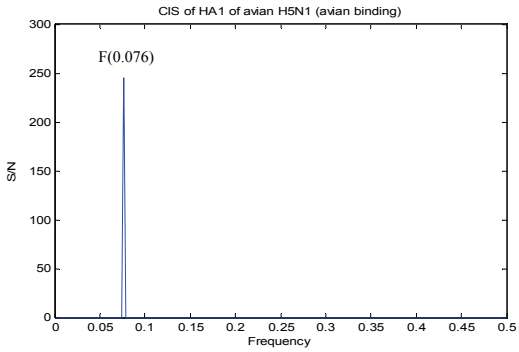
Although the whole set of HA1 sequences in avian H5N1 (n=1228) displayed the CIS dominant peak at frequency F(0.076) (**Figure 4(a)**), there were several



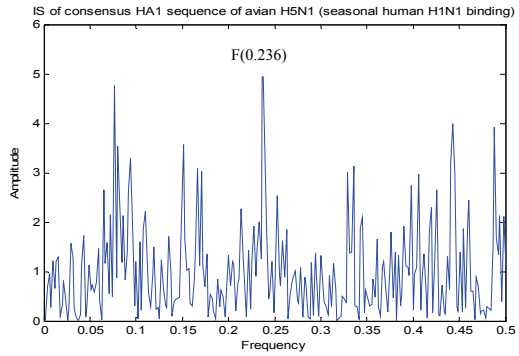
(a)



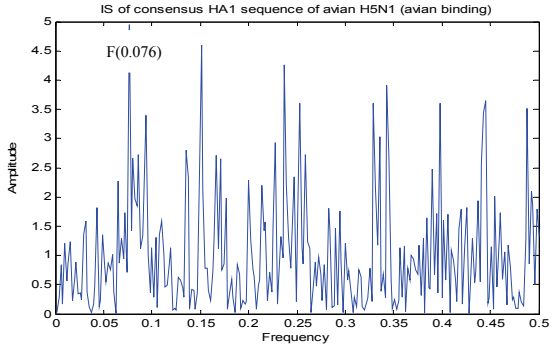
(b)



(c)



(d)



(e)

Figure 4. (a) CIS of consensus of all HA1 sequences of avian H5N1. (b) CIS of consensus HA1 sequence of avian H5N1 with seasonal human H1N1 binding. (c) CIS of consensus HA1 sequence of avian H5N1 with avian binding. (d) IS of consensus HA1 sequence of avian H5N1 with seasonal human H1N1 binding. (e) IS of consensus HA1 sequence of avian H5N1 with avian binding.

HA1 sequences in the dataset that had a higher IS peak at frequency F(0.236) than that at the frequency F(0.076). Bases on this observation, the whole set of HA1 sequences in avian H5N1 collected were divided into two subsets. One had the IS dominant peak at frequency F(0.076), referred to as avian binding subset (n=949), and the other had the IS dominant peak at frequency F(0.236), referred to as human binding subset (n=279). The CIS of these two subsets of HA1 sequences were plotted in (b) and (c) of **Figure 4**, and the IS of their consensus HA1 sequences were plotted in (d) and (e) of **Figure 4**, respectively. A total of 11 amino acid changes between the two consensus HA1 sequences of avian binding and human binding were found, and the resulting amplitude variation from each mutation was computed (**Table 5**). There were several mutations near the active site: D154N, N155S, A156T, and R189K.

3.1.3. Avian H1N1

For avian H1N1, we used the same strategy as for avian H5N1 to divide the whole set of HA1 sequences (n=78) in this subtype into two subsets according to different binding patterns (human (n=19) vs. avian (n=59)) as reflected by two different IS characteristic frequencies (F(0.295) vs. F(0.282))(**Figure 5**). Between the two consensus HA1 sequences of avian (F(0.282)) and human (F(0.295)) binding, there was only one mutation N121S (**Table 6**). The ISM was applied to quantify this mutation's contribution to the amplitude variation of the consensus HA1 sequence in each binding characteristics (**Table 6**).

For H1N1 viruses, the substitutions E190D/G225D were essential for avian virus HA to acquire human virus receptor specificity [13]. Here ISM was employed to verify this fact numerically (**Table 7**).

Table 5. Changes of amplitudes of IS frequencies by each mutation between avian and human binding patterns in HA1 of avian H5N1.

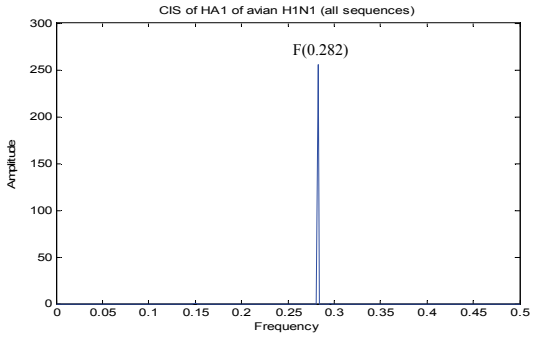
	Mutating Consensus HA1 Sequence of Avian Binding Patterns		Mutating Consensus HA1 Sequence of Human Binding Patterns	
Mutations	ΔA[F(0.076)]%	ΔA[F(0.236)]%	ΔA[F(0.076)]%	ΔA[F(0.236)]%
L71I	0	0	0	0
I83A	-2.7308	3.8849	2.6972	-3.2926
R140K	-7.8143	5.2052	8.2224	-4.3565
D154N	-14.371	5.9825	15.75	-3.3428
N155S	6.3122	9.7463	-5.5234	-8.0289
A156T	1.9481	4.9408	-1.3726	-3.7429
R189K	0.21422	-7.9896	0.36709	7.2805
N252Y	-5.9842	3.5318	6.173	-3.0255
T262A	0	0	0	0
I282M	4.7565	10.576	-4.6209	-8.6614
G323R	11.348	-8.0717	-10.859	7.048
Total	-6.3212	27.8086	10.8338	-20.1221

Table 6. Changes of amplitudes of IS frequencies by each mutation between avian and human binding patterns in HA1 of avian H1N1.

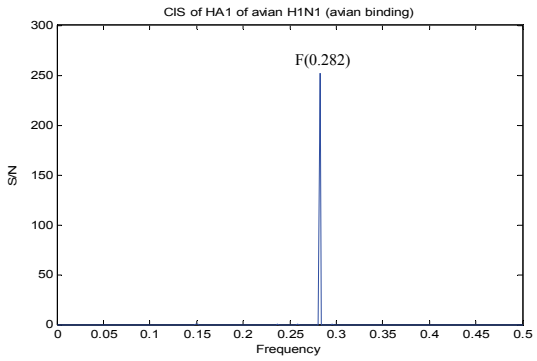
	Mutating Consensus HA1 Sequence of Avian Binding Patterns		Mutating Consensus HA1 Sequence of Human Binding Patterns	
Mutation	ΔA [F(0.282)]%	ΔA [F(0.295)]%	ΔA [F(0.282)]%	ΔA [F(0.295)]%
N121S	-15.7142	18.6251	18.6439	-15.7008

Table 7. Changes of amplitudes of IS frequencies by mutations E190D/G225D.

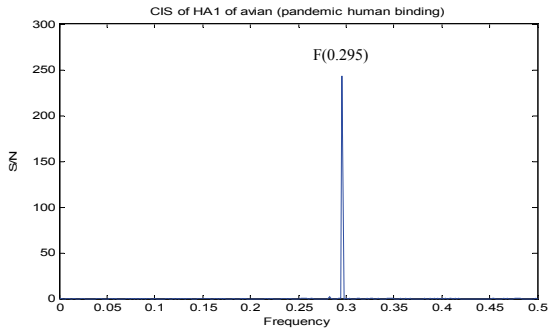
	Mutating Consensus HA1 Sequences of Avian H1N1		
Mutations	ΔA [F(0.282)]%	ΔA [F(0.295)]%	Dominant Peak Frequency
E190D	-13.2024	13.0862	F(0.282)
G225D	-10.4239	-0.6548	F(0.282)
E190D,G225D	-22.8997	11.7994	F(0.295)



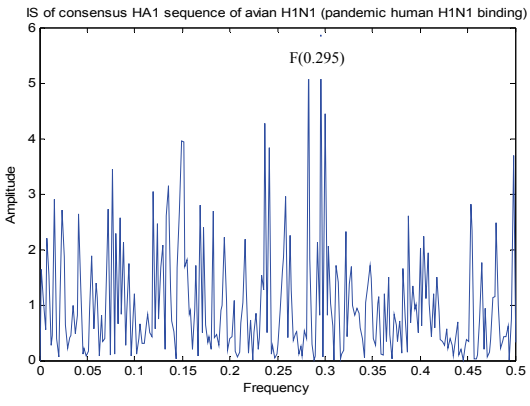
(a)



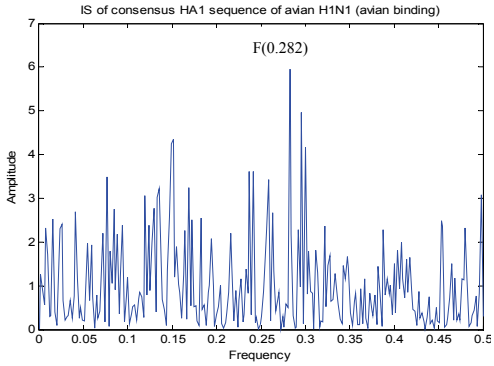
(b)



(c)



(d)



(e)

Figure 5. (a) CIS of consensus of all HA1 sequences of avian H1N1; (b) CIS of consensus HA1 sequence of avian H1N1 with pandemic human H1N1 binding; (c) CIS of consensus HA1 sequence of avian H1N1 with avian binding; (d) IS of consensus HA1 sequence of avian H1N1 with pandemic human H1N1 binding; (e) IS of consensus HA1 sequence of avian H1N1 with avian binding.

3.1.4. Swine H1N2

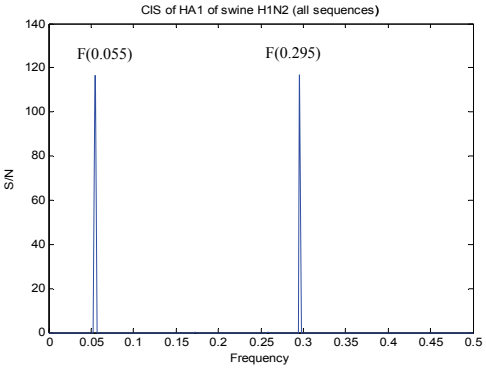
For swine H1N2, we used the same strategy as for avian H5N1 to divide the whole set of HA1 sequences (n=82) in this subtype into two subsets according to different binding patterns (human (n=45) vs. swine (n=37)) as reflected by two different IS characteristic frequencies (F(0.295) vs. F(0.055)) (**Figure 6**). The comparison of the two consensus HA1 sequences (human vs. swine) revealed 14 mutations, and the ISM was applied to quantify each mutation's contribution to the amplitude variation of the consensus HA1 sequence in each binding characteristics (**Table 8**). There were several mutations near the active site: N184T, R189Q, and T190S, and near the edges of the binding pocket: E127D, E224A, and R131K.

3.2. Mutations between Hosts

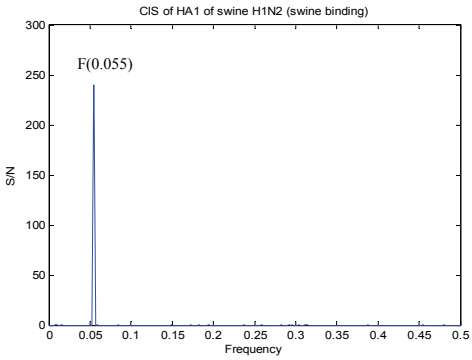
3.2.1. 2009 H1N1 and Swine H1N2

The 2009 H1N1 virus has its origin as a reassortant from a triple-reassortant virus circulating in North American swine and Eurasian avian-like swine H1N1, with its HA from the classical swine influenza viruses.

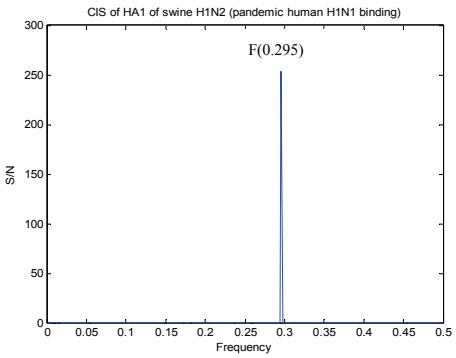
The comparison of the consensus HA1 sequences of 2009 H1N1 and swine H1N2 found 24 mutations. The ISM was applied to find the changes made by each mutation in the amplitude of IS of each consensus HA1 sequence (**Table 9**). Mutations K130R and E224A were near the edges of the binding pocket. There were three mutations at residues 71, 211, and 216 presented in Tables 3, 9, and 10, implying their importance in determining binding preferences of swine H1N2 and 2009 H1N1.



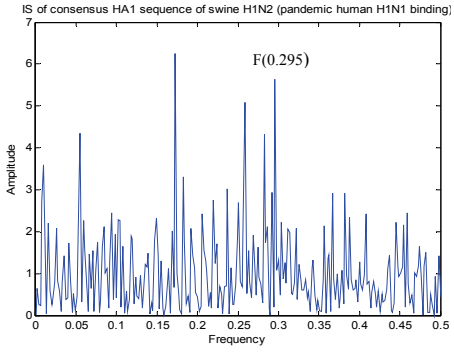
(a)



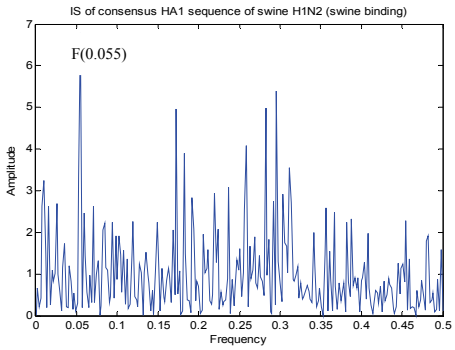
(b)



(c)

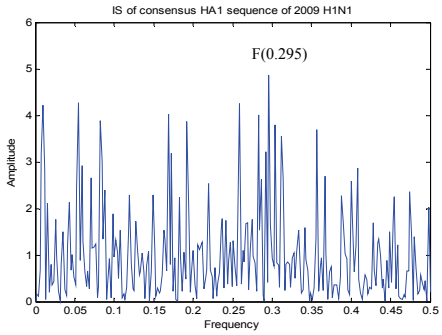


(d)

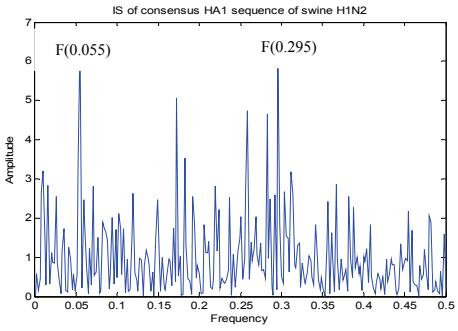


(e)

Figure 6. (a) CIS of consensus of all HA1 sequences of swine H1N2; (b) CIS of consensus HA1 sequence of swine H1N2 with pandemic human H1N1 binding; (c) CIS of consensus HA1 sequence of swine H1N2 with swine binding; (d) IS of consensus HA1 sequence of swine H1N2 with pandemic human H1N1 binding; (e) IS of consensus HA1 sequence of swine H1N2 with swine binding.



(a)



(b)

Figure 7. (a) IS of consensus HA1 sequence of 2009 H1N1. (b) IS of consensus HA1 sequence of swine H1N2.

Mutations in the binding site not only could modulate receptor specificity but also the antigenicity of the virus [27]. It was generally believed that the key residues at the binding site were not subject to selection, but recently a strong positive selection at position 190 in HA of H1N1 was detected in [13]. It was observed that there

Table 8. Changes of amplitudes of IS frequencies by each mutation between swine and human binding patterns in HA1 of swine H1N2.

	Mutating Consensus HA1 Sequence of Swine Binding Patterns		Mutating Consensus HA1 Sequence of Human Binding Patterns	
Mutations	ΔA [F(0.055)]%	ΔA [F(0.295)]%	ΔA [F(0.055)]%	ΔA [F(0.295)]%
V47I	-0.45253	-0.39918	0.61877	0.4217
I71F	-7.4498	3.7605	9.5403	-3.7394
E127D	-11.057	-1.2554	14.693	3.0055
R131K	0	0	0	0
G170E	0.014238	-0.07755	-0.01469	0.07289
N184T	-8.3652	-10.1	11.495	10.475
R189Q	-1.4448	-2.6783	1.6451	2.6875
T190S	-1.1476	0.1999	1.352	-0.09641
T211K	-6.1626	6.441	8.1349	-6.0602
K216T	-1.9443	7.4225	2.3502	-6.895
E224A	-0.10535	-1.9348	0.42226	1.7602
A250V	-1.1944	-3.5238	1.8901	3.4949
P271S	-2.8719	3.2365	3.4537	-3.3317
K278T	-0.95454	0.60359	2.0448	-0.81079
Total	-43.1358	1.6950	57.6254	0.9842

Table 9. Changes of amplitudes of IS frequencies by each mutation between consensus HA1 sequences of 2009 H1N1 and swine H1N2.

	Mutating Consensus HA1 Sequence of Swine H1N2		Mutating Consensus HA1 Sequence of 2009 H1N1	
Mutations	ΔA[F(0.055)]%	ΔA[F(0.295)]%	ΔA[F(0.055)]%	ΔA[F(0.295)]%
K36R	-5.6796	5.4297	7.2412	-4.7622
I61L	0	0	0	0
S71F	-0.75109	0.59143	0.90722	-0.85027
S84N	-7.0341	-8.3814	9.614	10.052
D97N	-4.8444	-9.6342	6.0451	9.9181
T120E	-8.911	5.4563	11.964	-7.0032
K130R	-4.3521	-4.4943	6.0754	5.9088
K142N	0.67563	-2.0369	-0.36048	2.8758
K146R	-5.7743	-2.0864	7.4097	1.4689
D168N	4.5193	-5.5381	-5.7447	4.824
G170E	0.011884	-0.07133	-0.01058	0.06883
R205H	1.7024	1.1275	-2.4361	-2.3933
K211T	6.5905	-5.7011	-7.8044	6.4193
I216K	-1.8394	3.6983	2.0037	-3.8566
E224A	-0.2376	-1.7644	0.61948	1.4082
K239T	-5.1394	0.52104	6.8924	0.89224
M257L	8.8688	-8.1637	-10.571	10.056
E258K	-2.8787	-2.2561	3.7825	2.0134
N260G	-0.05099	0.1242	0.071321	-0.12802
A261S	1.0216	0.76162	-0.593	-1.5978
I298V	0.093799	0.2947	-0.08087	-0.42787
K302E	-3.1322	1.8998	4.0108	-1.4074
L314M	-4.0997	6.3337	6.3078	-5.2087
V321I	-0.49821	-0.48678	0.6555	0.49537
Total	-31.7389	-27.3764	45.9990	28.7656

were two distinct evolutionary patterns in host-driven antigenic drift of human H1N1 HAs at positions 190 and 225, *i.e.*, the antigenic drift of 1918 pandemic HAs occurred at position 225, and that of epidemic HAs happened at position 190. In contrast to these two trends, the HAs in 2009 H1N1 took a different path, which were highly conserved at both positions 190 and 225, based on the 73 HA sequences of 2009 H1N1, as of July 10,

2009. In the present study, as of November 20, 2009, we had the following counts of various mutations at positions 190 and 225 in HA of 2009 H1N1 (**Table 10**). It appeared that the 2009 H1N1 HAs continued to keep this evolutionary pattern.

The alteration of receptor binding specificity was believed to be an essential step in host adaptation. Several studies on 1918 HA discovered that a single mutation D225G decreased the binding affinity of 1918 HA for α2,6 receptors and resulted in a mixed α2,6/α2,3 binding virus, furthermore, a double mutations D190E/D225G abolished the binding of 1918 HA to α2,6 and resulted in a α2,3 binding virus [28,29]. Our numerical analysis (**Table 11**) suggested that the mutations at positions 190 and 225 would produce similar binding affinity of 2009 H1N1 HA to that of 1918 HA. Mutation D190E decreased the amplitude at frequency F(0.295) more than that at frequency F(0.055), displaying the avian binding affinity, while mutation D225G produced the opposite effect, exhibiting the human binding specificity. The double mutations D190E/D225G showed avian binding preference.

Table 10. Mutation counts in HA1 sequences of 2009 H1N1.

Mutation	Counts
D190G	1
D190V	5
D225E	6
D225G	7
D225N	2

Table 11. Changes of amplitudes of IS frequencies by mutations observed in HA1 sequences of 2009 H1N1.

	Mutating Consensus HA1 Sequence of 2009 H1N1		
Mutations	ΔA [F(0.055)]%	ΔA [F(0.295)]%	Dominant Peak Frequency
D190E	-2.0061	-9.6206	F(0.282)
D190G	-2.0302	-9.6888	F(0.282)
D190N	-2.0789	-9.8255	F(0.282)
D190V	-2.0027	-9.6109	F(0.282)
D225E	-9.5672	-2.1233	F(0.295)
D225G	-9.6345	-2.1488	F(0.295)
D225N	-9.7694	-2.2003	F(0.295)
D190E,D225E	-10.5225	-12.3705	F(0.282)
D190E,D225G	-10.5836	-12.3981	F(0.282)
D190E,D225N	-10.7061	-12.4537	F(0.282)
D190G,D225E	-10.5392	-12.4420	F(0.282)
D190G,D225G	-10.6002	-12.4696	F(0.282)
D190G,D225N	-10.7226	-12.5253	F(0.282)
D190N,D225E	-10.5729	-12.5854	F(0.282)
D190N,D225G	-10.6338	-12.6131	F(0.282)
D190N,D225N	-10.7560	-12.6688	F(0.282)
D190V,D225E	-10.5201	-12.3603	F(0.282)
D190V,D225G	-10.5812	-12.3879	F(0.282)
D190V,D225N	-10.7037	-12.4435	F(0.282)

3.2.2. Avian H5N1 and Human H5N1

Most of the highly pathogenic avian H5N1 strains bind strongly to avian receptors. This specificity is normally a barrier to viral transmission from birds to humans. However, a few of them have been discovered to bind to human receptors as well as to avian receptors. The comparison of the consensus HA1 sequences of avian H5N1 and human H5N1 identified three mutations, R140K, R189K, and T263A, and their impact on the two characteristic frequencies F(0.076) and F(0.236) was illustrated in **Table 12**.

As demonstrated in the experiments conducted in references [5,6], substitutions Q192R and S223L could mediate a shift from avian to human binding preference in the H5N1 viruses. In references [30,31,32], mutations Q226L and G228S enhanced the human binding capacity while reduced avian binding capacity of the H5N1 viruses. Here the ISM was utilized to evaluate the outcome of these mutations (**Table 13**). **Figure 8** illustrated the effect of mutation S223L on the two characteristic frequencies F(0.076) and F(0.236).

4. DISCUSSION

Understanding the minimal adaptive changes necessary for viral adaptation to human host is of key importance in learning how pandemic influenza viruses emerge. Alteration of receptor recognition is a vital step in host adaptation, which is modulated directly by a subset of amino acids present in the HA protein. Other determinants of host adaptation include continued viral evolution to improve transmission and replication efficiency and optimize tissue tropism.

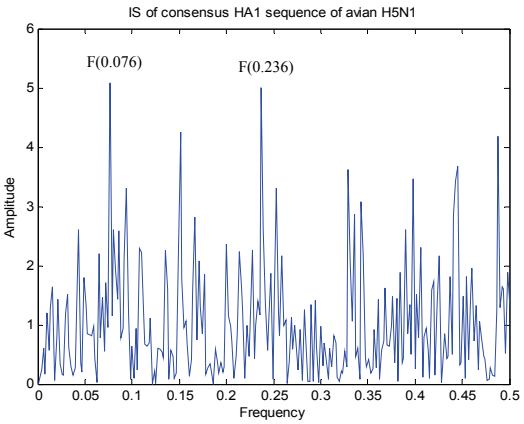
There are several well-known mutations in HA protein in different flu subtypes, playing critical roles in receptor binding preference shift. Besides mutations in HA, the glycosylation sites in HA might also impact the binding specificity of HA [32]. The mutations discovered in this study represent alternatives by which the HA can switch its substrate recognition. These mutations occur in nature, whereas those artificially engineered by wet lab techniques may have a very low probability to occur in nature. As evidenced by the study in [30] that mutations Q226L and G228S in HA proteins could alter the binding specificity of the H5N1 viruses. Nevertheless, it also pointed out that the likelihood to acquire the necessary nucleotide changes to produce these mutations in natural virus is small, which could explain in part why

Table 12. Changes of amplitudes of IS frequencies by each mutation between consensus HA1 sequences of avian H5N1 and human H5N1.

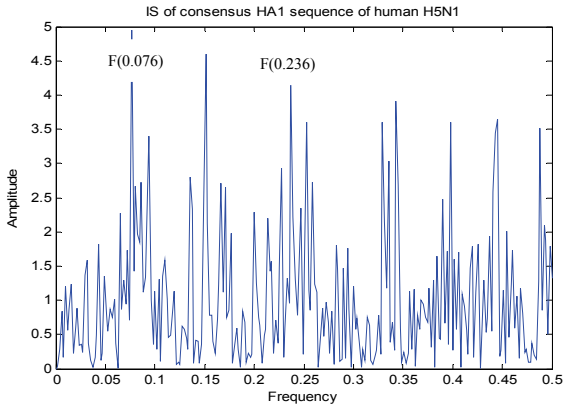
	Mutating Consensus HA1 Sequence of Avian H5N1		Mutating Consensus HA1 Sequence of Human H5N1	
Mutation	ΔA [F(0.076)]%	ΔA [F(0.236)]%	ΔA [F(0.076)]%	ΔA [F(0.236)]%
R140K	7.9425	-4.5900	-7.8143	5.2052
R189K	-0.5253	7.7370	0.2142	-7.9896
T263A	-3.7894	-6.0611	4.5619	6.8405
Total	3.6278	-2.9141	-3.0382	4.0561

Table 13. Changes of amplitudes of IS frequencies by mutations experimented in references [5,6,30,31,32].

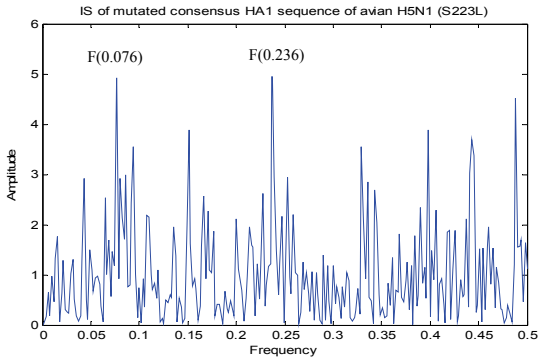
	Mutating Consensus HA1 Sequence of Avian H5N1	
Mutations	ΔA [F(0.076)]%	ΔA [F(0.236)]%
Q192R	-3.2185	1.5596
S223L	-4.0399	10.8075
Q192R, S223L	-6.5711	11.6823
Q226L	0.9498	17.4463
G228S	7.3516	15.7656
Q226L, G228S	8.0339	16.9010



(a)



(b)



(c)

Figure 8. (a) IS of consensus HA1 sequence of avian H5N1; (b) IS of consensus HA1 sequence of human H5N1; (c) IS of consensus HA1 sequence of avian H5N1 mutated by S223L.

viruses such as H5N1 have not yet evolved into human transmissible strains to cause a human pandemic.

It appears that increased binding affinity for human influenza receptors alone is not sufficient for efficient human transmission, and additional molecular determinants are required. In [33], it was proposed that binding to long-chain α2-6 sialosides is a necessary requirement for viruses to efficiently replicate and transmit in humans. There is inadequacy in assessing the adaption to human receptor affinity from the analysis of a few influenza strains [30]. It showed that the effects of the same mutations, such as Q226L/G228S, on the binding preference of one strain were not the same on others strains in H5N1. Because lab experiments are labor intensive and costly, bioinformatics approaches offer reasonable alternatives in the analysis of binding specificity given the large number of virus sequences, which can process all strains with any combination of mutations within a subtype efficiently.

5. CONCLUSIONS

The increasing trend of direct transmission of avian/swine influenza viruses to humans underscores the need to understand further the mechanism of glycan receptor recognition and specificity switch. In this study, muta-

tions in HA1 within or between hosts in various flu subtypes were identified, and their contribution to binding preference was measured using the ISM. Our numerical analysis implied that the mutations in HA1 of 2009 human H1N1 collectively tended to reduce the swine binding affinity in the seasonal H1N1 strains and to increase that in the pandemic H1N1 strains. At the same time, they increased the human binding affinity in the pandemic H1N1 strains and had little impact on that in the seasonal H1N1 strains. The mutations in HA1 of avian H5N1 and avian H1N1 exhibited reduced avian binding in the strains of avian binding propensity while showed enhanced avian binding affinity in the strains with human binding propensity. They displayed the opposite effects on human binding. The mutations in HA1 of swine H1N2 reduced the swine binding affinity in the strains with swine binding propensity and enhanced that in the strains with human binding propensity. They showed little impact on the human binding affinity, which was different from the mutations in HA1 of 2009 H1N1. The mutations between the consensus HA1 sequences of 2009 H1N1 and swine H1N2 decreased both the human and swine binding affinities in swine H1N2 and increased those in 2009 H1N1. The mutations between the consensus HA1 sequences of human H5N1 and avian H5N1 increased the avian binding affinity and decreased the human binding affinity in avian H5N1 while produced the opposite effects on those in human H5N1.

The mutations discovered in the present study confirmed the potential for influenza viruses to adapt to human host, and furthermore, our numerical analysis detailed the extent of binding preference changes induced by each mutation. These mutations and their corresponding contribution to the binding specificity alteration yielded new clues to the mechanism of receptor recognition switch within and between hosts. The ISM offered a complementary and efficient approach to investigate the binding affinities of all HA sequences in various subtypes, a task difficult to accomplish experimentally.

6. ACKNOWLEDGMENTS

We thank Houghton College for its financial support.

REFERENCES

[1] Webster, R.G. (1999) 1918 Spanish influenza: The secrets remain elusive. *Proceedings of the National Academy of Sciences,* USA, **96(4)**, 1164-1166.

[2] Li, Q., Kash, J.C., Dugan, V.G., Wang, R.X., Jin, G.Z., Cunningham, R.E. and Taubenberger, J.K. (2009) Role of sialic acid binding specificity of the 1918 influenza virus hemagglutinin protein in virulence and pathogenesis for mice. *Journal of Virology*, **83(8)**, 3754-3761.

[3] Matrosovich, M.N., Klenk, H. D. and Kawaoka, Y. (2006) Receptor specificity, host-range, and pathogenicity of influenza viruses. In Yoshihiro Kawaoka (ed.) *Influenza Virology: Current Topics*, Caister Academic Press, 95-137.

[4] Zambon, M. (2007). Lessons from the 1918 influenza. *Nature Biotech*, **25**, 433-434.

[5] Gambaryan, A. Tuzikov, A., Pazynina, G., Bovin, N., Balish, A. and Klimov, A. (2006) Evolution of the receptor binding phenotype of influenza A (H5) viruses. *Virology*, **344(2)**, 432-438.

[6] Yamada, S., Suzuki, Y., Suzuki, T., Le, M. Q., Nidom, C. A., Sakai-Tagawa, Y., Muramoto, Y., *et al*. (2006) Haemagglutinin mutations responsible for the binding of H5N1 influenza A viruses to human-type receptors. *Nature*, **444(7117)**, 378-382.

[7] Matrosovich, M., Tuzikov, A., Bovin, N., Gambaryan, A., and Klimov, A., *et al*. (2000) Early alterations of the receptor-binding properties of H1, H2, and H3 avian influenza virus hemagglutinins after their introduction into mammals. *Journal of Virology*, **74(18)**, 8502-8512.

[8] Bateman, A.C., Busch, M.G., Karasin, A.I., Bovin, N., and Olsen, C.W. (2008) Amino acid 226 in the hemagglutinin of H4N6 influenza virus determines binding affinity for alpha2, 6-linked sialic acid and infectivity levels in primary swine and human respiratory epithelial cells. *Journal of Virology*, **82(16)**, 8204-8209.

[9] Wan, H., Sorrell, E.M., Song, H., Hossain, M.J., Ramirez-Nieto, G., *et al*. (2008) Replication and transmission of H9N2 influenza viruses in ferrets: Evaluation of pandemic potential. *PLoS ONE*, **3(8)**, e2923.

[10] Rogers, G.N. and D'Souza, B.L. (1989) Receptor binding properties of human and animal H1 influenza virus isolates. *Virology*, **173**, 317-322.

[11] Matrosovich, M.N., Gambaryan, A.S., Teneberg, S., Piskarev, V.E., Yamnikova, S.S., *et al*. (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(1)**, 224-234.

[12] Reid, A.H., Janczewski, T.A., Lourens, R.M., Elliot, A.J., Daniels, R.S., Berry, C.L., Oxford, J.S. and Taubenberger, J.K. (2003) 1918 influenza pandemic caused by highly conserved viruses with two receptor-binding variants. *Emerg Infect Disease*, **9(10)**, 1249-1253.

[13] Shen, J., Ma, J. and Wang, Q. (2009) Evolutionary trends of A(H1N1) influenza virus hemagglutinin since 1918. *PLoS ONE*, **4(11)**, e7789.

[14] Srinivasan, A., Viswanathan, K., Raman, R., Chandrasekaran, A., Raguram, S., Tumpey, T.M., Sasisekharan, V. and Sasisekharan, R. (2008) Quantitative biochemical rationale for differences in transmissibility of 1918 pandemic influenza A viruses. *Proceedings of the National Academy of Sciences*, **105(8)**, 2800-2805.

[15] Soundararajan, V., Tharakaraman, K., Raman, R., Raguram, S., Shriver, Z., Sasisekharan, V. and Sasisekharan R. (2009) Extrapolating from sequence–the 2009 H1N1 "swine" influenza virus. *Nature Biotechnology*, **27(6)**, 510-513.

[16] Childs, R.A., Palma, A.S., Wharton, S., Matrosovich, T., Liu, Y., Chai, W.G., Campanero-Rhodes, M.A., Zhang, Y.B., Eickmann, M., Kiso, M., Hay, A., Matrosovich. M. and Feizi, T. (2009) Receptor-binding specificity of

pandemic influenza A (H1N1) 2009 virus determined by carbohydrate microarray. *Nature Biotechnology*, **27 (9)**, 797-799.

[17] Hu, W. (2009) Analysis of correlated mutations, stalk motifs, and phylogenetic relationship of the 2009 influenza A virus neuraminidase sequences. *Journal of Biomedical Science and Engineering*, **2(7)**, 550-558.

[18] Hu, W. (2010) The Interaction between the 2009 H1N1 influenza A hemagglutinin and neuraminidase: Mutations, co-mutations, and the NA stalk motifs. *Journal of Biomedical Science and Engineering*, **3**, 1-12.

[19] Veljkovic, V., Niman, H.L., Glisic, S., Veljkovic, N., Perovic, V. and Muller C.P. (2009). Identification of hemagglutinin structural domain and polymorphisms which may modulate swine H1N1 interactions with human receptor. *BMC Structural Biology*, **9**, 62.

[20] Veljkovic, V., Veljkovic, N., Muller, C.P., Müller, S. Glisic, S., Perovic, V. and Köhler, H. (2009) Characterization of conserved properties of hemagglutinin of H5N1 and human influenza viruses: Possible consequences for therapy and infection control. *BMC Structural Biology*, **7**, 9-21.

[21] Cosic, I. (1997) The resonant recognition model of macromolecular bioreactivity, theory and application. Berlin: Birkhauser Verlag.

[22] Hu, W. (2010) Identification of highly conserved domains in hemagglutinin associated with the receptor binding specificity of influenza viruses: 2009 H1N1, avian H5N1, and swine. *Journal of Biomedical Science and Engineering*, **3**, 114-123.

[23] Katoh, K., Kuma, K., Toh, H. and Miyata, T. (2005) MAFFT version 5: Improvement in accuracy of multiple sequence alignment. *Nucleic Acids Research*, **33(2)**, 511-518.

[24] MacKay, D. (2003) Information theory, inference, and learning algorithms. Cambridge University Press.

[25] KováccaronOVá, A., Ruttkay-Nedecký, G., Karol HaverlíK, I. and Janecccaronek, S. (2002) Sequence similarities and evolutionary relationships of influenza virus A hemagglutinins, *Virus Genes*, **24(1)**, 57-63.

[26] Gamblin, S.J., Haire, L.F., Russell, R.J., Stevens, D.J., Xiao, B., Ha, Y., *et al*. (2004) The structure and receptor binding properties of the 1918 influenza hemagglutinin. *Science*, **303**, 1838-1842.

[27] Daniels, R.S., Douglas, A.R., Skehel, J.J., Wiley, D.C., Naeve, C.W., Webster, R.G., Rogers, G.N. and Paulson, J. C. (1984). Antigenic analyses of influenza virus haemagglutinins with different receptorbinding specificities. *Virology*, **138**, 174-177.

[28] Srinivasan, A., Viswanathan, K., Raman, R., Chandrasekaran, A., Raguram, S., *et al*. (2008) Quantitative biochemical rationale for differences in transmissibility of 1918 pandemic influenza A viruses. *Proceedings of the National Academy of Sciences*, **105**, 2800-2805.

[29] Stevens, J., Blixt, O., Glaser, L., Taubenberger, J.K. , Palese, P., *et al*. (2006) Glycan microarray analysis of the hemagglutinins from modern and pandemic influenza viruses reveals different receptor specificities. *Journal of Molecular Biology*, **355(5)**, 1143-1155.

[30] Ayora-Talavera, G., Shelton, H., Scull, M.A., Ren, J., Jones, I.M., *et al*. (2009) Mutations in H5N1 influenza virus hemagglutinin that confer binding to human tracheal airway epithelium. *PLoS ONE*, **4(11)**, e7836.

[31] Stevens, J., Blixt, O., Glaser, L., Taubenberger, J.K., Palese, P., *et al*. (2006) Glycan microarray analysis of the hemagglutinins from modern and pandemic influenza viruses reveals different receptor specificities. *Journal of Molecular Biology*, **355(5)**, 1143-1155.

[32] Stevens, J., Blixt, O., Chen, L.M., Donis, R.O., Paulson, J.C., *et al*. (2008) Recent avian H5N1 viruses exhibit increased propensity for acquiring human receptor specificity. *Journal of Molecular Biology*, **381(5)**, 1382- 1394.

[33] Chandrasekaran, A., Srinivasan, A., Raman, R., Viswanathan, K., Raguram, S., Tumpey, T.M., Sasisekharan, V. and Sasisekharan, R. (1008) Glycan topology determines human adaptation of avian H5N1 virus hemagglutinin. *Nature Biotechnology*, **26(1)**, 107-113.

