INTERNATIONAL JOURNAL OF CURRENT RESEARCH IN CHEMISTRY AND PHARMACEUTICAL SCIENCES

(p-ISSN: 2348-5213: e-ISSN: 2348-5221)
www.ijcrcps.com

Research Article

PREDICTION OF NEW CONSERVED EPITOPES IN PROTEIN 3D MODEL TO NEUTRALIZE INFLUENZA A VIRUS STRAIN H3N2 CIRCULATING IN BANGLADESH

MODHUSUDON SHAHA[1], MOHAMMAD ARIFUL ISLAM[1], ABUL BMMK ISLAM[2], MD. FIROZ AHMED[3], MD. MAJIBUR RAHMAN[1], SABITA REZWANA RAHMAN[1]*

[1]Department of Microbiology, University of Dhaka, Dhaka-1000, Bangladesh.
[2]Department of Genetic Engineering and Biotechnology, University of Dhaka, Dhaka-1000, Bangladesh.
[3]Department of Microbiology, Jahangirnagar University, Savar, Dhaka, Bangladesh.

Corresponding Author: sabita.rahman223@gmail.com

Abstract

Influenza A virus H3N2 strain are simultaneously prevalent among human and bird population sometimes causing epidemic besides seasonal infections. This virus causes a substantial amount of morbidity and mortality in different parts of the world especially in the developing countries like Bangladesh. Although vaccine was developed against this virus strain, the activity of vaccine fails frequently due to accumulation of mutations in hemagglutinin (HA) gene. Here, we suggested an effective protein model with conserved epitope-based vaccine design which might be capable to neutralize that strain. After partial sequencing of HA gene of H3N2 isolated from Bangladeshi patients, we observed several mutations at different positions, some of which lies in existing epitopes or active sites which indicates possible resistance to existing vaccines, although experimental confirmation is needed. However, multiple sequence alignment with previously reported Bangladeshi and vaccine strains we have identified several conserved regions and some of these also fall in predicted and experimentally determined epitopes which may be useful for a new and potential vaccine development. We predicted a protein 3D model with the sequenced Bangladeshi H3N2 strain and identified conserved highly immunogenic epitopes and active sites in it which may be further evaluated experimentally for developing vaccine against it.

Keywords: Conserved epitopes, H3N2 strain, Hemagglutinin, Influenza A virus, Mutations.

Introduction

Influenza virus causes seasonal epidemics as well as occasional pandemics with substantial morbidity and mortality where seasonal influenza infections have been trying to be prevented by vaccination but not pandemic viruses (Xu et al., 2013). Different studies reported that Influenza A viruses (IAV) are responsible for millions of deaths causing acute respiratory illness (Chow et al., 2006; Mallia and Johnston, 2007; Tate et al., 2014). Being a negative stranded RNA virus, influenza A virus infects approximately one-fifth of human population per annum (Xiu et al., 2008). Treatment or vaccination against IAV is difficult for frequent changes of its genome through antigenic drift and shift. The two membrane bound surface glycoproteins, the neuraminidase (NA) and the hemagglutinin (HA) are expressed by IAV which are responsible for the variation of genome and helps to escape from the existing antiviral drugs and vaccines (Shen et al., 2013; Tate et al., 2014). Currently, the strategies against IAV in development include broadly neutralizing antiviral therapy, universal influenza vaccine and small-molecule inhibitors (Du et al., 2012; Ekiert and Wilson, 2012; Gilbert, 2012). Antigenic drift and shift occur in both NA and HA genes (Plotkin and Dushoff, 2003). As NA is less prevalent on virion surface and possess limited interaction with neutralizing antibodies, it is considered less relevant (Plotkin and Dushoff, 2003). Hence, HA proteins are mostly concentrated for the study of antigenic variability (Bush et al., 1999a; Bush et al., 1999b; Lee and Chen, 2004; Smith et al., 2004).

HA plays important functions in viral cell cycles like attachment of virus to sialic acid receptor of host cell membrane and internalization into the late endosome,

thus determining the host specificity (Skehel and Wiley, 2000; Webster et al., 1992; Wiley and Skehel, 1987). Of about 16 subtypes 3 (H1, H2 and H3) of viruses are known to be responsible for humans pandemics (Sahini et al., 2010). Some regions of HA are frequently drifted that may be responsible to evolve for some other reasons like to enhance post-translational modifications, facilitate protein folding or regulation of secondary as well as tertiary structures. HA of IAV contains different glycosylation sites where glycan can bind and thought to be important part which can contribute to antigenic drift. Number of glycosylation sites has a strong impact on the activity of HA, whether this will activated or deregulated (Wiley et al., 1981). Addition of glycan in HA protein sometime have selective advantage like hindering the binding of neutralizing antibody (Ab) to the epitopes of antigen. Epitopes, the conserved regions in different strains of influenza viruses which are recognized by neutralizing antibodies are elementary part to develop a proper treatment against influenza viruses (Iba et al., 2014). The globular head of HA contains epitopes which are well characterized and clustered in A, B, C, D and E sites of HA in H3 viruses (Underwood, 1982; Wiley et al., 1981). HA gene of influenza virus can acquire mutation readily and therefore make themselves resistance to these antibodies. Thus, preventive vaccination is the most preventive measure to control influenza A viruses and to remain effective, the strain selected for vaccine development need to be changed almost every year (Salzberg, 2008).

Herein, we determined the phylogenic relationship of IAV HA gene and amino acid variation comparing with native and foreign strains. We analyzed the conserved regions found in HA of IAV and checked its rate of mutation in these regions compared to others. We also develop a putative protein model which could be a target for designing a potential vaccine to prevent not only the strains of Bangladesh but also related all strains.

Materials and Methods

Sample collection and processing

To perform this study 400 nasopharyngeal swab samples were collected from the patients who were suffering from influenza like symptoms maintaining aseptic condition. The samples were transferred to the clinical laboratory using VTM and kept at -80^{o}C until use. This study was performed in the Clinical Laboratory of the Department of Microbiology, University of Dhaka and samples were collected after containing an informed written consent from the participants of this experiment. The study protocol was approved by the ethical committee of Dhaka Medical College, Dhaka-1000, Bangladesh (reference number: DMC-MEU/ECC/214/17).

Preparation of viral DNA and sequencing

Viral RNA was extracted from sample fluids by pathogen kit (Stratec molecular, Berlin, Germany) according to manufacturer's instruction. The extracted RNA was subjected to one step real-time reverse transcriptase PCR (rRT-PCR) using primers, forward: GAC CRA TCC TGT CAC CTC TGA C, reverse: AGG GCA TTY TGG ACA AAK CGT CTA and probe: TGC AGT CCT CGC TCA CTG GGC ACG. Amplification of HA gene by the above primers was done using a 20µl reaction mixture containing buffer (2x) 12.5µl, enzyme reverse transcriptase (25µM) 1µl, forward primer (40µM) 0.5µl, reverse primer (40µM) 0.5µl, probe (5µM) 0.5µl, enhancer 1µl and dH_2O. The thermal conditions, used for this amplification was 30 minutes reverse transcription at 50^{o}C, 15 min initial PCR activation step at 95^{o}C, 29 cycles of reaction containing denaturation 15 seconds, annealing 30 sec, extension 1 min 15 sec at 94^{o}C, 50^{o}C and 72^{o}C respectively along with a final extension of 10 min at 72^{o}C. Then the positive products of RT-PCR were purified using ExoSAP-IT (CA, USA) according to the manufacturer's instruction. Sequencing of HA gene was performed using a primer TGTAAAACGACGGCCAGT by BigDye®Terminator v3.1 Cycle Sequencing Kit (California) according to the instruction of manufacturer.

Phylogenetic analysis using HA gene of IAV

The obtained sequences were aligned by Clustal Omega method (Larkin et al., 2007). The reference sequences for this alignment were local and vaccine HA sequences collected from NCBI GenBank and Influenza Virus Database (IVR) (Bao et al., 2008) (Supplementary Table S1. Phylogenetic tree was formed by the bootstrap, distance and Neighbor-Joining (NJ) method to carry out phylogenetic analysis.

Comparison of experimental sequences with different conserved regions and epitopes

Different conserved regions and epitopes against B cell, T cell and MHC types in HA gene products were compared with the experiment sequences using MEGA 5 tool (Tamura et al., 2011) with Clustal algorithm. The epitopes and sequences required for this study were collected from Influenza Research Database (IRD) (Squires et al., 2012) and Influenza Virus Resource Database (IVR) (Bao et al., 2008) respectively.

Prediction of protein model

We used I-Tasser tool (Zhang, 2008) to model protein using first some 229 amino acid from our sequence and rest from closely matched sequence (accession no. ACZ05788.1). Protein 3D structure is visualized and epitopes are marked using UCSF Chimera (Pettersen et al., 2004). Epitopes present in our sequences along with surface accessibility were predicted using IEDB server (Kim et al., 2012).

Active sites or pockets of the protein were determined by using fPocket online tool (Schmidtke et al., 2010). Different binding sites, solvent accessibility and disulphide bridges were analyzed by PredictProtein server (Rost et al., 2004).

Sequence deposition

Partial CDS of HA sequences is deposited in GenBank under accession numbers: KP222533 (IAV_542), KP222534 (IAV_541), KP222535 (IAV_550), KP222536 (IAV_558) and KP238097 (IAV_473).

Results

Subtyping and sequencing of HA gene

HA gene, the antigenic determinant of IAV, was sequenced and BLAST searched to find the subtypes of that sequences. Of 400 samples, 48 (12%) was found positive to influenza A virus (IAV) by real time reverse transcriptase PCR (rRT-PCR). Of them, 5 samples (IAV_473, IAV_541, IAV_542, IAV_550 and IAV_558) were subjected to hemagglutinin (HA) gene sequencing and all of them were found to contain identical nucleotides although differ in length due to partial coding sequence (CDS) (data not shown). As all of the sequences were identical, IAV_542 (NCBI accession no: KP222533) – the largest partial CD was selected for further bioinformatics analysis. NCBI-BLAST search with IAV_542 revealed the subtype H3N2 (data not shown).

Evolutionary relationship of our IAV with other circulating IAVs of Bangladesh

The nucleotide sequence of IAV_542 was subjected to nucleotide BLAST search against NCBI database to observe the sequence similarity with the existing sequences in the NCBI GenBank and found that accession number KF598719 as the most closely matched sequence with 99.85% identity (Supplemental **Table S1**). The NCBI-BLAST result showed that only one nucleotide at position 209 (G209A) was mutated in our sequence.

Table 1: Observed mutations in IAV_542 comparing to existing epitopes collected from Influenza Virus Resource database.

Epitopes from IVR database	Observed Sequence in IAV_542	Mutation position
ALNVTMPNNEKFDKLYI	ALNVTMPNNEQFDKLYI	K189Q
ASGRVTVSTKRSQQTV	SSGRITVSTKRSQQTV	A214S, V218I
ATELVQSSSTGRICDS	ATELVQSSSTGEICDN	R66E, S70N
CKRRSNNSFFSRLNWLT	CIRRSNSSFFSRLNWLT	K156I, N161S
CYPYDVPDYASLRSLVASSGTLEFINEDFNWT	CYPYDVPDYASLRSLVASSGTLEFNNESFNWT	I137N, D140S
DQIEVTNATELVQSSSTGGI	DQIEVTNATELVQSSSTGGI	G66E
DYASLRSLVASSGTLEFINEGFNWTGVTQNGGSSAC	DYASLRSLVASSGTLEFNNESFNWTGVTQNGGSSAC	I137N, G140S
ELVQSSSTGRICDSPHQILD	ELVQSSSTGEICDNPHQILD	R66E, S70N
GTLVKTITNDQIE	GTIVKTITNDQIE	L41I
HHAVPNGTL	HHAVPNGTI	L41I
HHAVSNGTLVKTITNDQIEV	HHAVSNGTIVKTITNDQIEV	L41I
KSEYKYPALNVTMPNN	HLNFKYPALNVTMPNN	K72H, S73L, E74N, Y75F
KSFFSRLNWLTHLK	SSFFSRLNWLTHLN	K161S, K174N
LVKTITNDQIEVTNATELVQSSSTGRICDSPHRIL	IVKTITNDQIEVTNATELVQSSSTGEICDNPHQIL	L41I, R66E, S70N, R73Q
NFDKLYIWG	NFDKLYIWG	N189Q
NVTMPNNEKFDKLYIWGV	NVTMPNNEKFDKLYIWGV	N189Q
QIEVTNATELVQSSSTGRIC	QIEVTNATELVQSSSTGEIC	R66E
QKLPGNDNSTATLCLGHHAVPNGTLVKTITNDQIE	QKLPGNDNSTATLCLGHHAVPNGTIVKTITNDQIE	L41I
RLNWLTHLK	RLNWLTHLK	K174N
SACKRRSNKSFFSRLNWLTH	SACIRRSNSSFFSRLNWLTH	K156I, K161S
SFFSRLNWLHKSEYKY	SFFSRLNWLTHLNFKY	H171T, K172H, S173L, E174N, Y175F

SLYAQASGRITVSTKRS	FLYAQSSGRITVSTKRS	S209F, A214S
SSCKRRSNNSFFSRLNWLTH	SACIRRSNSSFFSRLNWLTH	S154A, K156I, N161S
VPNGTLVKTITNDQIE	VPNGTIVKTITNDQIE	L41I
VPNGTLVKTITNDQIEVTNAT	VPNGTIVKTITNDQIEVTNAT	L41I
VQSSSTGGICDSPHQIL	VQSSSTGEICDNPHQIL	G66E, S70N
YPALNVTMPNNGKFDKLYIWGVHHPS TDRDQTS	YPALNVTMPNNEQFDKLYIWGVHHPG TDKDQIF	G188E, K189Q, S202G, R205K, T208I, S209F

The sequence of IAV_542 was analyzed using EMBOSS Transeq tool of EMBL-EBI server (McWilliam et al., 2013) to find all six frames of in silico translation of proteins and subjected to protein BLAST (PSI) against NCBI protein database to find the correct frame (ORF) of the protein translation. The IAV_542 translated protein sequence was then aligned with HA protein of H3N2 of previously reported different Bangladeshi sequences **(**Supplementary **Table S2)** collected from Virus Resources Database (IVR) (Bao et al., 2008). We observed similarities of our partial HA protein sequence with other Bangladeshi sequences, however, there are several amino acid mutations like G66E (amino acid E at 66 position of HA protein instead of G), S70N, K156I, N/K161S, K174N, K/E189Q, S/N205K, S209F and A214S (**Figure 1a**). The above mutations were thought to have possible effect on structural and functional activity of HA protein which may be a significant cause of resistance to different drugs. Phylogenetic tree of above alignment revealed that our sequence was most closely related with ACC67740 with a similarity of 71% (**Figure 1b**).

Figure 1a: Alignment IAV_542 (KP222533) with different Bangladeshi strains. Matched amino acids were presented as dot and mismatched were shown in symbols. Alignment was done by MEGA 5 software and visualized using Jalview software.

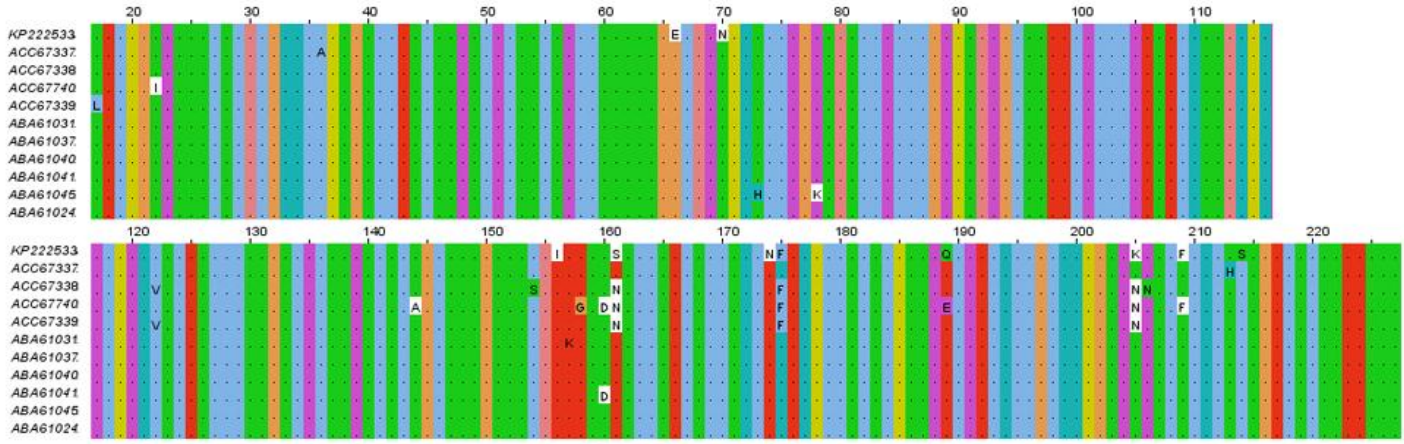
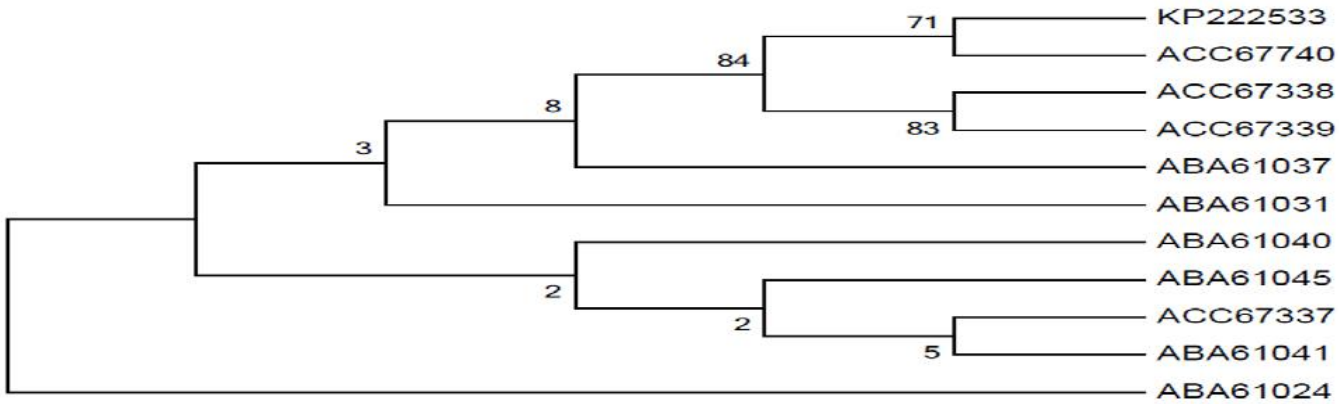


Figure 1b: Phylogenetic analysis of IAV_542 (KP222533) with different existing Bangladeshi strains collected from IVR database.



Mutations are observed in protein sequences of our IAV and current vaccine sequences

Multiple sequence alignment of our sequence (IAV_542) with different vaccine strains **(**Supplementary **Table S2)** used for developing existing vaccines are collected from IVR (Bao et al., 2008) database revealed several mutations, but the important one is at position 70 (S70N) (**Figure 2a**). This mutation position lies on 4 existing epitopes, which are ELVQSSSTGRICD**S**PHQILD, LVKTITNDQIEVTNATELVQSSSTGRICD**S**PHRIL,ATE LVQ SSSTGRICD**S** and VQSSSTGGICD**S**PHQIL **(Table 1)**. So, mutation S70N might cause those epitopes to be ineffective and might not be recognized by antibody or vaccine against that strain of IAV and thus may lead to vaccine resistance. This demands detail experimentation on the possible vaccine resistance due to our identification of such mutation. Vaccine strain ACS71642 is highly related (with 0.01 evolutionary distance) but not identical with our sequence that was confirmed by phylogenetic tree analysis (**Figure 2b**). As we did not find any vaccine strain which is totally matched with our sequences, it may be required to consider developing a new protein model which could be effective to decrease the burden of IAV, especially in slums of Dhaka city, Bangladesh.

Figure 2a: Alignment IAV_542 (KP222533) with vaccine strains. Amino acids were shown as dot and mismatched are presented in symbols. Alignment was done by MEGA 5 software and visualized using Jalview software.

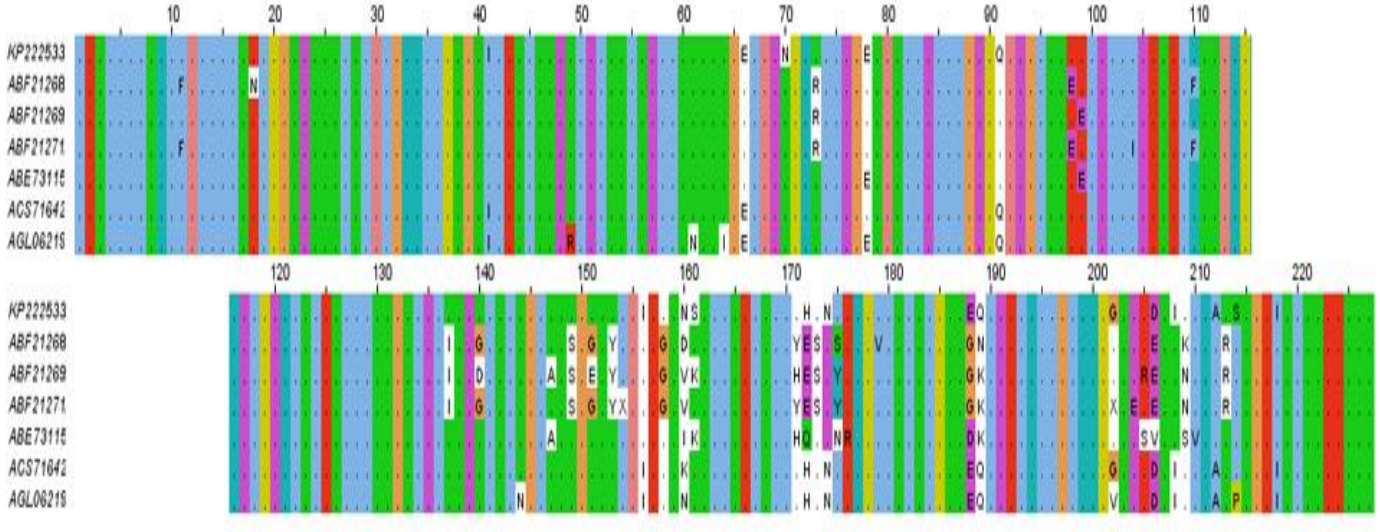


Figure 2b: Phylogenetic analysis of IAV_542 (KP222533) with different vaccine strains collected from IVR database.

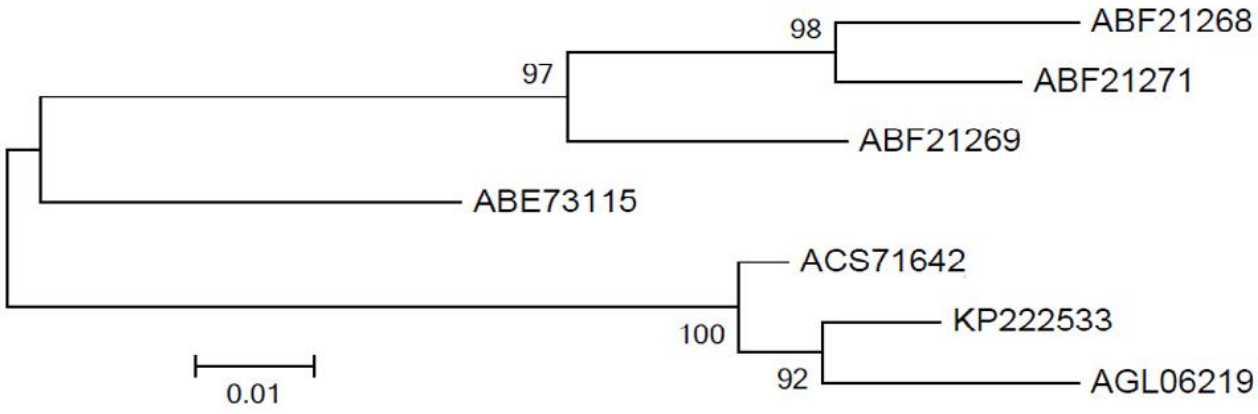
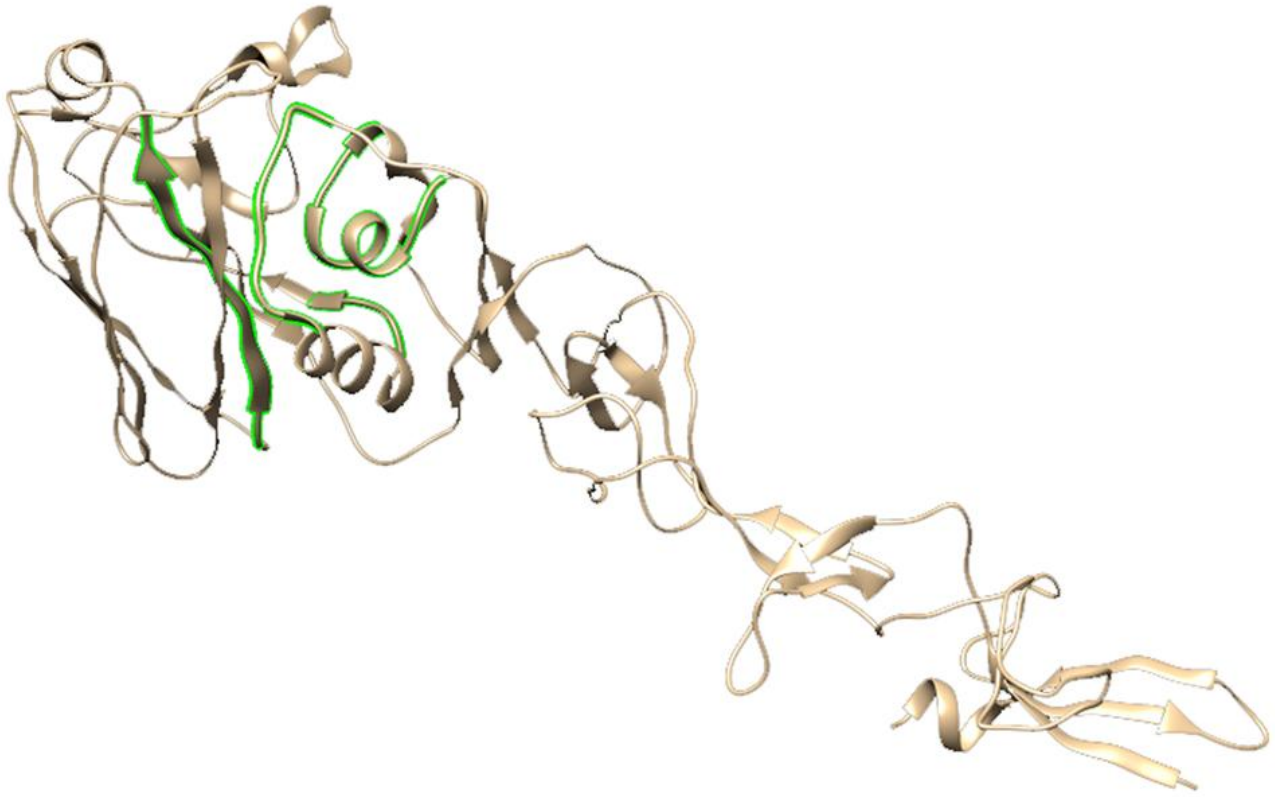


Figure 3: Protein 3D model of IAV_542 showing the predicted epitopes common among the conserved regions in sequence alignment with other Bangladeshi and vaccine strain (Supplementary Table S2). The highlighted regions with Green color demonstrates the predicted epitopes marked using UCSF Chimera software.



Multiple mutations are possible in known epitopes of IAVs in our sequence

Epitopes are the vital part of an organism to be identified by drugs and vaccines. More than 200 existing experimentally validated epitopes were collected from IRD (Squires et al., 2012) to be compared with our HA sequence. Analysis of our sequence with these existing epitopes revealed that 27 epitopes that may sighted in our sequence were mutated somehow (**Table 1**). These mutations in epitopes, which is the determinant of vaccine efficacy, might made the responsible virus resistant to different existing vaccine, and thus, demanding to design a new protein model that will be able to neutralize IAV (H3N2) circulating in Bangladesh.

Table S1: NCBI BLAST (BLASTn) search result of IAV_542. (Result shown only till one substitution).

Accession	Identity (%)	Coverage (%)	Substitution	E-value
KF598719.1	99.85	100	1	0
KF598718.1	99.85	100	1	0
KF598713.1	99.85	100	1	0
KF598714.1	99.85	100	1	0
KF598712.1	99.85	100	1	0
KF598708.1	99.85	100	1	0
KF598703.1	99.85	100	1	0
KF598705.1	99.85	100	1	0
KF598702.1	99.85	100	1	0
KF586764.1	99.85	100	1	0
KF586763.1	99.85	100	1	0
KF586760.1	99.85	100	1	0
KF586742.1	99.85	100	1	0
KF586740.1	99.85	100	1	0

KF586729.1	99.85	100	1	0
KC893081.1	99.85	100	1	0
KC892616.1	99.85	100	1	0
KC892303.1	99.85	100	1	0
CY120861.1	99.85	100	1	0

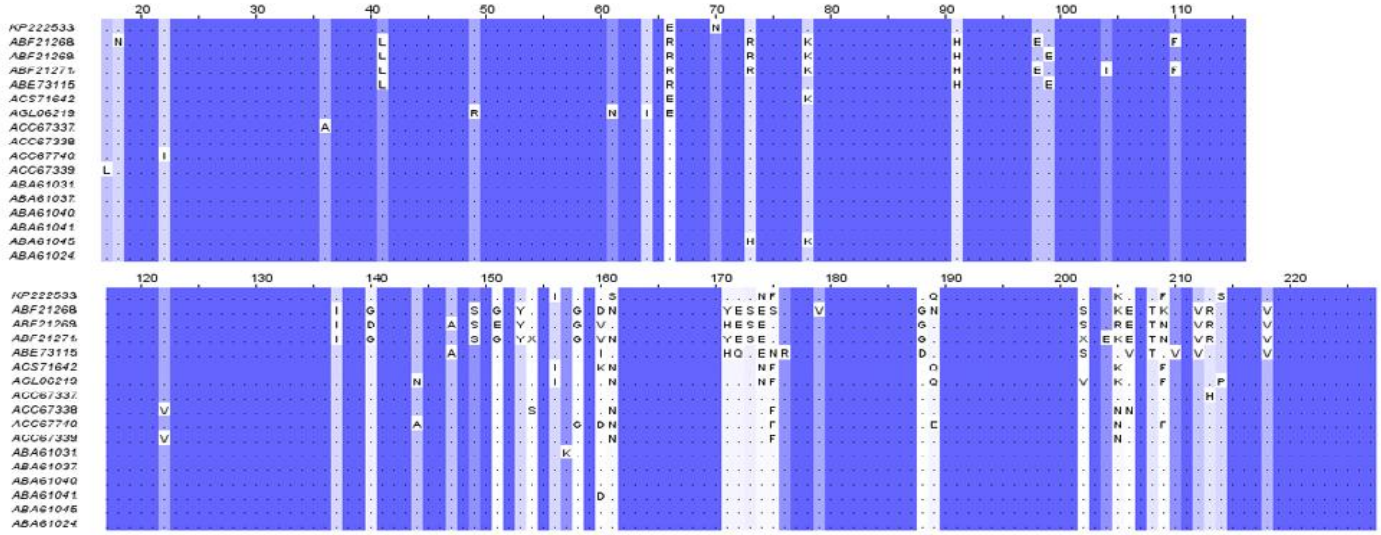
Table S2: Bangladeshi and "vaccine" IAV HA sequence information that were extracted from NCBI's Influenza Virus Resources (IVR) database.

Sequence type	Accession no.	Country	Year
Non-vaccine strain	ACC67740	Bangladesh	2006
	ACC67338	Bangladesh	2005
	ACC67339	Bangladesh	2005
	ABA61037	Bangladesh	2003
	ABA61031	Bangladesh	2003
	ABA61040	Bangladesh	2003
	ABA61045	Bangladesh	2003
	ACC67337	Bangladesh	2003
	ABA61041	Bangladesh	2003
	ABA61024	Bangladesh	2002
Vaccine strain	ABF21268	Bangkok	1979
	ABF21269	Beijing	1989
	ABF21271	Leningrad	1986
	ABE73115	Moscow	1999
	ACS71642	Perth	2009
	AGL06219	Texas	2012

Conserved locations among our and other strains could be targeted for epitope based vaccine development

Multiple sequence alignment of our experiment sequence with previously described Bangladeshi sequences and vaccine strain sequences revealed some conserved regions which could be the targets for designing vaccine. The observed conserved regions in both existing Bangladeshi strains, available vaccine strains and our experiment sequence are amino acid number 26 to 36 (STATLCLGHHAV), 51 to 61 (IEVTNATELVQ), 80 to 91 (N**CTLIDALLGDP**), 112 to 122 (SN**CYPYDVPDY**), 124 to 137 (SLRSLVASSG**TLEF**), 164 to 171 (FFSRLNWL), 191 to 202 (F**DKLYIWGVHHP**) (Supplementary **Figure S1).** Some of these conserved sites are potential immunogenic epitopes (shown in bold format) as predicted by IEDB server (Kim et al., 2012).

Figure S1: Evaluation of conserved regions in IAV_542 by aligning both Bangladeshi and Vaccine strains (Supplementary Table S2). Amino acids were shown as dot and mismatched are in symbols. Alignment was done by MEGA 5 software and visualized using Jalview software.



Predicted 3D model of the HA sequence with epitopes and active sites

Protein model of the compiled sequence (our partial HA sequence and rest from closely matched NCBI sequence ACZ05788.1) by I-TASSER bioinformatics server (Zhang, 2008) revealed a 3D structure in which we delineated the conserved predicted epitopes conformations (**Figure 3)**. Analysis of our partial HA protein sequence by fPocket online tool (Schmidtke et al., 2010) revealed 11 pockets in predicted tertiary protein structure (Supplementary **Figure S2**). The antigenicity and surface accessibility of the protein has been delineated in supplementary **Figure S3.** The developed protein by our sequence revealed that the partial protein contains 61.67% loop, 16.74% helix and 21.59% strand. It also analyzed the solvent accessibility where 54.19% were exposed, 39.65% were buried and the rest were intermediate. The disulphide bridges were found between amino acid 30 and 80, 68 and 92, 113 and 155 regions. It also revealed different binding sites at 21-23, 72, 78, 96, 98, 110-112, 117, 149, 156-159, 166, 186, 188, 213, 217 and 225 amino acid positions (Supplementary **Figure S4**). **Figure S2:** Presence of pockets in IAV_542 predicted using fPocket server. Figure A, B, C, D, E, F indicate the pockets of same sequence in different views. Different colors shows eleven pockets in the experiment sequence with a blackish background.

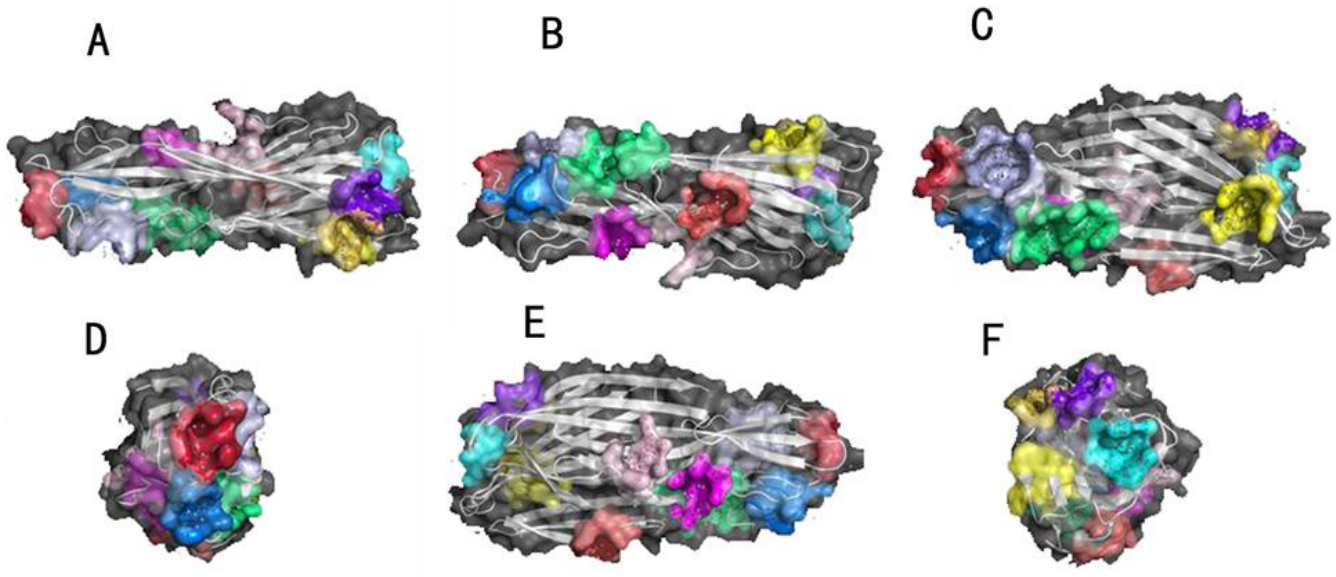


Figure S3: Prediction of (a) antigenicity and (b) surface accessibility of IAV_542 by IEDB server. The red line is the cut-off value; yellow and green color portion indicate above antigenicity (A), surface accessibility (B) cut-off and below antigenicity (A), surface accessibility (B) cut-off respectively.

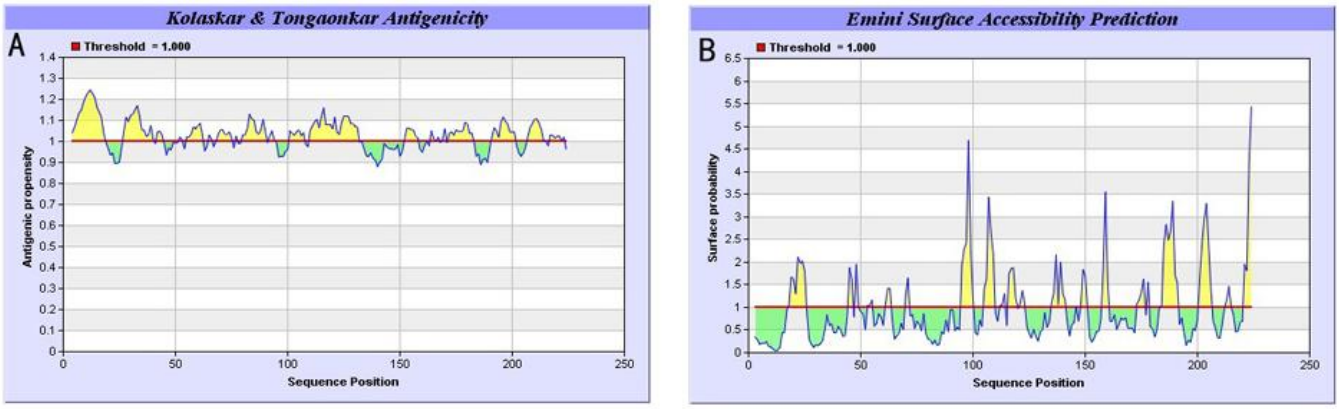
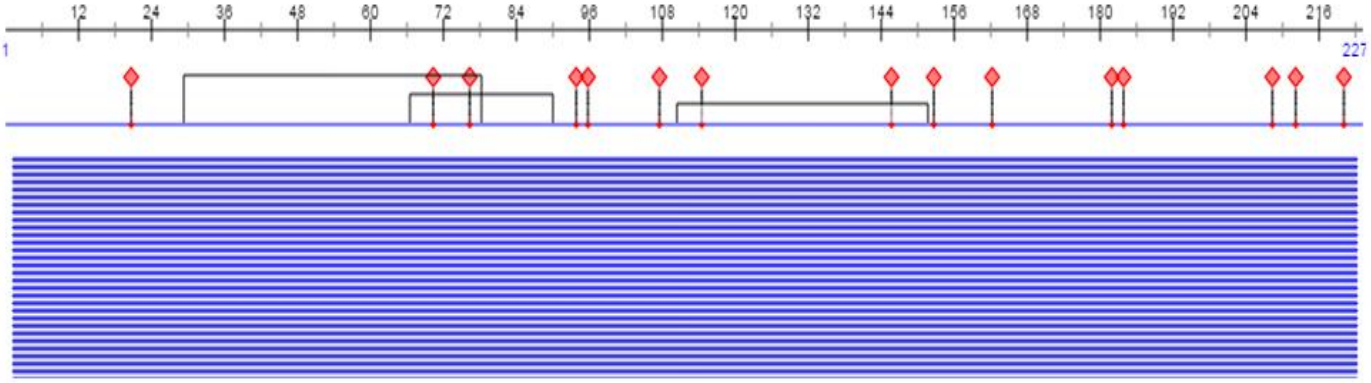


Figure S4: Protein binding sites and disulphide bridges of IAV_542 delineated by PredictProtein server. The scale on the top indicates amino acid numbers, the red color boxes are binding sites, the bridges between two points indicate disulphide bridge, the horizontal rectangular box contains different blue lines showing the alignment with other NCBI sequences.



Discussion

Seasonal influenza virus appears frequently with a significant reassortment of genome where existing drugs and vaccine are sometimes ineffective (Ghedin et al., 2005). Sometimes these reassortments by antigenic drift and shift make the virus to be epidemic although the virus origin is same to one previously determined. The mentioned antigenic drift and shift may be also initiated by frequent mutations which together can cause significant change in the common epitopes existing in the genome normally, abnormality in the conserved regions that were previously selected for vaccine development (Nicholls, 2006). Thus, we attempted to analyze phylogenetic analysis with native (Bangladeshi) and vaccine strains circulating in Bangladesh to observe the genetic and functional variability which were thought to have resistance to existing drugs and vaccine. In this study, above 200 experimental epitopes of IAV-H3N2 were collected from IVR (Bao et al., 2008) and compared with our sequences for mutations and developed a protein model that could be effective to neutralize IAV-H3N2 circulating in Bangladesh.

In this study, the prevalence of IAV-H3N2 was 12% which is higher than those of previous year 2012, 7% documented by Fally (Fally et al., 2012). The reason behind it may be the antigenic drift in the genome of IAV-H3N2 circulating in Bangladesh which have exposed to frequent mutations (Matrosovich et al., 2000; Shil et al., 2011). Our experiment sequences were found mutated compared to that of other existing Bangladeshi sequences those were documented before 2006. Mutations at different positions like G64E, S68N, K156I, K/N161S, K174N, K/E189Q, S/N205K and S209F when comparing with different Bangladeshi IAV-H3N2 sequences may affect the sequence likely to be distorted from the close phylogenetic relationship. Among above mutations, K/E189Q and S209F are responsible for changing the hydrophilic characteristics to hydroneutral and hydrophobic respectively which enable the protein to be insoluble (Phillips, 2013). The above mutations might have possible effect on structural and functional activity of HA protein which may be a significant cause of resistance to different drugs.

Comparison analysis with different vaccine strains of IAV-H3N2 revealed the phylogenetic analysis that our sequence was closely matched with ACS71642 strain isolated from Texas with an evolutionary distance of 0.0111859, which suggests the origin of the experiment sequence may be Texas and carried somehow in Bangladesh. The alignment with the vaccine strains showed one amino acid mutation at 70 no. position S70N, which may cause a change in HA protein and might lead to vaccine resistance.

Aligning and comparing with more than 200 epitopes of IAV-H3N2 strain revealed as high as 40 epitopes variations which contains mutations of 1 to 7 per epitope as shown in Figure 3. These mutations in epitopes, which is the determinant of vaccine efficacy, might make the responsible virus resistant to different existing vaccine, thus demand to design a new protein model that will be able to neutralize IAV (H3N2) circulating in Bangladesh (Munoz and Deem, 2005).

Comparing with both existing Bangladeshi sequences and vaccine strain sequences showed many conserved regions in our sequences, some of which contain potential epitopes as predicted by IEDB server (Kim et al., 2012) which could be targeted to develop a new and effective vaccine against IAV-H3N2.

The designed protein model by I-Tasser (Zhang, 2008) and visualized by UCSF Chimera software (Pettersen et al., 2004) (Figure 3) showed a 3D structure which delineated the loop, helix and sheet structure along with the highlighting of predicted epitopes. Predictprotein server (Rost et al., 2004) demonstrated solvent accessibility, protein disorder and flexibility, and different protein-protein and protein-nucleotide binding sites. These active sites can be tested for drug/small molecule binding capacity.

In conclusion, we have provided the importance of developing new vaccine targeting HA protein that would prevent influenza A virus H3N2 strain. As mutations in this region frequently generating vaccine resistance within a short interval, our identification of some antigentic conserved regions not only in existing Bangladeshi strains but also in current vaccine strains offers a potential target for new vaccine.

Acknowledgments

This study was funded by Higher Education Quality Enhancement Project (HEQEP) and University Grant Commission (UGC), People's Republic of Bangladesh.

Conflict of interest

Authors have declared no conflict of interest.

References

Bao, Y., P. Bolotov, D. Dernovoy, B. Kiryutin, L. Zaslavsky, T. Tatusova, J. Ostell, and D. Lipman. 2008. "The influenza virus resource at the National Center for Biotechnology Information." J Virol 82(2): 596-601.

Bush, R. M., C. A. Bender, K. Subbarao, N. J. Cox, and W. M. Fitch. 1999a. "Predicting the evolution of human influenza A." Science 286(5446): 1921-5.

Bush, R. M., W. M. Fitch, C. A. Bender, and N. J. Cox. 1999b. "Positive selection on the H3 hemagglutinin gene of human influenza virus A." Mol Biol Evol 16(11): 1457-65.

Chow, A., S. Ma, A. E. Ling, and S. K. Chew. 2006. "Influenza-associated deaths in tropical Singapore." Emerg Infect Dis 12(1): 114-21.

Du, J., T. A. Cross, and H. X. Zhou. 2012. "Recent progress in structure-based anti-influenza drug design." Drug Discov Today 17(19-20): 1111-20.

Ekiert, D. C. and I. A. Wilson. 2012. "Broadly neutralizing antibodies against influenza virus and prospects for universal therapies." Curr Opin Virol 2(2): 134-41.

Fally, M. A., M. Redlberger-Fritz, P. Starzengruber, P. Swoboda, H. P. Fuehrer, E. B. Yunus, W. A. Khan, and H. Noedl. 2012. "Characterization and epidemiology of influenza viruses in patients seeking treatment for influenza-like illnesses in rural Bangladesh." J Postgrad Med 58(4): 242-5.

Ghedin, E., N. A. Sengamalay, M. Shumway, J. Zaborsky, T. Feldblyum, V. Subbu, D. J. Spiro, J. Sitz, H. Koo, P. Bolotov, D. Dernovoy, T. Tatusova, Y. Bao, K. St George, J. Taylor, D. J. Lipman, C. M. Fraser, J. K. Taubenberger, and S. L. Salzberg. 2005. "Large-scale sequencing of human influenza reveals the dynamic nature of viral genome evolution." Nature 437(7062): 1162-6.

Gilbert, S. C. 2012. "Advances in the development of universal influenza vaccines." Influenza Other Respir Viruses 7(5): 750-8.

Iba, Y., Y. Fujii, N. Ohshima, T. Sumida, R. Kubota-Koketsu, M. Ikeda, M. Wakiyama, M. Shirouzu, J. Okada, Y. Okuno, Y. Kurosawa, and S. Yokoyama. 2014. "Conserved neutralizing epitope at globular head of hemagglutinin in H3N2 influenza viruses." J Virol 88(13): 7130-44.

Kim, Y., J. Ponomarenko, Z. Zhu, D. Tamang, P. Wang, J. Greenbaum, C. Lundegaard, A. Sette, O. Lund, P. E. Bourne, M. Nielsen, and B. Peters. 2012. "Immune epitope database analysis resource." Nucleic Acids Res 40(Web Server issue): W525-30.

Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson, and D. G. Higgins. 2007. "Clustal W and Clustal X version 2.0." Bioinformatics 23(21): 2947-8.

Lee, M. S. and J. S. Chen. 2004. “Predicting antigenic variants of influenza A/H3N2 viruses.” Emerg Infect Dis 10(8): 1385-90.

Mallia, P. and S. L. Johnston. 2007. “Influenza infection and COPD.” Int J Chron Obstruct Pulmon Dis 2(1): 55-64.

Matrosovich, M., A. Tuzikov, N. Bovin, A. Gambaryan, A. Klimov, M. R. Castrucci, I. Donatelli, and Y. Kawaoka. 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(18): 8502-12.

McWilliam, H., W. Li, M. Uludag, S. Squizzato, Y. M. Park, N. Buso, A. P. Cowley, and R. Lopez. 2013. “Analysis Tool Web Services from the EMBL-EBI.” Nucleic Acids Res 41(Web Server issue): W597-600.

Munoz, E. T. and M. W. Deem. 2005. “Epitope analysis for influenza vaccine design.” Vaccine 23(9): 1144-8.

Nicholls, H. 2006. “Pandemic influenza: the inside story.” PLoS Biol 4(2): e50.

Pettersen, E. F., T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin. 2004. “UCSF Chimera--a visualization system for exploratory research and analysis.” J Comput Chem 25(13): 1605-12.

Phillips, J. C. 2013. “Hierarchical hydropathic evolution of influenza glycoproteins (N2, H3, A/H3N2) under relentless vaccination pressure.” Cornell University Library (arXiv:1303.4383v1): 1-19.

Plotkin, J. B. and J. Dushoff. 2003. “Codon bias and frequency-dependent selection on the hemagglutinin epitopes of influenza A virus.” Proc Natl Acad Sci U S A 100(12): 7152-7.

Rost, B., G. Yachdav, and J. Liu. 2004. “The PredictProtein server.” Nucleic Acids Res 32(Web Server issue): W321-6.

Sahini, L., A. Tempczyk-Russell, and R. Agarwal. 2010. “Large-scale sequence analysis of hemagglutinin of influenza A virus identifies conserved regions suitable for targeting an anti-viral response.” PLoS One 5(2): e9268.

Salzberg, S. 2008. “The contents of the syringe.” Nature 454(7201): 160-1.

Schmidtke, P., V. Le Guilloux, J. Maupetit, and P. Tuffery. 2010. “fpocket: online tools for protein ensemble pocket detection and tracking.” Nucleic Acids Res 38(Web Server issue): W582-9.

Shen, X., X. Zhang, and S. Liu. 2013. “Novel hemagglutinin-based influenza virus inhibitors.” J Thorac Dis 5 Suppl 2: S149-59.

Shil, P., S. Chavan, and S. Cherian. 2011. “Molecular basis of antigenic drift in Influenza A/H3N2 strains (1968-2007) in the light of antigenantibody interactions.” Bioinformation 6(7): 266-70.

Skehel, J. J. and D. C. Wiley. 2000. “Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin.” Annu Rev Biochem 69: 531-69.

Smith, D. J., A. S. Lapedes, J. C. de Jong, T. M. Bestebroer, G. F. Rimmelzwaan, A. D. Osterhaus, and R. A. Fouchier. 2004. “Mapping the antigenic and genetic evolution of influenza virus.” Science 305(5682): 371-6.

Squires, R. B., J. Noronha, V. Hunt, A. Garcia-Sastre, C. Macken, N. Baumgarth, D. Suarez, B. E. Pickett, Y. Zhang, C. N. Larsen, A. Ramsey, L. Zhou, S. Zaremba, S. Kumar, J. Deitrich, E. Klem, and R. H. Scheuermann. 2012. “Influenza research database: an integrated bioinformatics resource for influenza research and surveillance.” Influenza Other Respir Viruses 6(6): 404-16.

Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar. 2011. “MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods.” Mol Biol Evol 28(10): 2731-9.

Tate, M. D., E. R. Job, Y. M. Deng, V. Gunalan, S. Maurer-Stroh, and P. C. Reading. 2014. “Playing hide and seek: how glycosylation of the influenza virus hemagglutinin can modulate the immune response to infection.” Viruses 6(3): 1294-316.

Underwood, P. A. 1982. “Mapping of antigenic changes in the haemagglutinin of Hong Kong influenza (H3N2) strains using a large panel of monoclonal antibodies.” J Gen Virol 62 (Pt 1): 153-69.

Webster, R. G., W. J. Bean, O. T. Gorman, T. M. Chambers, and Y. Kawaoka. 1992. “Evolution and ecology of influenza A viruses.” Microbiol Rev 56(1): 152-79.

Wiley, D. C. and J. J. Skehel. 1987. “The structure and function of the hemagglutinin membrane glycoprotein of influenza virus.” Annu Rev Biochem 56: 365-94.

Wiley, D. C., I. A. Wilson, and J. J. Skehel. 1981. “Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation.” Nature 289(5796): 373-8.

Xiu, W., Y. Wen, X. Shen, J. Xie, S. Yang, B. Wu, and M. Wang. 2008. “Molecular evolution of influenza A (H3N2) viruses circulated in Fujian Province, China during the 1996-2004 period.” Sci China C Life Sci 51(4): 373-80.

Xu, R., J. C. Krause, R. McBride, J. C. Paulson, J. E. Crowe, Jr., and I. A. Wilson. 2013. “A recurring motif for antibody recognition of the receptor-binding site of influenza hemagglutinin.” Nat Struct Mol Biol 20(3): 363-70.

Zhang, Y. 2008. “I-TASSER server for protein 3D structure prediction.” BMC Bioinformatics 9: 40.

