Mutations in Hemagglutinin of H5N1 Influenza That Switch Receptor Specificity from Avian to Human Types

Wei Hu

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

Received April 16, 2013; revised May 10, 2013; accepted May 17, 2013

Copyright © 2013 Wei Hu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT

Researchers have been searching for molecular features that could make avian H5N1 influenza transmissible among people since the first report of human infections with this virus in 1997. A recent study surprisingly demonstrated that only five mutations, fewer than previously estimated, are needed to make avian H5N1 influenza transmissible between ferrets through the air, raising fears that a human pandemic is possible if this virus escapes from the lab. Of the five mutations found, four of them are located in the HA gene that is responsible for the viral entry into the host cells. A crucial step for avian influenza to go across the species boundary to infect humans is the switch of its receptor binding specificity from avian to human types. The first task of this study was to quantify the individual as well as the collective effect of the known HA mutations from the previous research on receptor binding selection. Our second task was to identify new combinations of HA mutations that could change the receptor binding preference of H5N1 from avian to human types. Our findings thus deepened our understanding of the previous research and also extended its results by discovering new combinations of mutations that could enhance the binding of avian H5N1 to human type receptors while reduce that to avian types.

Keywords: Influenza; H5N1; Mutation; Receptor Binding Specificity; HA Gene

1. Introduction

Highly pathogenic avian H5N1 virus is contagious and lethal, normally only infecting birds. Outbreaks of this virus in poultry have been reported several times. First discovered in Hong Kong in 1997, human infections caused by infected poultry could result in serious illness or death. Yet, there is no evidence of direct transmission of this virus between people [1,2]. The relatively low occurrences of human infections to date imply a substantial species barrier. To initiate influenza virus infection, its hemagglutinin (HA) protein binds to receptors on the surface of host cells. The HA gene is a well-known major determinant of host range, since avian and human virus HAs bind preferentially to two different receptors on host cells [3]. However, the continuing evolution of this virus makes it possible to change genetically into a strain that is highly infectious for humans and allows for efficient human-to-human transmission, which could cause an influenza pandemic. The increasing serious public health threat posed by wide spread of this virus highlights the urgent call for full genetic and molecular characterization of its viral adaptation to human hosts.

A recent study confirmed that merely five mutations are required for H5N1 to become transmissible between ferrets via the air [4]. Herfst *et al.* [4] first introduced three known mutations: PB2-627K efficient for replication in mammals and HA-Q222L/G224S for human type receptor binding. Two additional mutations HA-T156A and HA-H103Y emerged during viral passages in ferrets, which were not seen in wild H5N1 before but occurred in the H2 and H3 viruses. Mutation HA-T156A could lead to the loss of a glycosylation site on the head of the HA and HA-H103Y is localized to the HA trimer interface. Taken together, their work proved that just five mutations, PB2-627K, HA-H103Y, HA-T156A, HA-Q222L, and HA-G224S, are necessary to make H5N1 transmissible among ferrets through the air [5].

Another study by Imai *et al.* [6] made use of an avian H5N1 HA gene and the remaining genes of pandemic 2009 (H1N1) virus. They randomly mutated the H5 HA and discovered two mutations, N220K and Q222L, which caused the virus to shift its receptor binding selection from avian to human types. Then this mutated HA gene was used to replace the original HA gene of pan-



demic 2009 H1N1. This re-assortment process in the lab produced a virus transmissible in ferrets. Furthermore, two additional mutations in HA, N154D and T314I, came out during viral passages in ferrets. It was note-worthy that mutation Q222L was analyzed in both studies [4,6], and mutations T156A and N154D found from these two separate experiments respectively both resulted in the loss of the same glycosylation site and therefore could influence the receptor binding preference. Mutation T314I turned out to affect HA stability [5].

A new development in the influenza research was the discovery of a novel reassortant avian-origin influenza A (H7N9) virus from several patients, with severe pneumonia and multiorgan dysfunction, in China in March 2013 [7]. Most likely by poultry-to-person transmission [8], this new virus has affected at least 108 people including 22 deaths as indicated in a news report from news (Xinhuanet.com on April 23, 2013). The study in [1] suggested that this novel 2013 H7N9 virus in China was of avian origin. Its NA gene is similar to that of another H7N9 virus from Korea (A/wild bird/Korea/A14/2011), its HA gene is close to that of an H7N3 virus from Zhejiang Province in China (A/duck/Zhejiang/12/2011), and all its internal gene segments are very similar to those from avian H9N2 viruses, particularly a virus isolated from a brambling in Beijing (A/brambling/Beijing/ 16/2012).

A follow-up study on this new virus can be found in [9], which took an early investigation of this virus by comparing the collected viral genome sequences of 2013 H7N9 in China against those of previous avian H7N9 and examined the receptor binding specificity of this new virus. This virus was found to be very different from the previous avian H7N9 viruses and many of the internal proteins of 2013 H7N9 from the avian and human hosts in China were either identical or similar. Furthermore, the current strains of 2013 H7N9 in China exhibited avian type receptors as their primary binding preference and human type receptors as secondary. Several HA mutations, including a few well known for host adaptation, of this virus were identified. Collectively these current HA mutations tended to decrease the HA binding affinity for avian type receptors and increase that for human type receptors, which could enhance the ability of this virus to infect humans [9].

In light of the emergence of this novel H7N9 virus in China, we were very curious to learn, employing the methodology developed in [10-25], the individual as well as the collective impact of the mutations identified in [4, 6] on the receptor selection of H5N1. At the same time, beyond the groups of HA mutations examined in these two studies we wanted to seek other possible groups of mutations that could render similar effect on HA receptor binding preference to those found in [4,6].

2. Materials and Methods

2.1. Sequence Data

Protein sequences of influenza viruses used in this study were retrieved from the EpiFlu Database (http://platform. gisaid.org) of GISAID. All sequences were aligned with MAFFT [26].

2.2. Informational Spectrum Method

The informational spectrum method (ISM) is a computational approach to analyzing protein sequences [27]. The idea is to transform the protein sequences into numerical sequences based on Electron-Ion Interaction Potential (EIIP) of each amino acid. Then the Discrete Fourier Transform (DFT) can be applied to these numerical sequences, and the resulting DFT coefficients are used to produce the energy density spectrum. The Informational Spectrum (IS) comprises the frequencies and the amplitudes of this energy density spectrum. According to the ISM theory, the peak frequencies of IS of a protein sequence encode its biological or biochemical functions. The ISM was successfully applied to quantify the effects of HA mutations on the receptor binding preference in [15,28,29]. The results of [29] revealed that the primary consensus IS (CIS) frequency is F(0.076) for avian H5N1 and that for human seasonal H1N1 is F(0.236).

3. Results

We report our analysis on the individual and collective impact of the HA mutations found in [4,6] and new combinations of mutations examined in [6,30]. Two HA genes of H5N1 were utilized to gauge the effect of HA mutations: The HA gene of A/Indonesia/5/2005 (H5N1) used in [4] and that of A/Viet Nam/1203/2004 (H5N1) in [6]. Although these two strains were isolated from human patients, their HA retains avian type receptors as their primary binding preference and human types as secondary (Figure 1). As observed in [19], influenza viruses tend to display dual binding preference, one as primary preference and the other as secondary. We computed the IS amplitude changes at the primary and secondary frequencies of the two HA genes of A/Indonesia/5/2005 and A/Viet Nam/1203/2004 caused by mutations (Tables 1 and 2).

3.1. Effect of HA Mutations Found in [4,6]

Our IS analysis showed that most groups of the mutations in [4] appeared to boost the binding to avian type receptors and all but one increased the binding to human types (**Table 1**). In contrast, most groups of the mutations in [6] seemed to reduce the binding to avian type receptors. It was interesting to note that both of these two sets of mutations from [4,6] demonstrated similar bind-

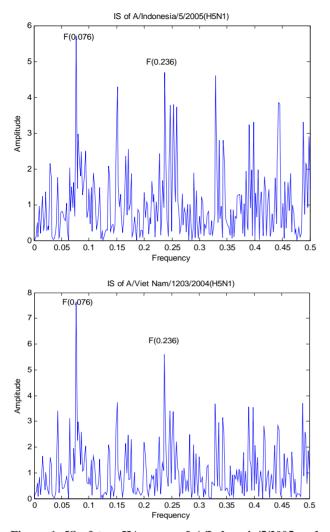


Figure 1. IS of two HA genes of A/Indonesia/5/2005 and A/Viet Nam/1203/2004, with the primary frequency being F(0.076) and the secondary being F(0.236).

Table 1. IS amplitude changes of Indonesia/5/2005 at the primary and secondary frequencies caused by different mutations from [4], where an up arrow indicates an increase and a down arrow suggests a decrease of the IS amplitude compared to the baseline.

Mutations (H3 numbering)	Amplitude at primary frequency F(0.076)	Amplitude at secondary frequency F(0.236)
	5.7112 (Base line)	4.6931 (Base line)
H103Y	5.6425↓	4.7968 ↑
T156A	5.8177 ↑	4.8876 ↑
N182K	5.6615↓	4.7891 ↑
Q222L	5.7235 ↑	4.757 ↑
G224S	6.1346 ↑	4.6767↓
H103Y T156A	5.7427 ↑	4.9980 ↑
Q222L G224S	6.1331 ↑	4.7627 ↑
Q222L G224S N182K	6.0782 ↑	4.8771 ↑
H103Y T156A Q222L G224S	6.1706 ↑	5.0258 ↑

Mutations (H5 numbering)	Amplitude at primary frequency F(0.076)	Amplitude at secondary frequency F(0.236)	
	7.6260 (Base line)	5.5980 (Base line)	
E119G	7.6295 ↑	5.5986 ↑	
V152I	7.5873↓	5.5832↓	
N158D	6.7276↓	5.7501 ↑	
N224K	7.4376↓	5.6954 ↑	
Q226L	7.6311↑	5.7068 ↑	
G228S	8.1158 ↑	5.6204 ↑	
T318I	7.2910↓	5.6667 ↑	
N224K Q226L	7.4368 ↓	5.8139 ↑	
Q226L G228S	8.1068 ↑	5.7509 ↑	
N158D N224K Q226L T318I	6.2309↓	5.9834 ↑	

Table 2. IS amplitude changes of Viet Nam/1203/2004 at the primary and secondary frequencies caused by different mu-

ing affinity for human type receptors (**Tables 1** and **2**). The famous four mutations, H103Y, T156A, Q222L, and G224S, in [4] actually boosted the binding to both avian and human type receptors (**Table 1**). On the other hand, the four highlighted mutations, N158D, N224K, Q226L, and T318I, in [6] reduced the binding to avian types and increased that to human types (**Table 2**). In addition to these four well studied mutations, three more mutations, E119G, V152I, and G228S, were also included for analysis in [6] (**Table 2**).

It was noted that each of the two well known mutations, Q226L and G228S, increased the binding to both avian and human type receptors with the exception of G228S having a little amplitude reduction at F(0.236) in **Table 1 (Tables 1** and **2**). The experiment in [6] suggested that three mutations, N158D, N224K and Q226L, contributed to human type receptor binding, while mutation T318I lowered the pH at which the HA protein undergoes a structural fusion process. Our calculation showed that T318I could also influence the HA binding specificity.

3.2. Effect of Other Groups of HA Mutations beyond Those Found in [4,6]

We tried to find out whether there were other combinations of mutations that exhibited similar receptor binding shifting capability. The candidate individual mutations included in our analysis were: E119G, V152I, N158D, N224K, Q226L, and G228L from [6] and D187G, E190G, Q196R, E225K, and S227N from [30]. The IS amplitudes of the mutated HA of A/Viet Nam/1203/2004 at F(0.236) by these mutations were sorted in **Table 3**. The groups of mutations in **Table 3** all increased the binding to human type receptors, and all but one reduced the binding to avian types, which is what is needed for a switch from avian to human hosts. **Figure 2** shows that three mutations, N158D, S227N, and E255K, were selected most of the time among all the 11 mutations considered for this purpose, implying their significance in the receptor binding preference.

The mutations N154D and T156A found from the two separate experiments [4,6] both resulted in the loss of the same glycosylation site. Mutation N154D changed the amplitudes of the IS of A/Viet Nam/1203/2004 at the primary and secondary from base line 7.6260 and 5.5980 to 6.7276 and 5.7501 respectively. Also T156A changed those of A/Indonesia/5/2005 from base line 5.7112 and 4.6931 to 5.8177 and 4.8876 respectively. Thus our calculation implied that N154D reduced the binding to avian type receptors and increased that to human types, and T156A increased the binding to both avian and human types.

4. Conclusions

The genomic signature of the H5N1 virus to switch from avian to human hosts can be used as a potentially predictive of an influenza pandemic. To this end, computational methods such as the one employed in this report could greatly increase the efficiency and effectiveness of this discovery process by calculating the potential indicators first and then verify them experimentally. In the present study, we examined the individual and collective impact of the mutations found in [4,6] on the H5N1 receptor binding preference. Our findings implied the four mutations, H103Y, T156A, Q222L, and G224S, from [4] increased the binding to both avian and human types, while the four mutations, N158D, N224K, O226L, and T318I, from [6] reduced the binding to avian types and enhanced the binding to human types. Mutation N182K was also analyzed in [4] and was found to favor human type receptors. Here our analysis of N182K indicated that it reduced the affinity for avian type receptors and increased that for human types, consistent with the experimental results found in [4]. Mutations T156A and N154D revealed in two separate experiments [4,6] both resulted in the loss of the same glycosylation site. Our analysis demonstrated that these two mutations both increased the binding to human type receptors whereas N154D decreased and T156A increased the binding to avian types.

Furthermore, our analysis suggested that most groups of the mutations in [4] enhanced the binding to avian type receptors and all but one increased the binding to human types. In contrast, most groups of the mutations in [6] reduced the binding to avian type receptors. Nonetheless both of these two sets of mutations from [4,6] presented increased binding selection for human type receptors. This subtle difference between these two sets of mutations, examined in two separate experiments, highlighted the benefits of our approach. The findings of this report offered new insight into the functions of the

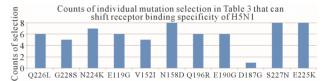
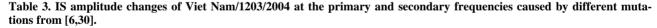


Figure 2. Counts of individual mutation selection in Table 3.



Mutations (H5 numbering)	Amplitude at primary frequency F(0.076) Amplitude at secondary frequency F(0.236)	
	7.6260 (Base line)	5.5980 (Base line)
N224K E119G V152I N158D D187G E190G Q196R S227N	6.0427↓	5.6903 ↑
E119G V152I N158D E190G Q226L N224K E255K	6.7311↓	6.0592 ↑
V152I E190G S227N G228S E255K	8.0359 ↑	6.2810 ↑
E119G V152I N158D E190G S227N E255K	6.6550↓	6.4433 ↑
E119G V152I N158D E190G Q196R N224K Q226L S227N G228S E255K	6.7573↓	6.6419 ↑
N158D Q196R N224K Q226L S227N E255K	6.3615↓	6.6555 ↑
N158D E190G Q196R N224K Q226L S227N G228S E255K	6.7905↓	6.6581 ↑
E119G N158D E190G Q196R N224K Q226L S227N G228S E255K	6.7938↓	6.6587 ↑
E119G N158D Q196R N224K Q226L S227N G228S E255K	6.8002↓	6.6604 ↑

HA mutations inspected in [4,6], and also identified new groups of HA mutations that displayed similar switch of receptor binding preference to those found in [4,6].

5. Acknowledgements

We thank Houghton College for its financial support.

REFERENCES

- E. C. Claas, A. D. Osterhaus, R. van Beek, J. C. De Jong, G. F. Rimmelzwaan, D. A. Senne, S. Krauss, K. F. Shortridge and R. G. Webster, "Human Influenza A H5N1 Virus Related to a Highly Pathogenic Avian Influenza Virus," *Lancet*, Vol. 351, No. 9101, 1998, pp. 472-477. doi:10.1016/S0140-6736(97)11212-0
- [2] K. Subbarao, A. Klimov, J. Katz, H. Regnery, W. Lim, H. Hall, M. Perdue, D. Swayne, C. Bender, J. Huang, M. Hemphill, T. Rowe, M. Shaw, X. Xu, K. Fukuda and N. Cox, "Characterization of an Avian Influenza A (H5N1) Virus Isolated from a Child with a Fatal Respiratory Illness," *Science*, Vol. 279, 1998, pp. 393-396. doi:10.1126/science.279.5349.393
- [3] M. Imai and Y. Kawaoka, "The Role of Receptor Binding Specificity in Interspecies Transmission of Influenza Viruses," *Current Opinion in Virology*, Vol. 2, No. 2, 2012, pp. 160-167. doi:10.1016/j.coviro.2012.03.003
- [4] S. Herfst, E. J. Schrauwen, M. Linster, S. Chutinimitkul, E. de Wit, *et al.*, "Airborne Transmission of Influenza A/H5N1 Virus between Ferrets," *Science*, Vol. 336, No. 6088, 2012, pp. 1534-1541. <u>doi:10.1126/science.1213362</u>
- [5] G. Neumann, C. A. Macken, A. I. Karasin, R. A. M. Fouchier and Y. Kawaoka, "Egyptian H5N1 Influenza Viruses— Cause for Concern?" *PLoS Pathogens*, Vol. 8, No. 11, 2012, p. e1002932. <u>doi:10.1371/journal.ppat.1002932</u>
- [6] M. Imai, T. Watanabe, M. Hatta, S. C. Das, M. Ozawa, et al., "Experimental Adaptation of an Influenza H5 HA Confers Respiratory Droplet Transmission to a Reassortant H5 HA/H1N1 Virus in Ferrets," *Nature*, Vol. 486, 2012, pp. 420-428.
- [7] R. B. Gao, B. Cao, Y. W. Hu, Z. J. Feng, *et al.*, "Human Infection with a Novel Avian-Origin Influenza A (H7N9) Virus," *The New England Journal of Medicine*, Vol. 368, No. 20, 2013, pp. 1888-1897. doi:10.1056/NEJMoa1304459
- [8] Y. Chen, W. F. Liang, S. G. Yang, N. P. Wu, H. N. Gao, et al., "Human Infections with the Emerging Avian Influenza A H7N9 Virus from Wet Market Poultry: Clinical Analysis and Characterisation of Viral Genome," *The Lancet*, 2013, Online.
- [9] W. Hu, "Receptor Binding Specificity and Sequence Comparison of a Novel Avian-Origin H7N9 Virus in China," *Journal of Biomedical Science and Engineering*, Vol. 6, No. 5, 2013, pp. 533-542.
- [10] W. Hu, "Analysis of Correlated Mutations, Stalk Motifs, and Phylogenetic Relationship of the 2009 Influenza A Virus Neuraminidase Sequences," *Journal of Biomedical Science and Engineering*, Vol. 2, No. 7, 2009, pp. 550-558. doi:10.4236/jbise.2009.27080

- [11] W. Hu, "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*, Vol. 3, No. 2, 2010, pp. 1-12.
- [12] W. Hu, "Novel Host Markers in the 2009 Pandemic H1N1 Influenza A Virus," *Journal of Biomedical Science and Engineering*, Vol. 3, No. 6, 2010, pp. 584-601. doi:10.4236/jbise.2010.36081
- [13] W. Hu, "Nucleotide Host Markers in the Influenza A Viruses," *Journal of Biomedical Science and Engineering*, Vol. 3, No. 7, 2010, pp. 684-699. doi:10.4236/jbise.2010.37093
- [14] W. Hu, "Identification of Highly Conserved Domains in Hemagglutinin Associated with the Receptor Binding Specificity of Influenza Viruses: 2009 H1N1, Avian H5N1, and Swine H1N2," *Journal of Biomedical Science and Engineering*, Vol. 3, No. 2, 2010, pp. 114-123. doi:10.4236/jbise.2010.32017
- [15] W. Hu, "Quantifying the Effects of Mutations on Receptor Binding Specificity of Influenza Viruses," *Journal of Biomedical Science and Engineering*, Vol. 3, No. 3, 2010, pp. 227-240. doi:10.4236/jbise.2010.33031
- [16] W. Hu, "Subtle Differences in Receptor Binding Specificity and Gene Sequences of the 2009 Pandemic H1N1 Influenza Virus," Advances in Bioscience and Biotechnology, Vol. 1, No. 4, 2010, pp. 305-314. doi:10.4236/abb.2010.14040
- [17] W. Hu, "Correlated Mutations in the Four Influenza Proteins Essential for Viral RNA Synthesis, Host Adaptation, and Virulence: NP, PA, PB1, and PB2," *Natural Science*, Vol. 2, No. 10, 2010, pp. 1138-1147. doi:10.4236/ns.2010.210141
- [18] D. King, Z. Miller, W. Jones and W. Hu, "Characteristic Sites in the Internal Proteins of Avian and Human Influenza Viruses," *Journal of Biomedical Science and Engineering*, Vol. 3, No. 10, 2010, pp. 943-955. doi:10.4236/jbise.2010.310125
- [19] W. Hu, "Highly Conserved Domains in Hemagglutinin of Influenza Viruses Characterizing Dual Receptor Binding," *Natural Science*, Vol. 2, No. 9, 2010, pp. 1005-1014. doi:10.4236/ns.2010.29123
- [20] W. Hu, "Host Markers and Correlated Mutations in the Overlapping Genes of Influenza Viruses: M1, M2; NS1, NS2; and PB1, PB1-F2," *Natural Science*, Vol. 2, No. 11, 2010, pp. 1225-1246. <u>doi:10.4236/ns.2010.211150</u>
- W. Hu, "Receptor Binding Specificity and Origin of 2009 H1N1 Pandemic Influenza Virus," *Natural Science*, Vol. 3, No. 3, 2011, pp. 234-248. doi:10.4236/ns.2011.33030
- [22] W. Hu, "New Mutational Trends in the HA Protein of 2009 H1N1 Pandemic Influenza Virus from May 2010 to February 2011," *Natural Science*, Vol. 3, No. 5, 2011, pp. 379-387. doi:10.4236/ns.2011.35051
- [23] W. Hu, "Characterization of Asian and North American Avian H5N1," American Journal of Molecular Biology, Vol. 1, No. 2, 2011, pp. 52-61. doi:10.4236/ajmb.2011.12007
- [24] W. Hu, "Molecular Features of Highly Pathogenic Avian

and Human H5N1 Influenza A Viruses in Asia," *Computational Molecular Bioscience*, Vol. 2, No. 2, 2012, pp. 45-59.

- [25] W. Hu, "Molecular Determinants for Receptor Binding in Hemagglutinin Protein of 2009 Pandemic H1N1," Proceedings of 6th International Conference Bioinformatics and Biomedical Engineering (ICBBE2012), Shanghai, 17-20 May 2012, pp. 629-632.
- [26] K. Katoh, K. Kuma, H. Toh and T. Miyata, "MAFFT Version 5: Improvement in Accuracy of Multiple Sequence Alignment," *Nucleic Acids Research*, Vol. 33, 2005, pp. 511-518. doi:10.1093/nar/gki198
- [27] I. Cosic, "The Resonant Recognition Model of Macromolecular Bioreactivity, Theory and Application," Birkhauser Verlag, Berlin, 1997. <u>doi:10.1007/978-3-0348-7475-5</u>
- [28] V. Veljkovic, H. L. Niman, S. Glisic, N. Veljkovic, V.

Perovic and C. P. Muller, "Identification of Hemagglutinin Structural Domain and Polymorphisms Which May Modulate Swine H1N1 Interactions with Human Receptor," *BMC Structural Biology*, Vol. 9, 2009, p. 62. doi:10.1186/1472-6807-9-62

- [29] V. Veljkovic, N. Veljkovic, C. P. Muller, S. Müller, S. Glisic, V. Perovic and H. Köhler, "Characterization of Conserved Properties of Hemagglutinin of H5N1 and Human Influenza Viruses: Possible Consequences for Therapy and Infection Control," *BMC Structural Biology*, Vol. 7, 2009, pp. 9-21.
- [30] L. M. Chen, O. Blixt, J. Stevens, A. S. Lipatov, C. T. Davis, B. E. Collins, N. J. Cox, J. C. Paulson and R. O. Donis, "*In Vitro* Evolution of H5N1 Avian Influenza Virus toward Human-Type Receptor Specificity," *Virology*, Vol. 422, No. 1, 2012, pp. 105-113. doi:10.1016/j.virol.2011.10.006